The Detection of Pedestrian Crossing Behaviour

Master's Thesis

By

Lars Kooijman

TU Delft

The Detection of Pedestrian Crossing Behaviour

Master's Thesis

By

Lars Kooijman

In partial fulfilment of the requirements

to obtain the degree

Master of Science

in *Biomedical Engineering*

at the Delft University of Technology

Student Number:	4502205
Project Duration:	March – September 2018
Thesis Committee:	Dr. ir J.C.F. de Winter (Chair of Committee)
	Dr. ir. R. Happee
	Dr. ir. A. L. Schwab

This work is supported by the NWO VIDI Project 'How should automated vehicles communicate with other road users?' Project number: 016.Vidi.178.047

Dr. ir. J.C.F. de Winter is supported by the NWO VIDI Project 'How should automated vehicles communicate with other road users?' Project number: 016.Vidi.178.047

Dr. ir. R. Happee is supported by the NWO TTW project 'Safe interactions with vulnerable road users' Project number HTSM2015-146667

There's no other species on Earth that does science. It is, so far, entirely a human invention, evolved by natural selection in the cerebral cortex for one simple reason: it works. It is not perfect. It can be misused. It is only a tool. But it is by far the best tool we have, self-correcting, on-going [and] applicable to everything.

-

Carl Sagan – *Cosmos (New York: Ballentine Books 2013), p.352*

Nothing in life is as important as you think it is when you are thinking about it.

-

Daniel Kahneman – *Thinking Fast and Slow* (Penguin Books Ltd. 2012) p. 400

Wie komm ich am besten den Berg hinan?
Steig nur hinauf und denk nicht dran!

-

Friedrich Nietzsche – *Die fröhliche Wissenschaft* (Reclams Universal-Bibliothek, 2000) p.19

There must be discussion, to show how experience is to be interpreted.

-

John Stuart Mill - *On Liberty* (Dover Publications, Inc. 2002) p.17

Elke relativering van het menselijke belang leidt vroeg of laat tot vormen van onmenselijkheid

-

Peter Venmans – *Over de zin van nut* (Uitgeverij Atlas, 2008) p.280

TU Delft

Abstract

The pedestrian is regarded to be one of the most vulnerable road users. Non-verbal communication between drivers and pedestrians seems to play an important role in the mitigation of collisions. The emergence of autonomous vehicles in traffic in the near future presses the need to investigate objective measures related to pedestrian crossing behaviour and the efficacy of communication devices on autonomous vehicles that might replace the nonverbal signals of the human driver. In order to objectively investigate the efficacy of communication devices on autonomous vehicles, 24 participants in this study were immersed in a virtual reality environment, via the use of an Oculus Rift and an Xsens Link motion tracking device. In this virtual reality environment, participants were presented with 18 series of autonomous vehicles. Each series represented one unique combination of independent variables and contained a total of five vehicles. The vehicles were either equipped with a Text display or Frontal Braking Lights that indicated the yielding intentions of the vehicle, or were without any external interface. Furthermore, the inter-vehicular distance between the second and the third vehicle in the series varied between 20, 30 or 40 meters. The participants were instructed to cross the road onto the zebra crossing in the virtual environment when they deemed it was safe to do so. The experiment was designed in such a way that the only crossing opportunity for the participants was between the second and third vehicle when the third vehicles yielded. The road crossing decision of the participants, operationalized by the objective measure of their forward gait velocity, was earlier in time when there was either a Text display or Frontal Braking Lights present on the third vehicle in the series, when the inter-vehicular distance between the second and third vehicle was 20 meters and the third and subsequent vehicles yielded. Congruently, the self-reported ability of participants to predict the behaviour of the oncoming vehicles was significantly better when the third vehicle had a Text display compared to when there was no external interface. However, no significant difference in self-reported ability to predict the behaviour of the oncoming vehicles was found for the Frontal Braking Lights. Furthermore, the forward gait velocity was significantly greater in the presence of a Text display compared to when there was no external interface present for the condition in which the inter-vehicular distance between the second and third vehicle was 30 meters and the third and subsequent vehicles yielded. This work shows that besides the current standard of subjective validation by pedestrians of external human-machine interfaces on autonomous vehicles these interfaces can objectively be validated through the recording and differentiation of body motions.

TU Delft

TU Delft

Acknowledgements

First of all, I'd like to thank Dr. ir. Joost de Winter for his sharp views and critical feedback throughout the various stages of this thesis project, whilst inspiring me to improve my work and providing me the opportunity to explore the fields of science. Also I'd like to thank Dr. ir. Riender Happee and Dr. Dimitra Dodou for their meticulous feedback and insights on the content of this thesis. Furthermore I'd like to thank M.Sc. Oliver Lee for sharing his knowledge on Unity and C#, which helped me in tremendously in the development and reordering of the experimental environment. Moreover, I'd like to thank M.Sc. Jork Stapel and M.Sc. Tugrul Imrak for all their advice and comments regarding the methodology of the research outlined in this thesis. Furthermore, I'd like to thank all the participants who went through the effort of participating in the experiment that is the foundation of this work; without you there would be no data and nothing to analyse.

The effort of creating a thesis relies on not only direct contributions to the work itself, but perhaps more on the indirect contributions to the author. In that line of thought, I'd like to thank my mother, father, sister, brother and sister-in-law for their unconditional love and moral support; you're the reason I've made it this far. All the while my heart goes out to my dear friends Manon Roubos, Renée van Noort, Marijn van den Berg and Anke Zweeris who have been there for each other and me, especially when times got rough. Also, I'd like to thank my dear friend David van den Brink for his friendship, moral support and sharp, witty views during our fruitful discussions regarding various topics of philosophy, politics, economics and technology. Certainly, I'd like to thank M.Sc. Matthijs Mazereeuw for his friendship and for being a great assignment partner in all the courses we took together, which inevitably went hand in hand with of the tons of laughter we had during our years as master students; we made it buddy! Furthermore, I'd like to thank M.Sc. Pavlo Bazilinskyy, whom I've met at the tail-end of his PhD journey and who turned out to be a true '*tovarish.'*

Table of Contents

List of Figures

List of Tables

Introduction

Pedestrians are regarded as one of the most vulnerable road users [1]. In both 2015 and 2016, about 8 per cent of all traffic fatalities in the Netherlands were pedestrians [2], while this was estimated to be 22% worldwide [3]. With the current rate of urbanization about 68% of the world population will be living in urban areas by 2050 [4], where pedestrians closely interact with other road users such as car drivers. Studies on the interactions between pedestrians and drivers have shown that, for example, eye contact between the driver and a pedestrian influences not only pedestrian crossing behaviour [5], but also the yielding behaviour of the driver [6]–[8]. In line to this fact, Rasouli et al [9] aptly stated that: "..*[i]n addition to official rules that govern the flow of traffic, humans often rely on some form of informal rules resulting from non-verbal communication among them and anticipation of the other traffic participants' intentions*." There is a high probability that, in the near future, autonomous vehicles (AVs) coexist with vehicles driven by humans [10]. When AVs are not able to interact with, for example, pedestrians in the informal way a driver would, it could debilitate the safety of the pedestrian. To mitigate the hazards of such a scenario, Nooij et al. [11] point out the necessity of externally visible feedback systems, also known as external Human-Machine Interfaces (eHMIs), on AVs that show the intention of the vehicle to vulnerable road users.

Current research on forms of non-verbal communication of AVs and the acceptance thereof by pedestrians ranges from field experiments that implement the Wizard of Oz (WoZ) method [10], [12] to simulator studies where the pedestrian interacts with AVs in a Virtual Reality (VR) environment [13]–[15]. For example, in the study of Rothenbücher et al. [12], researchers investigated how pedestrians interacted when they were deceivingly confronted with an AV. The vehicle was in fact operated by a human who was covered by a suit that looked like a car seat and thus could not be seen by the uninformed pedestrians. Other than stickers on the car stating it was an AV, no cues, via for example external human-machine interfaces, were provided to the pedestrians about the behaviour or intention of the car. The majority of the people interviewed thought that the car was an AV and were able to cross in front of the car without nonverbal communication from a human driver. However, the authors noted that even though the participants were able to successfully cross, pedestrians in general appreciate the acknowledgement from a driver that indicates they have been noticed when crossing the road. Therefore, it seems worthwhile to investigate the effects of some form of communication between AVs and other road users.

The effects of an AV communicating its intentions to pedestrians via the use of an external Human-Machine Interface (eHMI) has been investigated by Habibovic et al. [10], who also implemented a WoZ-type research methodology. They found that participants in their study appreciated AVs who communicated their intent more than AVs who were devoid of any external communication device. However, the WoZ-type research methodology requires some form of deception and lacks the controllability seen in, for example, lab simulator studies.

In the study of De Clercq [13], the investigation of the effects of eHMI was done by immersing the participants in a Virtual Reality environment (VR) via the use of a Head Mounted Display (HMD). In the VR environment, the participants were confronted with AVs which had a variety of external interfaces on them. The goal of De Clercq was to investigate whether an eHMI mounted on a car, influenced the participants' decision-making regarding his/her crossing intention.

The eHMI indicated whether the oncoming car would yield or not yield to the participants and subsequently the participants indicated, by continuously pressing a button during the approach of the car, that they felt safe enough to cross the road. One of the main findings was that participants, on average, did not feel safe enough to cross anymore when cars, devoid of an eHMI, yielded from 35 meters distance onwards and came within a distance of 20 to 30 meters away from them. The acceptance in the spatial region of 20–30 meters increased significantly when there was an eHMI present on the car that indicated the car its yielding behaviour from the moment of braking. However, as far as known to the author, all of the studies on AVs and eHMIs describe self-reported measures of the pedestrians' acceptance or preference of eHMIs. The question arises whether the subjective measures corroborate with objective measures on the acceptance of eHMIs by pedestrians.

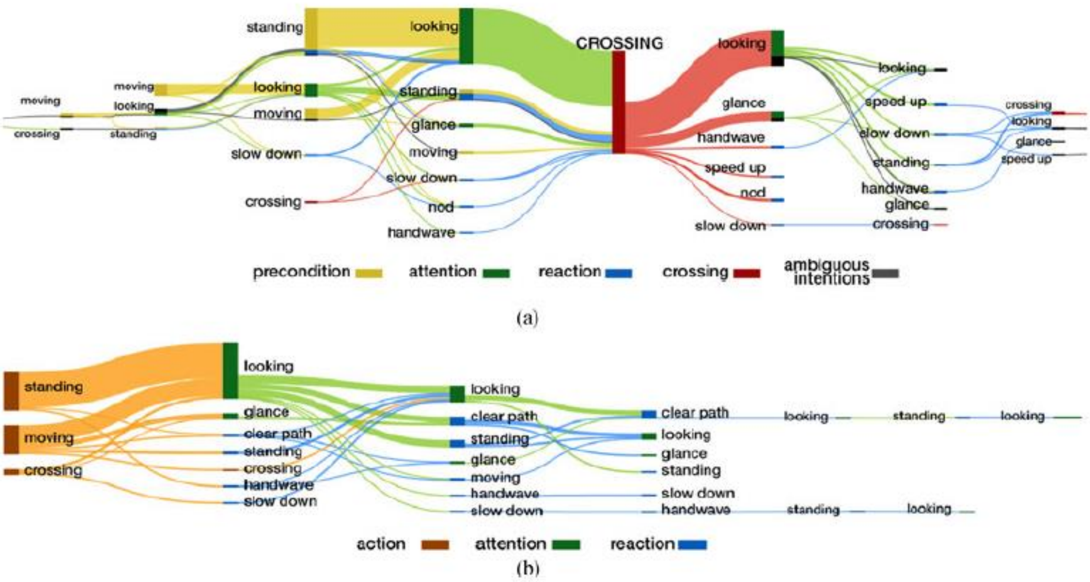


Figure 1: Observed behaviour from pedestrians. The upper diagram (a) shows the series of events before and after a crossing took place. The height of each connecting bar represents its frequency of occurrence. The length of the bars does not represent the duration of each event. A legend at the bottom of the graph defines the group the event belongs to; red and orange represent an action, green represents attention and blue represents reaction. The lower diagram (b) depicts the behavioural patterns which did not lead to a road crossing event. Copied from Rasouli et al. [6].

Objective measures, such as gait speed [16], [17] and head movements [18], resulting from studies on pedestrian crossing behaviour, have shown to be indicators on the urgency [16], [17] or confidence [18] of a pedestrian. Kalantarov et al. [16] and Morrongiello et al. [17] found in their simulator experiments that by asserting time pressure on their participants, the participants would resort to more risk-taking behaviour depicted by the acceptance of smaller spatial gaps between cars and by crossing the road earlier and with a higher average velocity than participants who were not pressed for time. Zito et al. [18] found that young participants showed secure crossing behaviour by alternating head movements between the left and right side of the street compared to older, less secure participants who looked more to the oncoming cars and the road in front of them to assess their crossing decision. Moreover, the extensive literature on affective body language contains numerous examples of objective measures that correspond to the subjective experience of participants. For example, Fawver et al. [19] and Naugle et al. [20] showed that factors of negative valence (i.e., the averseness of something) negatively impacted forward gait initiation. Roether et al. [21] and Crane and Gross [22] provided proof that factors of negative valance influenced average gait speed such that it was lower compared to the average gait speed of participants confronted with factors of positive valance. Gait speed or the initiation of motion could, therefore, be an informative objective measure of the pedestrians' perception of safety or acceptance when safety perception is seen as a factor of valance in the spectrum of human emotions (i.e., aversion when perceiving unsafe situations and approaching when perceiving safe situations).

It is critical to note however, that pedestrian behaviour is contextual, and therefore quite often seems ambiguous, as can be exemplified through the results of the study of Rasouli et al. [6]. The researchers of that study recorded which behavioural patterns (i.e., a sequence of body motions) pedestrians showed when they stood in the vicinity of a zebra crossing and the car containing a hidden video camera approached them. The majority of pedestrians observed in this study stood on the curb, looked at the oncoming car and then crossed. However, as the diagram in Figure 1 shows, this was not the only behavioural sequence shown by pedestrians that preceded a road crossing event, nor was it even singularly associated with road crossing itself. This observed variation in behavioural patterns could be ascribed to a variety of factors influencing the interaction between the pedestrian and driver, for example, the location (e.g., near a zebra crossing or at an unmarked crossing) and the time of interaction (e.g., during day or night-time) or the temporal gap between two vehicles where the pedestrian wishes to cross through. The temporal gap influences the pedestrians' safety margin which effectively is the difference between duration of a pedestrian crossing and the time it takes for the next vehicle to arrive at that crossing point [23]. In that sense, pedestrians could, for example, disregard communicating with the oncoming driver at all when they perceive a sufficiently large enough safety margin for them to cross. Therefore, the objective investigation of pedestrian crossing behaviour, specifically on the acceptance of AVs and eHMIs, requires a meticulous research methodology in a controlled context.

The use of a simulators, instead of conducting a field study, provides investigators with an advantage of greater experimental control [24], a possibility to conduct experiments which are deemed ethically challenging in the real world [25] and the possibility of testing future concepts of interactions with AVs, but at the cost of being less realistic than field studies [24]. Consequently, the benefits of using a simulator to investigate the efficacy of eHMI in the communication between AVs and pedestrians, through objective measures such as body motion, seem to outweigh the costs with respect to an alternative such as the Wizard of Oz experimental paradigm.

Concluding, the acceptance by pedestrians of eHMIs on AVs has subjectively been verified in various studies [10], [13], [26], but there is an absence in current literature of an objective verification of the efficacy of eHMIs in mediating the communication between AVs and pedestrians. An objective verification of the efficacy of eHMIs can help to identify which type of interface is able to facilitate an unambiguous line of communication between AVs and pedestrians and mitigate the future risk of vehicle-pedestrian collisions. We therefore set out to obtain the aforementioned objective verification by finding an answer to the following research question:

"*What is the influence of two different external human-machine interfaces, mounted on autonomous vehicles, and the distance of the vehicles to their predecessor, on the pedestrians' body movement and self-reflection as a measure of their crossing intention?"*

We hypothesized that when pedestrians interacted with an automated vehicle indicating its yielding behaviour through an eHMI they would show a greater acceptance to cross compared to when pedestrians interacted with an automated vehicle without an eHMI. This acceptance would objectively be reflected by:

1. A higher forward gait velocity
2. A higher rate of head movements in the horizontal plane

To investigate the crossing intention of the pedestrians and their hesitation in doing so, we hypothesized that the presence of an eHMI would influence the pedestrians' crossing intention by:

3. A shorter time between the last failed and the first successful gait initiation
4. A lower frequency of gait initiation attempts

The last two hypotheses arose from the interpretation of Fawver et al. [19], Kalantarov et al. [16] and Naugle et al. [20]. We expected the pedestrian to make approaching steps towards the zebra crossing, but not defer to continuously walking during the moments in which they were deciding to cross when interacting with AVs devoid of eHMI.

Complimentary to the objective measures, a higher acceptance of an eHMI would subjectively be reflected by:

5. Lower self-reported feeling of fear
6. Higher self-reported ability to predict the behaviour of the car

Methods

Participants

A group of 24 participants (18 males, 6 females, mean age = 25.4 years, SD = 2.5 years) volunteered to participate in the study. The participants were Bachelor, Master, and PhD students at the Delft University of Technology. The study was approved by the Human Research Ethics Committee of TU Delft, and written informed consent was given by each participant at the start of the experiment. The data for all participants was anonymised and can be downloaded for replication purposes using the link provided in Appendix K.

Materials

A desktop computer with Windows 10 Pro 64 bit (Windows), Unity version 5.5.0f3 Personal 64Bit (Unity) and MVN Analyze 2018.0.3b (*Xsens, Enschede, Twente, NL*) software was used for conducting the experiment. The sample rate used in MVN Analyze was approximately 240 Hz. To capture the motion data of the participant the Xsens Link motion tracking device (*Model, Xsens, Enschede, Twente, NL*), consisting of seventeen accelerometers, a battery pack and a data storage/transmitting device, was used. The transmitting device of Xsens sent its data via an Asus Router to the desktop. The MVN Analyze 2018 software was also used to send the recorded motion data to an avatar in Unity. The Xsens was thus used to drive the head motion in VR and, through the projection of motion onto the avatar, provided the participants with a schematic representation of their own body.

Table 1: The specifications of devices and accessories used in this experiment.

Device	Specification	
Desktop	CPU:	Intel(R) Core™ i7-6700 CPU @3.4 GHz
	RAM:	16 GB Single Channel @ 1064 MHz
	MOBO:	MSI H110M Pro-D (MS-7996)
	GPU:	NVIDIA GeForce GTX 1070 4GB
	Storage:	500 GB Samsung SSD 850 EVO (SATA SSD)
		1 TB Toshiba DT01ACA100 (SATA)
Monitor	Type:	27" Asus VE278h – Black
Router	Type:	Asus RT-AC68U
Oculus Rift	Display:	OLED 1080x1200 per eye @ 90 Hz
Consumer Version	Cables:	4 meters USB 3.0 and HDMI
Extension Cables	HDMI:	1 meter DeLOCK HiSpeed (m/f)
	USB 3.0:	1 meter DeLOCK@ 5 GB/s
Xsens Link	Sensors:	MTX2 – 4A7G6
	BodyPack:	Serial - 00A002E8
	Battery:	Serial - NC2040XS31

An Oculus Rift (*Facebook, Inc., California, USA*) was used to visually and audibly immerse the participant in a VR environment. The way the Xsens was worn along with the Oculus Rift is depicted in Figure 23 in Appendix A. The HDMI and USB cables of the Oculus Rift were extended by 1 meter with a DeLOCK 1.4 HDMI extension cable and a DeLOCK USB 3.0 cable. The specifications of the devices and accessories are summarised in Table 1.

Experimental Environment

The experimental environment for this project was two-fold: one environment in the physical dimension and one Virtual Reality (VR) environment. In the VR environment, participants were situated on a curb near a zebra crossing, and series of vehicles driving on the street approached the participants. A visualization of the VR environment can be found in Figure 2. The VR environment was a rebuilt version of the VR environment made in Unity by De Clercq [13]. Assets used in De Clercq were directly copied, and scripts running various processes in the simulation were adapted, omitted or re-used in their original form. The original scripts associated with the avatar asset provided by Xsens,which can be downloaded free of charge from the Unity Asset Store, were adapted to solve a bug in the rendered rotational motions of the avatar in the VR environment. The positive x-axis of the left-hand coordinate system used in Unity was aligned to the positive x-axis of the right-hand coordinate system used in the MVN Analyze software. The positive x-axes in both programs were parallel the length of the zebra crossing in the VR environment (i.e., the longitudinal direction of the zebra crossing). The orientation of the axes adhered to during the experiment, and during off-line processing of the data, can be found in Figure 3. Furthermore, the VR environment was exported to an executable format (i.e., .exe), which was used during the experiments instead of running the environment directly via the Unity editor, in order to reduce the processing workload on the desktop. The VR environment, including all the scripts and assets needed to run this experiment, can be downloaded using the link provided in Appendix K.

Figure 2: Cars approaching participant in the VR Environment. The starting location of the participant during the experiment is visible here as the grey circle behind the blue leg of the avatar. Cars only approached the participants from their left-hand side. Participants were able to cross the road up until the third zebra stripe counting from the curb.

The physical environment was the GAIT-laboratory at the TU Delft which was about 4.4 by 11 meters in size. The area of the lab used in this experiment was approximately 4.4 by 6 meters. The participants' starting location, which corresponded to the marker location in the VR environment, was near the centre of the room. Participants stood facing a wall which was approximately 6 meters away from them. On the participants' left-hand side there were walls and on their right-hand side windows. Both the windows and the walls were approximately 2 meters away from the participants.

A standing table was about 3 meters at front-right of the participants' starting location. The standing table was used by the experimenter to keep the desktop which ran the software for this experiment. A rope hung suspended from the ceiling, ran in a diagonal fashion above the ground and was tied to a pipe that was in front of the windows. A hook tied to the rope hung about 2.5 meters above the ground and was used to hold the cables from the oculus above the ground in order to prevent the participants from tripping over them. A schematic overview of the physical environment can be found in Figure 3.

