Eye-tracking for identification of effective External Human Machine interfaces

S. van Bergen (4585933),
S.M. ter Brake (4275446),
M.T.T. Hensen (4573293),
W.J. Tempelaar (4378385)
Cognitive Robotics
University of technology Delft
Delft, the Netherlands

s.vanbergen@student.tudelft.nl
s.m.terbrake@student.tudelft.nl
m.t.t.hensen@student.tudelft.nl
w.j.tempelaar@student.tudelft.nl

Y.B. Eisma,
J.C.F. de Winter
Cognitive Robotics
University of technology Delft
Delft, the Netherlands

Abstract—

This paper aims to find the influences of external Human Machine Interfaces (eHMI's) mounted on an automated vehicle on different positions under variable stresses on the visual attention and interpretation speed a subject acting on said eHMI.

An eye-tracking experiment was set up showing 36 different 3D animations consisting of 6 scenarios with 3 levels of stress and 6 different eHMI setups per scenario: displays on the grill, roof, and above the wheels; projection on the windscreen and on the road; and no eHMI. This eHMI would show 'Waiting' when stopping and 'Driving' when continuing to drive combined with appropriate symbols. While a video was playing the participant was asked to press and hold the spacebar when they felt like it was safe to cross. After each video there was a question about the clarity of the situation. The eyes of each participant were tracked during each video.

The majority of the eHMI's performance scores are close together. The automated vehicles without any eHMI perform the worst, followed up by the eHMI positioned above the wheels. On the second place is the projection on the road. The eHMI on the roof, on the windscreen and on the grill have tied as the most effective place for an eHMI on an automated vehicle. Moreover, as the stress level increases, the effectivity of the eHMI's gets worse compared to the effectivity when having no eHMI.

Keywords—Autonomous vehicle, communication, eHMI, eye-tracking, traffic

I. Introduction

The realisation of autonomous vehicles (AV's) is getting ever closer. With each new generation of cars it seems like more and more of tasks that used to be the responsibility of the driver are handled instead by automatic systems in the car itself. The popularity of developments like adaptive cruise control, lane assist, recognition of traffic signs and many others show that cars do not only drive more autonomously than they used to, but they will most likely also continue to become more and more autonomous in the future.

However, this increase in autonomy brings its fair share of problems, many of which have nothing to do with the technical aspects of the innovation. One of the big problems this development faces is related to communication between road users. In normal traffic, drivers often use non-verbal communication to give signals to other road users which improves behaviour in traffic (Kitazaki 2015). This form of communication will disappear when autonomy of cars increases and car passengers will no longer take active part in traffic. If this form of communication is not replaced, a decrease of efficiency and potentially even safety in traffic can be expected, especially as it relates to Vulnerable Road Users (VRU's) like cyclists or pedestrians (Vlakveld 2018).

1.1 external Human Machine Interface

A large amount of options have been considered to substitute this form of communication, many of which take the form of external Human-Machine Interfaces (eHMI's). An eHMI is any signal that a machine, in this case an autonomous car, gives off to communicate a message directly to a human. These signals can take many forms. An overview of many possibilities in this field is given by Sergiu C. Stanciu (Stanciu 2018).

Communication is important in traffic, in particular between drivers of cars and Vulnerable Road Users (VRU's). (Kitazaki 2015, Pillai 2017, Maag 2012). To add to this, it has also been shown that autonomous vehicles are generally perceived to be less trustworthy than normal cars with a driver. However, when an autonomous vehicle uses an eHMI to communicate with the outside world, it is generally perceived to be more trustworthy than a regular car being driven (Vlakveld 2018).

Many car manufacturers and independent researchers have suggested ways to design an eHMI to be as effective as possible in real traffic situations. Several studies agree that an AV should communicate their own intent whenever possible, rather than giving a VRU a command or suggestion (Vinkhuyzen 2016, Schaudt 2018). It has also been shown that a VRU values knowing that they have been seen by the AV and that acknowledging a VRU can already make an AV come across as more trustworthy (Florentine 2015, Schaudt 2018).

To make the intended communication possible, there are several aspects to take into account: crafting a comprehensible signal, drawing the attention of a VRU, and finally the location of the eHMI.

‘Among the eHMIs presented by the industry, the eHMIs that combined (simple) textual instructions (‘Go ahead’, ‘Waiting for You to Cross’, ‘Safe to Cross’) with icons (‘>’ sign/eyes, pedestrian on zebra) received the highest clarity ratings’ (Bazilinskyy, 2019). For this reason the decision was made to use this type of eHMI in the experiment.

Drawing someone's attention is achieved well by using a snapping motion, for example the quick change of a sign from stop to go (Bockler 2014). This means that it would be a good idea to turn the signal on abruptly the moment communication is required.

As for the final aspect, research has shown that there are several locations on an AV where an eHMI could be effective. When looking for communication with a car, the windscreen, the grill, top of the vehicle and projecting on the road in front of the vehicle are all possibilities that can grab the attention of a VRU (Ackermann 2019, Bazilinskyy 2019). Furthermore, an experiment showed that VRU’s tend to start by looking at the wheels of a vehicle to see whether the vehicle is in motion (Ter Borg & Foorthuis & Tas & Van Zee 2019). This makes near the wheels an interesting location to place an eHMI.

1.2 Research question

As mentioned above, research surrounding eHMI’s for AV’s has uncovered many different locations to place an eHMI. However, not much has been done to compare these locations on how well they grab people’s attention and how fast people can interpret their signals, nor to stress test these possibilities. This paper aims to fill this gap in knowledge by answering the following research question:

‘How do participants distribute their visual attention and how are their crossing intentions affected when viewing animations of automated cars with eHMIs mounted at different locations under influence of different stresses and distractions?’

1.3 Experimental Setup

The different locations mentioned in section 1.1 will be tested in different traffic scenarios to understand how they compare under different levels of stress. A test subject will be asked to indicate when they think it is safe to cross a road as autonomous vehicles pass by or temporarily stop to let the VRU cross. This road will either be a calm 2-way street, a somewhat busy T-junction or a busier intersection. This traffic scenario will be displayed using a 3D animation where the test subject is shown the perspective of a pedestrian. During the experiment the test subjects eye positions will be tracked. The eHMI used will consist of passive text combined with a symbol as proposed in section 1.1.

II. Study aim and Hypotheses

The aim of the research is to get a better understanding of the possibilities regarding locations of external Human Machine Interfaces on autonomous vehicles under variable stresses. In particular how they relate to the communication between an AV and a VRU. Therefore, the research question is as follows:

‘How do participants distribute their visual attention and how are their crossing intentions affected when viewing animations of automated cars with eHMIs mounted at different locations under influence of different stresses and distractions?’

By using an eye-tracker as well as a simulated 3D traffic environments, which will be elaborated upon in section 3, this research aims to find correlations between a person’s behaviour in traffic and the location of an eHMI on a car. Several hypotheses will be tested to draw conclusions from the resulting data. The hypotheses are as follows:

2.1 Hypothesis A - Reaction time

The expected reaction times of the test subjects are plotted in figure 2.1. The square titled “eHMI on” is representative of the moment the eHMI shows “Waiting”. Depending on the effectiveness of the eHMI, this message will be interpreted at different speeds, leading to different reaction times. A faster response signifies easier interpretation of the eHMI.

