

The relation between humans' interactive behavior and fixation behavior in a coupled virtual reality driving simulator.

By

Loran Bogaart

in partial fulfilment of the requirements for the degree of

Master of Science

in Mechanical Engineering (Biorobotics)

at the Delft University of Technology,

to be defended publicly on Tuesday January 31, 2023 at 10:00 AM.

Student number: 5145279

Thesis committee: Dr. ir. A. Zgonnikov,

TU Delft, Supervisor

Ir. O. Siebinga,

TU Delft, Supervisor

Dr. ir. J.C.F. De Winter,

TU Delft, External member



The relation between humans' interactive behavior and fixation behavior in a coupled virtual reality driving simulator.

Loran Bogaart

Department of Cognitive Robotics, Delft University of Technology, Faculty of 3mE

Abstract

In order to design safe and effective interactions between autonomous vehicles (AVs) and human road users, it is essential to understand the mechanisms underlying human-human merging behavior. Driving simulator experiments can be used to study these mechanisms, but previous research has primarily focused on the behavior of individual drivers rather than the dynamics of interactions. In addition, current experimental scenarios and data analysis tools do not adequately capture interactive human-human merging behavior. To address these issues, I propose an experimental framework featuring a simplified highway-merging scenario that can facilitate human factors research on merging interactions. In a case study with fourteen participants, I used the framework in a coupled virtual reality driving simulator to show a relation between participants' interactive behavior and fixation behavior. This work shows how to better understand human-human merging interactions, which is essential for developing AVs that can safely and successfully interact with other road users.

1. INTRODUCTION

Massive interest is going out to autonomous vehicles (AVs), and work on various algorithms for AVs is published regularly [1, 2, 3]. Work contains, for instance, path planning, decision-making, and perception. Allowing an AV to drive on a highway or through urban cities will involve solving highly complex scenarios with much uncertainty [4]. It is a difficult challenge to mitigate the complexity and uncertainties to ensure safety, and until this day, numerous problems remain unsolved for autonomous driving. For instance, when AVs are integrated into the daily traffic, they will encounter interactions with human drivers. In order to guarantee safety (i.e., avoid collisions, perform safe actions) and make well-estimated predictions on humans, a new technology, named interaction-aware controller (IAC), is proposed [5]. IACs incorporate two prediction models: a dynamical model to predict future states and a human behavior model that predicts the human driver's future actions. However, modeling natural human behavior is a difficult challenge, and much research is done to find optimal human behavior models.

1.1 Related work. Work in non-interactive traffic scenarios model human behavior as car-followers with constant velocity [6, 7, 8, 9, 10, 11]. In such control approaches, surrounding 'human-driven' vehicles are programmed in swarms with fixed positions to the vehicle ahead. The AV may react to upcoming vehicles, yet it cannot influence the other vehicles. Predefined traffic scenarios are solved. However, such models do not simulate natural human behavior and mainly focus on controlling the AV. More advanced car-following models allow the interacting vehicles to change accelerations [12, 13, 14, 15, 16, 17, 18]. As a result, a more natural human behavior is realized, yet strong motion assumptions are made. In this work, the car-following models are classified as one-way interacting models that obey a leader-follower rule

[19]; the follower (human driver) may react upon the leader (AV), but not vice versa. In other words, the main focus is controlling the AV, and human behavior is modeled as an individual with simple motion heuristics.

On the contrary, game theoretical-based approaches [20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30] incorporate a two-way interacting model. Interesting scenarios involving two vehicles are solved, and optimal solutions to a game are found. In game theoretical-based approaches, taken actions influence both the AV and the other vehicle's future actions. However, future actions taken by the human driver are based on a rational utility-maximizing model. This simplification made on human behavior does not hold in interactive driving.

A new framework to model human behavior is proposed by Siebinga [19]. This approach does not model the vehicles as individual drivers or rational agents but models a two-way human-human interaction with communication. Two interacting vehicles are both given a plan and a belief. The belief about what the other will do is interchanged through communication. This new interacting model is validated with a case study in a coupled driving simulator involving participants [31]. In the case study, the scenery is presented in a 2D top-down view, and participants are tasked to merge in a simplified merging scenario. However, it is assumed that participants always observe the opponent, which does not hold in 3D driving scenarios.

Besides mathematically modeling human behavior, interesting research is done in quantifying humans' gaze behavior during driving scenarios [32, 33, 34, 35]. Analyzing gaze behavior allows researchers to gain additional insights into one's area of interest at specific times during driving scenarios. Furthermore, existing literature studies human's merging behavior [36, 37, 38]. Nevertheless, most work on gaze and merging behavior focused on the behavior of single drivers, which does not hold in interactive driving scenarios.

1.2 Aim and hypotheses. Little is known about how humans behave in a two-way interacting driving scenario since previous work mainly focused on the behavior of single drivers. The approach by Siebinga [19] does model the two-way interacting approach. However, doing a case study in a 2D scenery is limited, and participants are expected to behave differently in realistic 3D driving scenarios since they need to observe the other vehicle.

This research addresses that gap by presenting a framework for conducting human factors experiments using a coupled virtual reality (VR) driving simulator. The framework investigates interactive human-human driving behavior during a simplified merging scenario in a 3D scenery. The 3D space is realized with new VR technologies. This space makes studying participants' areas of interest possible. The framework relates this VR data to participants' interactive behavior (i.e., the behavior of controlling a vehicle). By doing this research, I want to better understand the underlying mechanisms of interactive human-human merging interactions, which is crucial for creating AVs that can safely interact with other road users. Furthermore, understanding the interaction dynamics better might help future researchers who incorporate human behavior models. This work, therefore, aims to find answers to the following research question:

What is the relation between participants' interactive behavior and fixation behavior?

In order to find answers to the aimed research question, I constructed five sub-questions which are chronologically revised to studiable hypotheses in Table I:

- 1) What is the effect of velocity differences on participants' merging behavior?
- 2) What is the effect of velocity differences on the conflict duration?
- 3) What is the effect of a different conflict duration on the number of participants' head rotations?
- 4) What is the effect of velocity differences on the number of participants' head rotations during the interactive section?
- 5) What is the effect of velocity differences on the number of participants' head rotations before and after the conflict?

A different relative velocity implies that the vehicles are positioned differently in the tunnels. These various locations might influence participants' merging behavior since the vehicles exit the tunnel at different times. For example, the vehicle that exits the tunnel first is ahead of the opponent and gets control earlier. As a result, this ahead vehicle might reach the merging point earlier than its opponent. The first question investigates whether this assumption holds.

The second question studies the effect of different velocity conditions on the conflict duration. Participants are forced to be in conflict and tasked to solve it naturally. I expect that the different velocities influence the duration of the conflict because the vehicles exit the tunnel at different times. In other words, the earlier a vehicle exits the tunnel, the shorter

the conflict last due to one vehicle getting control earlier.

The third question investigates how often participants observe the other vehicle during a different conflict duration. If the conflict is longer, I assume that participants are more aware of the opponent since they need more time to solve it.

The fourth question studies the fixations of participants. In this case, participants are expected to be differently focused on each other during the interactive section. I assume that the participant who exits the tunnel first tries to observe the opponent more often since this participant expects an upcoming vehicle.

The fifth question investigates the number of fixations before and after solving the conflict. I assume that participants fixate less on the opponent when the conflict is resolved since the chance of a collision diminishes.

2. METHOD

2.1 Participants. Fourteen participants (3 females) between 22 and 65 ($M = 33.6$, $SD = 15.2$) volunteered in this research. Everyone had good vision and was in possession of a driver's license. All participants signed an informed consent, and the Human Research Ethics Committee (HREC) of the Delft University of Technology approved this research.



Fig. 1: Overview of the experiment setup

2.2 Apparatus. Figure 1 shows the experiment setup. Two desks are centered within a truss system that is placed in a square with curtains attached, separating both desks and minimizing light interference. An Alienware Aurora R9 computer with an NVIDIA GeForce RTX 3090 graphics card is connected to a monitor on each desk. The experimenter only uses these monitors. Each computer has a Varjo VR-3 head mount connected. In order to orientate the head mounts, two HTC steamVR base stations 2.0 are mounted to the truss system above each desk. Activating VR is done with the SteamVR plugin in Steam [39]. The head mounts have eye trackers integrated and broadcast the scenery with a 115-degree viewing range. The scenery is built with Roadrunner [40] and integrated into Unreal Engine, which simulates vehicle dynamics using PhysX [41]. Two commercial Logitech Driving Force GT steering wheels with corresponding pedals

TABLE I: Hypothesized sub-questions with the aim to find the relation between participants’ interactive behavior and fixation behavior.