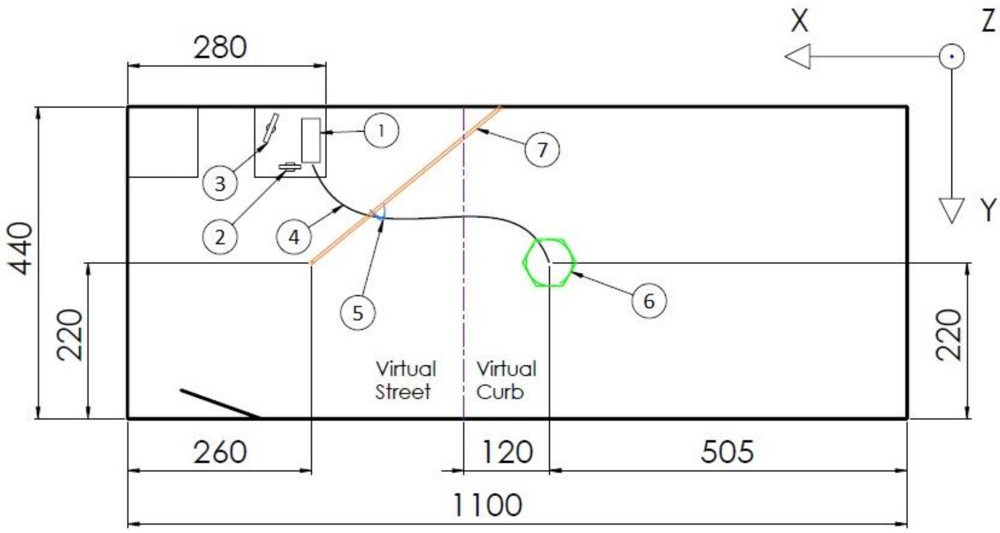


Figure 3: Schematic overview of the lab environment. Location of the participant is indicated with a green hexagon. Participants stood facing towards the positive x-axis. 1 = Desktop, 2 = Asus Router, 3 = Monitor experimenter. 4= Cable Oculus. 5 = Hook. 6 = participant. 7 = Rope. All measures are in cm.

Independent Variables

There were three independent variable categories in this experiment with each having either two or three sublevels. The total amount of independent variables was 3 inter-vehicular distances X 2 types of yielding behaviour X 3 eHMI possibilities = 18 independent variables. The variable categories are summarized in table 2. Each trial in the experiment represented one condition which consisted out of a unique combination of the independent variables. Therefore, a total of 18 trials were conducted, and all combinations were tested. The order in which the conditions were presented to the participants was random. Each of the following three paragraphs describes each category of independent variables in detail.

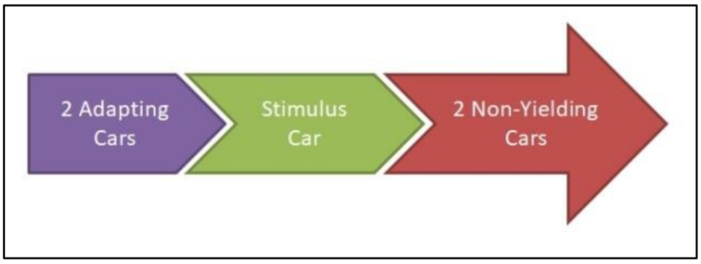


Figure 4: The order of vehicles in each train as presented to participants during one trial.

In each trial, a train of five cars would turn the corner about 90 meters away from the participant's left hand-side perspective in the VR environment. There was no traffic coming from the participants' right-hand side, and there were no drivers present in any of the cars. The first two cars in the train never yielded. The third car in the train represented the stimulus vehicle onto which the independent variables were applicable. Depending on the composition of the independent variables, the last two cars in the train would adapt their behaviour to that of the stimulus vehicle. A depiction of the order of cars presented to the participants can be found in Figure 4. The cars approached the participant with a speed of 50 km/h. When the third car yielded, it would start braking at 35 meters away from the zebra crossing with a deceleration rate of 3.5 m/s^2 and come to a standstill right before the zebra crossing. Consequently, the fourth and the fifth car would also be yielding at 40 and 50 meters from the zebra crossing and would come to a standstill behind the third car and a few meters separated from one another. The third car would slowly pick up speed after standing still for 5 s. However, when the third car did not yield, all the cars in the train would pass the participants and disappear by turning left around the corner on the participant's right-hand-side perspective.

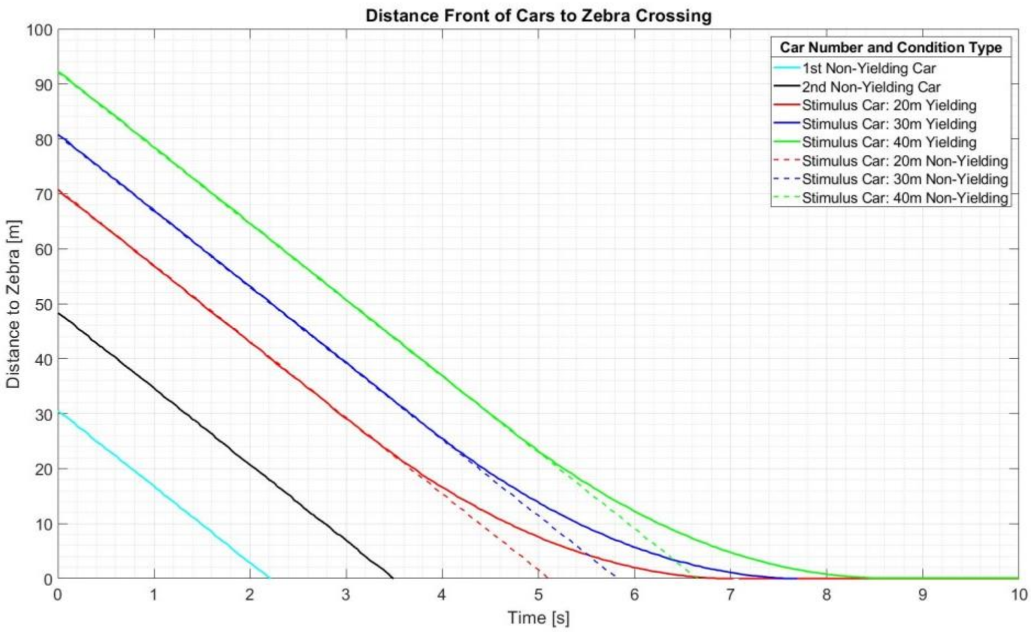


Figure 5: Distances of the fronts of cars to zebra crossing during a trial. T = 0 is when the first non-yielding vehicle in the train of cars was approximately 30 meters away from the participant. The inter-vehicular distance remained constant for the conditions in which the stimulus car did not yield, but changed in the conditions when the stimulus car did start braking.

The distance between the front of the first and front of the second car was set at 18 meters. The same distance was used between the third (i.e., the stimulus vehicle) and the fourth vehicle (i.e., the first adapting vehicle) as well as between the fourth vehicle and the fifth vehicle (i.e., the first adapting vehicle and the second adapting vehicle). However, the distance between the front of the stimulus vehicle and the back of its predecessor (i.e., the second vehicle) varied between 20, 30 and 40 meters. Therefore, the participants were only presented with a potential crossing opportunity between the second non-yielding vehicle and the stimulus car or they had to wait until all the cars had passed. For example, in the condition where the stimulus vehicle did not yield and had a 40 meter inter-vehicular distance to its predecessor, there was a temporal gap of about 2.88 seconds. This temporal gap could be perceived as sufficiently large enough for participants when they started crossing directly after the second vehicle in the train passed in front of them. A visualization of distances to the zebra crossing as a function of elapsed time for the first three cars can be found in Figure 5.

Lastly, the types of eHMI used in this experiment were the Frontal Braking Lights and Text eHMI copied from the experiment of De Clercq [13]. These two eHMIs were chosen since the participants in the study of De Clercq considered these to be the least (i.e., the Text eHMI) and the most (i.e., Frontal Braking Lights) ambiguous of all of the eHMIs presented to them. De Clercq determined the ambiguity of the eHMIs via a post-hoc questionnaire in which images were shown of eHMIs during yielding and non-yielding conditions. In this questionnaire participants indicated for each depiction of the eHMIs whether they felt safe to cross.

The conditions in which no eHMI was present are referred to as baseline conditions. The presence of eHMIs varied randomly among the four of the cars in each train, meaning that there was not a homogeneous presentation of eHMIs each train. Only the presence of an eHMI on the stimulus car was controlled so that it was possible to investigate what its effect was on the participant since this was the first car in the train that would indicate its yielding behaviour. The external appearance of the cars used in this experiment can be seen in Figure 6.

Figure 6: External appearance of the cars used in this simulation. From left to right, top to bottom: Truck Text eHMI Yielding, Smart Text eHMI Yielding, Truck Frontal Braking Lights Yielding, Smart Frontal Braking Lights Yielding, Truck No eHMI Yielding, Smart No eHMI Yielding, Truck Text eHMI Non-Yielding, Smart Text eHMI Non-Yielding, Truck Frontal Braking Lights Non-Yielding, Smart Frontal Braking Lights Non-Yielding, Truck No eHMI Non-Yielding, Smart No eHMI Non-Yielding

The type of car (i.e., Ford f150 or Smart Fortwo) was also randomized separately from the independent variables throughout the experiment leading to mixed platoons. In resemblance with the experiment of De Clercq, the Smart Fortwo's had their interior visible while the Ford f150's did not. During the experiment, the visuals of the eHMI and/or the pitch of the cars provided cues to the participants as to whether or not the car was yielding. The text eHMI would change in appearance, once the car was yielding, from stating “Don’t Walk” to “Walk.” The Frontal Braking Lights would start emitting a bright green colour from the location of the front lights once the car was yielding. A summary of the conditions (i.e., independent variables), which were presented to the participants, can be found in Table 2 and links to YouTube videos, showing examples of pedestrians crossing in this experiment and an overview of the eHMIs presented to participants, can be found in Appendix K.

The conditions of this experiment are abbreviated by means of the variance of independent variables in that condition in the following order: inter-vehicular distance – yielding behaviour – eHMI type. An example of such an abbreviation would be ‘20MY-None’ which would indicate that the stimulus vehicle in that condition had 20 meters distance to its predecessor, yielded and had no eHMI.

Procedure

A timeline of the experiment can be found in Figure 8. Furthermore, a more detailed overview can be found in Appendix B, which contains all the steps and task instructions of this experiment.

Upon arrival, participants were briefed in short about the experimental set-up through the use of the informed consent. Participants were instructed about their goal in this experiment, namely to cross the road as safe as possible. Furthermore, the participants were instructed about the behaviour and the possible variability in behaviour of each car in the train. An emphasis was made on the fact that crossing the road before the first two cars in the train had passed was not allowed. Also, the participants were informed about the content of the three questions/statements that would be asked after each trial; the first answer being their subjective MISC rating, the second a numeral response reflecting their feeling of fear and the third a numeral response reflecting their ability to predict the behaviour of the oncoming cars. Lastly, the participants were told that they could stop the experiment at any time or that the experiment would be paused or terminated if the participant indicated a MISC rating of 4 or higher. The details of each integer on the MISC scale that was used in this experiment, and explained to the participant, are summarized in table 3.

Table 2: Independent Variables. Three categories of independent variables were used in this experiment and each variation of the category is stated in the column next to it.

Independent Variables	
Yielding Behaviour	Yielding at 35 meter, Not yielding
Inter-vehicular distance	20, 30 and 40 meter
Type of eHMI	None, Frontal Braking Lights and Text

After the briefing, the participants signed the consent form, which can be found in Appendix A, and were asked to complete a questionnaire containing statements about pedestrians and motorists. These statements were copied from the questionnaire about pedestrian crossing behaviour of Papadimitriou et al. [27]. This questionnaire was used to investigate the way pedestrians perceived the behaviour of motorists and perceived the way motorists interact with pedestrians in general. Section E of the original questionnaire was omitted, to reduce the chance of behavioural priming (i.e., that the participants adopted a certain style of walking), since it contained statements about various tactics of road crossing (e.g., I cross diagonally, I cross even though there are oncoming vehicles). The adapted questionnaire used in this experiment can be found in Appendix C.

Table 3: MISC scale and the associated experiences. If a participant indicated a 4 or higher during the experiment, the experiment would be paused and eventually terminated if the participant indicated wanting so.

MISC Rating	Experienced Sensation
1	No problems
2	Slight discomfort but no specific symptoms
3	Vague dizziness, warm sensation, headache, stomach awareness, sweating
4	Some dizziness, warm sensation, headache, stomach awareness, sweating
5 and upwards	Medium to severe dizziness, slight to severe nausea

Once finished with the questionnaire, anthropometrics of the participants' arms, ankles, hip, knees, feet, and shoulders were taken. After these measurements, participants were familiarized with the use of the Xsens Link. The Xsens Link motion tracking sensors were then attached to the participant at the locations depicted in Figure 7. The sensors were either incorporated in an Xsens stretch shirt or placed on the participant using Velcro straps supplied by Xsens. Anthropometrics were entered into MVN Analyze, and the program was used to calibrate the Xsens Link using the program its N-Pose function. After calibration, the participants were familiarized with the use of the Oculus Rift Virtual Reality goggles.

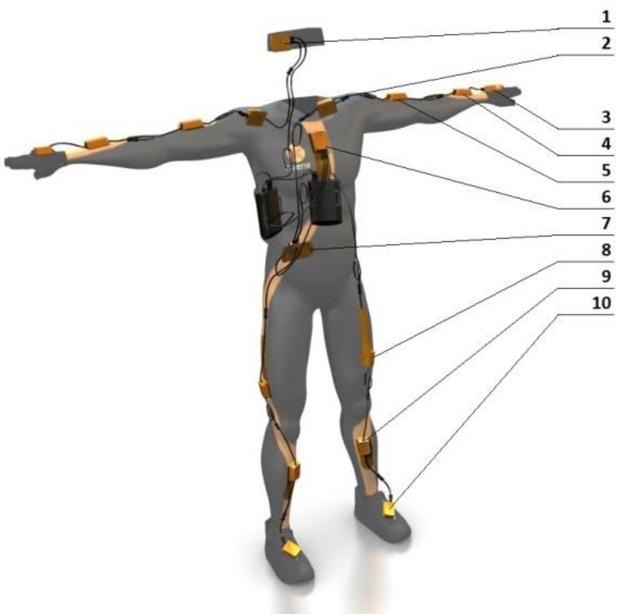


Figure 7: Sensor locations of Xsens Link. 1= Head, 2= Shoulder left/right, 3= Hands left/right, 4= Forearm left/right, 5= Upper arm left/right, 6= Sternum, 7= Pelvis, 8= Upper leg left/right, 9= Lower leg left/right, 10= Instep of foot left/right. The image was amended from the original version which was downloaded from https://www.xsens.com/wp-content/uploads/2013/04/unity_pr-1024x782.jpg.

Using the Oculus Rift, the participants were immersed in a city-like environment in Unity, which in Unity was called the 'Main Scene.' While the participants accustomed themselves to the environment by moving around on the curb, they were reminded that during the experiment, trains of cars would pass by the sidewalk on which the participants now stood and that the first two cars of each train would never yield. However, the third until the fifth car would provide the participants with cues regarding their yielding behaviour. It would be up to the participants to assess the situation and determine whether it was safe for him/her to cross the road. When they felt safe enough to cross, they could do so by walking from the sidewalk towards the curb and onto the zebra crossing, but no further than the third white zebra line. The restriction of moving up until the third white zebra line was applied due to spatial limitations of the lab where the experiment was conducted. The spatial dimensions of the lab can be found in Figure 3. If the participant stood on the third white zebra line, there would be about 1.5 meters of free space in front of the participant and to the participant's left hand-side perspective until the walls of the lab would be reached. To the participant's right-hand-side perspective there was about 1 meter of free space until the table was reached on which the experimenter ran the software for the experiment. The participants were provided with the opportunity to explore the spatial limitations of the lab under the guidance of the experimenter to reduce the fear of collision with any walls while residing in the VR environment. Not all of the participants indicated the need to explore the limitations of the lab. No cars were present during the exploration in the Main Scene.

Figure 8: Timeline of the experiment

Once the participants indicated that they were comfortable walking around in the VR environment, they were requested to step onto the circular marker in the VR environment. The following scene was loaded in Unity and was called the 'Test Scene.' In order to accustom the participants to the presence of cars in the environment, trains of non-yielding cars were created, which passed by the sidewalk on which the participants were standing. During this Test Scene participants were not allowed to try out any form of road crossing. Each train contained five cars, like the train of cars in the real experiment, and all the cars in the Test Scene were devoid of any eHMi.

The participants were once again reminded that their goal during the experiment was to start crossing the road once they felt it was safe to do so. When the participants indicated that they understood the course and goal of the experiment, the VR environment was reset, and the experimental trials were initiated via the use of the Trial Scene.

For each participant, a total of 18 trials were conducted in every experiment. After each trial, the experimenter noted down the participants' answers to the statements regarding the participants' motion sickness, feeling of fear and ability to predict the car behaviour on an answer sheet. This answer sheet can be found in Appendix D. During most of the experiments after a few trials, the questions and statements were eventually substituted by the questions; "*Level of motion sickness*?", "*Level of fear*?", "*Predictability*?" or ruled out altogether because the participants themselves would call out the numbers. The participants had the opportunity to have a break after each block of five trials or whenever he/she indicated needing one. Once all 18 trials had been completed, the oculus and the sensors were removed from the participants, and the participants were asked to fill in a final questionnaire. This questionnaire contained questions about the VR experience and can be found in Appendix E. The questions were selected for relevance (e.g., no haptic feedback was present in this experiment; therefore related questions were not incorporated) from the questionnaire developed by Witmer et al [28].

Dependent Variables

The hypothesized changes in bodily movements related to crossing intention were operationalized by mean of four dependent objective measures and two dependent subjective measures. The objective measures were defined to be the participants':

1. Forward gait velocity
2. Absolute angular velocities of the head
3. Time between the last failed and the first successful gait initiation
4. Frequency of gait initiation attempts

The forward gait velocity of the pedestrians was measured as the velocity of the centre of mass (C.o.M.) of the pedestrian, in world coordinates, where the forward direction is in the X-direction. Using the Xsens Link, the pelvis sensor was the closest approximation to the C.o.M., and therefore those velocities were used.

We obtained angular position and velocity data of the head using the head sensor of the Xsens Link to determine the absolute angular velocity of the head in the horizontal plane in world coordinates. Due to the physiological and physical limitations of our body, there is a limit to which frequency band, and correspondingly rotational velocities, can be ascribed to human motion. Grossman et al. [29] found that the rotational velocities of the head in the horizontal plane during walking did not exceed 90 deg/sec (i.e., 0.25 Hz) and remained below 170 deg/sec (i.e., 0.45 Hz) during running. We expected the head rotational velocities to remain below the 380 deg/sec.

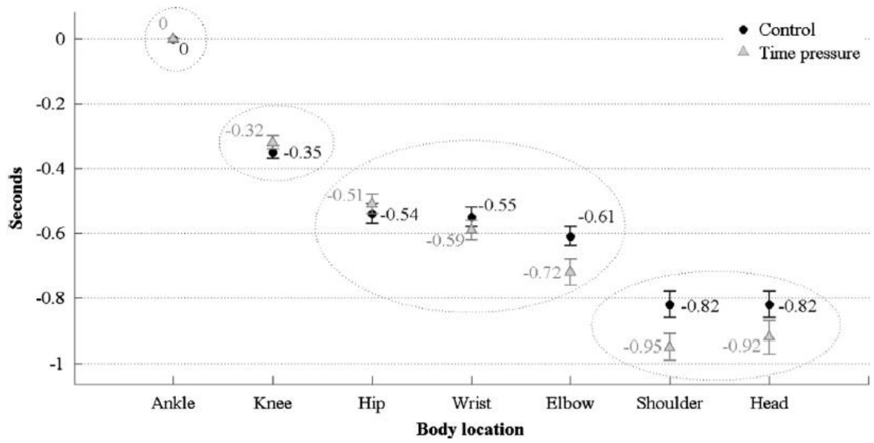


Figure 9: Mean and Standard Error of the time in which the forward movement was recognized for various body parts These results are from the study of Kalantarov et al. [16] **and depict the timing of the sequential motion of body parts during the initiation of gait. T= 0 is defined as the start of continuous gait. From the graph is can be derived that the start of gait is preceded by the sequential motion of the head, shoulder, elbow, wrist, hip knee and lastly the ankle of the stance leg.**

A gait initiation attempt was defined to be the fully sequential movement of four body-part groups (i.e., group 1: head and shoulders, group 2: elbow, wrist and hip, group 3: knee, and group 4: ankles), as defined by Kalantarov et al. [16], not directly followed by continuous bipedal locomotion within a time-span of 0.6 seconds. This time-span was inferred from the findings of Kalantarov et al. [16] where they showed that a complete gait initiation took about 0.80–-0.95 seconds from the moment the head moved forward until the forward motion of ankle of the stance leg as can be seen in Figure 9. Furthermore, they measured that the time between the onset of ankle motion and hip motion was approximately 0.55 seconds. We therefore assumed that a failed initiation would be quantifiable as singular motion event of the hip and ankles, singled out through a temporal gap larger than 0.6 seconds until the next motion event.

The two dependent subjective measures incorporated in this experiment were the participants':

5. Self-reported feeling of fear
6. Self-reported ability to predict the behaviour of the car

To assess these measures we would ask each participant after each trial to reply, by stating a number, to the following two statements: '*When you were considering crossing the road, you experienced a feeling of fear*' and '*it was difficult for you to predict the behaviour of the oncoming cars.*' Each of the two statements would be rated with an integer number by the participant, which would range from 1 to 10 where answering 1 would represent the answer *'strongly disagree'* and answering 10 would represent the answer '*strongly agree.*'

Data Reduction

The collected data were processed off-line using custom written Matlab scripts (The MathWorks inc. R2018a, Natick, MA, USA), which can be downloaded using the link provided in Appendix K. A thorough description of how the data was processed can be found in Appendix E. In short, the motion data recorded via Unity were interpolated to match the sampling frequency of the data recorded in MVN Analyze. Furthermore, the two sets of data were cross-correlated to find the delay between the signals and were compensated for the delay. All the MVN data were then low-pass filtered using a 10th order zero phase Butterworth filter with a cut-off frequency of 8 Hz except for the angular velocity data of the head which was filtered with the same filter but with a cut-off frequency of 3 Hz. The absolute velocity was adopted representing the magnitude of head rotation activity in a certain period, rather than the specific motion pattern. Hence the 3Hz filter on head rotation velocity was applied after taking the absolute value. The Unity data were filtered using a Moving Average filter with a cut-off frequency of 8 Hz. The data were then sorted by stimulus type.

Statistical Methods

Grouping of Conditions

To assess whether an eHMI affected the crossing intention of a pedestrian, the data were clustered by conditions with equal inter-vehicular distance and yielding behaviour for statistical and qualitative analysis (e.g., 20MY-None compared to 20MY-Text, 20MY-None compared to 20MY-FBL).

Quantitative Measures

The participants Center of Mass (C.o.M.) velocities, measured along the x-axis[1], were used to quantify the forward gait velocity. To assess whether and when significant differences occurred when an eHMI was present on the car compared to when no eHMI was present, a paired sample t-test was conducted between conditions with similar inter-vehicular distances and yielding behaviour of the cars for each time sample. The effects of inter-vehicular distance will be described qualitatively.

Angular velocities of the head were obtained by taking the absolute value of the angular velocities relative to the x-axis of the lab, which were measured using the Xsens. The resulting data were used as a quantitative measure for the rate of head movements in the horizontal plane. Similar to the C.o.M. velocities, paired sample t-test was conducted for each time sample.

Furthermore, a paired sample t-test was conducted on two different pairs of white noise signals to simulate the analysis of the C.o.M. and the rate of head movements. This was done to ensure that the results found through the method of conducting a paired sample t-test for each time sample were not statistically significant due to random chance.