According to studies done the expectation is that the position will be ranked based on efficiency in the following order:

1. Windscreen
2. Front bumper
3. Roof
4. Above the wheels
5. Projection on the road
6. No eHMI

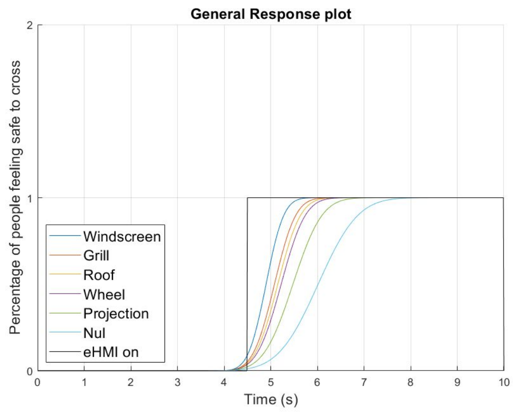


Fig. 2.1. Expected response plot

2.2 Hypothesis B - Varying stresses

As the traffic scenarios vary in complexity, participants will be faced with more distractions like objects or events that draw their focus. It is expected that as complexity increases, the amount of saccades will also increase and the average fixation time will go down as participants will be required to look at more things in the same timespan. As complexity increases it is also expected that the performance of participants, how often they make the correct crossing choice, will decrease since they will not be able to focus on the task at hand as effectively.

2.3 Hypothesis C - Fixations

How long a test subject fixates on an eHMI will correlate with how effective an eHMI is. Longer fixation times indicate that a participant needs more time to interpret the message that the eHMI is giving off which makes the eHMI less effective. Conversely, eHMI's which correspond to lower fixation times are most likely more effective. The ranking of eHMI's by this form of effectiveness is expected to be the same as in hypothesis A, for the same reasons.

III. Methods

3.1 Participants

51 males and 10 females between 19 and 27 years old (Mean = 23.0, SD = 1.83) participated in this research. They were all students of Bachelor and Master studies in the faculty 3ME at the TU Delft. A written informed consent form was signed by all participants before the start of the experiment.

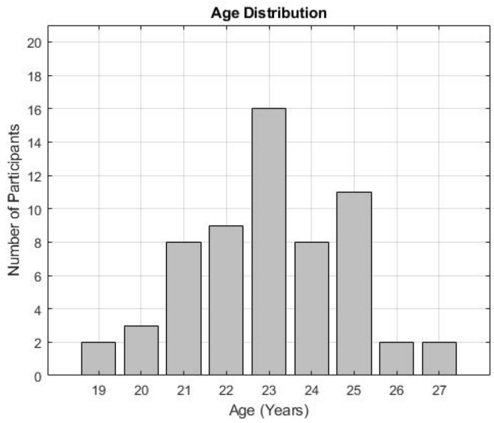


Fig. 3.1. Age distribution of participants

3.2 Apparatus

Eye movements were recorded at 2000 Hz using the SR-Research Eyelink 1000 Plus. Participants were asked to place their head in the head support during the entire experiment. The stimuli were displayed on a 24-inch BENQ monitor with a resolution of 1920 x 1080 pixels (531 x 298 mm). The refresh rate of the monitor was 144 Hz. The distance between the monitor and the head support was 95 cm, and the distance between the eye-tracking camera/IR light source and the head monitor was 68 cm.

3.3 Participant's task

The participant was tasked with watching 36 videos displaying virtual traffic scenarios. While watching, they were tasked to press and hold the spacebar whenever they felt it was safe to cross the road. After each scenario the participants were asked to rate their agreement on a scale from 0-10 with the statement: 'It was clear when I could cross the road' where 0 represented complete disagreement and 10 represented complete agreement.

3.4 Independent variables

Two main independent variables can be identified in this experiment:

- The locations of the eHMI.

As discussed in section 1.1, the different locations on an AV where an eHMI was placed were display screens on the roof, grill and above the wheels and projections on the windscreen and on the road in front of the AV. As control test the experiment also used an AV without any eHMI.

- The traffic scenarios.

As mentioned in section 1.3 the traffic scenarios have been divided into three different situations: a straight road, a T-junction and a intersection. These situations have a different level of stress due to a different level of complexity and traffic density. The straight road is the simplest, the T-junction is busier and more complex and the intersection is the busiest and most complex. Two different traffic scenarios have been created for each situation where the complexity and traffic density have been kept as constant as possible.

3.5 Design of the stimuli

The experiment consists of different virtual traffic scenarios in the form of 3D animations. These animations are 25 seconds long and are played at 60 frames per second. The videos perspective is from the eyes of a pedestrian waiting to cross the road at a crossing with a traffic island. The cars will all have the same eHMI within the same video which can show two different messages: one for if the car decides to stop and one for if the car decides to keep driving shown in figure 3.2. Furthermore the decision has been made to use a white vs a black background to achieve the biggest contrast, mainly to achieve clarity and because colour will not be the main subject of this research.

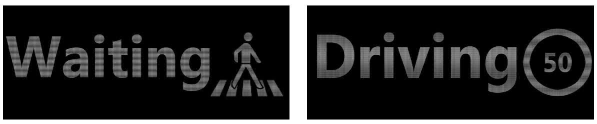


Fig. 3.2. The eHMI message displayed when the car is going to stop (left) or continue driving (right)

There are three different settings in which the scenarios take place: a straight road, a T-junction and a intersection. The environment for each of these scenarios is very similar. They use the same types of buildings, streetlights, backgrounds and other objects. The AV's carrying the eHMI is also kept constant. In other words, during the entire experiment, the participants will be looking at 1 type of car, albeit with a differently placed eHMI. An example of this scenario is shown in figure 3.4. This figure shows a special AV which carries all different types of eHMI as discussed in section 3.3. This "supercar" was not used in experiments, but does illustrate the possibilities within the actual experiment. Each of the 6 trials within each scenario was shifted in time according to table 3.1. The scenarios in this table are according to the table in Appendix B.