Hypothesis
1. The vehicle that is ahead has the highest probability of merging first.
2. Increasing the relative velocity decreases the conflict resolution time.
3. Participants’ average fixations on the opponent increase during longer conflicts.
4. The vehicle that is ahead will have higher average fixations on the opponent compared to the vehicle that is behind during the interactive section.
5. Participants’ average fixations on the opponent decrease when the conflict is resolved.

are used for controlling the vehicles. Vehicles are spawned in the simulator by using an open-source software framework called JOAN [42], developed at the TU Delft. This software is built upon the CARLA open-source simulator [43] (version 9.13). Furthermore, two Sony WH-1000XM3 noise-canceling headphones are provided to the participants to prevent communication.

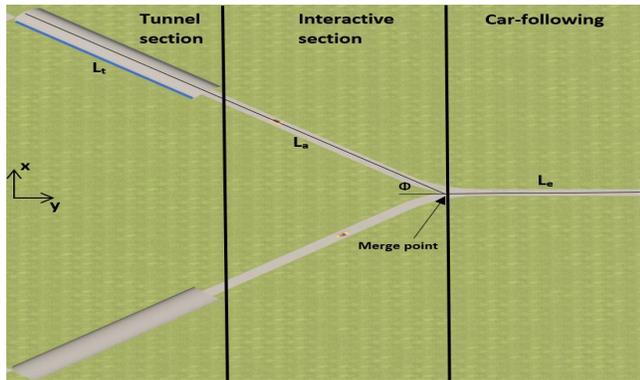


Fig. 2: A top-down view of the simplified merging scenario. Two vehicles approach a merge point that fluently merges into a straight section. The track is divided into three sections: tunnel section, interactive section, and car-following section. Positive x is upwards, and positive y is rightwards.

TABLE II: Measures of the track in Figure 2.

Parameter	Value
l_t	125 m
l_a	325 m
l_e	120 m
θ	45°
Track width	6 m
Vehicle length	4.7 m
Vehicle width	2 m

2.3 Environment design. In the simplified merging scenario depicted in Figure 2, two vehicles of the same size approach a predetermined merging point. The track is divided into three sections: the tunnel section, the interactive section, and the car-following section. The tunnels in the tunnel section function as barriers that block participants’ vision to see the other vehicle. In addition, the track has two approach sections (l_a) with equal dimensions under an approach angle (θ). The interactive section captures the actual merging behavior. If

the course of the road is followed, both vehicles will reach a predefined merge point. The taken assumption on this merge point is explained in section 3.2. At the merging point, the track fluently merges into a straight car-following section (l_e). All track dimensions are summarized in table II.

The merging scenario simplifies environmental factors. For example, this scenario has no right of way, so the collected data is symmetrical. Furthermore, a speedometer is integrated into each vehicle’s dashboard, but other visual cues for velocity (e.g., speed signs and trees) are not included. The reason for this is that participants can better focus on the interaction.

The track is built with Roadrunner, which provides an easy and efficient way to design a track. Moreover, it is integrative into Unreal Engine. Appendix A shows a complete overview of the used track. This double-sided track is a mirrored version of the simplified merging scenario visualized in Figure 2. Section 2.4 explains the motivation for mirroring the track. A wall is placed in the middle of the double-sided track to make the other side invisible to the participants. More walls are placed along the outer side of the track to instruct participants on which side they will exit the tunnel. Furthermore, they function as a boundary.

2.4 Experiment design. The experiment included seven sessions with two participants each. During each session, two vehicles were continuously spawned at various locations with different initial velocities in opposite tunnels visualized in Figure 2. As argued in section 2.3, the two tunnels function as visual barriers. The predefined spawn locations with the corresponding velocities are computed such that if participants maintain their initial velocity, they will collide at the merging point. This forces the participants to adapt to the current situation and solve the conflict by avoiding a collision at the merge point. Participants were exposed to three initial velocity conditions:

- 1) 50-50 km/h,
- 2) 55-45 km/h,
- 3) 60-40 km/h.

For instance, condition 2 implies that the vehicle of participant A gets an initial velocity of 55 km/h and the vehicle of participant B an initial velocity of 45 km/h. This also implies that participant A is seated in a vehicle that is behind, whereas participant B is seated in a vehicle that is ahead. Figure 3 illustrates the conditions and being ahead or behind. The participant is able to steer the vehicle at any

time. However, the initial velocity of each vehicle is fixed until it reaches the end of the tunnel. By fixing the velocity until this moment, I assume that the interaction starts when the first vehicle exits the tunnel.

The conditions were repeated 48 times in random order during each session. Each condition is repeated 16 times, with a particular case for conditions 2 and 3. In these conditions, the 16 trials are split into two sets of 8. For example, in the first set of condition 2, participant A is seated in a vehicle with a velocity of 55 km/h, and participant B is seated in a vehicle with a velocity of 45 km/h. In the second set of condition 2, participant A is seated in a vehicle with a velocity of 45 km/h, and participant B is seated in a vehicle with a velocity of 55 km/h. Additionally, participants were randomly swapped between the tunnels. This ensures that one does not notice the three conditions. In order to create more randomness, five trials ($\approx 10\%$) of condition 1 were added. However, during these trials, each vehicle is spawned on one side of the double-sided track. Consequently, participants will never have an interactive conflict since they were alone on the track.

If a vehicle reaches the endpoint, the trial is stopped automatically, and another is manually started directly after through JOAN. Participants may experience motion sickness in virtual reality [44]. Therefore, I integrated collision detection into JOAN, which stops the trial immediately when a collision occurs.

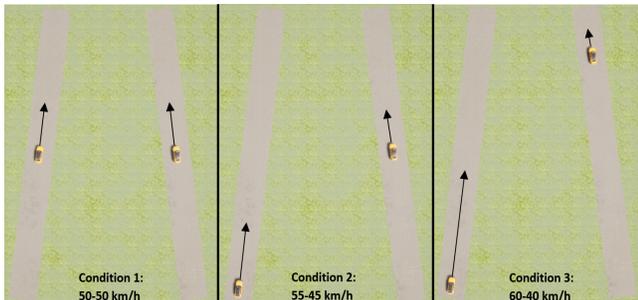


Fig. 3: A top-down view of the velocity conditions. The length of the roads illustrates the length of the tunnels. The different velocities are indicated with velocity vectors. In condition 1, neither vehicle is ahead or behind. In conditions 2 and 3, the right vehicle with a lower initial velocity is ahead of the left vehicle that has a higher initial velocity.

2.5 Procedure. To instruct participants, a similar top-down view as Figure 2 was shown. The instruction mainly explained how the track is designed and that participants’ vehicles were being spawned inside one of the two tunnels at predefined positions with various initial velocities. Furthermore, participants were instructed to maintain these initial velocities and that right-of-way is not applicable. The participants were then requested to take their seats. A short setup period started to orientate the Varjo head mounts and to correctly place the participants in the driver seat. Before the start of the actual session, participants performed

several training runs (≈ 5) to familiarize themselves with the scenario. Condition 1 was used during each training run. Finally, participants were provided with noise-canceling headphones that played non-copyrighted music [45].

Each session took approximately 1 hour. This included a 15-minute setup period, a 15-minute break, and 30 minutes of driving the trials. The length of this experiment is chosen to be relatively short due to the possibility of experiencing motion sickness in VR. I conducted an extra break when participants experienced any form of motion sickness.

2.6. Data collection. Besides spawning the vehicles in Unreal Engine through JOAN, it can log various data. In this research, the data logger was used to collect the current epoch in nanoseconds, position $[x, y]$ of the vehicles, and the velocity of the vehicles. The data collection for the head mounts was done by a python wrapper that is based on Varjo’s open-source C++ API [46]. After calibration, this wrapper collected the current epoch in nanoseconds, head mount rotation, and participants’ gaze. Only the head mount rotations are used in this work since it is assumed to give a good enough estimation of participants’ areas of interest at specific times. The term fixation behavior is used throughout this work and describes participants’ head rotations. Fixation behavior is described in detail in section 3.4. Both data collectors ran simultaneously during each session.