Gait initiation attempts and the onset of walking were determined from the position of both ankles. To find the onset of walking the function 'ischange' in Matlab was utilized to find abrupt changes in the mean of the data. The timing of the attempts and the start of walking was grouped over time bins of 0.5 seconds (i.e., averages from 0.25–0.75, 0.75–1.25 and so on). The size of the time bins was estimated to be sufficiently large to contain motion data (e.g., the onset of walking) for every participant while at the same time to be small enough to contain a single attempt per participant, based on the findings detailed in the introduction. Paired sample t-tests were conducted to determine whether significant differences in the number and timing of the initiation attempts occurred at different time bins.

Lastly, to assess the subjective influence of the eHMI's, the responses to the two statements regarding experienced fear and perceived predictability of on-coming car behaviour were also compared among the conditions with an equal inter-vehicular distance and yielding behaviour. This comparison was done by means of a paired sample t-test.

The alpha level for statistical significance (i.e., α) was set at 0.005, which can be motivated by the goal to reduce the possibility of false positive findings (i.e., type I error) [30]. If we assume, for a two-tailed test for differences between two dependent means, a large effect size (i.e., d ≈ 0.8) for our sample of N = 24 with the conventionally used α of 0.05, the power of this study (i.e., 1-β) would approximately be ± 0.96.

$$False\ Positive\ Rating \approx \frac{\alpha\phi}{\alpha\phi+(1-\beta)(1-\phi)} \tag{1}$$

Using Equation 1, obtained from Benjamin et al. [30], for false positive rates and estimating the prior odds of 1:10 in favour of the null hypothesis (i.e., ϕ = 0.9), the rate of false positive findings for this study would approximately be 0.32. Lowering the significance level to α = 0.005 for an equal sample size with similar effect size and prior odds assumptions would result in a lower but acceptable power of ± 0.78 while reducing the rate of false positive findings to ± 0.05.

[1] See the section 'Experimental Environment' of this chapter for the definition of the x-axis.

Qualitative Measures

Responses to the opening questionnaire were compared to the responses obtained in the study of Papadimitriou et al. [27] in a qualitative manner, and therefore no statistical test was involved. In order to assess learning and possible changes in subjective fear, linear regressions were performed to mean responses along the trails of the experiment of which the first two and last two mean responses were compared using a paired sample t-test for significant differences.

The responses for each question of the Virtual Reality Presence Questionnaire (VRPQ) were factorized, following Witmer et al [28], along four factors, namely; Involvement, Sensor Fidelity, Adaptation/Immersion and Interface Quality. The mean response of each question in the group were summed and compared to the summed mean responses obtained in the study of De Clercq [13], for the same factor, using a two-sample t-test.

Results

The condition 40MY-None of this experiment was presented to each participant as the condition 30MY-None due to an error in the 'CarSpawner.cs' script. Subsequently, no comparison could be made to the conditions 40MY-Text and 40MY-FBL. The data of condition 40MY-None has not been used in the analysis. Furthermore, of the total of 216 trials in which the cars did not yield it was not possible to correlate the data recorded in MVN Analyze to the data recorded in Unity for 23 trials. The amount of the participants that were included in the analysis of those conditions therefore varied and has been emphasized in figure legends where applicable.

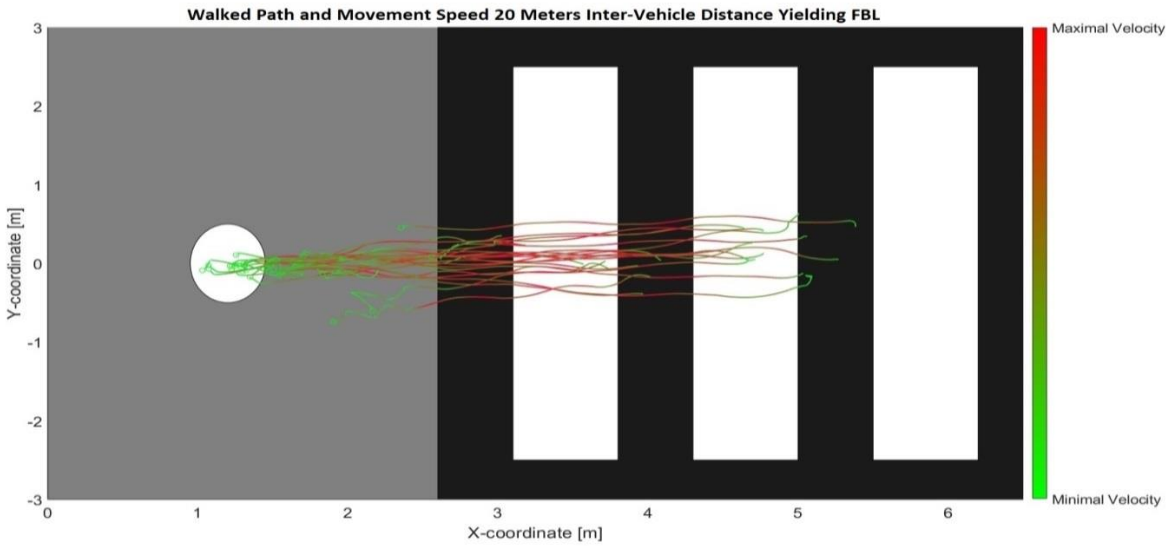


Figure 10: Walked Path and Movement Speed. Abbreviation: FBL = Frontal Braking Lights. Movement speed is visualized by means of a varying colour patterns ranging from green for the lowest velocities and red for the highest velocities. Participants were instructed that in each trial they had to start walking from the circular marker up until the third zebra stripe. A few of the participants moved closer to the curb over the course of the experiment, hence the variation in starting location.

Overall, participants were able to comply with the instructions and successfully finish the experiment. However, we observed some variations of the participant starting positions, which have been depicted in Figure 10. This figure also depicts the variation of movement speed in relation to the participants' location during the experiment. Furthermore, it can be seen that, in accordance with the instructions given, no participant moved further than the third zebra stripe. The figure only depicts the motion trajectories of the participants from the moment they started walking away from the starting location until the moment they stopped walking and began moving from the zebra back towards the starting location. Only this segment of the participants' motion data was used for the specified quantitative measures.

Centre of Mass Velocities

For all the conditions in which the cars did not yield, no significant differences in C.o.M. velocities of the participants were found when comparing the presence of an eHMI to the absence of one. However, for the condition where the cars did yield and the stimulus car had an inter-vehicular distance of 20 meter, significant differences were found in C.o.M. velocities of the participants during the pre-defined temporal region of interest when comparing between the presence of a text eHMI and no eHMI as well as the presence of Frontal Braking Lights and no eHMI. In both cases, the C.o.M. velocities of the participants were greater when there was an eHMI present on the stimulus car compared to when there was no eHMI present. The change in C.o.M. position over time can be found in Figure 11 while the change in C.o.M. velocity and the change in p-value over time have been visualized in Figure 12 for the conditions with the independent variable of 20M for both yielding and non-yielding conditions. The upper graph depicts the mean C.o.M. velocities of the participants for the variations in the presence of eHMIs for both yielding and non-yielding conditions and the bottom graph the varying p-value over time of the comparison between the presence of an eHMI and the absence of one.

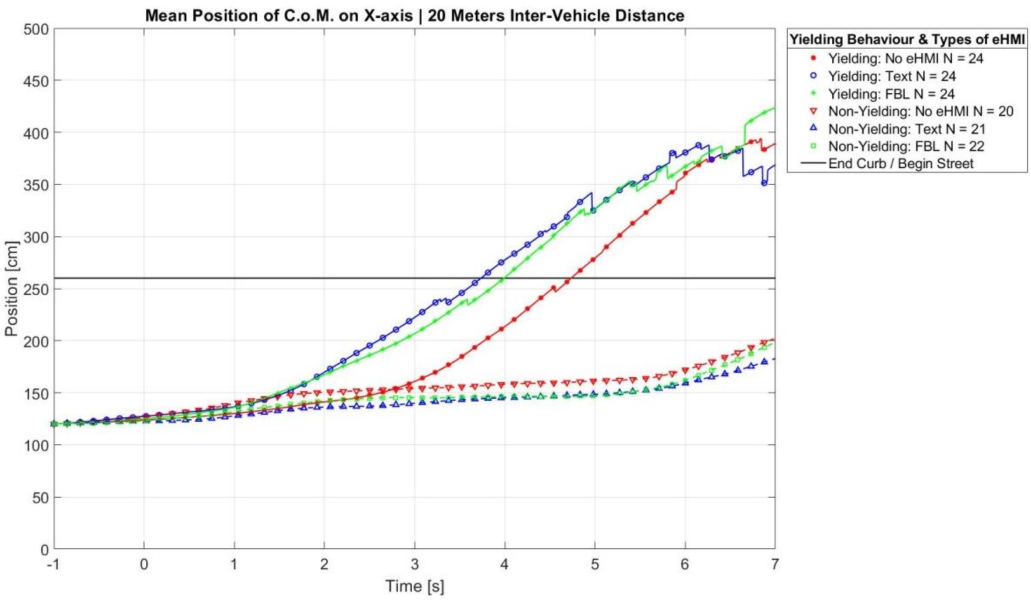


Figure 11: Mean C.o.M. positions of the participants on x-axis over time for the conditions with independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. Participants were faster in moving towards the zebra during the conditions where an eHMI was present on the stimulus car and the cars yielded than when there was no eHMI present on the stimulus car and the cars yielded.

For the condition where the cars did yield, and the stimulus car had an inter-vehicular distance of 30 meters, there were significant differences in the C.o.M. velocities of the participants when there was a text eHMI present compared to when there was no eHMI. Interestingly, earlier in the temporal region of interest the C.o.M. velocities were smaller in the presence of a text eHMI as compared to no eHMI, if only for a very brief amount of time (i.e., 0.15 seconds), but later on this was the other way around as can be seen in the upper graph in Figure 13.

No significant differences were found for the condition with the independent variable of 40 meters inter-vehicular distance in which the cars yielded due to the absence of a baseline to which we could compare the presence of an eHMI. Furthermore, no significant differences were found for the yielding condition with the independent variable of 40 meters inter-vehicular distance. An overview of the time periods of significant difference with the corresponding mean C.o.M. velocity values can be found in Table 4. The graph for the non-yielding condition with the independent variable of 40 meters inter-vehicular distance can be found in Appendix H.

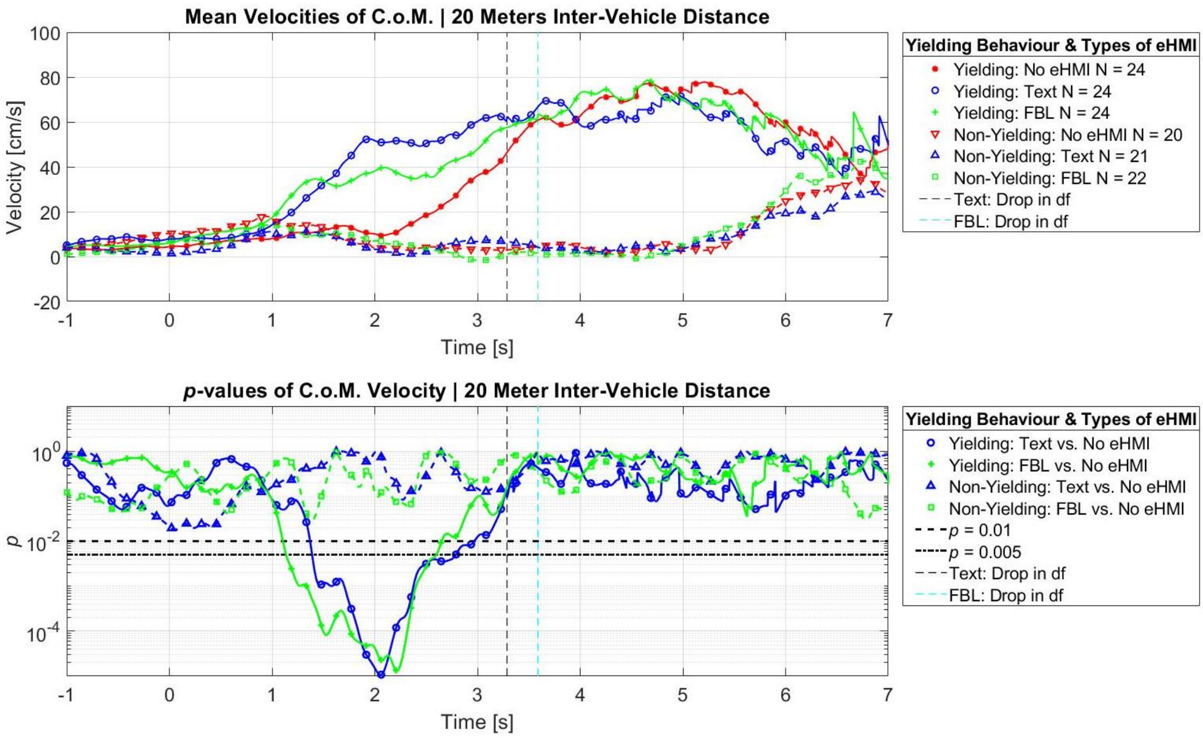
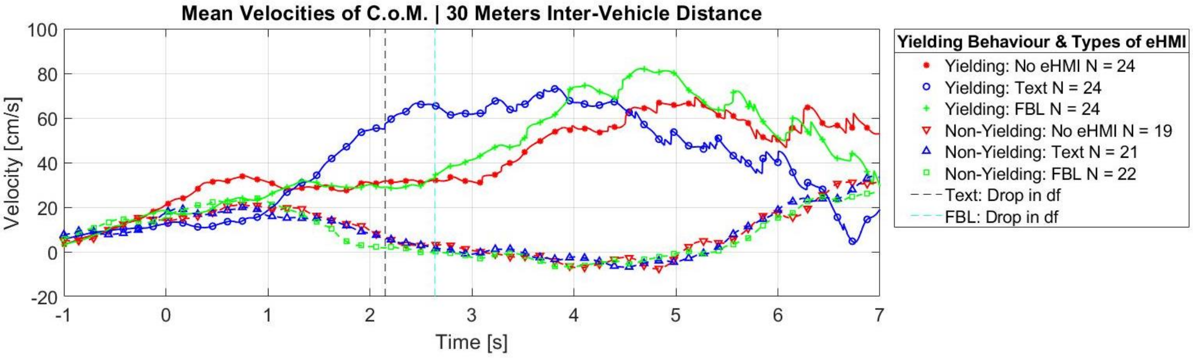


Figure 12: Mean C.o.M. forward velocities of the participants and corresponding p-values over time for the conditions with independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: Greater mean C.o.M. velocities were found between the 1st and the 3rd second from when a stimulus car with an eHMI started braking. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the C.o.M. velocities of the participants differed significantly in the presence compared to the absence of an eHMI.



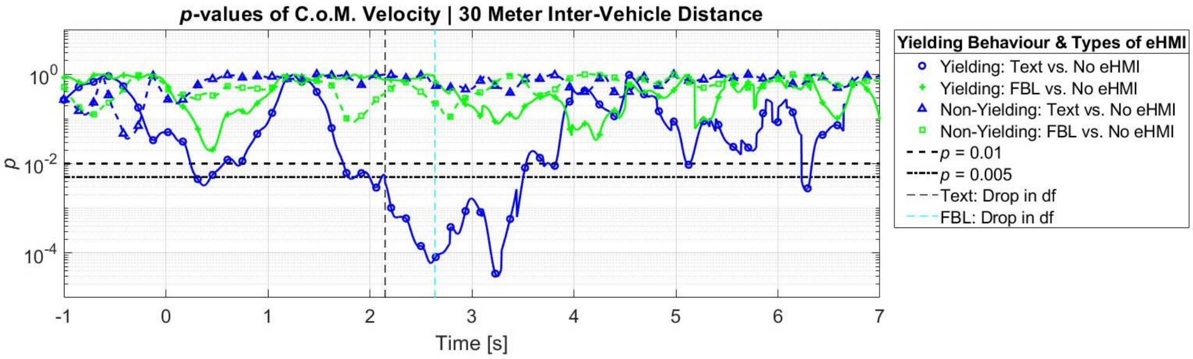


Figure 13: Mean C.o.M. forward velocities of the participants and corresponding p-values over time for the conditions with the independent variable 30 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: Between the moment the stimulus car started braking and the 1st second after that, the C.o.M. velocities in the absence of an eHMI were shortly significantly greater than the C.o.M. velocity in the presence of a text eHMI. However, C.o.M. velocities in the presence of a text eHMI were shown to be significantly greater after the 2nd second. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference, and the intersections of the boundaries with the varying p-values were used to estimate the moment when the C.o.M. velocities of the participants differed significantly.

Table 4: Centre of Mass velocities. Results of paired sample t-tests over time between conditions where the stimulus car had an eHMI compared to no eHMI. The table depicts the time ranges wherein the p-value resulting from the comparison of eHMI to Baseline was lower than p = 0.005. For every time-sample of the conditions compared, a paired sample t-test was conducted to investigate at which moment significant differences would occur and what the corresponding C.o.M. velocities were. The time ranges wherein significant differences occurred are indicated by their respective start and end time including the mean and standard deviation of the centre of mass velocities at those moments in time for both the condition in which an eHMI was present as well as when it was absent.

			Comparison eHMI to Baseline				
			eHMI		None		
Independent Variables	Time [s]		M ($^{cm}/_{s}$)	SD ($^{cm}/_{s}$)	M ($^{cm}/_{s}$)	SD ($^{cm}/_{s}$)	df
20MY - FBL	Start	2.14	24.67	4.75	9.20	3.51	23
	End	3.60	36.57	5.25	20.51	4.37	23
20MY - Text	Start	2.40	28.32	4.52	10.70	3.79	23
	End	3.79	54.12	5.75	27.14	4.81	23
30MY - Text	Start	1.30	11.26	4.03	27.76	4.92	23
	End	1.45	11.28	4.01	31.18	5.36	23
	Start	2.80	47.81	5.18	27.64	5.53	23
	End	2.86	50.76	5.39	28.87	5.61	23
	Start	3.00	54.59	5.79	31.08	5.71	23
	End	3.12	55.16	5.45	31.28	5.67	23
	Start	3.15	57.54	5.14	31.75	5.71	22
	End	4.50	66.00	4.43	43.57	4.50	18
	Start	7.24	28.98	2.11	57.36	3.25	3
	End	7.32	30.58	1.60	64.83	2.91	2

Abbreviations: T = Text eHMI, FBL = Frontal Braking Lights, df = Degrees of Freedom.

Angular Velocities of the Head

Similar to the analysis of the C.o.M. velocities, for the condition where the cars did yield, and the stimulus car had an inter-vehicular distance of 20 meters, significant differences were found in the angular velocities of the participants' heads when comparing between the presence of a text eHMI and no eHMI. The angular velocities of the participants' heads were greater during multiple time events when there was an eHMI present on the stimulus car compared to when there was no eHMI present. Furthermore, for the condition where the cars did yield, and the stimulus car had 30 meters inter-vehicular distance and a text eHMI, a single time event was found where the participants' head angular velocities were significantly greater compared to when there was no eHMI present on the stimulus car. The mean change in angle of the participants' heads over time can be found in Figure 14, while the mean change in angular velocity of the participants' heads and the corresponding change in p-value over time have been visualized in Figure 15 for the conditions with the independent variables 20MY. The corresponding temporal start and end of the significant differences, including the mean value and standard deviation for both of the conditions that were compared, are summarized in table 5. The graphs of the mean angular velocities of the head over time, including the corresponding p-values over time for all conditions can be found in Appendix I.

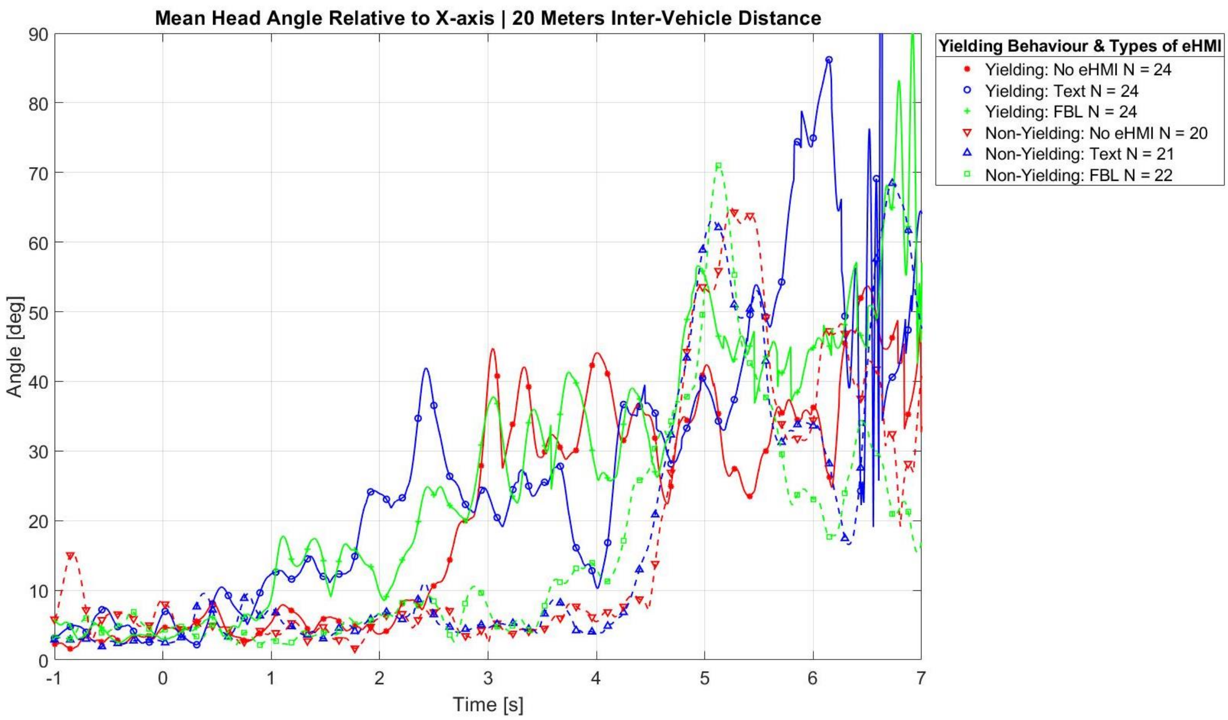


Figure 14: Mean head angle relative to x-axis of the participants over time for the conditions with independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. Yielding conditions: Between the first and the third second after yielding, participants presented with a textual eHMI showed a greater deviation in head angle relative to the x-axis compared to participants presented with a stimulus car devoid of an eHMI. Non-yielding conditions:

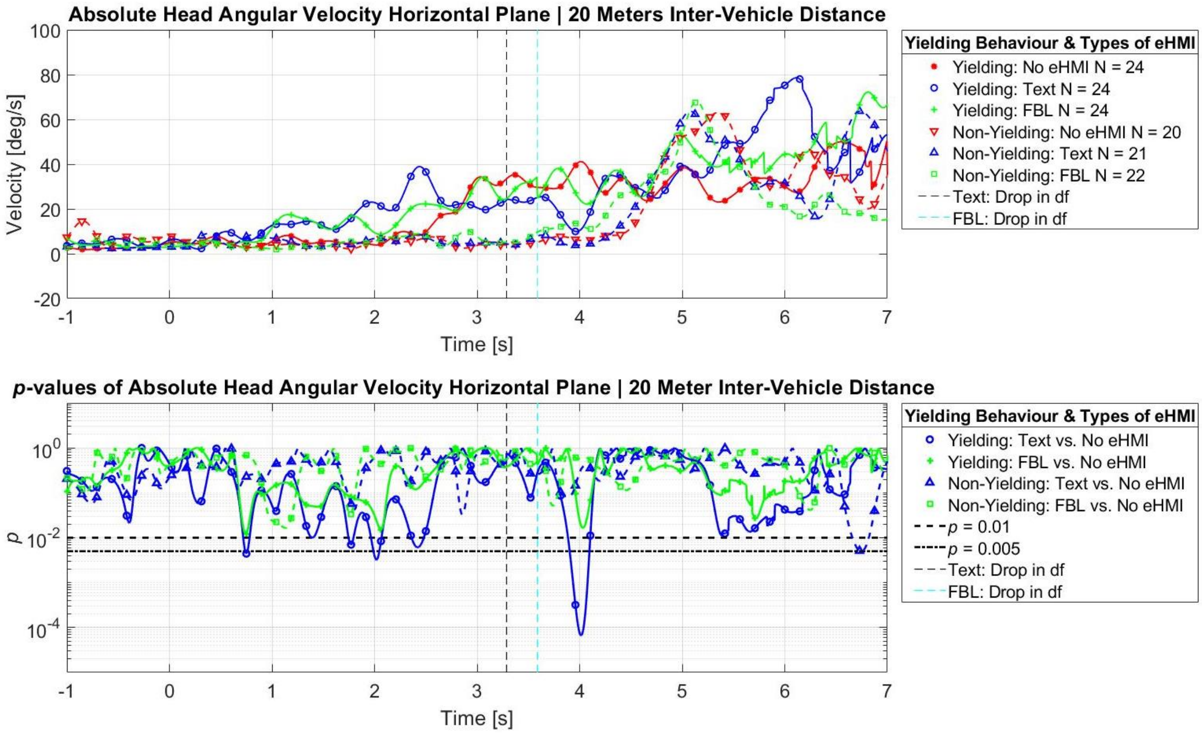


Figure 15: Absolute angular velocities of the participants and corresponding p-values over time for the conditions with the independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: Greater absolute mean angular velocities were found for multiple temporal events from the moment the stimulus vehicles with eHMI started braking to the 7th second thereafter. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the absolute mean angular velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.