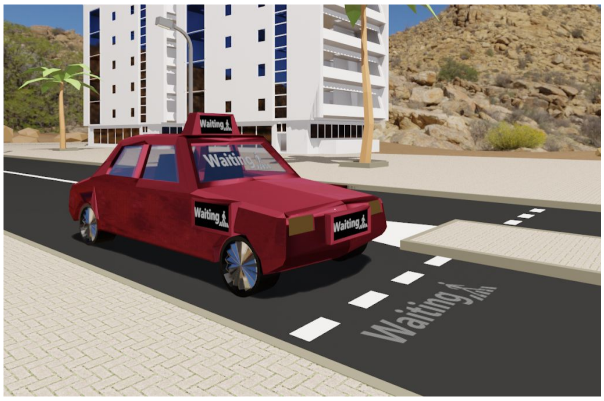


Fig. 3.3. The "supercar" showing all possible eHMI's

Table 3.1. Timeshifts for each different scenario

Scenario 1	Timeshift (s)	Scenario 2	Timeshift (s)	Scenario 3	Timeshift (s)
Grill	0	Grill	0	Grill	0
Projection	+1	Projection	+2	Projection	+1
Roof	+2	Roof	+4	Roof	+2
Wheels	+3	Wheels	+1	Wheels	+3
Windscreen	+4	Windscreen	+3	Windscreen	+4
Nul	+5	Nul	+6	Nul	+5
Scenario 4	Timeshift (s)	Scenario 5	Timeshift (s)	Scenario 6	Timeshift (s)
Grill	0	Grill	0	Grill	0
Projection	+1	Projection	+2	Projection	+1
Roof	+2	Roof	+4	Roof	+2
Wheels	+3	Wheels	+6	Wheels	+3
Windscreen	+4	Windscreen	+8	Windscreen	+4
Nul	+5	Nul	+10	Nul	+5

3.6 Dependent variables

Each participant gives dependent variables obtained through keyboard input and eye-tracker data. The variables used in the analysis are:

- Eyegaze X & Y-coordinates (in pixels)
- Pupil size (in µm)
- Key pressed (boolean)
- Time key pressed (ms)
- Subjective clarity (integer 0-10)

3.7 Data analysis

The dependent variables Eyegaze X & Y-coordinates and Pupil size are acquired in a string of 2000 data points per second. The saccades and fixations of the participants are retrieved by this string and missing gaze points have been interpolated. Plotting the Eyegaze coordinates over the videos will give an understanding of how participants distribute their visual attention during the various scenarios.

To make crossing intentions of the participants clear, the duration of the key press and the status of the spacebar (key pressed or released) are measured. A bonus or malus will be assigned to the output of the participants. The bonus is applicable when the input of the participant is valid, for example: it is safe to cross and the spacebar is depressed. A malus will be applied when the participant has a false input, for example: the spacebar is pressed but crossing the road is unsafe.

The rate of stress is obtained by the subjective clarity and can clarify the fixation duration of the participants.

IV. Results

This section will sum up all different relevant results gathered during the experiments.

4.1 Clarity

Subjective clarity is visualised in figures 4.1 & 4.2. The input with range 0 (= unclear for me to cross) to 10 (= very clear for me to cross) are normalised and plotted in two figures. The first figure shows the clarity over the three road layouts and the second figure is a representation of the clarity per eHMI location. A t-test was done to determine if the differences in clarity were significant for both the eHMI's and the scenarios. In table 4.1 the p-values of these t-tests are displayed for the comparison between the different eHMI's and for comparison between the different scenarios.

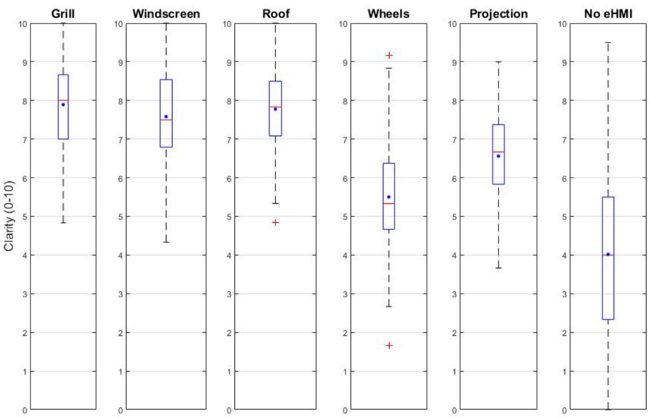


Fig. 4.1. Clarity plot per eHMI

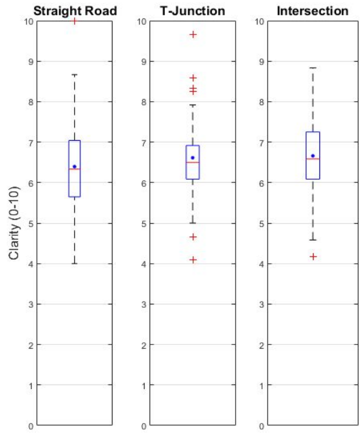


Fig. 4.2. Clarity plot per traffic layout

Table 4.1. t-test p-values clarity (Green cells indicate a significant difference, red cells indicate no significant difference)

	Grill	Roof	Screen	Wheels	Projection	No eHMI
Grill	NaN	2.03E+00	4.10E-01	3.56E-15	2.89E-09	2.39E-17
Roof	2.03E+00	NaN	1.52E+00	3.82E-15	2.31E-08	1.01E-17
Screen	4.10E-01	1.52E+00	NaN	5.16E-12	1.68E-06	3.05E-16
Wheels	3.56E-15	3.82E-15	5.16E-12	NaN	2.53E-05	4.72E-08
Projection	2.89E-09	2.31E-08	1.68E-06	2.53E-05	NaN	2.18E-12
Nul	2.39E-17	1.01E-17	3.05E-16	4.72E-08	2.18E-12	NaN

t-Test (p>0.05)	Straight	T-Junction	Intersection
Straight	NaN	0.06279	0.029
T-Junction	0.06279	NaN	1.57028
Intersection	0.02992	1.57028	NaN

4.2 Crossing behaviour

The plots in figures 4.3 & 4.4 are a normalized score of the performance. The performance is a comparison of the ground truth and the response per eHMI as shown in Appendix B. A bonus and a malus are allocated to the performance score as the response of the participant and the ground truth are compared. There is a bonus per time step of +1 when they agree and a malus of -1 when they do not. As mentioned in section 3.5, each trial has a timeshift. This has to be taken into account while calculating the performance scores. Therefore, a table has been set up with start and end times which mark the boundaries between which data was collected, shown in table 4.2. These boundaries have been determined by looking at where the lines in the plots in appendix B were closest to each other. Figure 4.3 compares the performance scores of different eHMI's and the table 4.3 shows the p-values of the t-test between the different eHMI's. The figure 4.4 shows the comparison of the performance score between the straight road, T-junction and intersection and the table 4.3 shows the p-values of the t-test between the different traffic settings. The figure 4.5 shows the standardized mean difference effect size for within-subject designs (aka cohen's dz), which was calculated via the test-statistics by dividing this value with the square root of the number of participants (= 61).