After the data collection, the data from JOAN and the wrapper were merged on the closest epoch with a neglectable small error since both epochs were captured at nanoseconds. Collected data in which vehicles collided at the merging point or at the car-following section were removed from the data set. The five random trials during each session were also removed from the data set since this does not include interactive data. The entire data set included 315 trial files.

3. RESULTS

This chapter starts with analyzing participants’ merging behavior, and findings show additional insights into who is reaching the merge point first.

Next, two metrics to study interactive merging behavior are used in this work [31]: 1) a visual representation of the pair-wise behavior and 2) conflict resolution time. After showing the visualization of the pair-wise behavior, the conflict resolution time is explained and captured in a metric for all trials.

Finally, various visualizations show participants’ fixation behavior. These visualizations show the average fixations at an area of interest during the interactive approach and before-after the conflict.

3.1 Merging behavior. Figure 4 shows an in-depth analysis of the probability of merging first for vehicle 1 depending on the initial velocity for each session individually. The various velocities entail that one vehicle is ahead of the opponent (40 and 45 km/h), behind the opponent (55 and 60 km/h), or

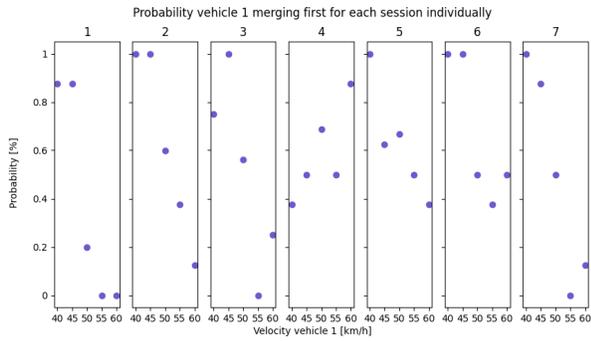


Fig. 4: Probability of merging first for vehicle 1. The blue dots indicate the probability of merging first at the exposed velocity.

TABLE III: Probability of merging first for vehicles 1 and 2 combined, computed for each session individually and as an average.

Session	Probability
1	0.94
2	0.88
3	0.88
4	0.38
5	0.69
6	0.78
7	0.94
Average	0.78

neither (50 km/h), as argued in section 2.4. The probability of merging first is visualized for vehicle 1, and $1 - P_{vehicle1}$ symbolizes the probability of merging first for vehicle 2. The probabilities on condition 1 (50 km/h) indicate participants' aggressiveness in merging. For example, if the probability is higher than 0.5, it indicates that this particular participant is more aggressive than the opponent since the interaction starts simultaneously. Table III shows the average probability of merging first for vehicles 1 and 2 when they are ahead.

The analysis shows that the vehicle ahead merges first 0.78 of the time. To further investigate this result, a logistic regression [47] is performed on the dependent binary data (i.e., reaching the merge point first or not) with velocity as an independent predictor variable for each trial. The regression model ($\beta = -0.17$, $z\text{-score} = -7.43$, $p = 1.09e^{-13}$) is visualized in Figure 5 and shows a negative coefficient with a statistical difference, indicating that the ahead vehicle will most likely merge first.

3.2 Visualization of the pair-wise behavior. Understanding what happens during each trial between participants is necessary to study human behavior. Therefore, an illustration is given in Figure 6 and shows the pair-wise behavior for a random trial. The triangles in the panels indicate the moment the vehicles exit the tunnel, and the squares indicate the moment the vehicles reach the merge point. The cross indicates when the conflict is resolved and is explained in section 3.3.

Panel A shows the positions of both vehicles during the trial. The light gray connected dots illustrate the positions in time. As described in section 2.3, the roads in the interac-

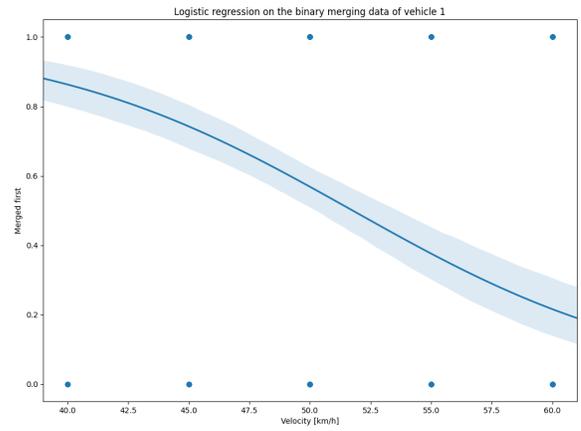


Fig. 5: The logistic regression shows the relation between various initial velocities and the merging behavior of vehicle 1. The dots indicate the binary data (i.e., reaching the merge point first or not)

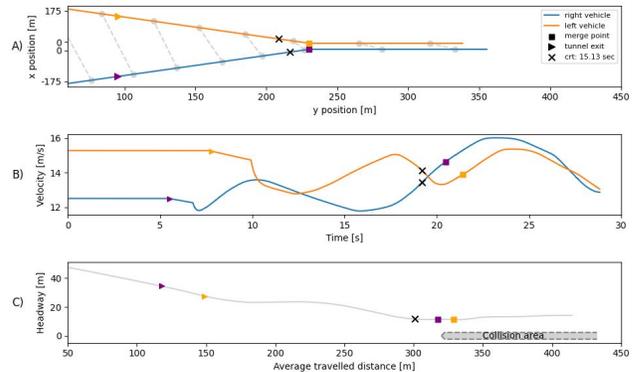


Fig. 6: Pair-wise behavior visualization of trial 44, session 4, and condition 3. Panel A shows the vehicles' positions individually, and the connected markers visualize the relative positions in time. Panel B shows correlating velocities. The headway against the average traveled distance is illustrated in panel C. The trajectory trace reflects the gap and relative velocity. When vehicles exit the tunnel it is indicated with a triangle and reaching the merge point with a square. The conflict resolution time is indicated with a cross and describes the final moment participants solve the conflict.

tive section merge fluently with curvatures into the straight section, but in this panel, it is assumed that the curvatures at the merging point are straight section parts, resulting in a fixed merge point. This assumption is made to efficiently define the collision area later explained in this section. An illustration is given in Appendix B. However, in practice, the vehicles arrive differently at the merge point during each trial due to free steering. Panel B shows the corresponding velocity changes. Panels A and B illustrate information about the vehicles' dynamics, yet how participants react during the interactive section is still vague.

Panel C is, therefore, used and captures the gap between the vehicles, relative velocity, and safety margins. For a detailed explanation, see [31] (p. 6). What is important regarding this work is the trajectory trace and the collision area. The trajectory trace arises by computing the distance between the front bumpers (i.e., headway) against the average traveled

distances for both vehicles. The trace is positive if vehicle 1 is ahead. All conditions are designed such that if the vehicles maintain their initial velocity, they collide at the merge point. Vehicles cannot collide during the interactive section, yet participants can be on a collision course. Therefore, the collision area is defined to visualize a possible collision. If the trajectory trace orientates towards the collision area, it indicates that the vehicles are on a collision course. When the trajectory trace enters the collision area, a collision occurs. The collision area only exists at the merging point and at the car-following section.

A brief elaboration on the trajectory trace and collision area is needed to explain section 3.3. The pair-wise visualizations are mainly used to spot corrupted data and collision data. Furthermore, it is used to validate the CRT algorithm.

3.3 Conflict resolution time. Both vehicles are positioned at predefined locations in the tunnel with a corresponding initial velocity, as described in section 2.4. By doing so, participants are initially heading towards a collision (i.e., trajectory trace pointing towards the collision area) and must change their velocity during the interactive approach to solve the conflict. During analyzing the trials, it is recognized that participants often alternate between being on a collision course and having a conflict solved, indicating that participants regularly get back into conflict after solving it initially. Appendix C emphasizes this. The work of Siebinga et al. [31] quotes: "We define the CRT as the time between the start of the interaction (tunnel exit) and the first moment the vehicles are no longer on a collision course.". Since the CRT often occurred multiple times during a trial in this work, I define the CRT as the time between the start of the interaction (first vehicle exiting the tunnel) and the last moment the vehicles are no longer on a collision course. The CRT algorithm [31] computes if the vehicles will end up in the collision area if they maintain their current velocities for each time step in the trial data. At the last time step during the interactive section, when vehicles are no longer on a collision course, it is assumed that the conflict is resolved.