Table 5: Absolute angular velocities of the head. The table depicts the time ranges wherein the p-value resulting from the comparison of eHMI to Baseline was lower than p = 0.005. For every time sample of a trial, a paired sample t-test was conducted to investigate at which moment significant differences would occur and what the corresponding mean values of the absolute angular velocities were.

			Comparison eHMI to Baseline				
			eHMI		None		
Independent Variables	Time [s]		M ($^{deg}/_{s}$)	SD ($^{deg}/_{s}$)	M ($^{deg}/_{s}$)	SD ($^{deg}/_{s}$)	df
20MY – Text	Start	1.75	5.92	2.36	2.79	1.65	23
	End	1.76	5.81	2.32	2.76	1.63	23
	Start	3.00	22.95	4.79	3.98	1.83	23
	End	3.05	21.82	4.47	4.02	1.65	23
	Start	4.91	11.59	2.90	35.67	5.60	21
	End	5.10	17.19	3.78	37.40	5.10	21
30MY – Text	Start	2.05	20.36	4.53	4.71	1.94	23
	End	2.09	20.91	4.48	4.79	2.07	23
30MN - Text	Start	0.78	2.14	1.27	6.98	2.20	23
	End	0.84	2.22	1.31	5.65	1.89	23

Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, df = Degrees of Freedom.

Noise Simulation

A simulation was made of white noise signals which have been subjected to the same method of statistical analysis as the C.o.M. velocities and head movements. In total three signals (i.e., representing no eHMI, Text and FBL) have been made which varied randomly between 0 and 1, were 10 seconds in length and had a sampling frequency of 240 Hz. The noise signal representing no eHMI was compared to the noise signal representing the Text eHMI and the Frontal Braking Lights. Significant differences between the two sets of compared simulated white noise signals, with a p-value lower than $p = 0.005$, have been found at various instances over time. In the comparison of Text to None a total of 14 samples out of 2400 were equal to or lower than $p = 0.005$ and in the comparison of FBL to None 15 samples out of 2400 were equal to or lower than $p = 0.005$. The variation of the signals and the variation of the *p*-values over time can be found in Figure 16.

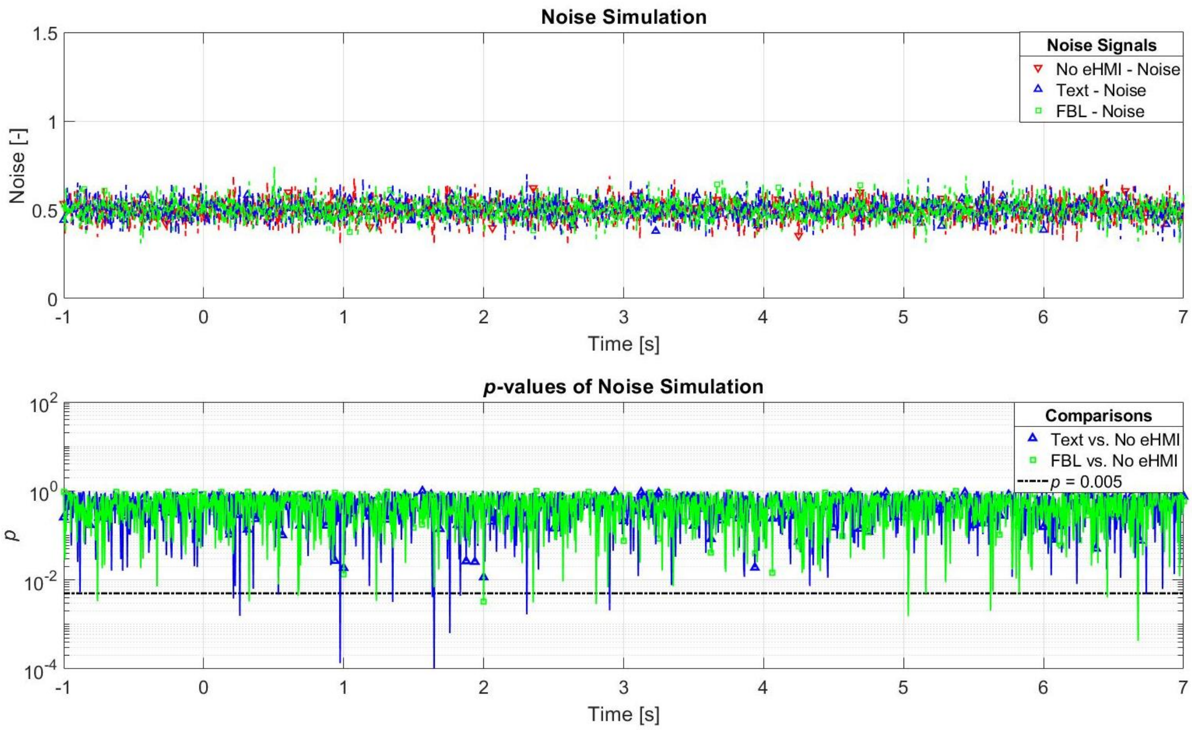


Figure 16: Noise simulation of three random signals. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights. T = 0 represents the moment where the stimulus car started braking. Top graph: Dimensionless white noise signals which varied between 0 and 1. Bottom graph: The horizontal boundaries correspond to p-values for the level of strong significant difference (i.e., *p* = 0.005). At various moments in time, the compared white noise signals differed significantly due to chance.

Gait Initiation Frequency and Initiation Times

An effort to quantify gait initiation frequencies and gait initiation times was made, but the distribution of the onsets of ankle motions was found to be non-uniform over time. Such non-uniformity influenced the statistical testability of the data, since, for example, not every time bin for the investigated conditions contained an equal amount of data for every participant. This non-uniformity can be derived from the following figures, namely Figure 17 and 18.

It can, for example, be seen in Figure 17 that this particular participant made some steps in the positive x-direction prior to the deceleration of the vehicle. However, Figure 18 shows another participant who refrained from the approaching steps to the curb seen in Figure 17, and initiated locomotion from standstill. Clearly, the fact that the approaching steps seen in Figure 17 are not present in Figure 18 as well as the fact that the timing of motion, and thus the timing of steps, differed between participants underlines the reason why the attempt to quantify gait initiation frequencies and gait initiation times was unsuccessful. Furthermore, our algorithm found multiple instances of a motion onset, for example, for the left ankle between 0 and 2 seconds as can be seen in Figure 17.

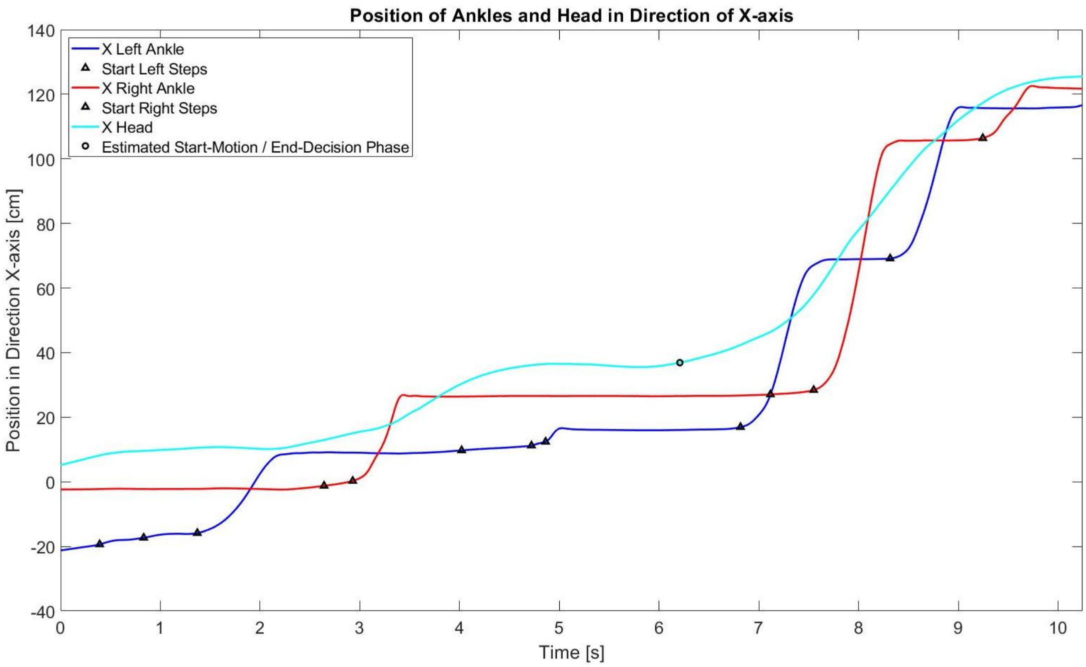


Figure 17: Position of ankles and head along the x-axis of participant 2. Triangles represent the start of motion found by the custom algorithm made for this experiment. The triangles depicted between the 4th until the 5th second on the line representing the left ankle position show that the algorithm was not robust enough to detect a single moment that reflected the onset of motion. The head position was used as a heuristic in an attempt to differentiate between the event of continuous bipedal locomotion and the event of non-sequential steps.

The aforementioned reasons were deemed sufficiently critical to decide to omit the gait initiation frequencies and gait initiation times from statistical testing. However, an analysis on the ankle accelerations as well as on the thorax angle was done to search for potential indicators of hesitation behaviour of pedestrians. The results of this analysis, including an interpretation of these results, can be found in Appendix J.

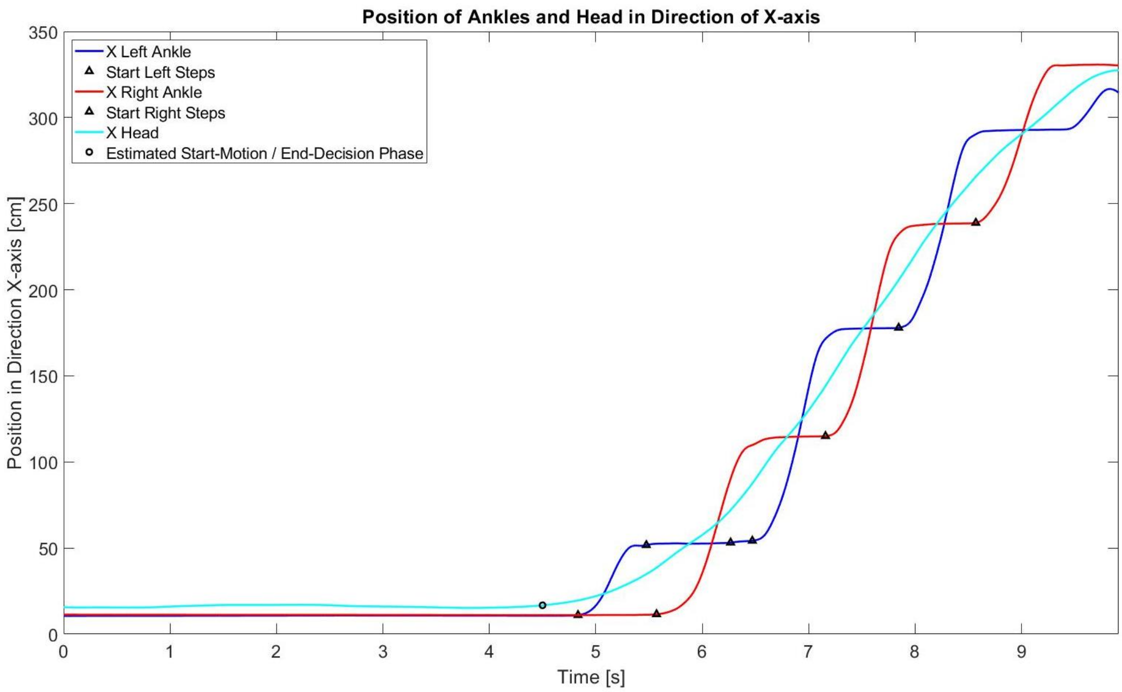


Figure 18: Position of ankles and head along the x-axis of participant 16. Triangles represent the start of motion found by the custom algorithm made for this experiment. In this particular case the custom algorithm was almost successful in detecting the onset of motion. However, the triangles between the 5th and the 7th second on the line representing the left ankle position show that the algorithm was not robust enough to detect a single moment that reflected the onset of motion. The head position was used as a heuristic in an attempt to differentiate between the event of continuous bipedal locomotion and the event of non-sequential steps.

Opening Questionnaire

A large percentage of the respondents in this experiment disagreed with the statements that reflected motorists negatively. For example, none of the respondents strongly agreed with the statement that drivers are not respectful to pedestrians and most of the respondents disagreed with the statement that drivers are aggressive and careless (i.e., disagreed: 50% strongly disagreed: 8.3%) as can be seen in Figure 19. The percentage of respondents that would let a car go by even if they had right of way (i.e., agreed: 25%, strongly agree: 8.3%) is about equal to the percentage of respondents that would not (i.e., disagreed: 33.3%) or were not certain (i.e., neither disagree nor agree: 33.3%). An overview of all the responses given to each statement in the questionnaire can be found in Appendix G.

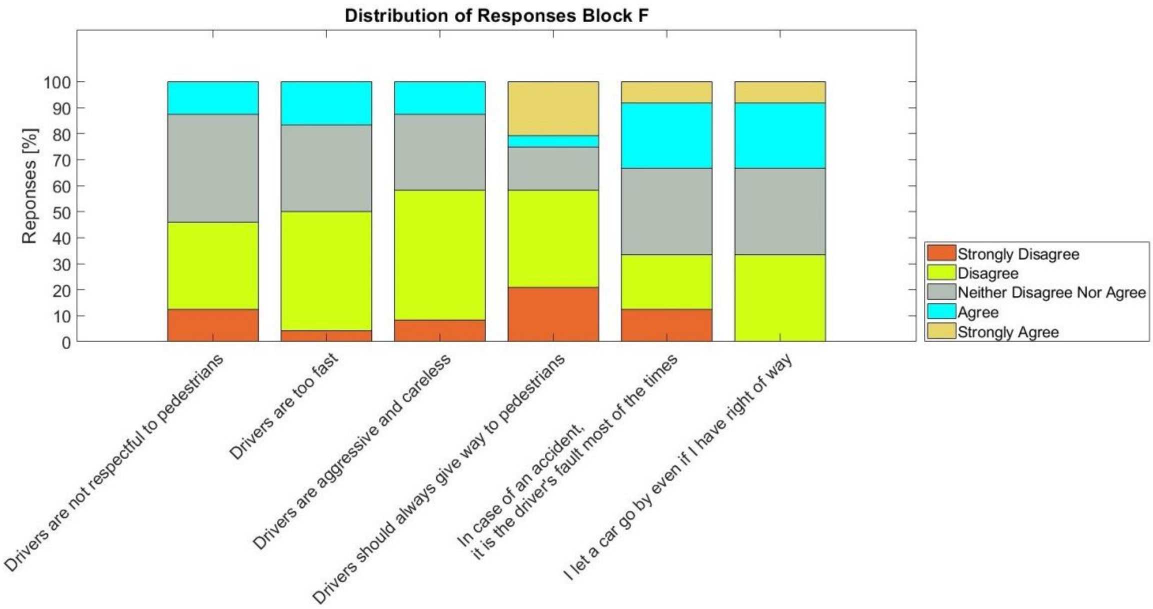


Figure 19: Distribution of responses to the statements of block F. These statements were concerned with how pedestrians interact with and perceive other road users.

Statements during experiment

A simple linear regression was calculated to predict the change of the mean response over the course of the 18 trials in the experiment. Significant regression equations were found for the predictability of oncoming car behaviour ($F(1,16) = 12.279$, $p < 0.005$), the subjective experience of fear ($F(1,16) = 34.575$, $p < 0.001$) and the level of motion sickness ($F(1,16) = 45.224$, $p < 0.001$), which are shown in Figure 20 along with the corresponding mean responses . To determine whether the mean response changed significantly over the course of the experiment, we conducted paired sample t-test comparing the mean response of the first trial to the mean responses of the last two trials as well as comparing the mean response of the second trial to the mean responses of the last two trials. No significant differences were found between the first and the last two mean responses as well as the second and the last two mean responses for the predictability of the oncoming car behaviour and the level of motion sickness. Only for the subjective level of fear there was a weak significant difference between the first response (i.e., 3.08 ± 0.13) and the second last response (i.e., 1.83 ± 1.01, $t(23) = 3.05$, $p < 0.006$).

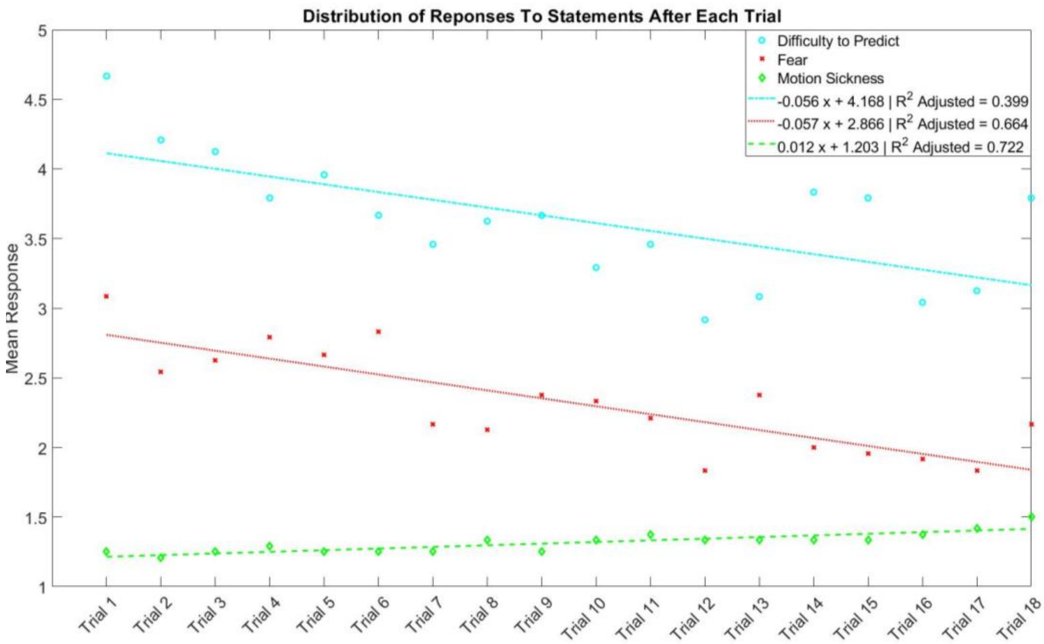


Figure 20: Mean responses of the participants to the three statements after each trial. Significant linear regressions were found for the MISC response rates and the feeling of fear response rates. Predictability: Responses to the enquiry whether it was difficult to predict the behaviour of the oncoming vehicles. The value 1 represented 'strongly disagree' and 7 'strongly agree.' Fear: Responses to the enquiry about whether they experienced a feeling of fear when considering crossing the road. The value 1 represented 'strongly disagree' and 7 'strongly agree.' Motion Sickness: Self-reported MISC ratings by participants after each trial. The value 1 reflected that participants experienced no problems, 2 slight discomfort, 3 and 4 slight and mild nausea and 5 and greater indicated more severe symptoms of sickness.

The distribution of responses to the statement regarding the predictability of the oncoming car behaviour were significantly different for the yielding and non-yielding trials where the stimulus car had a text eHMI (20M Yielding - Text: 2.71 ± 1.72, $t(23) = 4.38$, $p < 0.001$; 20M Non-Yielding - Text: 2.13 ± 1.52, $t(23) = 3.98$, $p < 0.001$;) compared to the conditions with the same inter-vehicular distance and yielding behaviour in which the stimulus car had no eHMI (20M Yielding - None: 4.67 ±

2.26; 20M Non-Yielding - None: 4.04 ± 2.10). However, no significant differences were found for the responses to the statement regarding the feeling of fear. The distribution of the responses per condition for the self-reported ability to predict the car behaviour and feeling of fear are depicted in Figure 21 and Figure 22.