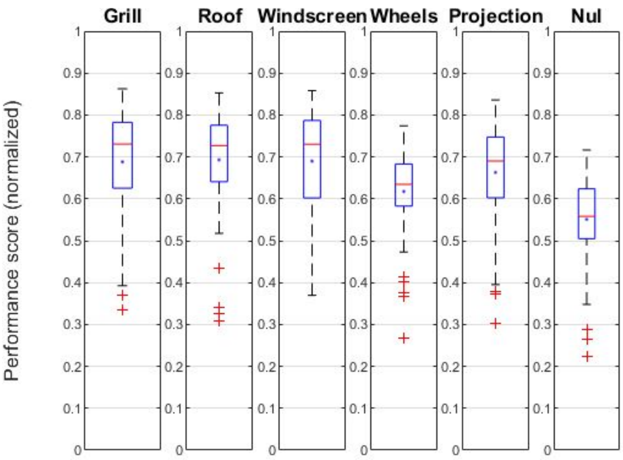


Fig. 4.3. Performance score per eHMI

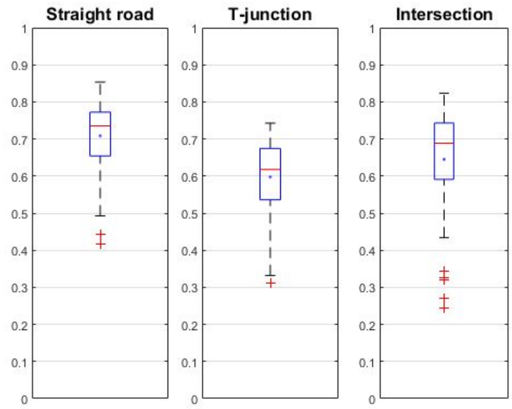


Fig. 4.4. Performance score per traffic layout

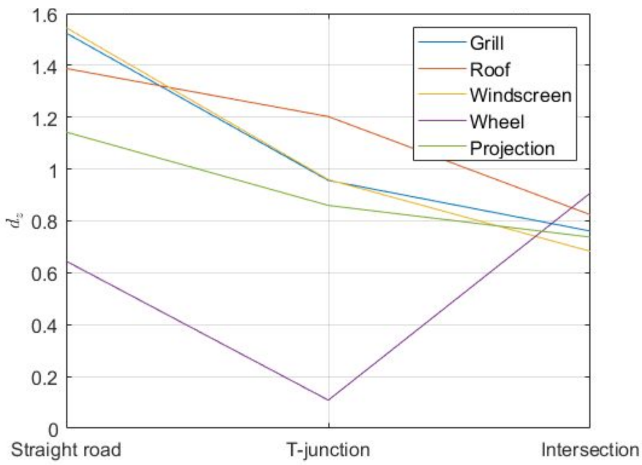


Fig. 4.5. Cohen's dz for each eHMI compared to "No eHMI" vs increasing workload

Table 4.2. Time intervals of used data per trial for performance calculations

Scenario 1	Boundaries (s)	Scenario 2	Boundaries (s)	Scenario 3	Boundaries (s)
Grill	6-25	Grill	8-25	Grill	9-25
Projection	5-24	Projection	6-23	Projection	8-24
Roof	4-23	Roof	4-21	Roof	7-23
Wheels	3-22	Wheels	7-24	Wheels	6-22
Windscreen	2-21	Windscreen	5-22	Windscreen	5-21
Nul	1-20	Nul	2-19	Nul	4-20
Scenario 4	Boundaries (s)	Scenario 5	Boundaries (s)	Scenario 6	Boundaries (s)
Grill	7-25	Grill	13-20	Grill	8-25
Projection	6-24	Projection	11-18	Projection	7-24
Roof	5-23	Roof	9-16	Roof	6-23
Wheels	4-22	Wheels	7-14	Wheels	5-22
Windscreen	3-21	Windscreen	5-12	Windscreen	4-21
Nul	2-20	Nul	3-10	Nul	3-20

Table 4.3. t-Test p-values performance (Green cells indicate a significant difference, red cells indicate no significant difference)

T-test (p>0.05)	Grill	Roof	Windscreen	Wheels	Projection	Nul
Grill	NaN	2.93E+00	4.89E+00	2.60E-10	7.53E-03	4.87E-16
Roof	2.93E+00	NaN	4.17E+00	2.49E-10	1.10E-03	1.77E-16
Windscreen	4.89E+00	4.17E+00	NaN	3.59E-10	6.15E-03	3.34E-17
Wheels	2.60E-10	2.49E-10	3.59E-10	NaN	5.68E-07	4.29E-09
Projection	7.53E-03	1.10E-03	6.15E-03	5.68E-07	NaN	2.34E-14
Nul	4.87E-16	1.77E-16	3.34E-17	4.29E-09	2.34E-14	NaN

t-Test (p>0.05)	Straight road	T-junction	Intersection
Straight road	NaN	1.85E-25	3.78E-09
T-junction	1.85E-25	NaN	3.81E-08
Intersection	3.78E-09	3.81E-08	NaN

4.3 Gaze plot

The eye-tracking data of all 61 participants has been plotted onto the corresponding videos. An example of these plots can be seen in figure 4.6 where screenshots from three of these plots are shown. More screenshots can be found in appendix C of all eHMI's on a straight road. These screenshots are all taken at approximately the same three points in the animation for the same eHMI at different traffic settings.

Fig. 4.6. Gaze plot of eHMI positioned on the grill in different traffic settings

V. Discussion

5.1 Interpretations

When looking at the plots and data in section 4, they seem to suggest significant patterns within the independent variables. These patterns will be further discussed in the following section.

5.1.1 Research question

The research question as posed in section 1.2 can be divided into 4 separate sub-questions that together answer the full question. These 4 questions are based on how different parts of the question can be connected. On one side the visual distribution of participants and the crossing intentions of participants, and the locations of the eHMI's and the different stresses/distractions on the other side. Combining these two sides gives four possible combinations which will be discussed in the following section.

'How do participants distribute their visual attention with eHMIs mounted at different locations?'

As can be seen in the gaze plots (Appendix C), the attention of participants quickly focusses to the eHMI the moment an eHMI turns on. The full videos that these figures are taken from also show that once the participants know what eHMI they're dealing with in a video they tend to fixate on the position of the eHMI before it is turned on or even visible. Participants quickly learn and adapt to the position of the eHMI. Differences of this effect between eHMI's are small as can be seen in appendix C. The spread of the data points does grow significantly in the absence of an eHMI. This indicates that, in general, eHMI's draw visual attention from the participant.

'How do participants distribute their visual attention under influence of different stresses and distractions?'

As shown in the previous question, once participants spot an eHMI turning on they generally fixate on the message displayed until it is no longer relevant for their crossing behaviour. This behaviour does not change when different stresses or distractions are introduced. This is not to say that participants are completely unaffected by the added stimuli. According to figure 4.4, the T-junction is shown to be the most difficult environment for the participants to fulfill their task. Interestingly, the participants did not seem to experience this difficulty, as the T-junction is not necessarily the least clear environment, according to participants. It can only be significantly shown that the participants felt it was most clear when they could cross on the straight road, which does correspond with the achieved performance.

'How are participants' crossing intentions affected when viewing animations of automated cars with eHMIs mounted at different locations?'

From the performance plots in figure 4.3 and the t-test run over these values a general ranking for the locations of the eHMI can be found on the basis on how they perform during the experiment. This ranking shows that the roof, windscreen and grill are the best choices for eHMI locations. It can not be significantly stated that either the roof, windscreen or grill is better than the other. Clearly the worst of the tested eHMI locations is the absence of an eHMI, followed by the eHMI positioned above the wheels. The projection on the road is the second best after the shared first place of the roof, windscreen and grill.

'How are participants' crossing intentions affected when under influence of different stresses and distractions?'