This research aims to find the effect of the velocity differences on the conflict duration. Therefore, Figure 7 shows the distribution of all CRTs between the conditions captured in a boxplot. Table IV summarizes the medians of Figure 7 and the number of alternations of being on a collision course or not. To statistically show a distinction between conditions, a one-way ANOVA test [48] ($F = 40.84$, $p = 1.75e^{-16}$) is performed that compares the means of each condition on the dependent CRT variable. The ANOVA test can tell if the results are statistically significant, yet it does not answer the differences between conditions. Therefore, Table V shows the results of a Tukey-HSD test that compares all possible pairs. Based on the results, I conclude that the CRT changed significantly between conditions, being lowest in condition 1 (50-50 km/h) and highest in condition 3 (60-40 km/h), indicating that a higher initial relative velocity leads to more prolonged conflicts.

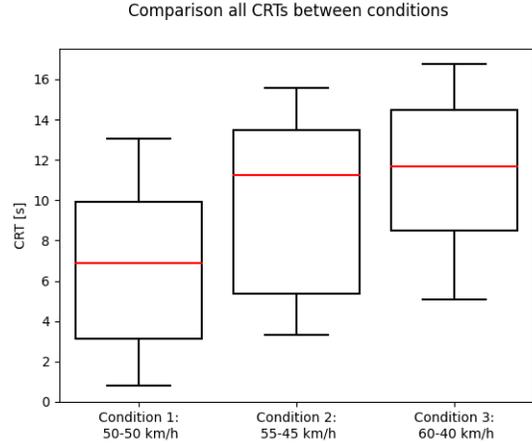


Fig. 7: Boxplot of all CRTs between conditions.

TABLE IV: Summary of the CRT results for each condition.

Condition	# of alternates	Median CRT
1	4.8	6.3
2	6.2	10.3
3	5.5	10.8

TABLE V: The results of the tukey-HSD test.

Comparison	T	p
1&2	-6.68	$3.25e^{-10}$
1&3	-9.12	0
2&3	-2.56	$2.91e^{-2}$

3.4 Fixation behavior. During the interactive section, participants are in conflict and alternate between focusing on the road and focusing on the other vehicle. Therefore, the simulator has two areas of interest: 1) on the road and 2) on the opponent. Two assumptions are made during the design of the areas of interest. The first assumption is that fixations on the speedometer are classified as fixations on the road since participants were not focusing on the interaction at this moment. The second assumption is that fixations to the other side of the road are neglectable small compared to fixations on the other vehicle. This assumption is made since the walls indicate on which side the vehicle exits the tunnel. Furthermore, it is expected that participants are not interested in this area due to being tasked to focus on the interaction.

As argued in section 2.6, head rotations are used to describe when participants fixate on the areas of interest instead of gaze. In order to determine when participants switch between the areas of interest, a threshold needs to be found. Therefore, Appendix D shows an analysis that visualizes the head rotations against time during the interactive section for multiple trials. This analysis is based on the python wrapper, described in section 2.6, that returns a value of 1 for looking straight and a reduced value for every head rotation. Rotating the head 90 degrees results in a value of 0. Based on the analysis, I assume that the threshold is at 0.95 since most peaks in head rotations crossed this value. In the following sections, head mount rotations above the threshold

are classified with a value of 1 (i.e., fixation on the road), and head mount rotations below the threshold are classified with a value of 0 (i.e., fixation on the other vehicle). This classification of fixations on the areas of interest is defined as fixation behavior.

This chapter separates the combined analysis of the fixation behavior in condition 1 and the individual analysis of the fixation behavior in conditions 2 and 3. This is because in condition 1 both vehicles enter the interactive section simultaneously, whereas one vehicle is ahead of the other in conditions 2 and 3.

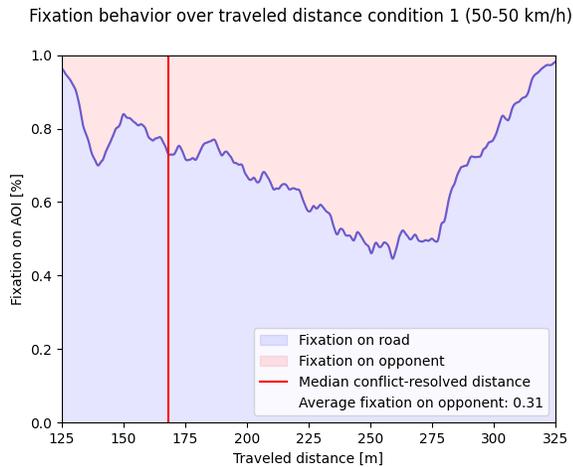


Fig. 8: Aggregated fixation behavior of the participants in condition 1. The blue area illustrates fixations on the road, and the red area represents fixations on the opponent during the interactive area. The median conflict-resolved distance is visualized with the red vertical line.

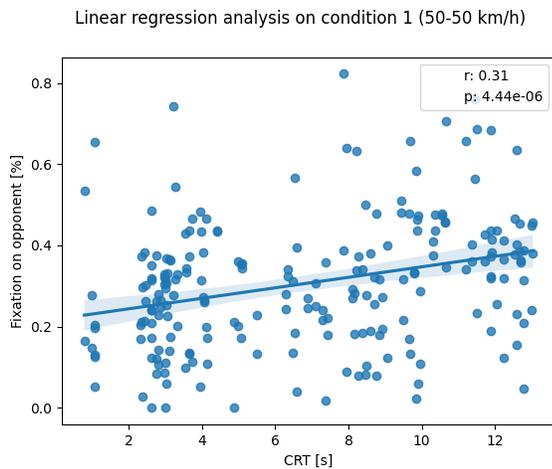


Fig. 9: Linear regression visualization for condition 1. The regression indicates the trend based on the average fixations at a certain CRT.

3.4.1 Combined fixation behavior analysis. In condition 1, both vehicles exit the tunnel simultaneously. Therefore, it can be assumed that collected data can be combined since

both participants are positioned in the same situation (i.e., a vehicle is not ahead nor behind).

Summing the threshold values over all trials leads to a fixation percentage on a particular area of interest which can be visualized against traveled distance or time. It is chosen to visualize the average fixations against the traveled distance for one crucial reason: the distance traveled in time differs for each trial since velocities differ over time, whereas the distance of the interactive section is always the same. The traveled distance contains only the distances of the interactive section. Figure 8 visualizes the aggregated fixation behavior during the interactive section. The fixation behavior and traveled distances of each trial are linearly interpolated to merge all concerning trials easily.

Besides having a figure that captures the fixation behavior of all trials, individual traces (i.e., data containing fixations, traveled distance, and CRT for one vehicle during a trial) can show additional insight into the number of fixations during the interactive section. Figure 9 shows the linear regression [49] analyses on the individual average fixations at a certain CRT. In this figure, the average of the individual fixation trace describes all fixations between the tunnel exit and merge point. This average is matched to the specific CRT, indicated with the dots.

The CRT is also related to an index of the individual trace and can be used to obtain the traveled distance. Appendix E shows the histogram and kernel density estimation of all traveled distances at which the conflict is resolved. The median conflict-resolved distance is visualized with the red vertical line in Figure 8.

3.4.2 Individual fixation behavior analysis. In Conditions 2 and 3, one vehicle is ahead of the other. Therefore, it is expected that participants' fixation behavior differs between being ahead or behind. Figure 10 divides the participants into two classifications: 1) a participant is ahead, and 2) a participant is behind. For instance, the vehicle in condition 2 with a velocity of 55 km/h is initially positioned at the tunnel's beginning and thus classified as behind, whereas the vehicle with a velocity of 45 km/h is positioned near the tunnel's exit and is classified as ahead, as visualized in Figure 3. Figure 11 visualizes the linear regression analysis for both conditions. The same approach as in section 3.4.1 is used to complete the figures.

3.4.3 Results of the fixation behavior analysis. The spike in average fixations at 275 meters in Figure 8 and Figure 10 indicate the beginning of the curvatures, suggesting that the participant starts focusing on steering more.