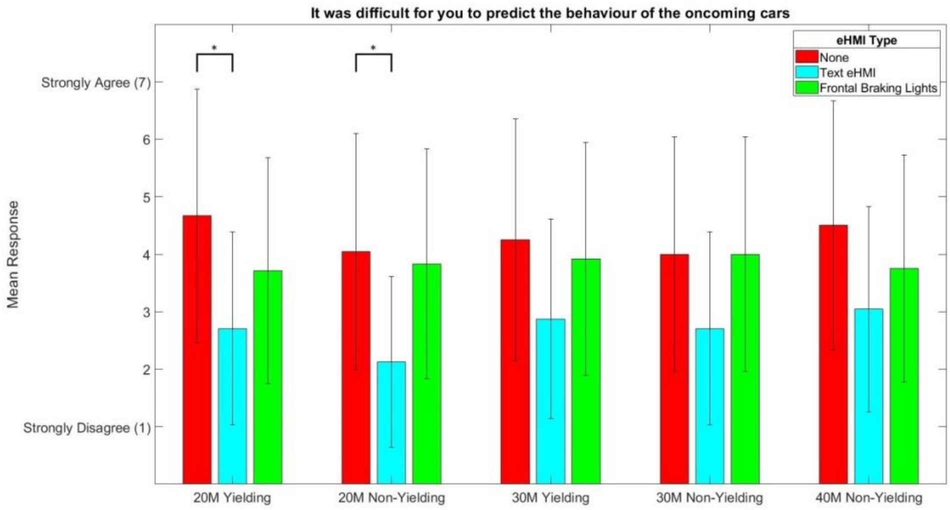


Figure 21: Responses to the statement regarding the predictability of the behaviour of the oncoming cars for each condition. Responses are bar-plotted as mean values with the error bars representing Mean – SD and Mean + SD. Abbreviations: 20M = 20 meters inter-vehicular distance, 30M = 30 meters inter-vehicular distance, 40M = 40 meters inter-vehicular distance. Significant differences were found for the responses regarding the participants' ability to predict the behaviour of the cars of the condition 20M Yielding between the presence of Text eHMI and the absence of an eHMI (*p*< 0.001), as well as for the condition 20M Non-Yielding between the presence of Text eHMI and the absence of an eHMI (*p*< 0.001). These significantly different pairs are indicated by a bracket with an asterisk.

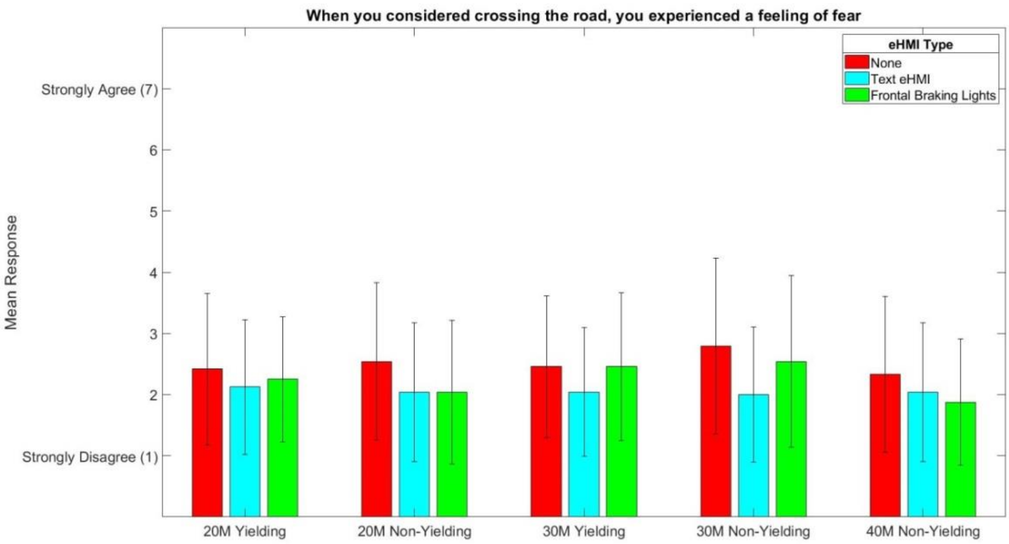


Figure 22: Responses to the statement about the feeling of fear for each condition. Responses are bar-plotted as mean values with the error bars representing Mean – SD and Mean + SD. Abbreviations: 20M = 20 meters inter-vehicular distance, 30M = 30 meters inter-vehicular distance, 40M = 40 meters inter-vehicular distance. No significant differences were found between the responses regarding the participants' feeling of fear for any of the conditions with equal inter-vehicular distance and yielding behaviour.

Virtual Reality Presence Questionnaire

We compared presence in the current experiment including walking motion to the study of de Clercq including standing pedestrians. Two questions (i.e., question 24 and 25, see Appendix C) from the VRPQ were excluded from the statistical analysis, since they were not present in the questionnaire in the study of De Clercq. However, participants reported on average, to be well focussed on the experimental task or environment (i.e., question 24; M = 5.5, SD =0.88) and agreed on the consistency of information through different senses (i.e., question 25; M = 5.6, SD = 1.06). Nonetheless, no significant differences were found between the responses to the VRPQ of this study and the study of De Clercq. The results of the two sample t-test for each factor are summarized in table 6.

Table 6: Virtual Reality Presence Questionnaire. Results of two sample t-test including descriptive statistics of the comparison of our study to that of De Clercq. For each factor of the questionnaire no significant differences were found between the responses of the participants in our study and the responses of the participants in the study of De Clercq.

	Responses Per Study									
	Kooijman			De Clercq			95% CI for Mean Differences	*p*	T	df
	M	SD	N	M	SD	N				
Involvement	51.25	1.26	24	50.14	1.27	28	-0.12, 0.34	0.349	0.94	518
Sensor Fidelity	29.58	1.35	24	30.75	1.25	28	-0.50, 0.11	0.206	-1.27	310
Adaption / Immersion	16.38	1.07	24	16.61	1.02	28	-0.45, 0.30	0.684	-0.41	154
Interface Quality	9.58	1.55	24	8.96	1.29	28	-0.30, 0.71	0.421	0.81	154

Discussion

The results found in this experiment confirm our initial hypothesis that the presence of an eHMI, mounted on a vehicle front, would positively contribute to pedestrians' crossing intention. The main finding to support this statement is the fact that we found higher forward C.o.M. velocities when there was either a textual display or there were frontal braking lights present on the yielding stimulus car compared to when no display was present and the inter-vehicular distance was 20 meters. More specifically, the significant difference occurred during the time window when the stimulus car indicated it was yielding. Furthermore, we found the same effect for the textual display for the condition in which the inter-vehicular distance was 30 meters. The literature on affective body language describes that higher forward velocities are associated with approaching behaviour [19]–[22] and accordingly we can objectively conclude, based on this bodily measure, that the presence of an eHMI positively contributed to the crossing decision of the participants during the conditions in which the cars yielded and had either an inter-vehicular distance of 20 meters or 30 meters.

We found that during the time window in which the stimulus car was yielding, and had an inter-vehicular distance of 20 meters to its predecessor, there were multiple time events where the absolute angular velocities of the participants' heads was greater when the stimulus car had a text eHMI compared to when it did not. This greater absolute rate of change in head angle could be associated with the previously mentioned finding of Zito et al. [18] who related more secure pedestrian behaviour to a higher frequency of alternating head movements of the participants between their left and right-hand side. Moreover, the total length of the temporal ranges in which these significant differences occurred was considerably larger than the total amount of samples that were found to be significantly different in the noise simulation. However, the measure of the absolute rate of change in head angle must be seen in a temporal context, meaning that the rate of change in head angle is related to or limited by the temporal distance between the stimulus car and its predecessor as well as the cars' velocities. Furthermore, as was shown by Grossman et al. [29], angular velocities of the head in the horizontal plane also arise due to the mere fact that someone is walking or running. In contrast to the study of Zito et al. [18] participants in our study were able to move freely prior to their crossing decision i.e., in the experiment of Zito et al. [18] participants were instructed to take a step forward once they felt it was safe to cross but otherwise participants stood still. The fact that the combination of gait and observation was present in our study could potentially confound the results of the analysis of the angular velocities of the head and thereby negatively impact the interpretation of the results. It, therefore, seems that utility of the measure of alternating head movements can be ambiguous and thus requires extensive research before any claims based upon this measure can be made, certaintly in the specific context of our experiment.

One would expect the self-reported fear to correspond to the objective measure of the participants' velocities and rate of change in head angle, but this appeared not to be the case. The reported feelings of fear did not differ significantly between conditions in their respective comparison cluster and even declined over the course of the experiment. Apparently, the participants did not associate the feeling of fear with crossing the road within the context of this experiment or with the hesitation in doing so. This nonassociation can be ascribed to two possible factors, namely the context of the experiment and the participants' attitude towards motorists.

First, the context of this experiment was, besides the five approaching vehicles in each trial, stressor free; no time constraints were applied, and the participants were presented with the binary choice of either crossing at some point of their own choice between the passing cars or waiting until the cars had all passed. Specifically, the instructions to the participants were that they should only cross when they deemed it was safe to do so. These contextual factors might have been of influence on the participants' perception of fear. Secondly, the participants' attitudes can be obtained from their responses to the opening questionnaire. Overall, participants were indecisive to or leaned more towards disagreeing with statements that reflected motorists negatively, for example, the statement that drivers are aggressive and careless. More specific, the larger portion of the participants indicated that drivers should not always give way to pedestrians. However, when asked directly whether the participants would let a motorist pass by, even if they had the right of way, the distribution of the responses centred around neither disagreeing nor agreeing. We can infer from this subjective data that in general the pedestrians were not fearful of the particular interactive scenario of this experiment, had a rather respectful view of motorists and relied on the context of the interaction with the motorist before deciding whether they would give up their right of way or not.

The self-reported ability to predict the behaviour of the oncoming cars seemed to corroborate more with the context of the experiment as an subjective measure than the self-reported feeling of fear. The inability to predict the future behaviour of something/one could cause the observer to hesitate in taking action and thereby find its reflection in the expressed behaviour of the observer. We found that the self-reported ability of participants to predict the behaviour of the stimulus car with a text display was significantly better compared to when there was no display present. However, the finding was limited only to the yielding and non-yielding conditions in which the inter-vehicular distance was 20 meters. The quality of the HDM can explain the limitation of this finding when we take into account the safety margins pedestrians assess when making road crossing decisions. The optimal point for a pedestrian to cross the road between two moving vehicles is directly after the first vehicle of the two has passed in front of the pedestrian. This grants the pedestrians the longest temporal region to perform their road crossing. Since the distance between the stimulus car and its predecessor was larger than 20 meters in the other conditions, participants had the opportunity to cross the road earlier. However, participants were not able to decide for certain what the textual display stated due to the 'screen door effect,' which is a known limitation of the HMD, and therefore they could have indicated a lower ability to predict the behaviour of the oncoming vehicles.

This, in turn, can be qualitatively inferred from Figure 13 depicting the mean C.o.M. velocity of the participants in the conditions with an inter-vehicular distance between the second and third car of 30 meter. Participants who were presented with a stimulus vehicle devoid of an eHMI showed a greater forward velocity, compared to the condition where the stimulus vehicle had a textual eHMI, from the moment the stimulus vehicle started braking to the first second after braking. This could mean that in that case, the decision to cross, directly after the second vehicle in the train had passed the participant, was made before the third car in the train started braking (e.g., at T = -1). Nonetheless, the temporal region between the second and third vehicle was not large enough for the participants to cross without the stimulus vehicle braking, which the participants, after their initial decision to cross, re-estimated. This is reflected by the relatively constant, instead of an increasing, velocity of the participants from the first second after braking until the third second after braking when there was no eHMI present on the stimulus vehicle. For the condition in which a textual display was present on the stimulus vehicle, participants were aware of the presence of the display but could not read yet what it stated. Thus participants did not decide to continue walking at, for example, T = 0, but waited until they saw that the display changed state which is reflected by their increasing velocity at T = 1. Furthermore, it can be seen in Figure 13 that the trends of the C.o.M. velocities for the non-yielding conditions between T = -1 and T = 0 are similar to the yielding conditions. However, in the non-yielding conditions participants were not presented with a confirmatory signal to cross and thus aborted their crossing decision which can be inferred from the decreasing forward C.o.M. velocity seen in Figure 13 that starts at T = 1. The fact that no significant effect was found for the condition in which the inter-vehicular distance was 40 meters and the cars did not yield, can be interpreted using the same reasoning as for the non-yielding condition in which the inter-vehicular distance was 30 meters. More so, since the similarity of the trends of the non-yielding conditions with 30 meters inter-vehicular distance can be found in Figure 28 in Appendix H for the non-yielding conditions with 40 meters inter-vehicular distance.

Conclusion

The goal of this study was to investigate the influence of two different external human-machine interfaces, mounted on autonomous vehicles, and the distance of the vehicles to their predecessor, on the pedestrians' body movement and self-reflection as a measure of their crossing intention. The objective measure of the participants' forward gait velocity showed to effectively be an operationalization of a greater acceptance by the participants to cross the road when the autonomous vehicles had an external human-machine interface indicating its yielding behaviour compared to when there was no external interface present. Furthermore, the findings can be interpreted in a qualitative manner, in such, that the Text eHMI was less ambiguous than the Frontal Braking Lights, since for two out of the three possible inter-vehicular distances the Text eHMI facilitated a significantly greater forward gait velocity compared to no eHMI, whereas the Frontal Braking Lights only showed a significant effect for one out of three possible inter-vehicular distances. This interpretation corroborates with the findings of De Clercq [13] where the subjective reports on eHMI acceptance by participants showed that the Text eHMI was perceived to the least ambiguous and the Frontal Braking Lights the most ambiguous. Concluding, it can be deduced that even though a road crossing decision can be made using car dynamics alone, an external human-machine interface positively contributes to the non-verbal communication between pedestrians and autonomous vehicles and thus proves to be an adequate replacement for non-verbal signals from humanoid drivers.

Validation of Our Research Method

The validation of our research method relied on two measures; the MISC ratings the participants reported throughout the experiment and the VRPQ they completed after the experiment. Primarily, no significant increase or fluctuation was found in the reported MISC rating, which remained on average well below the proposed critical value of 4 throughout the entire experiment. The majority of the participants indicated at some point during the experiment that they experienced some form of discomfort; however none of the participants indicated experiencing any of the sensations that would reflect a MISC score of 4 or higher. In addition, participants reported in the VRPQ of a relatively high consistency of the information presented to them through different senses and of the ability to be well focussed on the experimental task or environment. From this information, we can infer that the participants were able to perceive the environment and effectively perform the proposed tasks of this experiment without any negative influence on their subjective well-being.

Although our study had a few significant differences compared to the study of De Clercq (i.e., the visualization of the participants' own body motion via the use of an avatar, the ability of participants to cross the street), no significant differences were found between the subjective VRPQ reports of our participants and those in the study of De Clercq. However, exactly these contextual differences in our study could have influenced the subjective experience of the participants. For example, a genderless avatar was used in our study which could have caused our participants to rate the immersiveness of the VR experience lower compared to when there would be no avatar present. Nonetheless, we can state, based on the low MISC rates and the more than average VRPQ ratings comparable to the study of De Clercq, that our research method is a valid method to investigate pedestrian crossing behaviour in a fixed context.

Outlook and Limitations of Our Research Method

As mentioned at the beginning of the result section, we failed to record the condition in which vehicles yielded and the inter-vehicular distance between the second and third vehicle was 40 meters. Furthermore we were not able to correlate all of the data during the conditions in which the cars did not yield. Effectively, this did not influence the quality of the results that were used to confirm our hypothesis. However, for future research purposes, we would propose the use of a more rigorous method of time synchronization, such as recording the system time, when measuring human motion in two programs such as MVN Analyze and Unity.

Besides the obvious implications that, for example, the skewed sample size we had for our experiment might have, the experimental set-up of this study was shown to have some limitations. First, the size of the experimental physical environment was a limiting factor in our study. The fact that the participants were presented with a limited spatial region that could be used to cross the road could have influenced the planning of their walking trajectory and effectively their gait (i.e., more specifically the pacing of their steps). Secondly, although the context of this experiment was clear and unambiguous, the results might not be directly extrapolated to real life. In real life, the zebra crossing near a T junction or crossroad has traffic lights that regulate the moments during which the pedestrians are allowed to cross or not (i.e., not taking into account illegal behaviour such as jaywalking). In such a case, pedestrians need not rely on confirmatory signals of motorists. The explicit regulatory measure of a traffic light was not present in our study. Thirdly, the lack of surrounding traffic in our experiment might have also influenced the behaviour of the participants. The focus of the participants was mainly on their left-hand side, while no information was presented to them from different areas in their surroundings. This spatial focus could have effectively biased their motion and decision-making of road-crossing. Future research on the influence of eHMI on the crossing decision of pedestrians at a zebra crossing could perhaps investigate this at a zebra crossing that is located in the middle of a street, with a refugee island between the two lanes and presenting a bi-directional traffic flow. Contextually this would be more in line with a real-life scenario. Nonetheless, we deem that the findings presented in this thesis can be used for further improvements on the research on pedestrian crossing behaviour as well as research on the efficacy of eHMIs in communicating the intentions of autonomous vehicles to pedestrians' who are intent on crossing.

References

[1] C. Bila, F. Sivrikaya, M. A. Khan, and S. Albayrak, “Vehicles of the Future: A Survey of Research on Safety Issues,” *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 5, pp. 1046–1065, 2017.

[2] Centraal Bureau voor Statistiek (CBS), “Overledenen; doden door verkeersongeval in Nederland, wijze van deelname,” *http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=71936NED*, 2017.

[3] World Health Organization (WHO), “Global status report on road safety,” *World Heal. Orgainisation*, p. 318, 2015.

[4] United Nations, “World Urbanization Prospects: The 2018 Revision, Key Facts,” p. 2, 2018.

[5] S. Schmidt and B. Färber, “Pedestrians at the kerb - Recognising the action intentions of humans,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 12, no. 4, pp. 300–310, 2009.

[6] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, “Understanding Pedestrian Behavior in Complex Traffic Scenes,” *IEEE Trans. Intell. Veh.*, pp. 1–1, 2018.

[7] M. Sucha, D. Dostal, and R. Risser, “Pedestrian-driver communication and decision strategies at marked crossings,” *Accid. Anal. Prev.*, vol. 102, pp. 41–50, 2017.

[8] Z. Ren, X. Jiang, and W. Wang, “Analysis of the Influence of Pedestrians’ eye Contact on Drivers’ Comfort Boundary during the Crossing Conflict,” in *Procedia Engineering*, 2016, vol. 137, pp. 399–406.

[9] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, “Agreeing to cross: How drivers and pedestrians communicate,” in *IEEE Intelligent Vehicles Symposium, Proceedings*, 2017, pp. 264–269.

[10] A. Habibovic *et al.*, “Communicating Intent of Automated Vehicles to Pedestrians,” vol. 9, no. August, 2018.

[11] M. V. S. Nooij *et al.*, “Definition of necessary vehicle and infrastructure systems for Automated Driving,” *Eur. Comm.*, pp. 1–111, 2011.

[12] D. Rothenbücher, J. Li, D. Sirkin, B. Mok, and W. Ju, “Ghost Driver: A Platform for Investigating Interactions Between Pedestrians and Driverless Vehicles,” *Adjun. Proc. 7th Int. Conf. Automot. User Interfaces Interact. Veh. Appl. - AutomotiveUI ’15*, pp. 44–49, 2015.

[13] G. K. de Clercq, “The Effects of External Human-Machine Interfaces of Automated Vehicles on the Crossing Behaviour of Pedestrians,” *MS Thesis TU Delft*, 2017.

[14] C. Chang, K. Toda, D. Sakamoto, and T. Igarashi, “Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian,” in *Proceedings of the 9th ACM International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI ’17)*, 2017.

[15] S. Deb, L. J. Strawderman, and D. W. Carruth, “Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 59, pp. 135–149, 2018.

[16] S. Kalantarov, R. Riemer, and T. Oron-Gilad, “Pedestrians’ road crossing decisions and body parts’ movements,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 53, pp. 155–171, 2018.

[17] B. A. Morrongiello, M. Corbett, J. Switzer, and T. Hall, “Using a Virtual Environment to Study Pedestrian Behaviors: How Does Time Pressure Affect Children’s and Adults’ Street Crossing Behaviors?,” *J. Pediatr. Psychol.*, vol. 40, no. 7, pp. 697–703, 2015.

[18] G. A. Zito *et al.*, “Street crossing behavior in younger and older pedestrians: An eye- and head-tracking study,” *BMC Geriatr.*, vol. 15, no. 176, 2015.

[19] B. Fawver, G. F. Beatty, K. M. Naugle, C. J. Hass, and C. M. Janelle, “Emotional state impacts center of pressure displacement before forward gait initiation,” *J. Appl. Biomech.*, vol. 31, no. 1, pp. 35–40, 2015.

[20] K. M. Naugle, C. J. Hass, J. Joyner, S. A. Coombes, and C. M. Janelle, “Emotional state affects the initiation of forward gait,” *Emotion*, vol. 11, no. 2, pp. 267–277, 2011.

[21] C. L. Roether, L. Omlor, A. Christensen, and M. A. Giese, “Critical features for the perception of emotion from gait,” *J. Vis.*, vol. 9, no. 6, pp. 15–15, 2009.

[22] E. Crane and M. Gross, “Motion Capture and Emotion: Affect Detection in Whole Body Movement,” *LNCS*, vol. 4738, pp. 95–101, 2007.

[23] X. Zhuang and C. Wu, “The safety margin and perceived safety of pedestrians at unmarked roadway,” *Transp. Res. Part F Traffic Psychol. Behav.*, 2012.

[24] J. Brade, M. Lorenz, M. Busch, N. Hammer, M. Tscheligi, and P. Klimant, “Being there again – Presence in real and virtual environments and its relation to usability and user experience using a mobile navigation task,” *Int. J. Hum. Comput. Stud.*, vol. 101, pp. 76–87, 2017.

[25] S. Deb, D. W. Carruth, R. Sween, L. Strawderman, and T. M. Garrison, “Efficacy of virtual reality in pedestrian safety research,” *Appl. Ergon.*, vol. 65, pp. 449–460, 2017.

[26] L. Fridman, B. Mehler, L. Xia, Y. Yang, L. Y. Facusse, and B. Reimer, “To Walk or Not to Walk: Crowdsourced Assessment of External Vehicle-to-Pedestrian Displays,” 2017.

[27] E. Papadimitriou, S. Lassarre, and G. Yannis, “Human factors of pedestrian walking and crossing behaviour,” in *Transportation Research Procedia*, 2017, vol. 25, pp. 2007–2020.

[28] B. G. Witmer, C. J. Jerome, and M. J. Singer, “The factor structure of the Presence Questionnaire,” *Presence: Teleoperators and Virtual Environments*, vol. 14, no. 3. pp. 298–312, 2005.

[29] G. E. Grossman, R. J. Leigh, L. A. Abel, D. J. Lanska, and S. E. Thurston, “Frequency and velocity of rotational head perturbations during locomotion,” *Exp. Brain Res.*, vol. 70, no. 3, pp. 470–476, 1988.

[30] D. J. Benjamin *et al.*, “Redefine statistical significance,” *Nat. Hum. Behav.*, 2017.