That which can be concluded from the Cohen's dz value, which compares each eHMI to the scenario with no eHMI, confirms the conclusion that the straight road was the clearest environment as it relates to participants' crossing decisions. The first increase of stress between a straight road and T-junction sees all of the eHMI's dropping in performance compared to the absence of an eHMI. This makes sense given the increased complexity of the scenario. When comparing the T-junction and the intersection the eHMI's show mixed results for the increase in stress. All the eHMI's decrease in performance, except for the one placed above the wheels. This can be explained by the additional angle (from the left) from which cars can enter the traffic scenario. While the eHMI above the wheels is hard to see head-on, it does give have an advantage when observing a car from the side, which becomes relevant when they turn into the crossing from the left. Most of the eHMI's lose effectiveness compared to the absence of eHMI. Even though most eHMI's perform worse at the cross-section than the T-junction, the increase in performance of the eHMI above the wheels is so large that the T-junction seems to be the hardest environment to perform the given task in. However, the truth is that this varies with eHMI location and for most locations the cross section is the most detrimental to the task.

5.1.2 Hypotheses

Hypothesis A claims that the eHMI projected on the windscreen will be the best eHMI. In the results there is no one eHMI which excels over the others. The eHMIs mounted on the roof, grill or projected on the windscreen are not significantly different from each other (table 4.3) and are the most effective (see figure 4.3).

Hypothesis B claims that performance will decrease when a traffic scenario becomes more complex. In figures 4.4 and 4.5 can be seen that this does not always hold true. The eHMI positioned above the wheels shows a deviation. While all eHMI's decrease in effectiveness as the traffic becomes more complicated, the eHMI above the wheels becomes much more effective in the cross section. An explanation for this is given in section 5.1.1 and is most likely not because of the change in complexity, but due to the introduction of cars coming around corners. This means it can be concluded that increasing the stress reduces the usefulness of an eHMI compared to the absence of an eHMI. Hypothesis B seems to hold true. Though it is worth noting that the absence of an eHMI will still perform significantly worse.

Fixation duration is not necessarily a measure of stress. The length of the duration can have multiple reasons to differ per trial. So hypothesis C will remain unanswered after this experiment.

5.2 Error analysis

During the design and deployment of the described experiment, several problems were encountered that could not be fully fixed and may have caused inaccuracies in the results. These problems and the measures taken to minimize them will be described here.

5.2.1. Learning effects

The videos used in the experiments repeat so that all the different locations of eHMI can be tested on each traffic configuration. This means that it is possible for a participant to recognize a traffic configuration and perform their task of pressing the spacebar when they want to cross the road based on their memory rather than on the information provided by the eHMI's.

Several steps have been taken to diminish this learning effect. First, Each participant was shown the videos in a semi-random order that was different for each participant in which the same traffic setting does not repeat twice. This means that the learning effects that would show should be evenly distributed over the trials. This will normalize the results as a whole. Second, each traffic scenario has been given a second video with different traffic to increase the variety within the experiment and to not let the road alone be enough to recognize a video. Third, while each video used in the experiment was 25 seconds long, longer animations were created so that each video could start at a different point in the animation. This means the start of each video was different and could not be used to recognize the traffic configuration. However, this measure can also be a cause of errors, because it changes the context of a participant's decision, even if the environment when making the crossing decision has remained constant.

Another learning effect sets in at the start of the experiment, when the participant still has to come to grips with their task. Because their task is new to them, more mistakes will be made when the participant is just starting out with the experiment. To minimize this the task is explained to the participant in different ways, including an example video showing the participant some of the stimuli used in the rest of the experiment. The randomization of trials as mentioned before will help normalize what is left of this effect in the results.

5.2.2 Sight

Five participants had to take off their glasses, because they conflicted with the eye-tracker, and could not see very well without them. Their data has not been excluded because this reason was not deemed significant enough to ignore their data.

5.2.3 Calibration

At the start of the experiment, the eye-tracker is calibrated to achieve the most accurate results possible. This calibration was graded by the eye-tracking software as "GOOD", "FAIR" or "POOR". For a minority of participants calibration proved difficult, and a "FAIR" rating was accepted. Occasionally a "POOR" rating was given due to a single outlier. This calibration was also accepted if improvement was not possible. A full summary of the errors relating to participants can be found in Appendix A.

5.2.4 Errors in the simulations

Upon data analysis two errors were found within the scenarios. In trial 5 (scenario 1, wheels), the timing of the first car to stop and signal "waiting" is one second too early. Because this error was found early on it has been accounted for in the calculation of the performance. This means it should not have an effect in the processed data, but since the scenario is slightly different the error could have had an effect on the participants while running the experiment.

Furthermore, the video with trial number 25 (scenario 5, grill) had an error concerning its duration. All the videos were meant to take 25 seconds exactly in order to stay consistent, this video's duration is only 20 seconds. This error has also been accounted for and should have no effect on the performance score data.

VI. Conclusion

6.1 Conclusions

Of the eHMI's tested the projection on the windscreen and the eHMI positioned on the roof and grill have the highest performance score. Conversely, "no eHMI" achieves the worst score, followed by the eHMI placed above the wheels. The eHMI above the wheels does seem to perform better when the AV is turning into the crossing from the left as can be seen in figure 4.5 (Cohen) where the dz rises when going from the T-junction to the cross section. It has also been shown that eHMI's generally work better under less stress when compared to "no eHMI".

6.2 Recommendations

The scope of this experiment has purposefully been limited to only compare several versions of visual eHMI's. However, as shown in the literature, there are more possibilities including audio, colour and vehicle behaviour (Bazilinskyy, 2019). Further research could be done to compare these other types of eHMI's in a similar way as was done in the described experiment. Other than comparing individual eHMI's in this manner, further research could also explore the possibilities of combining multiple eHMI's. In discussions after the experiments, participants especially noted how the eHMI above the wheels was unclear when the car was coming head-on, but very clear when the car was turning. Combining it with a complementing eHMI could make for an interesting option for AV's.

References

Kitazaki, S., Myrhe, N.J., 2015. Effects of non-verbal communication cues on decisions
and confidence of drivers at an uncontrolled intersection. In: Proceedings of the
Eighth International Driving Symposium on Human Factors in Driver Assessment,
Training and Vehicle Design, pp. 113–119.

Stanciu, S. C., Eby, D. W., Molnar, L. J., St. Louis, R. M., Zanier, N., & Kostyniuk, L. P. (2018). Pedestrians/Bicyclists and Autonomous Vehicles: How Will They Communicate? Transportation Research Record: Journal of the Transportation Research Board, 2672(22), 58–66. https://doi.org/10.1177/0361198118777091

Anantha Pillai. 2017. Virtual Reality based Study to Analyse Pedestrian attitude towards Autonomous Vehicles. In Proceedings of Aalto Master Thesis, Espoo, Oct 2017, 14 pages.