Table VI summarizes the results of the aggregated combined and individual fixation behavior analysis. This table captures the average fixations on the opponent over the traveled distance in a percentage with corresponding statistical Pearson values [50] of the linear regressions. No clear statistical differences can be argued. Therefore, a multi-linear regression [51] is performed between conditions on the de-

pendent average fixation variable with independent variables: CRT and relative velocity. The relative velocity describes the condition. Based on the results shown in Table VII, I found that participants' number of fixations on the opponent increased during more prolonged conflicts. Furthermore, the relative velocity variable tells us that the number of fixations decreases if the relative velocity is increased. Hence, the most fixations occurred in condition 1 (50-50 km/h) and the fewest fixations occur in condition 3 (60-40 km/h)

Figure 12 illustrates whether the participant ahead or behind fixates more on the opponent. This boxplot captures all the data from Figure 11 and enables the possibility of performing a 2-sample T-test [52] to show statistical differences between the behind and ahead participant within condition 2 and 3. The T-test outcomes are summarized in Table VIII and show that the participant ahead fixates more on the opponent than the participant behind. A multi-linear regression is performed on the same data to further investigate this result and is summarized in Table IX. In this multi-linear regression, an extra independent variable, named behind or ahead, is added. The results also indicate that the participant ahead fixates more on the opponent than the participant behind.

My findings can be summarized as 1) a higher CRT results in higher average fixations, 2) increasing the relative velocity results in fewer average fixations, and 3) the participant ahead fixates more on the opponent.

TABLE VI: Summary of the results of the fixation behavior analysis.

Condition	Fixation on opponent [%]	r	p
1	0.31	0.31	$4.44e^{-6}$
2 participant behind	0.24	0.02	0.84
2 participant ahead	0.31	0.43	$3.28e^{-6}$
3 participant behind	0.18	-0.22	0.03
3 participant ahead	0.31	0.15	0.13

TABLE VII: Results of the multi-linear regression on all conditions with dependent variable: average fixations and independent variables: CRT and relative velocity.

Independent variable	Regression coefficient	T	p
CRT	$6.5e^{-3}$	3.63	$5.74e^{-4}$
Relative velocity	$-4.4e^{-3}$	-4.59	$5.47e^{-6}$

TABLE VIII: Two-sample T-test summary of Figure 12.

Condition	t	p
2	3.56	$4.61e^{-4}$
3	5.77	$3.02e^{-8}$

TABLE IX: Results of the multi-linear regression on conditions 2 and 3 with dependent variable: average fixations and independent variables: CRT and behind or ahead.

Independent variable	Regression coefficient	T	p
CRT	$5.0e^{-3}$	2.31	$2.14e^{-2}$
Relative velocity	$-7.1e^{-3}$	-3.62	$3.30e^{-4}$
Behind or ahead	$7.3e^{-2}$	3.29	$1.09e^{-3}$

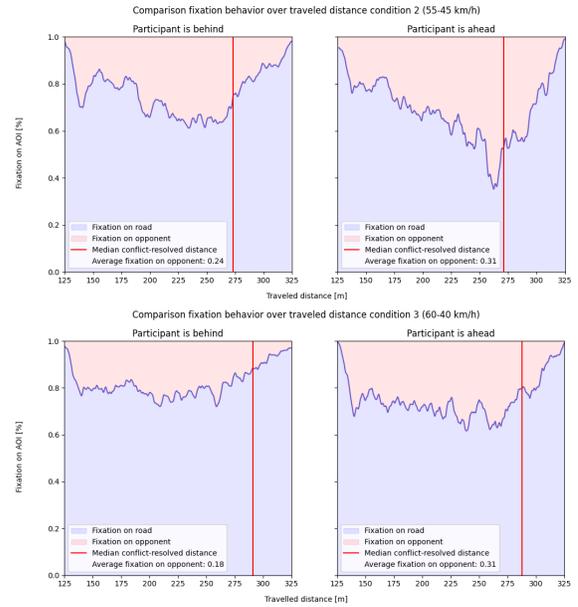


Fig. 10: The aggregated individual fixation behavior is divided into a participant being ahead or behind for conditions 2 and 3. The blue area illustrates fixations on the road, and the red area represents fixations on the opponent during the interactive area. The red line illustrates the median conflict-resolved distance

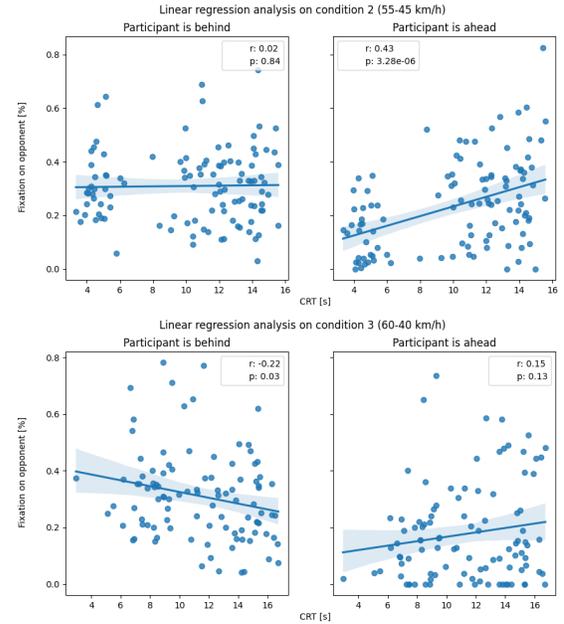


Fig. 11: Linear regression visualizations for conditions 2 and 3. The average fixations are computed for each CRT individually, indicated by the dots.

3.5 Before-after CRT analysis. Intuitively it is expected that participants' fixation behavior differs between being in conflict and after resolving it. Accordingly, this section analyses the fixation behaviors before and after solving the conflict within conditions. Multiple linear regressions are performed on the before-after data to indicate an increase or decrease in average fixations.

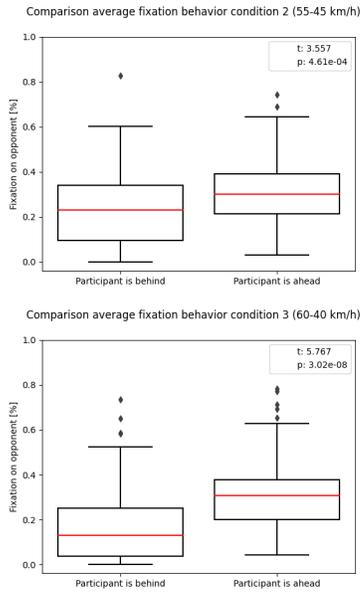


Fig. 12: Boxplot visualization of the average fixations on the opponent for conditions 2 and 3, divided into being ahead or behind.

3.5.1 Combined before-after CRT analysis. In section 3.4, it is argued that the aggregated data can best be visualized over the traveled distance. In this chapter, it is best to visualize the aggregated data over time. The reason behind this is that a specific CRT is related to an individual trace. If one subtracts the CRT from the particular time trace, the trace can be divided into two parts: 1) negative time representing fixation behavior before the CRT and 2) positive time representing fixation behavior after the CRT. Furthermore, the CRT gets aligned at time zero for each individual trace in the data set. This enables the visual comparison of the fixation behavior before-after the CRT over all trials and is shown in Figure 13. In this figure, the data is linearly interpolated so that the aggregated data only contains individual traces with an equal number of data points.

3.5.1 Individual before-after CRT analysis. Figure 15 shows the aggregated fixation behavior of participants before and after the CRT for conditions 2 and 3, divided into a participant being ahead and behind. The same approach as in section 3.5.1 is used to complete the figure.

3.5.3 Results before-after CRT analysis. As argued in section 1.2, I presume that participants fixate more during a conflict than after. Therefore, this before-after CRT analysis aims to show differences in participants' average fixation behavior during and after solving the conflict. To compare equal amounts of data, it is chosen to study the fixation behavior 5 seconds before and after the CRT. I chose the 5 seconds since Figures 13 and 15 show the most fixations around this period.

Table X summarizes the fixation behaviors of Figure 13 and Figure 15 in a percentage. This table suggests that

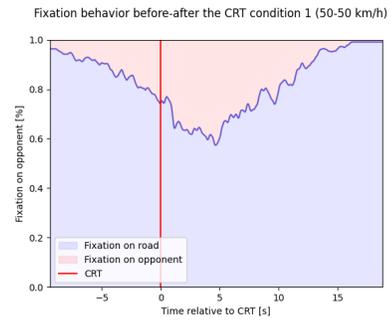


Fig. 13: Before-after CRT analysis for condition 1. The figure shows the fixation behavior in percentages for both areas of interest, the CRT at zero, negative time corresponding to fixation behaviors before the CRT, and positive time representing fixation behaviors after the CRT.