[31] S. Schreven, P. J. Beek, and J. B. J. Smeets, “Optimising filtering parameters for a 3D motion analysis system,” *J. Electromyogr. Kinesiol.*, 2015.

Appendix A: Informed Consent

Department of BioMechanical Engineering
Faculty of Mechanical, Maritime and Materials Engineering

Participant Informed Consent

TITLE: **"Detection of Pedestrian Crossing Behaviour"**

INVESTIGATORS: - B.Eng. Lars Kooijman - MSc Student
- Dr. ir. Joost C.F. de Winter – Supervisor

L. Kooijman Contact: - 0031-640140722
- L.Kooijman@student.tudelft.nl

This consent form gives you the basic idea of what the research is about and what your participation will involve. If you would like more detail, please ask the experimenter. Please take the time to read and understand this form carefully. You will receive a copy of this form.

WHAT IS THE PURPOSE OF THE STUDY?

The purpose of this study is to determine whether the crossing intention of human pedestrians can be detected from body motion. This knowledge might, for example, be helpful in the detection of a human's crossing intentions by self-driving vehicles near zebra crossings.

WHAT WOULD I HAVE TO DO?

You will be asked to attend one testing session of less than 1 hour. Upon your arrival, you will be asked to fill in a questionnaire about your behaviour as a pedestrian in traffic. After completing the questionnaire, you will be familiarized with the 3D motion tracking device, called Xsens. You will also be familiarized with the Oculus Rift, which are Virtual Reality goggles used to immerse you in a virtual reality world. The complete set-up is depicted in Figure 1.

After manually measuring the lengths/heights of your arms, ankles, hip, knees, feet, and shoulders, the experimenter will attach the sensors of the Xsens on seventeen of your body parts. Some of these sensors are incorporated into a t-shirt, which you will be asked to wear on top of your clothing. A few sensors need to be attached to your body using Velcro straps. After the attachment of the sensors, you will be instructed for the calibration of the Xsens to perform an N-pose, after which you will shortly need to walk back and forth in the experimental environment. Once a successful calibration has been obtained, you will be asked to wear the Oculus Rift and walk around to accustom yourself to the Virtual Reality environment, which is shown in Figure 2.

Figure 23: Xsens and Oculus Set-up

The goal of the testing sessions for you is to cross the road in the Virtual Reality environment as safely as possible. During the testing session, you will be standing on the side walk near a zebra crossing, which is indicated by a red dot on the right side of Figure 1. During each trial of the testing session, a train of cars will pass you by. Each train of cars will consist out of five cars. Your goal in this experiment is to start crossing the road whenever you think it is safe enough to do so, whilst avoiding any collisions with the oncoming cars. The distance between cars in the train, their yielding behaviour, the presence and type of a display indicating their yielding behaviour, and lastly the type of car, will vary. The first two cars of each train will never yield, and therefore you can start choosing to cross after these two cars have passed. If you do not see a safe possibility to cross the road, wait until the last car has passed you, and the trial will end.

Figure 24: Virtual Reality Environment

Before the starting of a new trial, you will be asked to a few questions regarding motion sickness and your level of comfort in the environment. If you report a 4 or higher level of motion sickness, the experiment will be stopped or paused. After five trains of cars, the experiment will pause; you can take off the Oculus Rift and take a break. A total of 18 different trains of cars will be presented to you throughout the entire experiment. At the end of the experiment, you will be asked to fill in a final questionnaire about the virtual environment and your experience in it.

WHAT ARE THE RISKS?

In some persons, there is a possibility that motion sickness occurs due to the Oculus Rift. If at any point you begin to experience any discomfort, disorientation, or nausea, please notify the experimenter and the experiment will be paused or ceased entirely. Please do not engage in potentially hazardous activities (driving, cycling) in case you continue to feel nauseous. There is also a possibility that unsafe crossing decisions may be experienced as genuinely stressful or frightening, due to the high level of visual immersion.

WILL I BENEFIT IF I TAKE PART?

This study will provide information in terms of which nuances can be detected from human movement. As such, this study provides basic scientific information on movement and implicitly on human behaviour. There is no direct benefit to the participants involved.

DO I HAVE TO PARTICIPATE?

Participation in this study is voluntary, and you may withdraw from the study at any time without negative consequence. If you wish to stop the testing at any time, please notify the experimenter to stop the experiment immediately.

WILL MY RECORDS BE KEPT PRIVATE?

Information obtained during this research project is made anonymous. The information may be used for statistical analysis, used in an MSc thesis and research article, and stored in a public repository in an anonymous form.

SIGNATURES

Your signature on this form indicates that you have understood to your satisfaction the information regarding your participation in the research project and agree to participate as a participant. You are free to withdraw from the study at any time. If you have further questions concerning matters related to this research, please contact: B. Eng. L. Kooijman.

Participant's Name	Signature and Date
Investigator/Delegate's Name	Signature and Date

A signed copy of this consent form has been given to you to keep for your records and reference.

Appendix B: Steps and instructions

1. Arrival participant
2. Briefing experiment using informed consent
3. Signing informed consent
4. Opening questionnaire by participant
5. Familiarization with equipment
6. Anthropometrics
7. Placement of Xsens sensors and Oculus
8. Acclimation Virtual Environment
9. Reminder to participant regarding:
    a. Yielding behaviour first two cars
    b. Random behaviour of third until fifth car
    c. Spatial limitations of the lab
10. Familiarization with cars in VR
11. Reminder to participant regarding:
    a. Questions / statements after each trial
    b. Total number of trials
    c. Goal of the experiment
12. Ask participant whether they are comfortable to start
13. Start experimental trials
14. Take a break after 5 trials
15. During break propose: a break each 5 trials / continue until participant asks for break
16. Resume experimental trials
17. After 18 trials remove equipment
18. Virtual Reality Questionnaire by participant
19. Departure participant

Appendix C: Opening Questionnaire

Opening Questionnaire

Thank you for participating in this study! Before starting the experiment, we'd like you to answer the following questions about yourself. Please answer them as truthfully as possible. This information will be made anonymous to ensure your privacy.

1. **What's your nationality?**

____

2. **What's your age?**

____

3. **What's your gender?**

Mark only one oval.

- ◯ Female
- ◯ Male
- ◯ Prefer not to say
- ◯ Other:

4. **How often did you commute to work or school by foot in the last 12 months on average?**

Mark only one oval.

- ◯ Daily
- ◯ 4 to 6 days a week
- ◯ 1 to 3 days a week
- ◯ Once a month to once a week
- ◯ Less than once a month
- ◯ Never

As a pedestrian, how much would you agree with each one of the following statements:

5. **I walk for the pleasure of it**

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

6. **I walk because it is healthy**

Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

7. **In short trips I prefer walking over other modes of transportation**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

8. **Crossing roads is difficult**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

9. **Crossing roads outside designated locations increases the risk of an accident**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

10. **Crossing roads outside designated locations is wrong**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

11. **Crossing roads outside designated locations saves time**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

12. **Crossing roads outside designated locations is acceptable because other people do it**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

13. **I prefer routes with signalised crosswalks**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

14. **I try to make as few road crossings as possible**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

15. **I try to take the most direct route to my destination**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

16. **I try to take the route to my destination on which I encounter the least amount of traffic**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

17. **I am willing to make a detour to find a protected crossing**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

18. **I am willing to take any opportunity to cross**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

19. **I am willing to make dangerous actions as a pedestrian to save time**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

Compared to other pedestrians, how much do you agree that:

20. **I am less likely to be involved in a road crash than other pedestrians**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	Strongly agree

21. **I am faster than other pedestrians**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

22. **I am more careful than other pedestrians**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

As a pedestrian, how much would you agree with each one of the following statements:

23. **Drivers are not respectful to pedestrians**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

24. **Drivers are too fast**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

25. **Drivers are aggressive and careless**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

26. **Drivers should always give way to pedestrians**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

27. **In case of an accident, it is the driver's fault most of the times**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

28. **I let a car go by even if I have right of way**
Mark only one oval.

	1	2	3	4	5	
Strongly disagree	◯	◯	◯	◯	◯	Strongly agree

TU Delft

Appendix D: Experimenter Form

Interim Questionnaire Participant

Motion Sickness		Fear		Predict		
Trial	Level	Trial	Level	Trial	Level	Body Height
1		1		1		Foot Size
2		2		2		Arm Width
3		3		3		Ankle Height
4		4		4		Hip Height
5		5		5		Hip Width
6		6		6		Knee Height
7		7		7		Shoulder Width
8		8		8		Shoulder Height
9		9		9		
10		10		10		
11		11		11		
12		12		12		
13		13		13		
14		14		14		
15		15		15		
16		16		16		
17		17		17		
18		18		18		

Appendix E: Virtual Reality Presence Questionnaire

Now you are finished with the experiment , we'd like to ask you some questions about your experience in the virtual environment. Your answers can help to improve future research on traffic safety. Thanks in advance

1. **How much were you able to control events?**
   Mark only one oval.

	1	2	3	4	5	6	7	
Not at all								Completely

2. **How responsive was the environment to actions that you initiated (or performed)?**
   Mark only one oval.

	1	2	3	4	5	6	7	
Not responsive								Completely responsive

3. **How natural did your interactions with the environment seem?**
   Mark only one oval.

	1	2	3	4	5	6	7	
Extremely artificial								Completely natural

4. **How much did the visual aspects of the environment involve you?**
   Mark only one oval.

	1	2	3	4	5	6	7	
Not at all								Completely

5. **How much did the auditory aspects of the environment involve you?**
   Mark only one oval.

	1	2	3	4	5	6	7	
Not at all								Completely

6. **How natural was the mechanism which controlled movement through the environment**
   Mark only one oval.

	1	2	3	4	5	6	7	
Extremely artificial								Completely natural

TU Delft

7. **How compelling was your sense of objects moving through space?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Very compelling

8. **How much did your experiences in the virtual worlds seem consistent with your real world experiences?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not consistent	◯	◯	◯	◯	◯	◯	◯	Very consistent

9. **Were you able to anticipate what would happen next in response to the actions that you performed?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Completely

10. **How completely were you able to actively survey or search the environment using vision?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Completely

11. **How well could you identify sounds?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Completely

12. **How well could you localize sounds?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Completely

13. **How compelling was your sense of moving around inside the virtual environment?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not compelling	◯	◯	◯	◯	◯	◯	◯	Very compelling

14. **How closely were you able to examine objects?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Very closely

15. **How well could you examine objects from multiple viewpoints?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Extensively

16. **How involved were you in the virtual environment experience?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not involved	○	○	○	○	○	○	○	Completely engrossed

17. **How much delay did you experience between your actions and expected outcomes?**
Mark only one oval.

	1	2	3	4	5	6	7	
No delays	○	○	○	○	○	○	○	Long delays

18. **How quickly did you adjust to the virtual environment experience?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Less than one minute

19. **How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not proficient	○	○	○	○	○	○	○	Very proficient

20. **How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	○	○	○	○	○	○	○	Prevented task performance

21. **How much did the control devices interfere with the performance of assigned tasks or with other activities?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Interfered greatly

22. **How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Completely

23. **Were you involved in the experimental task to the extent that you lost track of time?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Completely

24. **Were there moments during the virtual environment experience when you felt completely focused on the task or environment?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not at all	◯	◯	◯	◯	◯	◯	◯	Constantly

25. **Was the information provided through different senses in the virtual environment (e.g., vision, hearing) consistent?**
Mark only one oval.

	1	2	3	4	5	6	7	
Not consistent	◯	◯	◯	◯	◯	◯	◯	Very consistent

26. **Do you have any comments or suggestions about the experiment? If not, you can leave this question empty**

____
____
____
____
____

You are finished! Thank you for participating in the experiment!

TUDelft

Appendix F: Data Reduction

The data obtained recorded in Unity was first interpolated to match the sample frequency of the MVN data, namely 240 Hz. Since Unity renders the real-time motion on the basis of 'best-fit', it has a non-consistent sampling rate, as shown in Fig. 25, and therefore we recorded the participant's motion twice, using Unity and MVN Analyze.

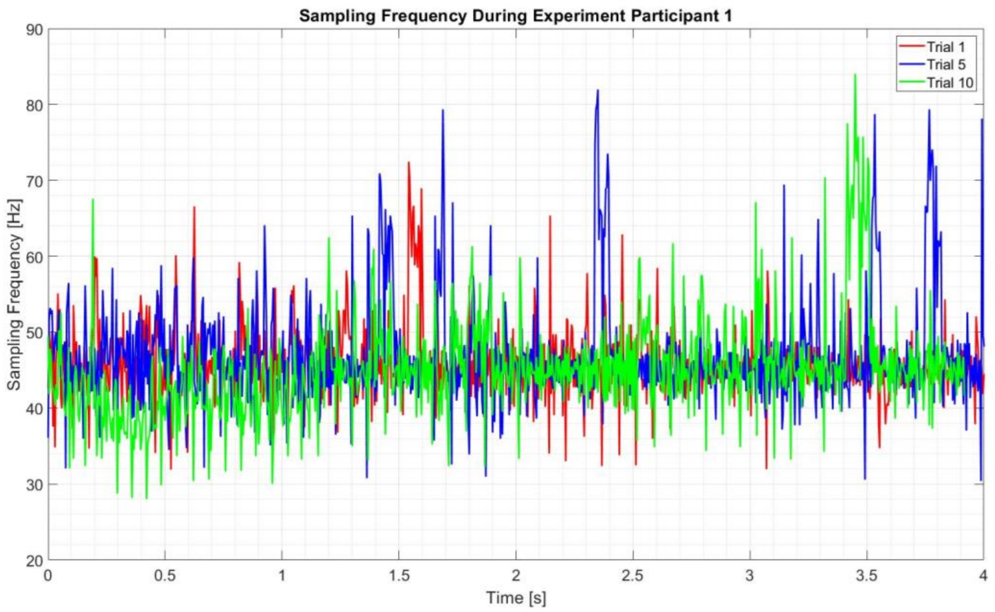


Figure 25: Sampling frequency over time in Unity for three trials of participant 1. The graph depicts a four second period of the fluctuation of the sampling frequency. Unity renders the motion in the simulation on basis of best fit and therefore will skip rendering a frame when the data it needs to obtain from various inputs, for example the Xsens, is incomplete or not present.

After interpolation, the lag between the Unity and MVN motion data of the participant was determined by computing either finding the time lag for which the Root Mean Square Error, using formula 1, was the smallest (RMSE) or by finding the time lag that corresponded to the maximum value for the cross-correlation of the two signals using formula 2.

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}\left(x_{mvn,t}-x_{unity,t}\right)^2}{N}} \quad (1)$$

$$\phi_{xy}(\tau) = E[x(t-\tau)y(t)] \quad (2)$$

The lag associated with the highest correlation between the two signals was chosen and the MVN signal was cropped from this sample onwards. The region of interest in this experiment was determined to be from the moment when the first car was approximately 35 meters away from the participant until the moment when the fourth car passed the participant because this time window contained the motion information where the participant prepared him/herself for crossing the road (a.i., moving closer to the curb) as well as the temporal information of the previously mentioned crossing opportunity. Therefore, the MVN data and Unity data were cropped to contain only the information of the region of interest. After cropping, the motion data were low-pass filtered using a zero phase 10^{th} order Butterworth filter with a cut-off frequency of 8 Hz [31], except for the angular velocity data of the head which was low-pass filtered using a zero phase 10^{th} order Butterworth filter with a cut-off frequency of 3 Hz. This 3Hz was inferred from Grossman et al. [29] who showed that the head rotation frequency in the horizontal plane during walking did not exceed 1 Hz and during running did not exceed 2.7 Hz. The predominant frequency during walking was shown to be 0.8 Hz while during running it was shown to be 1.5Hz. Since the participants in our experiment performed a gait which was not similar to running, but more to walking fast, we used the cut-off frequency of 3 Hz to ensure that a partial band of the rotation frequencies found in running were incorporated in the signal. The Unity data of the cars were filtered with a zero-phase moving average filter with a cut-off frequency 8 Hz. The window width of the zero-phase moving average filter was the average duration of the steps in the piece-wise continuous signal.

Appendix G: Results Opening Questionnaire

Question	Strongly disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
Block B					
I walk for the pleasure of it	4.2	12.5	29.2	41.7	12.5
I walk because it is healthy	0.0	29.2	12.5	50.0	8.3
In short trips I prefer walking over other modes of transportation	0.0	12.5	20.8	37.5	29.2
Block C I					
Crossing roads is difficult	41.7	33.3	16.7	8.3	0.0
Crossing roads outside designated locations increases the risk of an accident	4.2	16.7	4.2	54.2	20.8
Crossing roads outside designated locations is wrong	12.5	37.5	20.8	25.0	4.2
Crossing roads outside designated locations saves time	4.2	12.5	12.5	41.7	29.2
Crossing roads outside designated locations is acceptable because other people do it	29.2	50.0	12.5	4.2	4.2
Block C II					
I prefer routes with signalised crosswalks	8.3	25.0	25.0	20.8	20.8
I try to make as few road crossings as possible	8.3	16.7	33.3	37.5	4.2
I try to take the most direct route to my destination	0.0	8.3	8.3	54.2	29.2
I try to take the route to my destination on which I encounter the least amount of traffic	8.3	41.7	20.8	20.8	8.3
I am willing to make a detour to find a protected crossing	8.3	62.5	16.7	12.5	0.0
I am willing to take any opportunity to cross	4.2	25.0	29.2	41.7	0.0
I am willing to make dangerous actions as a pedestrian to save time	25.0	33.3	33.3	8.3	0.0
Block D					
I am less likely to be involved in a road crash than other pedestrians	0.0	12.5	41.7	29.2	16.7
I am faster than other pedestrians	0.0	12.5	25.0	37.5	25.0
I am more careful than other pedestrians	0.0	12.5	58.3	25.0	4.2

Results of the questionnaire continue on next page

Question	Strongly disagree	Disagree	Neither Agree Nor Disagree	Agree	Strongly Agree
Block F					
Drivers are not respectful to pedestrians	12,5	33,3	41,7	12,5	0,0
Drivers are too fast	4,2	45,8	33,3	16,7	0,0
Drivers are aggressive and careless	8,3	50,0	29,2	12,5	0,0
Drivers should always give way to pedestrians	20,8	37,5	16,7	4,2	20,8
In case of an accident, it is the driver's fault most of the times	12,5	20,8	33,3	25,0	8,3
I let a car go by even if I have right of way	0,0	33,3	33,3	25,0	8,3

Appendix H: Centre of Mass - p-value plots

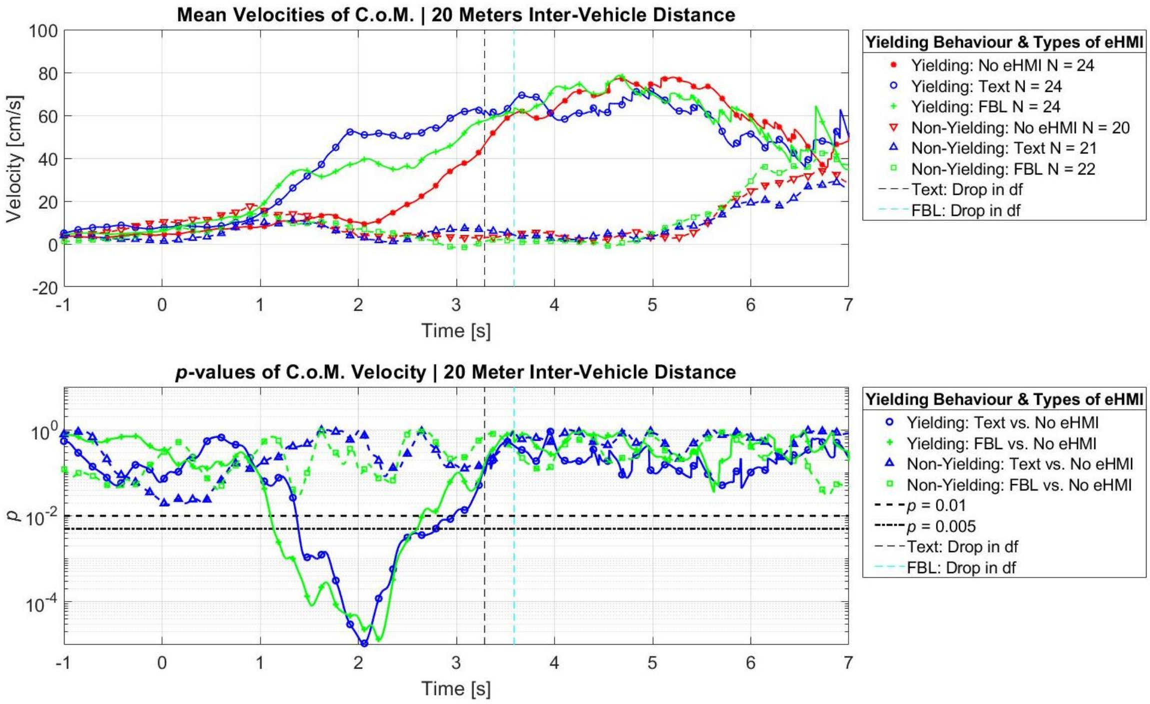


Figure 26: Mean C.o.M. velocities of the participants and corresponding p-values over time for the conditions with independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: Greater mean C.o.M. velocities were found between the 1st and 3rd second from when a stimulus car with an eHMI started braking. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the C.o.M. velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.

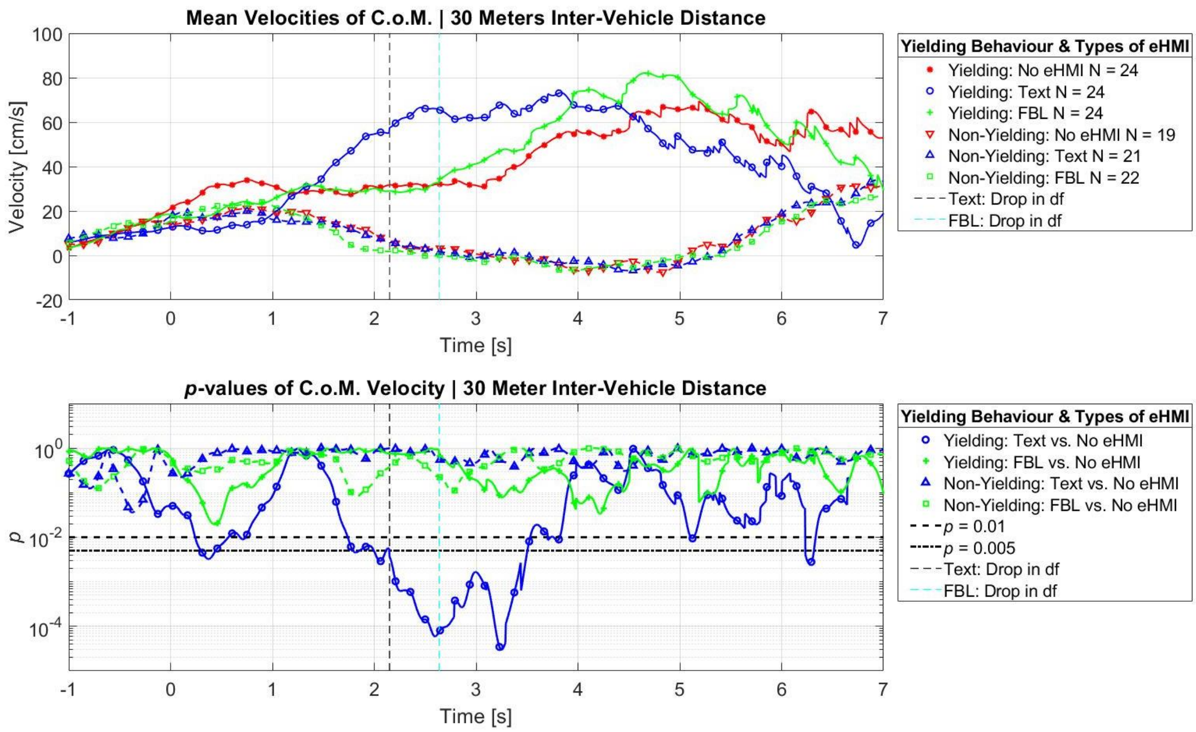


Figure 27: Mean C.o.M. velocities of the participants and corresponding p-values over time for the conditions with the independent variable 30 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: Between the moment the stimulus car started braking and the 1st second thereafter, the C.o.M. velocities in the absence of an eHMI were shortly significantly greater than the C.o.M. velocity in the presence of a text eHMI. However, C.o.M. velocities in the presence of a text eHMI were shown to be significantly greater after the second second. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference and the intersections of the boundaries with the varying p-values were used to estimate the moment when the C.o.M. velocities of the participants differed significantly.