Maag, C., Muhlbacher, D., Mark, C., & Kruger, H. P. (2012). Studying Effects of Advanced Driver Assistance Systems (ADAS) on Individual and Group Level Using Multi-Driver Simulation. IEEE Intelligent Transportation Systems Magazine, 4(3), 45–54. https://doi.org/10.1109/MITS.2012.2203231

Vlakveld, W.P., Van der Kint, S., et al. (2018). Hoe reageren fietsers op zelfrijdende auto's?. R-2018-21. SWOV, The Hague.

Vinkhuyzen, E., & Cefkin, M. (2016). Developing Socially Acceptable Autonomous Vehicles. 2016 EPIC Proceedings, 522–534. https://www.epicpeople.org/developing-socially-acceptable-autonomous-vehicles/

Schaudt, W. A., & Russell, S. (2018). Judging a Car by its Cover: Human Factors Implications for Automated Vehicle External Communication. Lecture Notes in Mobility, 69–76. https://doi.org/10.1007/978-3-319-94896-6_6

Florentine, E., Andersen, H., Ang, M. A., Pendleton, S. D., Fu, G. M. J., & Ang, M. H. (2015). Self-Driving Vehicle Acknowledgement of Pedestrian Presence Conveyed via Light-emitting Diodes . International Conference Humanoid, Nanotechnology, Information Technology Communication and Control, Environment and Management (HNICEM). https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7393208

Bazilinskyy, P., Dodou, D., & De Winter, J. C. F. (ter perse). Survey on eHMI concepts: The effect of text, color, and perspective. Department of Cognitive Robotics, Department of Biomechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering, Delft .

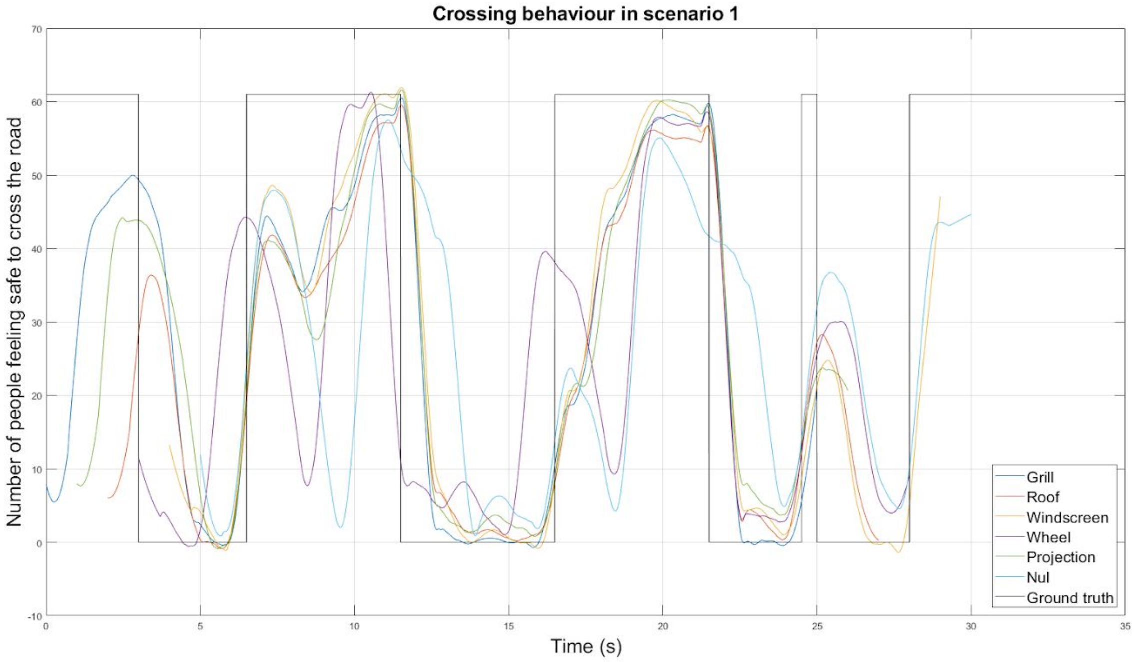
Appendix A - Calibration quality

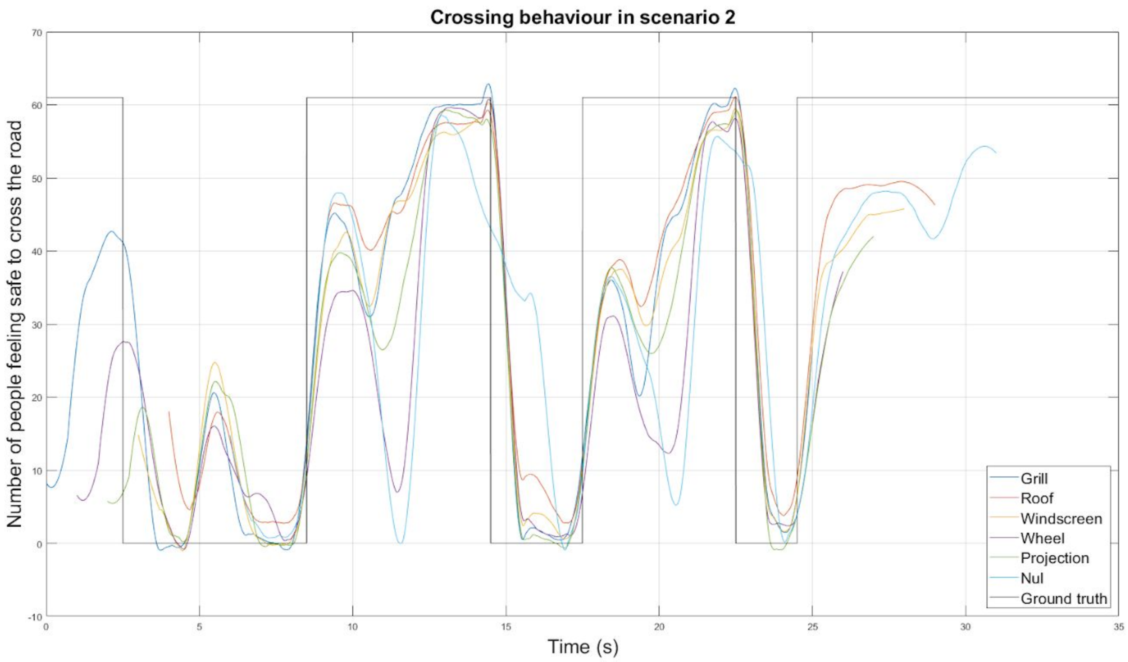
	Validation (GOOD)	Validation (FAIR)	Validation (POOR)	Poor Eyesight	Comments
p01			x		
p02	x				
p03	x				
p04			x		
p05			x		
p06	x				
p07	x				
p08	x			x	Unable to read the eHMI's
p09	x				
p10	x				
p11	x				
p12	x				
p13		x		x	Unable to read the eHMI's
p14	x				
p15	x				
p16	x			x	Unable to read the eHMI's
p17	x				
p18	x				
p19	x				
p20	x			x	Unable to read the eHMI's
p21	x				
p22	x				
p23	x				
p24	x				
p25	x				Left eye was tracked
p26	x			x	Unable to read the eHMI's
p27	x				
p28	x				
p29	x				Left eye was tracked
p30	x				
p31	x				