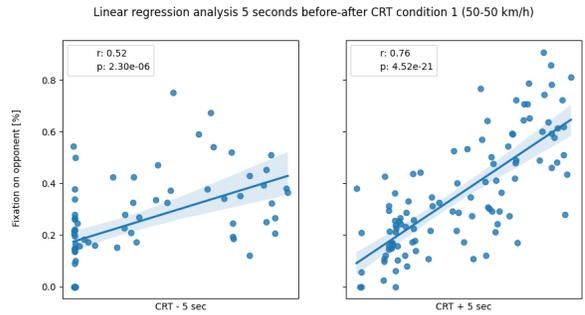


Fig. 14: Linear regression analysis 5 seconds before and after the CRT for condition 1.

TABLE X: Summary of 5 seconds before-after CRT analysis for each condition. The values represent percentages.

Condition	CRT - 5 sec	CRT + 5 sec
1	0.16	0.35
2 participant behind	0.14	0.29
2 participant ahead	0.19	0.38
3 participant behind	0.12	0.1
3 participant ahead	0.19	0.17

participants fixate more on the opponent after the CRT than before the CRT in conditions 1 and 2, whereas this is the opposite for condition 3.

To statistically dive into the numbers of Table X, linear regressions are performed in Figure 14, Figure 16, and Figure 17 for condition 1, 2, and 3 respectively. In these visualizations, the average fixation of 5 seconds before and after the CRT is computed for each individual trace. In each individual trace, the time located at the index of negative or positive 5 seconds is used to match the average fixation, resulting in the dots. Table XI summarizes the Pearson values of the linear regressions. Based on these results, I cannot statistically prove clear differences in fixations 5 seconds before or after the CRT.

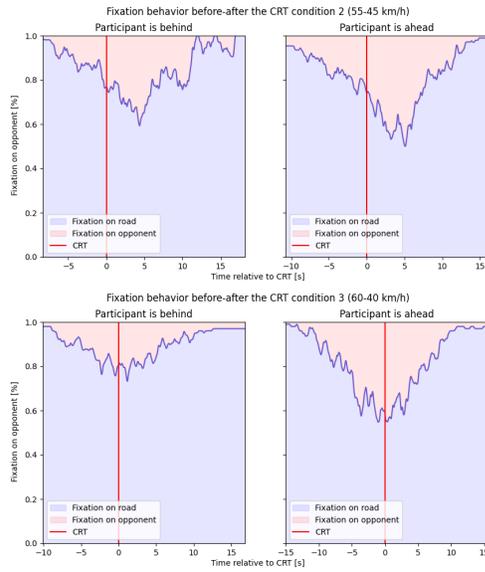


Fig. 15: Individual fixation behavior before-after CRT for conditions 2 and 3, divided into being ahead or behind. The fixation behavior is visualized against time. The CRT is located at zero, negative time represents fixation behavior before the CRT, and positive time represents fixation behavior after the CRT.

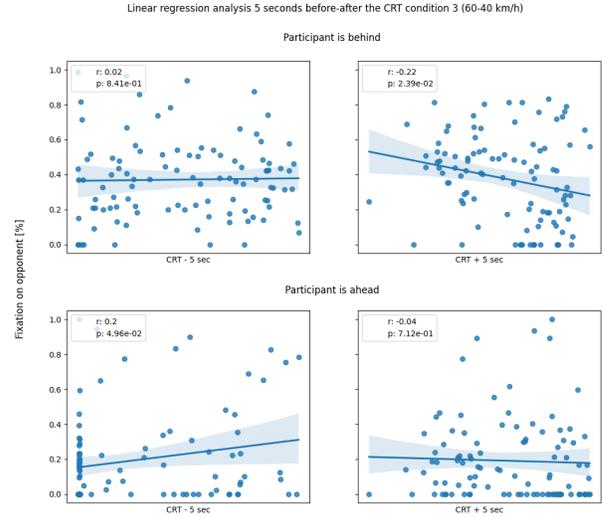


Fig. 17: Linear regression analysis 5 seconds before and after the CRT for condition 3. The top two figures illustrate the linear regression for the participant being behind. The lower two figures illustrate the linear regression for the participant being ahead.

TABLE XI: Summary of 5 seconds before-after CRT analysis divided into combined, being behind, or being ahead.

Conditon	Peason value	CRT - 5 sec	CRT + 5 sec
1	r	0.52	0.76
	p	$2.30e^{-6}$	$4.52e^{-21}$
2 participant behind	r	0.18	0.47
	p	$9.69e^{-2}$	$3.96e^{-7}$
2 participant ahead	r	0.3	0.54
	p	$8.65e^{-3}$	$2.18e^{-9}$
3 participant behind	r	0.02	-0.22
	p	$8.41e^{-1}$	$2.39e^{-2}$
3 participant ahead	r	0.20	-0.04
	p	$4.96e^{-2}$	$7.12e^{-1}$

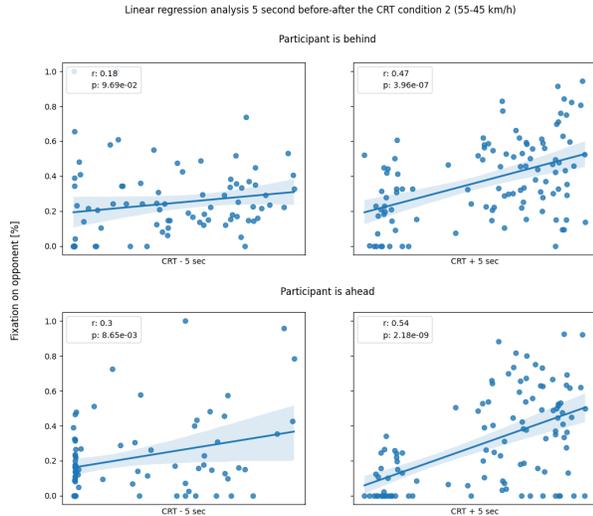


Fig. 16: Linear regression analysis 5 seconds before and after the CRT for condition 2. The top two figures illustrate the linear regression for the participant being behind. The lower two figures illustrate the linear regression for the participant being ahead.

4. DISCUSSION

This chapter starts with a discussion of the methods. Next, it discusses the findings and if the hypotheses can be confirmed or denied. Finally, a glance is given at the implications of this work.

4.1 Discussion of methods. In order to let the participants behave as naturally as they would in real driving scenarios, I

considered two factors during the environment- and experiment design. The first consideration concerns disregarding velocity cues (e.g., trees, speed signs). The speedometer gives the participant some visual feedback on the velocity, but it is limited and makes speeding easy. Better velocity indicators are useful in virtual reality since perceiving speed is difficult [53, 54]. Literature shows various solutions to improve participants' speed perception. For instance, sound effects of the vehicle's engine are proven to diminish speeding [55]. Advanced auditory and visual feedback is not included in this work, despite the benefits it could have on the studied human behavior. The second consideration is about attaching mirrors to the vehicles. The Varjo head mounts could only be used in the newest version of Unreal Engine, which consequently led to the newest version of Carla. This Carla version is based on a forked Unreal Engine. In this forked version, the option to use mirrors is not available. A workaround is found, yet this led to a tremendous decrease in process performance. Accordingly, mirrors are excluded, and an increased approach angle is chosen, so mirrors are not

needed during the interactive section. Nevertheless, mirrors are presumably important in future research when studying participants' natural behavior in different traffic scenarios.

Another limitation is that participants are free to steer on a 6-meter wide road, which leads to deviations from the middle of the road. This could have led to unrealistic traveled distances, resulting in wrongful related data. Furthermore, the curvatures around the merging point are needed to merge the interactive sections into the straight section smoothly. However, a considerable simplification is made here. The merge point is taken as a predefined point, but in practice, it should be a merge line since participants always arrive differently at the merge point due to the curvatures in the road and free steering. Hence, it remains to be determined how the collision area should be defined based on the curvatures and merge line. It is recommended to do further research in correctly converting the merge point into a line. Having a more realistic collision area will probably result in different CRTs.

Appendix F shows another limitation which is the negative velocity bump recurring every time one takes control (i.e., uses the pedals). In some cases, this has a negative effect on the CRT since the velocity of the vehicles is changed unknowingly to a different velocity (i.e., a higher initial velocity results in a more significant velocity drop). Hence, in some trials, it is not reliable that the participant solved the conflict, but it was solved due to this random error. Therefore, in future work, one must investigate which software (JOAN or Unreal Engine) causes this error.