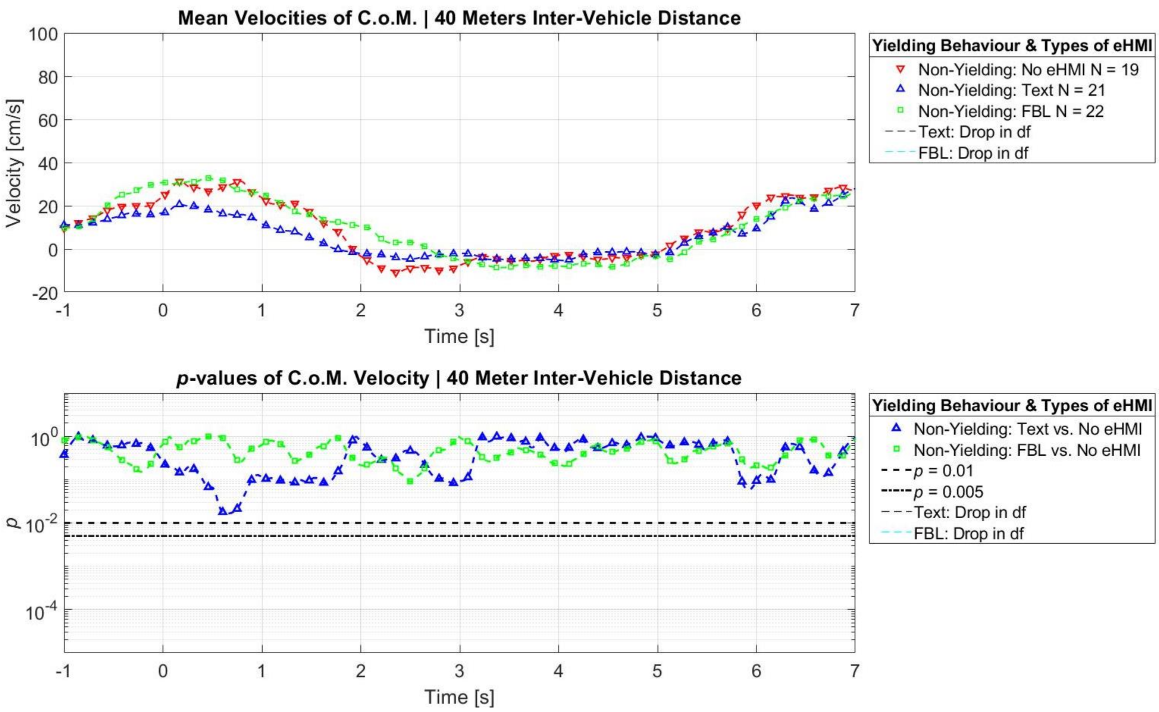


Figure 28: Mean C.o.M. velocities of the participants and corresponding p-values over time for the conditions with the independent variable 40 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. No significant differences were found between conditions in which an eHMI was present and in which one was absent.

Appendix I: Angular Velocities of the Head – p-value plots

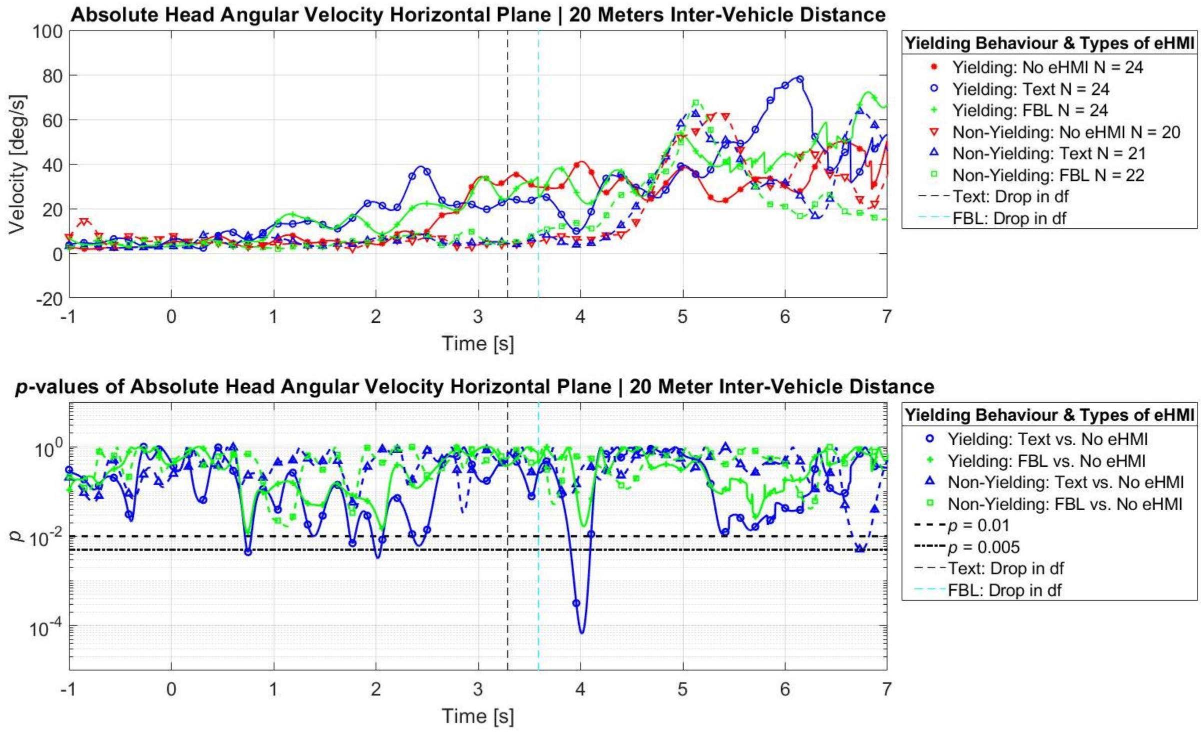


Figure 29: Absolute angular velocities of the participants and corresponding p-values over time for the conditions with the independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: Greater absolute mean angular velocities were found for multiple temporal events from the moment the stimulus vehicles with eHMI started braking to the 7th second thereafter. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the absolute mean angular velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.

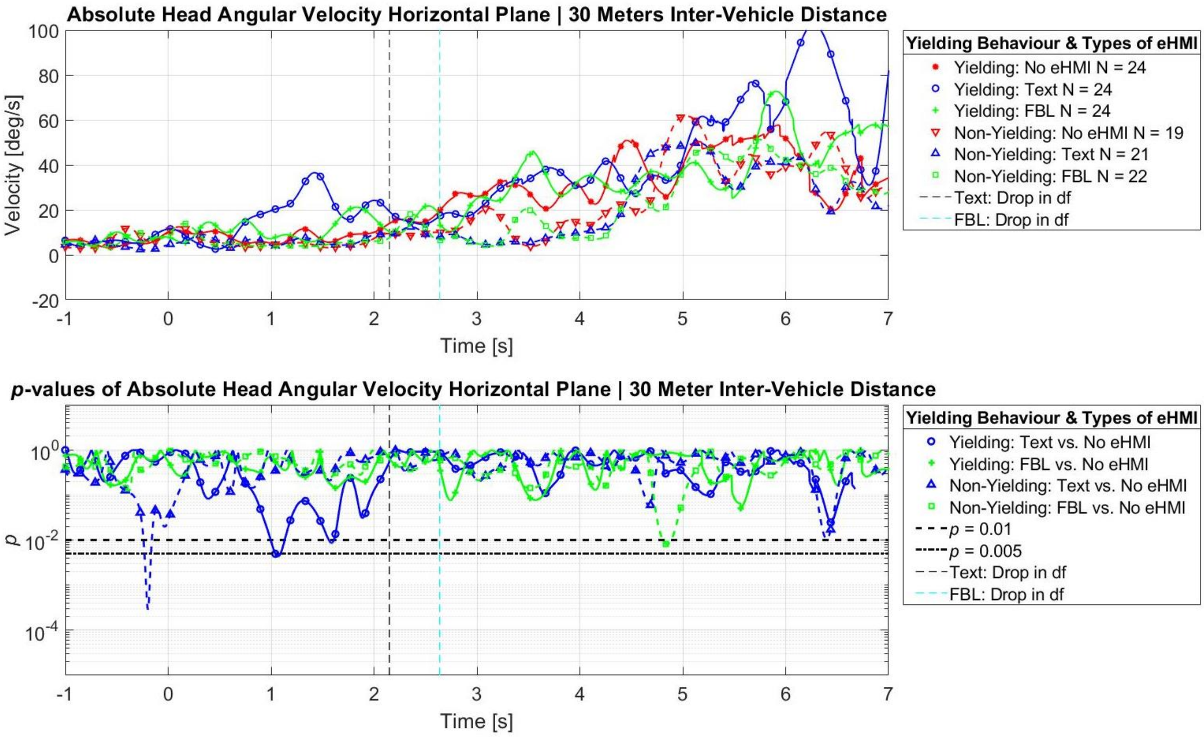


Figure 30: Absolute angular velocities of the participants and corresponding p-values over time for the conditions with the independent variable 30 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: Greater absolute mean angular velocities were found for multiple temporal events from the moment the stimulus vehicles with eHMI started braking to the 7th second thereafter. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the absolute mean angular velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.

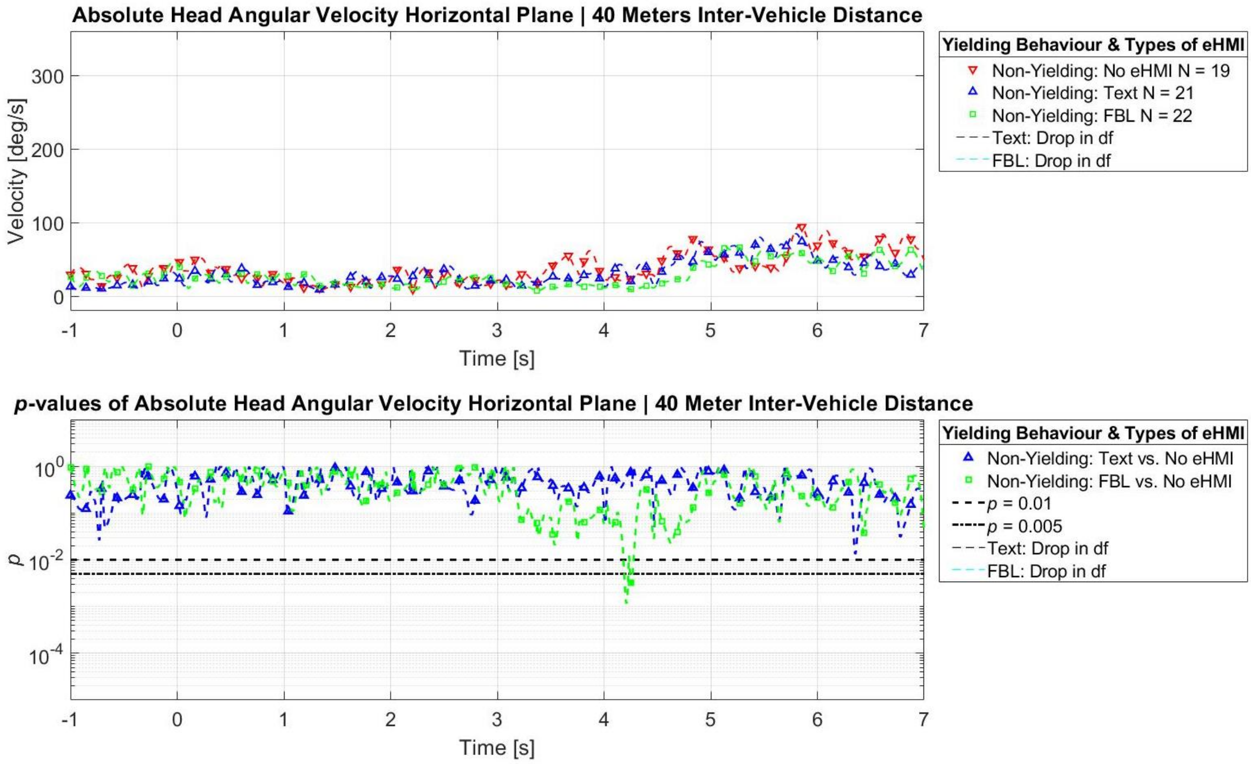


Figure 31: Absolute angular velocities of the participants and corresponding p-values over time for the conditions with the independent variable 40 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend.

Appendix J: Analysis of Thorax Angle and Ankle Accelerations

An exploratory analysis was done on the thorax angle and ankle accelerations of the pedestrians in order to see if these could be indicators of hesitation behaviour, since our attempt to analyse the gait initiation of pedestrians was unsuccessful. The following Appendix has been subdivided into the following three sections, namely:

1. The Analysis of the Thorax Angle
2. The Analysis of the Ankle Accelerations
3. Interpretation

The first and second section contains a brief hypothesis with a qualitative motivation, a concise summary of the method of analysis and a description of the found results on the respective subject. In the third section an interpretation of the results from the first and second section can be found.

The Analysis of the Thorax Angle

The thorax angle in respect to the x-axis was estimated to be an interesting indicator of hesitation in terms of gait initiation, since the thorax is part of the first body group to start moving. This might not directly be visible from the results of Kalantarov et al. [16], which can be found in Figure 9, who showed that this group contains only the shoulders and the head. However, the thorax is situated between the shoulders and rotates in accordance with the shoulders. Moreover, the thorax might be a great suitor for the proposed hesitation analysis as it is unilateral like the head but is not confounded by factors such as gaze direction or walking speed. Furthermore, the benefit of analysing the thorax instead of the shoulders can be exemplified by the fact that shoulder motion is confounded, for example, by arm motion and differs depending on which side of the body one is applying the analysis.

We hypothesized that in the yielding conditions in which an eHMI was present, the angle of the thorax relative to the X-axis would decrease earlier in time than it would in the conditions in which an eHMI was absent. We expected to find no significant differences between conditions in which the vehicles did not yield. The underlying reasoning behind this hypothesis is that the pedestrians would be able to, due to the message conveyed by the eHMI, decide to cross the road earlier in time and thus rotate their body, from the direction of the oncoming vehicles they were observing, towards the direction in which they were going to walk (i.e., the zebra crossing). This would result in an objective measure of certain or approaching behaviour as compared to uncertain or avoiding behaviour which could be interpreted as such that the pedestrian might continue to observe the oncoming vehicles searching for clues which would help their decision-making of whether to cross or not. The thorax angle was obtained from the joint T8 from the biomechanical model used in MVN Analyze. These data were low-pass filtered using a zero phase 10th order Butterworth filter with a cut-off frequency of 8Hz. Utilizing the same methods as in the analysis of the Centre of Mass velocities and angular velocities of the head, paired sample t-tests were conducted for every time sample of the thorax angle relative to the x-axis. The x-axis, as defined in the method subsection “experimental environment”, was parallel to the walking direction of the pedestrians on the zebra crossing (i.e. the longitudinal direction).

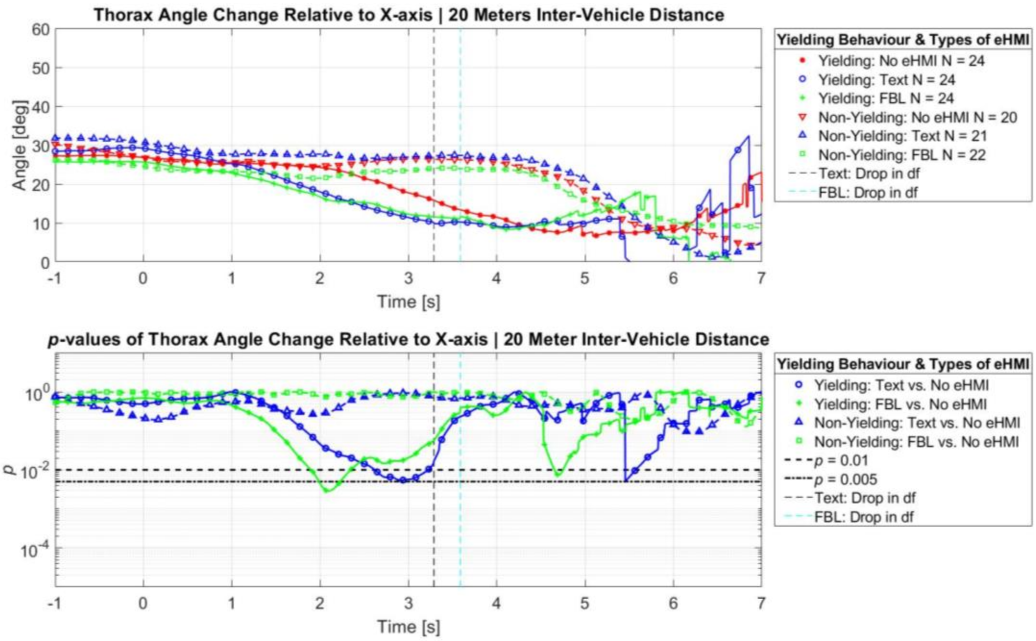


Figure 32: Mean thorax angle of the participants and corresponding p-values over time for the conditions with the independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: In the yielding conditions in which the stimulus vehicle had an eHMI, participants showed lower mean angles relative to the axis, between the 1st and the 4th second after braking, compared to when the stimulus vehicle was devoid of an eHMI. In the conditions in which the vehicles did not yield, the trends of the thorax angle follow similar patterns compared to each other. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the mean thoracic angle of the participants differed significantly when comparing between the presence and absence of an eHMI.

Only the conditions with independent variables 20 and 30 meters inter-vehicular distance were included in the analysis. The conditions with the independent variable 40 meters inter-vehicular conditions only included trials where the vehicles did not yield and from the analysis of the C.o.M. and angular velocities of the head it was concluded that this conditions was uninformative.

For the non-yielding conditions with the independent variable 20 meters inter-vehicular distance no significant differences were found when comparing between the presence and absence of an eHMI. The angle of the thorax seemed to follow a similar trend in the presence and the absence of an eHMI on the stimulus vehicle. However, for the conditions in which the vehicles did yield, the mean thorax angle in the presence of eHMIs deviated from the mean thorax angle when there was no eHMI present on the stimulus vehicle which can be seen in Figure 32. For the Frontal Braking Lights, a brief period of strong significant difference was found two seconds after braking, while for the Text eHMI a short period of weak significant difference was found before the third second after braking.

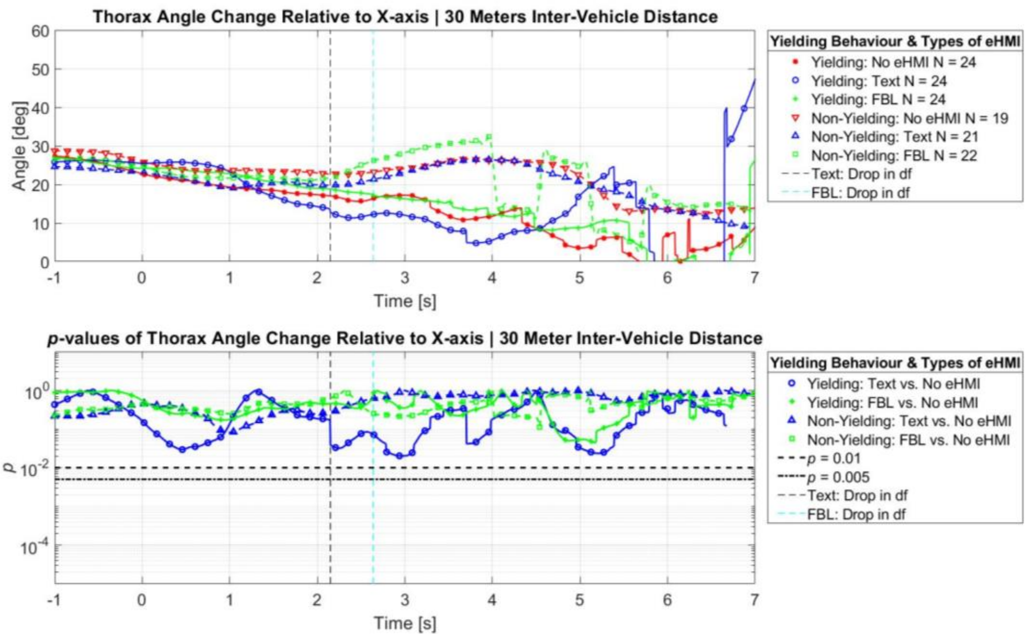


Figure 33: Mean thorax angle of the participants and corresponding p-values over time for the conditions with the independent variable 30 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: In the yielding conditions in which the stimulus vehicle was devoid of an eHMI, initially lower mean angles relative to the axis were found, prior to the moment of braking until the 1st second after braking, compared to when the stimulus vehicle had an eHMI. In the conditions in which the vehicles did not yield, the trends of the thorax angle follow similar patterns compared to each other. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the mean thoracic angle of the participants differed significantly when comparing between the presence and absence of an eHMI.