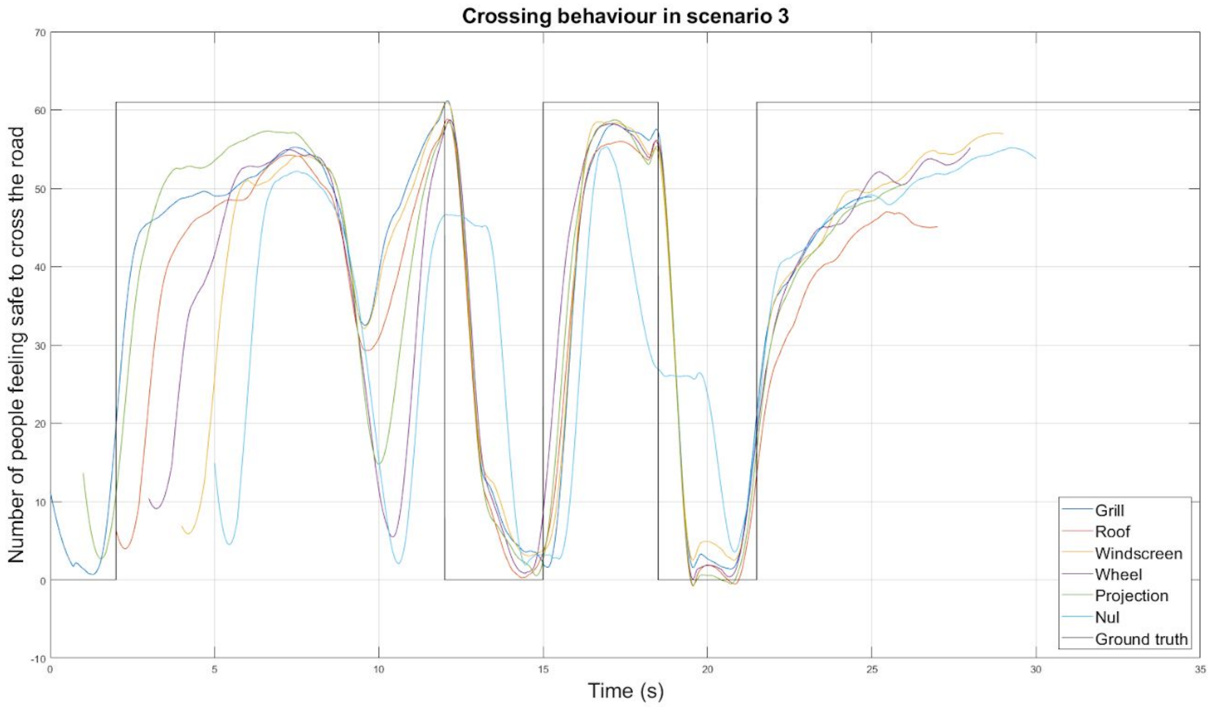
p32	x				
p33	x				
p34	x				
p35	x				
p36	x				
p37	x				
p38	x				
p39	x				
p40	x				
p41	x				Eye-tracker had trouble tracking the light, turning on artificial light on fixed the problem
p42	x				
p43	x				
p44	x				
p45	x				
p46			x		
p47	x				
p48	x				
p49	x				
p50	x				
p51	x				
p52	x				
p53		x			
p54	x				
p55	x				
p56	x				
p57	x				
p58	x				
p59	x				
p60	x				
p61	x				