After evaluating the experiment, the interactive section might be too long. This is due to several reasons. First, the CRT is defined as the final moment in time participants deviate from the collision course. This definition arose since multiple conflicts are generally solved during a trial, indicating that participants have much time to alternate between being on a collision course or not. Second, as argued above, it is difficult for participants to perceive speed. The long interactive section only amplifies this effect since the other vehicle is often quite far away in the participants' viewing range. Therefore, I expect more realistic interactions with a shorter interactive section. Also, fewer alternations between being on a collision course or not are expected, which probably results in different CRTs.

Another significant limitation needs to be discussed. The fixation data for the participant behind might be heavily influenced by the fact that the participant needs less steep head rotations due to almost or already (i.e., this depends on the velocity profile of the ahead vehicle) having the other vehicle in the viewing range when exiting the tunnel. Consequently, one does not cross the threshold used to indicate when one switches attention between the areas of interest, while in practice, the participant is fixating on the opponent. As described in section 3.4.3, the ahead vehicle fixates more on the opponent, yet this result might be different since the average fixation for the participant behind are probably higher. In future research, it is advised to include the forward $[x, y]$ gaze data combined with head rotations instead of head

rotations only.

A wide variety of statistical techniques exist to empower the statistical significance of data in behavioral science [56, 57]. Even within specific techniques, more in-depth approaches can be used to prove relations. For example, statistical interactions [58] investigate an independent variable's effect on another independent variable. In this work, more straightforward techniques are used to illustrate relations. Nevertheless, these techniques still suffice in answering the hypotheses statistically.

A final limitation of the method is the number of participants. Fourteen participants volunteered in this work, which resulted in an interesting analysis with quite some data. However, the results will be strengthened if more participants are recruited (i.e., increasing the statistical power) [59].

4.1 the effect of velocity differences on the merging behavior of participants. Section 3.1 shows the probability of merging first. The probability is statistically justified with a logistic regression on the binary merging data, and results show a negative regression coefficient with a $p < 0.05$, indicating that the ahead vehicle most likely merges first with a statistical significance. Hence, the null hypothesis can be rejected, confirming hypothesis 1 in Table I.

4.2 The effect of velocity differences on the CRT. The metric in section 3.3 captures all the CRTs per condition, and the ANOVA test statistically compares the differences in medians. The ANOVA test shows a $p < 0.05$, suggesting that increasing the relative velocity increases the CRT. Thus, hypothesis 2 in Table I can be denied based on this finding.

A possible explanation that substantiates the findings is proposed in the work of Siebinga et al. [31]. They approach conditions according to the level of conflict. The level of conflict concept with respect to this work could be summarized as follows: smaller deviations in relative velocity during the interactive section are needed in condition 1 to solve the conflict compared to conditions 2 and 3. This is because the initial headway is zero in condition 1, resulting in a flat trajectory line toward the collision area. In contrast, in conditions 2 and 3, the trajectory trace starts under an angle. This difference holds that changes in relative velocity in condition 1 will solve the conflict more easily compared to conditions 2 and 3. The level of conflict concept is excluded in this work, yet it might have shown interesting findings that enhance answering the hypothesis.

Another possible effect of this result could be that increasing the relative velocity leads to unawareness in combination with the limitation on speed perception. This assumption mainly leads to a limitation on CRTs found in conditions 2 and 3. The CRT is computed based on the assumption that the interaction starts when the first vehicle exits the tunnel. In condition 1, both vehicles exit the tunnel simultaneously, so it is known that the interaction starts immediately. Contrary to conditions 2 and 3. In these conditions, there is a difference in time between the vehicles exiting the tunnels. Appendix G shows the effect of different initial relative velocities by

visualizing the velocity traces of the first vehicle exiting the tunnel over all trials for conditions 2 and 3. Results show that participants in almost all cases start speeding shortly after exiting the tunnel. This has the most effect on condition 3. Herefore, the assumption of starting the interaction when the first vehicle exits the tunnel might heavily influence the found CRTs. In future research, decreasing the relative velocity differences between conditions is advised.

4.3 the effect of different conflict durations on participants' fixation behavior. Hypothesis 3 in Table I aims to find an answer to the relation between the CRT and participants' fixation behavior. The analysis in section 3.4 shows the trends within conditions with corresponding Pearson values. To reject the null hypothesis within a condition where multiple samples are compared, a Bonferroni correction [60] needs to be applied to decrease the false positive rate. Hence, the value to reject the null hypothesis is $\alpha < 0.01$ (i.e., 5 samples). Based on this correction, the null hypothesis can be rejected for condition 1 and for the participant ahead in condition 2. The failure to reject the null hypothesis for the participant being behind in conditions 2 and 3 may be due to the limitation of not exceeding the threshold, as argued in section 4.1. Based on these findings, I conclude that the individual linear regression analyses within the conditions do not immediately indicate a clear trend with statistical significance. Therefore, a multi-linear regression analysis between the conditions is performed to show a clear relation with statistical significance. The model indicates that higher CRTs lead to more fixations on the opponent, confirming hypothesis 3.

4.4 The differences in participants' fixation behavior during the interactive section. Hypothesis 4 in Table I investigates differences in fixation behavior between being ahead or behind. There is no difference in fixation behaviors in condition 1 since both participants are positioned in the same situation. Hence, this hypothesis applies to the results of conditions 2 and 3. The null hypothesis can be rejected based on comparing the p-values of the T-test to the Bonferroni correction and by the multi-linear regression, confirming hypothesis 4.

4.5 The differences in participants' fixation behavior before and after the CRT. It is chosen to study the fixation behaviors 5 seconds before and after since the most average fixations occurred during this period. This shows an interesting finding which indicates that participants focus heavily around the CRT. In conditions 1 and 2, the trend 5 seconds after shows a more substantial r value than before, indicating that participants fixate more after the CRT than before with statistical significance. On the contrary, the Pearson values in condition 3 show the opposite result, suggesting that participants fixate less after the CRT than before. Furthermore, the statistical power is insignificant in two out of four cases.

Based on these findings, no clear answers are found to hypothesis 5 in Table I, and more extensive statistical techniques are advised to find answers to this hypothesis.

4.6 Discussion of implications. Results show that the proposed metrics by Siebinga et al. [31] can also be used for human behavior analyses in the coupled virtual reality simulator. The framework shows that traditional velocity and position visualization can be extended. Moreover, using the framework can aid in comparing human-human behavior in interactive scenarios across trials and between conditions in a meaningful and realistic way.

Integrating fixation behavior can be valuable for researchers incorporating human behavior models since it reveals humans' attention during driving scenarios. For example, Siebinga et al. [31] assume that participants constantly observe the other vehicle during a trial, which is clearly not the case based on the shown results.

5. CONCLUSION

This work showed a framework to study interactive human-human merging behavior in a coupled virtual reality simulator. The analyses are done on participants' merging behavior, interactive pair-wise behavior, conflict resolution time, and fixation behavior. Furthermore, the conflict resolution time is related to humans' fixation behavior. Various statistical techniques are used to justify answers regarding the relation between participants' interactive behavior and fixation behavior during a simple merging scenario. The main findings can be summarized as follows:

- 1) Humans who are ahead of the opponent most likely merge first.
- 2) Increasing the relative velocity yield longer conflicts.
- 3) Humans focus more on the opponent during longer conflicts.
- 4) Humans observe the opponent more often when they are initially ahead compared to humans positioned behind.

No clear answer is found to the following sub-question:

- 1) What is the effect of velocity differences on the number of participants' head rotations before and after the conflict?

This work may have utility to researchers that incorporate human behavior models since it shows that integrating fixation behavior is essential to understand human-human merging behavior. The simulator is believed to be a valuable tool for studying interactive merging behavior between a pair of human drivers. Moreover, future work can extend the simulator to study interactive human-human behavior in various traffic scenarios. The ultimate goal is to understand humans better and improve traditional human behavior models, which will enhance the safety and predictions of autonomous vehicles.

DATA AVAILABILITY

The experiment data, analysis code, and figures can be downloaded via the following link: <https://figshare.com/s/65650f7761e3a4389e08>.