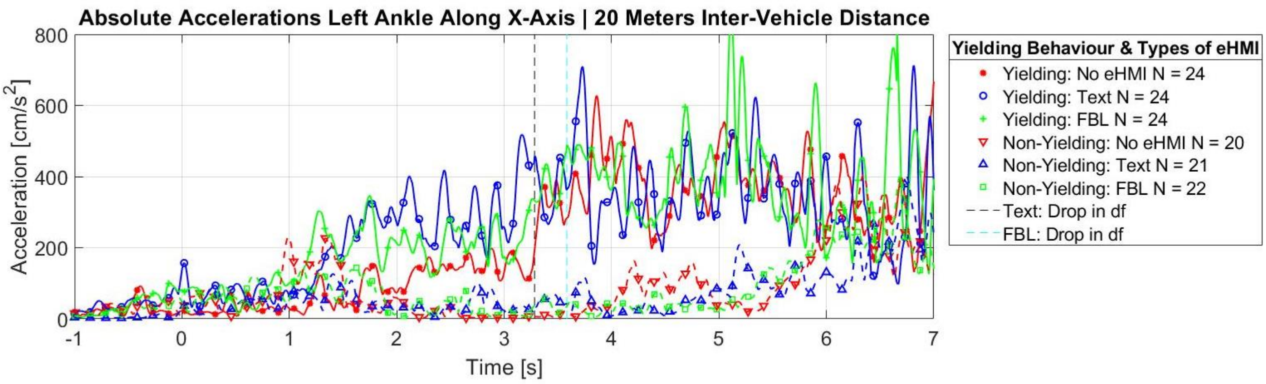
For the conditions with the independent variable of 30 meters inter-vehicular distance, no significant differences were found both when the vehicles did and did not yield. Figure 33 shows a variety of jumps in the mean value after three seconds of braking. This effect occurred due to the fact that the number of participants incorporated in the mean at those points in time changed since participants had already finished crossing the road. Similar effects can be found in Figure 32 from about five seconds after braking.

The Analysis of the Ankle Accelerations

Initially, we attempted to use the position of the ankles to determine motion onset and thereby investigate the effects of eHMIs on successful and failed gait initiations. However, this turned out to be more complicated than expected due to the behavioural variability of our participants. Nonetheless, knowing from the analysis of the C.o.M. velocity that there was an effect of eHMI on average gait speed we deemed it worthy to analyse the motion data of the ankles via a different approach than previously proposed.

Kalantarov et al. [16] found that the ankle of the stance leg was the last body part group to move in the motion sequence prior to gait initiation, but there was no information in their study on the ankle of the swinging limb. However, Figure 17 in the Result section of this thesis showed that while the position of the participants' head remained relative constant, the ankle of the left leg changed. This information led to the idea that the swinging ankle could potentially be more informative regarding the hesitation behaviour of a pedestrian rather than searching for the motion onset of the stance leg ankle, used to define an attempt to initiate gait, and thereby analyse the successful and unsuccessful gaits as a measure of hesitation.

We hypothesized that in the yielding conditions in which an eHMI was present, the acceleration of the ankle in the direction of the X-axis would increase earlier in time than it would in the conditions in which an eHMI was absent. Furthermore, we expected to find no significant differences between conditions in which the vehicles did not yield. Since no information was recorded regarding the participants' preference of leg, both the left and the right ankle were investigated. The ankle accelerations were obtained by differentiating the velocities from the ankle joints from the biomechanical model used in MVN Analyze. These velocity data were first rectified, then low-pass filtered using a zero phase 10^{th} order Butterworth filter with a cut-off frequency of 8Hz, then differentiated and lastly rectified again to obtain the accelerations of the ankles. Utilizing the same methods as in the analysis of the angle of the thorax, Centre of Mass velocities and angular velocities of the head, paired sample t-tests were conducted for every time sample of the ankle accelerations in the direction of the x-axis. The results were grouped according to which side of the body the ankle was on, where the sides were derived from the perspective of the participant (i.e., left ankle on the participants' left hand-side).



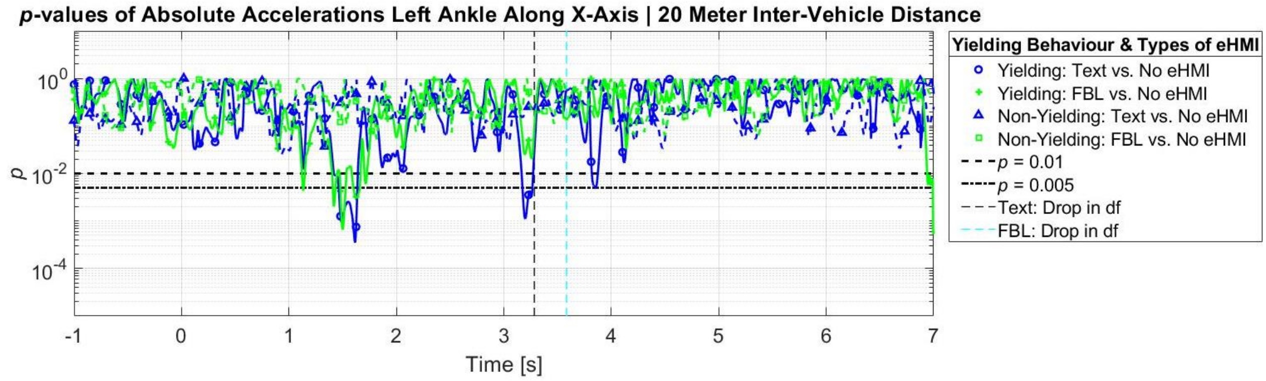
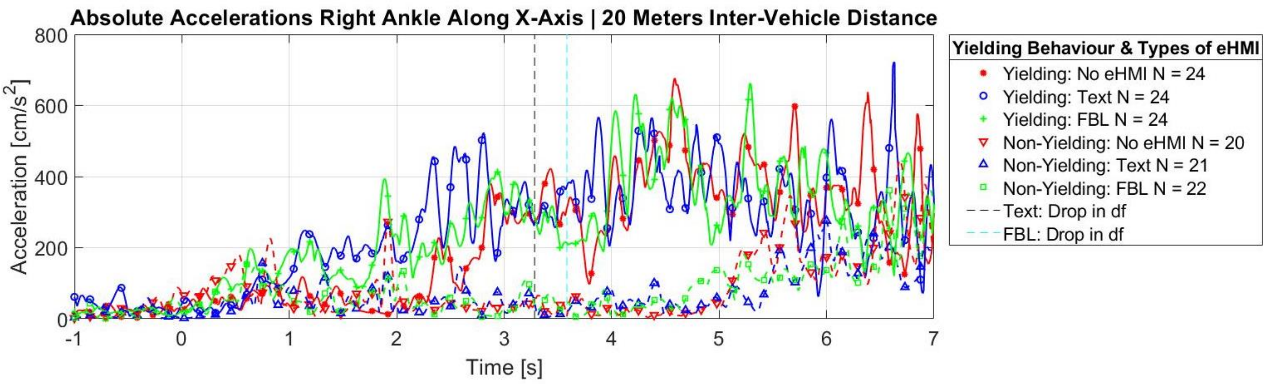


Figure 34: Mean absolute accelerations of the participants' left ankle and corresponding p-values over time for the 1conditions with the independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: In the yielding conditions in which the stimulus vehicle had an eHMI, periods of significantly greater mean absolute accelerations in the direction of the axis were found, between the 1st and 2nd second after braking, compared to when the stimulus vehicle was devoid of an eHMI. In the conditions in which the vehicles did not yield, no significant differences were found when comparing between the presence and absence of an eHMI. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the absolute mean angular velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.



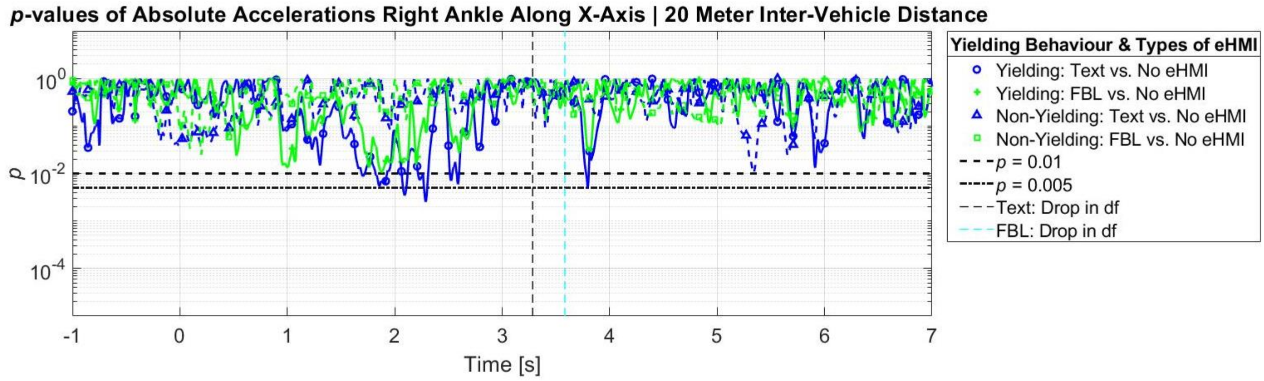


Figure 35: Mean absolute accelerations of the participants' right ankle and corresponding p-values over time for the conditions with the independent variable 20 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: In the yielding conditions in which the stimulus vehicle had an eHMI, periods of significantly greater mean absolute accelerations in the direction of the axis were found, between the 1st and 2nd second after braking, compared to when the stimulus vehicle was devoid of an eHMI. In the conditions in which the vehicles did not yield, no significant differences were found when comparing between the presence and absence of an eHMI. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the absolute mean angular velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.

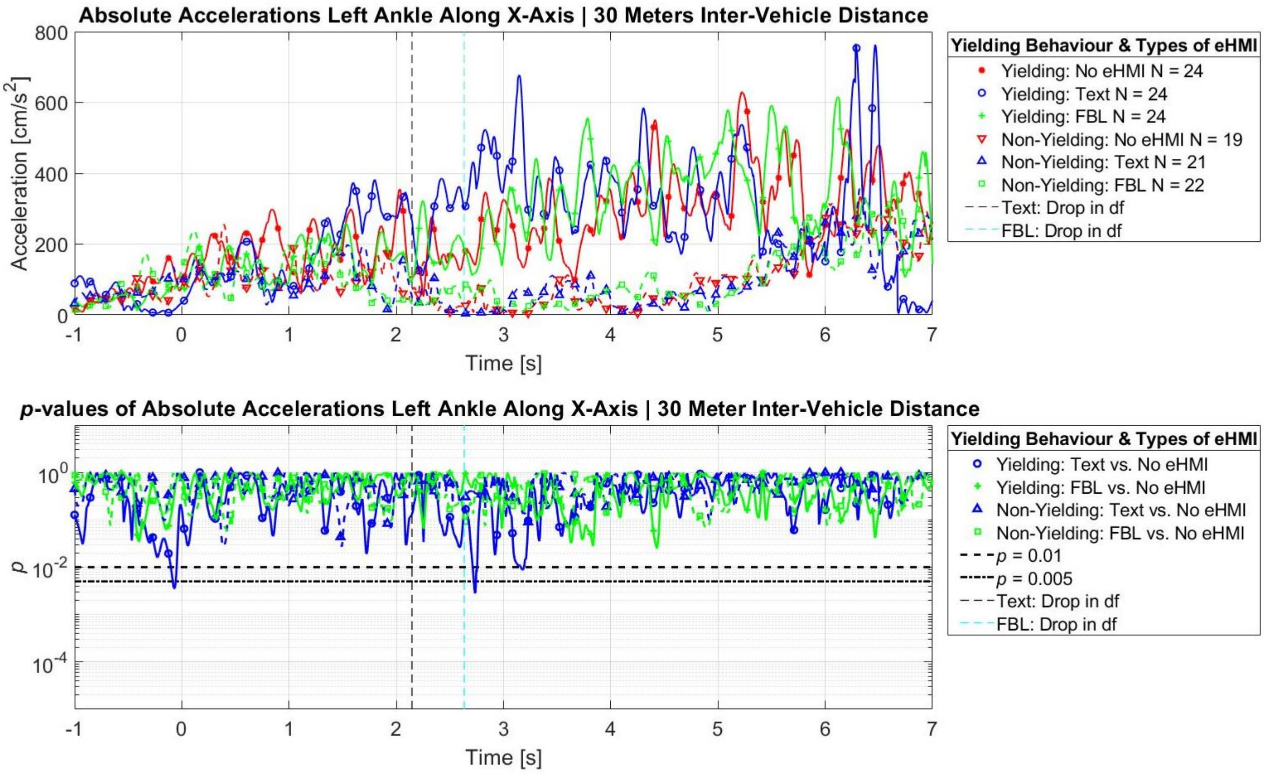


Figure 36: Mean absolute accelerations of the participants' left ankle and corresponding p-values over time for the conditions with the independent variable 30 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: In the yielding conditions in which the stimulus vehicle had an eHMI, brief periods of significant differences in mean absolute accelerations in the direction of the axis were found, prior to braking and between the 2nd and 3rd second after braking, compared to when the stimulus vehicle was devoid of an eHMI. Prior to braking the mean absolute accelerations were lower in the presence of a text eHMI compared to when there was no eHMI on the stimulus vehicle. Two seconds after braking, the absolute mean accelerations were greater in the presence of an eHMI compared to when there was no eHMI on the stimulus vehicle. In the conditions in which the vehicles did not yield, no significant differences were found when comparing between the presence and absence of an eHMI. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the absolute mean angular velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.

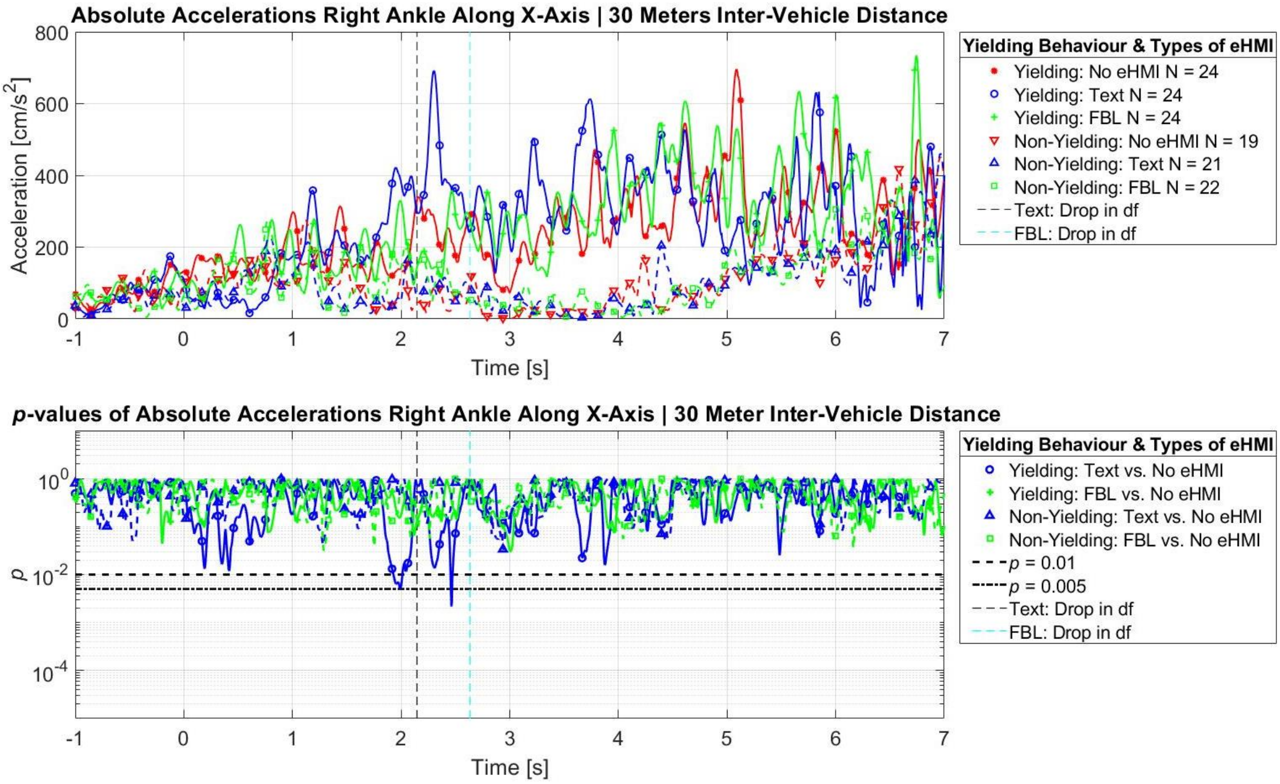


Figure 37: Mean absolute accelerations of the participants' right ankle and corresponding p-values over time for the conditions with the independent variable 30 meters inter-vehicular distance. Abbreviations: Text = Text eHMI, FBL = Frontal Braking Lights, N = Sample size at T = 0. T = 0 is the moment where the stimulus car started braking. The drop in df represents the moment when the amount of people incorporated in the analysis dropped by one from the initial sample size indicated in the legend. Top graph: In the yielding conditions in which the stimulus vehicle had an eHMI, a brief period of significant difference in mean absolute acceleration in the direction of the axis was found, between the 2nd and 3rd second after braking, compared to when the stimulus vehicle was devoid of an eHMI. The mean absolute accelerations were higher in the presence of a text eHMI compared to when there was no eHMI on the stimulus vehicle. Bottom graph: The horizontal boundaries correspond to p-values for the level of weak and strong significant difference. The intersections of the strong significant difference boundary with the varying p-values were used to estimate the moment when the absolute mean angular velocities of the participants differed significantly when comparing between the presence and absence of an eHMI.

In the conditions in which the vehicles did not yield, and the inter-vehicular distance was either 20 or 30 meters, no significant differences were found when comparing between the presence and the absence of an eHMI. However, periods of significant differences were found in the conditions where the vehicles did yield. Specifically, Figure 34 shows that, for the left ankle for the condition with the independent variable of 20 meters inter-vehicular distance, a substantial period of significantly greater mean absolute accelerations in the direction of the axis was found, between the 1st and 2nd second as well as between the 3rd and 4th second after braking, for when the stimulus vehicle had a Text eHMI compared to when the stimulus vehicle was devoid of an eHMI. In all other cases, only peaks of significant differences were found.

Interpretation

Even though no strong evidence, in the quantitative form of continuous bands of significant differences, was found to support the hypothesis of both analyses, the results were nonetheless informative in a qualitative manner. Looking at the findings obtained from the analysis of the thorax angle for the yielding conditions with the independent variable 20 meters inter-vehicular distance, it can be seen in Figure 32 that the thoracic rotation towards the x-axis for both cases in which there was an eHMI present occurred at an earlier point in time compared to the thoracic rotation towards the x-axis when the stimulus vehicle did not have an eHMI. In other words, from the graph it can be derived that the participants' reaction time to start crossing was faster for the conditions where an eHMI was present compared to the condition in which there was not one present. This interpretation corroborates with our quantitative findings on, and interpretation of, the C.o.M. velocities for similar conditions. Furthermore, the change in thoracic angle for the yielding conditions with the independent variable of 30 meters inter-vehicular distance shows that on average participants reacted to the Text eHMI one second after braking as can be seen in Figure 33. For the condition where there was no eHMI present, participants, on average, started their thoracic rotation prior to braking but kept it at a relative constant angle for period between the 1st and the 2nd second after braking. This behaviour can be interpreted as such that the gap of 30 meters between the second and the third vehicle influenced the crossing decision of participants in the case where they saw that there was no eHMI on the front of the vehicle.

In a temporal context, it is more efficient for pedestrians crossing between two vehicles to cross directly after the first vehicle has passed in front of them. In that case, the pedestrian has the largest safety margin and thereby the longest period available to them to cross the road. In the case where the inter-vehicular distance was 20 meters, participants apparently disregarded the option to cross between cars. The observed difference in the case where the inter-vehicular distance was 30 meters seems to underlie the gap accepting behaviour of pedestrians. However, on average, participants refrained from pushing their decision since it was not completely clear whether the stimulus vehicle was going to yield or not and therefore the thoracic angle remained at a relative constant angle during the 1st and 2nd second after braking. In the condition where there was a Text eHMI present, participants waited until they observed the eHMI to start conveying the message that participants could walk before they started their thoracic rotation (i.e., about 1 second after braking).

This behavioural approach can also be deduced from the absolute mean accelerations of the ankles. For example, Figure 37 shows that the trend of the absolute mean acceleration of the right ankle, in the yielding condition with the independent variable of 30 meters inter-vehicular distance where an Text eHMI was present, initially increased, decreased and then increased again between T = -1 and T = 1, while Figure 36 shows a similar trend for the left ankle but one that is shifted a few hundred milliseconds along the x-axis. This trend could be interpreted as such that pedestrians started to initiate their gait, but refrained from continuous bipedal locomotion, or even aborted their gait cycle, due to the fact that no confirmatory signal was perceived from the Text eHMI. Looking at the trends of the absolute mean accelerations of the ankles between T = -1 and T = 1, for the same independent variable of 30 meters inter-vehicular distance, but with the presence of Frontal Braking Lights or the absence of an eHMI, it can be observed that the trends in both of these cases are quite similar to each other and different to the case there a Text eHMI was present. In both of these cases, it was not clearly visible what the intention of the stimulus vehicle was, due to the display quality of the Oculus Rift. More specifically, the external appearance of the stimulus vehicle with Frontal Braking Lights indicating it is not yielding is quite similar to the external appearance of the non-yielding stimulus vehicle without any external display. Pedestrians saw that there was no Text eHMI present that would unambiguously inform them of the intention of the stimulus vehicle and therefore started searching for other cues indicating the yielding behaviour of the stimulus vehicle. The fact that not even weak significant differences were found between the presence of Frontal Braking Lights and the absence of an eHMI in this condition seems to corroborate with this line of reasoning. However, the peaks of significant differences we found when comparing between the Text eHMI and the absence of an eHMI could have arisen due to the rectification of the acceleration data. Moreover, due to the fact that the participants' preference of leg was not recorded in our study, it was not possible to sort out which limb was the swing limb and which the stance limb and thereby sort out which participants to include in the analysis of, for example, the left ankle. Therefore, the interpretation of the analysis of the absolute mean ankle accelerations remains rather tentative by nature.

Concluding, our findings on, and interpretation of, the C.o.M. velocities for similar conditions seem to corroborate with the interpretation of the analysis on the thoracic angle and absolute mean accelerations of the ankles. Even though no compelling evidence was found to support the hypotheses of this exploratory analysis, future analyses on the motion of various body parts could uncover robust objective measures related to hesitation behaviour of pedestrians.

Appendix K: Link to data and scripts

The following link contains the data of every participant included in this study in the following formats: .MVN, .MVNX and .TXT. All Matlab scripts, including the history of revisions, have been included in a folder found under this link. Furthermore, the data manipulated in Matlab have also been stored for every step taken during post-processing.

https://www.dropbox.com/sh/bsdnvh0h7l4lmmk/AADUpptHW8cdtoSMPjXlHKqLa?dl=0

A visualization of some of the conditions presented to the participants and examples of how the trials were conducted in this study can be found by following the YouTube links below:

1. Example of a trial with non-yielding and yielding vehicles: https://www.youtube.com/watch?v=iigSHEmPBPw
2. Examples of the different interfaces used in the experiment and how they differed in yielding and non-yielding conditions: https://www.youtube.com/watch?v=lDglly8hQFo