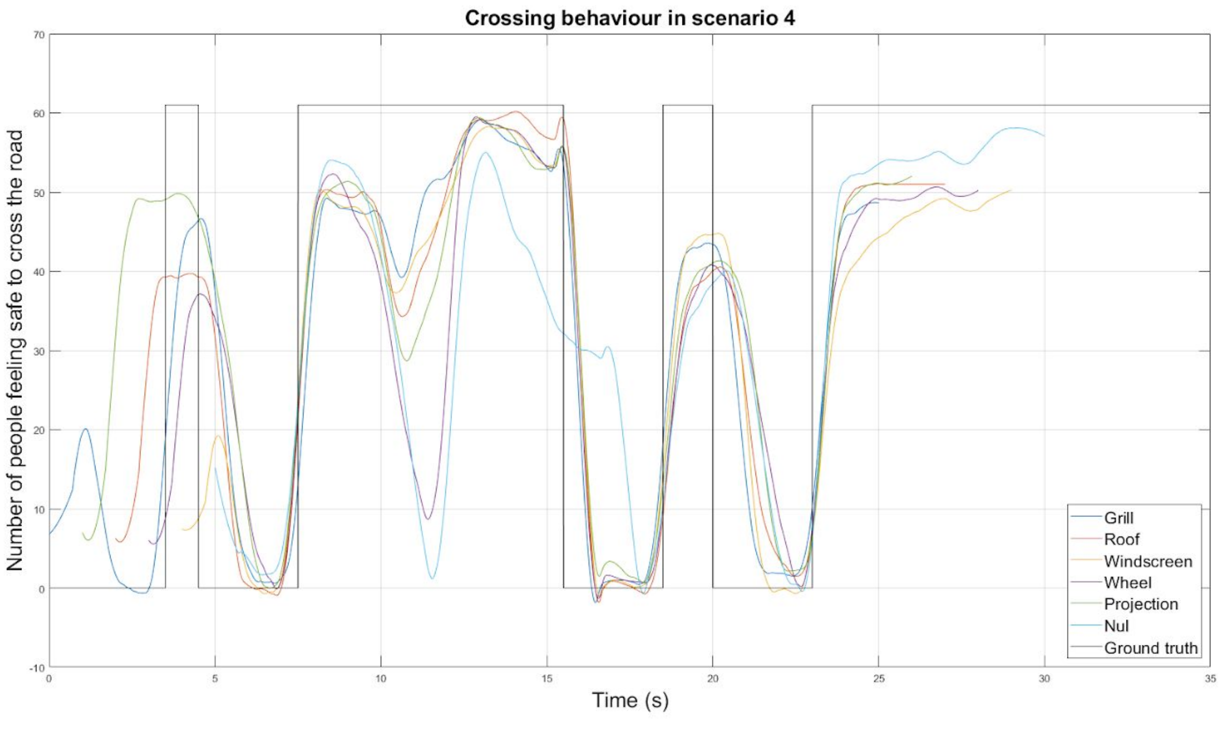
APPENDIX B - CROSSING BEHAVIOUR PLOTS

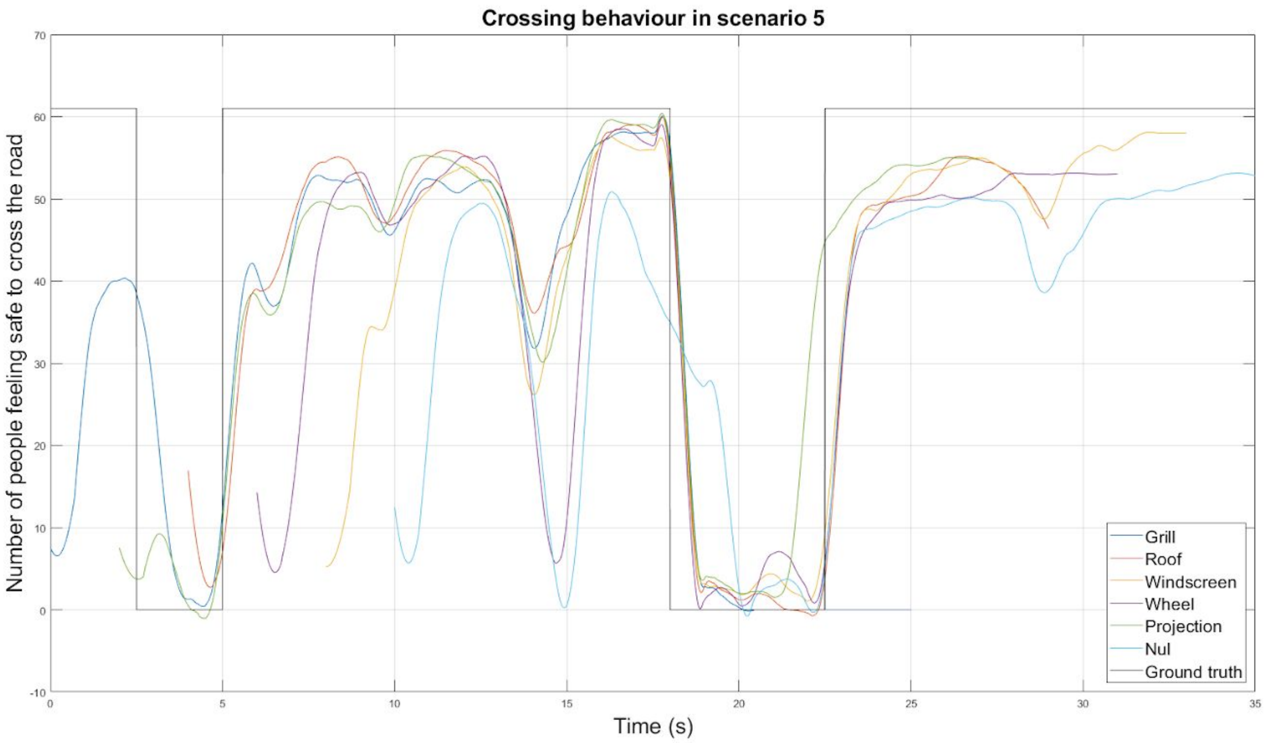
Scenario 1 & 4	Scenario 2 & 5	Scenario 3 & 6
Straight road	T-Junction	Intersection



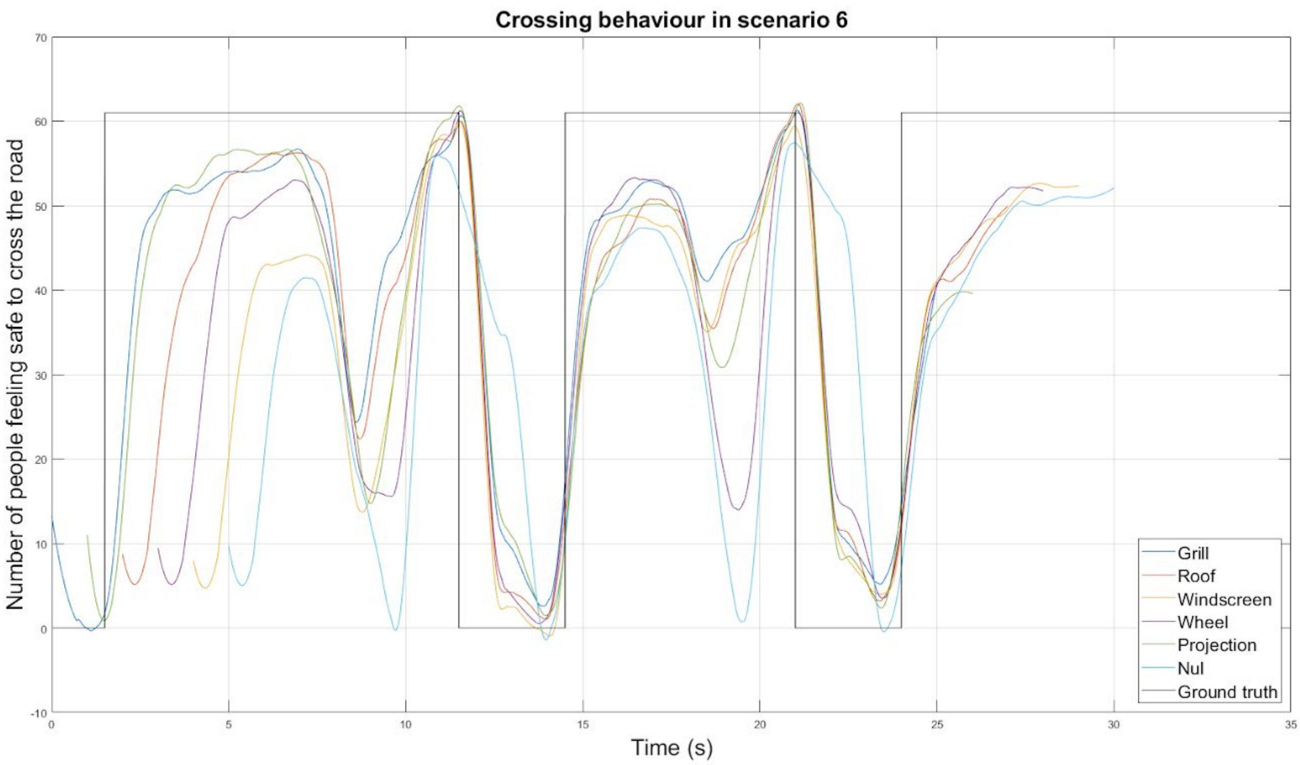



Crossing behaviour in scenario 3
Number of people feeling safe to cross the road
Time (s)
Grill
Roof
Windscreen
Wheel
Projection
Nul
Ground truth


Crossing behaviour in scenario 4
Number of people feeling safe to cross the road
Time (s)
Grill
Roof
Windscreen
Wheel
Projection
Nul
Ground truth



Crossing behaviour in scenario 5
Number of people feeling safe to cross the road
Time (s)
Grill
Roof
Windscreen
Wheel
Projection
Nul
Ground truth



Crossing behaviour in scenario 6
Number of people feeling safe to cross the road
Time (s)
Grill
Roof
Windscreen
Wheel
Projection
Nul
Ground truth

Appendix C - Image stills of video plotting gaze coordinates per eHMI

Windscreen(0 - 0.5 - 1.0 s)	Grill (0 - 0.5 - 1.0 s)	No eHMI (0 - 0.5 - 1.0 s)
		
		
		

Projection road (0 - 0.5 - 1.0 s)	Roof (0 - 0.5 - 1.0 s)	Wheels (0 - 0.5 - 1.0 s)