ACKNOWLEDGEMENTS

I would like to thank Olger Siebinga and Arkady Zgonnikov for their excellent assistance during the project. Each meeting led to interesting new insights, and they were always happy to help. Further acknowledgment goes to Olger for his help in building the simulator. Also, many thanks to all participants that volunteered in this research.

REFERENCES

- [1] "Incorporating human-factors in car-following models: A review of recent developments and research needs," pp. 379–403, Nov. 2014, [Online]; accessed 3, Jan. 2023].
- [2] X. Di and R. Shi, "A Survey on Autonomous Vehicle Control in the Era of Mixed-Autonomy: From Physics-Based to AI-Guided Driving Policy Learning," *arXiv*, Jul. 2020.
- [3] W. Schwarting, J. Alonso-Mora, and D. Rus, "Planning and Decision-Making for Autonomous Vehicles," *Annu. Rev. Control Rob. Auton. Syst.*, May 2018.
- [4] "Research Advances and Challenges of Autonomous and Connected Ground Vehicles," Jan. 2023, [Online]; accessed 3, Jan. 2023]. [Online]. Available: <https://ieeexplore-ieee-org.tudelft.idm.oclc.org/abstract/document/8936542?casa.token=KYioLmJEJ90AAAAA:udnrODTI02f.4nSHcYP8dOCjRTDfy5UHK85k8.Wzo48GAQzR-C4BEkMFLDSSu4I2ggEY0MAR2w>
- [5] O. Siebinga, A. Zgonnikov, and D. Abbink, "Validating human driver models for interaction-aware automated vehicle controllers: A human factors approach," *arXiv*, Sep 2021. [Online]. Available: <https://arxiv.org/abs/2109.13077v1>
- [6] M. Werling, T. Gindele, D. Jagszent, and L. Groll, "A robust algorithm for handling moving traffic in urban scenarios," in *2008 IEEE Intelligent Vehicles Symposium*, 2008, pp. 1108–1112.
- [7] V. Raman, A. Donzé, D. Sadigh, R. M. Murray, and S. A. Seshia, "Reactive synthesis from signal temporal logic specifications," ser. HSCC '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 239–248. [Online]. Available: <https://doi-org.tudelft.idm.oclc.org/10.1145/2728606.2728628>
- [8] J. E. Naranjo, C. Gonzalez, R. Garcia, and T. de Pedro, "Lane-change fuzzy control in autonomous vehicles for the overtaking maneuver," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 3, pp. 438–450, 2008.
- [9] J. Nilsson, J. Silvlin, M. Brännström, E. Coelingh, and J. Fredriksson, "If, When, and How to Perform Lane Change Maneuvers on Highways," *IEEE Intell. Transp. Syst. Mag.*, vol. 8, pp. 68–78, Jan 2016.
- [10] J. Eggert, F. Damerow, and S. Klingelschmitt, "The foresighted driver model," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, 2015, pp. 322–329.
- [11] E. Julian and F. Damerow, "Complex lane change behavior in the foresighted driver model," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 1747–1754.
- [12] J. Wei, J. M. Dolan, and B. Litkouhi, "Autonomous vehicle social behavior for highway entrance ramp management," in *2013 IEEE Intelligent Vehicles Symposium (IV)*, 2013, pp. 201–207.
- [13] N. Evestedt, E. Ward, J. Folkesson, and D. Axehill, "Interaction aware trajectory planning for merge scenarios in congested traffic situations," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 465–472.
- [14] E. Ward, N. Evestedt, D. Axehill, and J. Folkesson, "Probabilistic model for interaction aware planning in merge scenarios," *IEEE Transactions on Intelligent Vehicles*, vol. 2, no. 2, pp. 133–146, 2017.
- [15] O. Speidel, M. Graf, A. Kaushik, T. Phan-Huu, A. Wedel, and K. Dietmayer, "Trajectory Planning for Automated Driving in Intersection Scenarios using Driver Models," *arXiv*, Oct 2020. [Online]. Available: <https://arxiv.org/abs/2010.03345v1>
- [16] S. Hoermann, D. Stumper, and K. Dietmayer, "Probabilistic long-term prediction for autonomous vehicles," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, 2017, pp. 237–243.
- [17] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic Decision-Making under Uncertainty for Autonomous Driving using Continuous POMDPs," *2014 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014*, Oct 2014.
- [18] M. Bouton, A. Nakhaei, K. Fujimura, and M. J. Kochenderfer, "Cooperation-Aware Reinforcement Learning for Merging in Dense Traffic," *arXiv*, Jun 2019. [Online]. Available: <https://arxiv.org/abs/1906.11021v1>
- [19] O. Siebinga, "Driver knowledge progress report," Oct 2021.
- [20] D. Isele, "Interactive Decision Making for Autonomous Vehicles in Dense Traffic," *ResearchGate*, pp. 3981–3986, Oct 2019.
- [21] Q. Zhang, D. Filev, H. E. Tseng, S. Szwabowski, and R. Langari, "Addressing mandatory lane change problem with game theoretic model predictive control and fuzzy markov chain," in *2018 Annual American Control Conference (ACC)*, 2018, pp. 4764–4771.
- [22] H. Yu, H. E. Tseng, and R. Langari, "A human-like game theory-based controller for automatic lane changing," *Transportation Research Part C: Emerging Technologies*, vol. 88, pp. 140–158, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0968090X18300494>
- [23] D. A. Lazar, K. Chandrasekher, R. Pedarsani, and D. Sadigh, "Maximizing Road Capacity Using Cars that Influence People," *arXiv*, Jul 2018. [Online]. Available: <https://arxiv.org/abs/1807.04414v2>
- [24] R. Tian, S. Li, N. Li, I. Kolmanovsky, A. Girard, and Y. Yildiz, "Adaptive Game-Theoretic Decision Making for Autonomous Vehicle Control at Roundabouts," *arXiv*, Oct 2018. [Online]. Available: <https://arxiv.org/abs/1810.00829v1>
- [25] N. Li, D. W. Oyler, M. Zhang, Y. Yildiz, I. Kolmanovsky, and A. R. Girard, "Game theoretic modeling of driver and vehicle interactions for verification and validation of autonomous vehicle control systems," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 5, pp. 1782–1797, 2018.
- [26] N. Li, I. Kolmanovsky, A. Girard, and Y. Yildiz, "Game theoretic modeling of vehicle interactions at unsignalized intersections and application to autonomous vehicle control," in *2018 Annual American Control Conference (ACC)*, 2018, pp. 3215–3220.
- [27] M. Garzón and A. Spalanzani, "Game theoretic decision making for autonomous vehicles' merge manoeuvre in high traffic scenarios," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 3448–3453.
- [28] M. Wang, S. P. Hoogendoorn, W. Daamen, B. van Arem, and R. Happee, "Game theoretic approach for predictive lane-changing and car-following control," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 73–92, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0968090X15002491>
- [29] F. Meng, J. Su, C. Liu, and W.-H. Chen, "Dynamic decision making in lane change: Game theory with receding horizon," in *2016 UKACC 11th International Conference on Control (CONTROL)*, 2016, pp. 1–6.
- [30] S. Coskun, Q. Zhang, and R. Langari, "Receding horizon markov game autonomous driving strategy," in *2019 American Control Conference (ACC)*, 2019, pp. 1367–1374.
- [31] D. A. A. Olger Siebinga, Arkady Zgonnikov, "Interactive merging behavior in a coupled driving simulator," 2022. [Online]. Available: <https://repository.tudelft.nl/islandora/object/uuid:c18ff282-a14a-4bc1-8003-993fca6a8579>
- [32] D. R. Mestre, F. Mars, S. Durand, F. Vienne, and S. Espié, "Gaze Behavior During Simulated Driving: Elements for a Visual Driving Aid," *Driving Assessment Conference*, vol. 3, no. 2005, Jun. 2005.
- [33] A. Löcken, F. Yan, W. Heuten, and S. Boll, "Investigating driver gaze behavior during lane changes using two visual cues: ambient light and focal icons," *J. Multimodal User Interfaces*, vol. 13, no. 2, pp. 119–136, Jun. 2019.
- [34] "Gaze behavior when approaching an intersection: Dwell time distribution and comparison with a quantitative prediction," pp. 60–74, Nov. 2015, [Online]; accessed 4, Jan. 2023].
- [35] J. A. Abbasi, D. Mullins, N. Ringelstein, P. Reilhac, E. Jones, and M. Glavin, "An Analysis of Driver Gaze Behaviour at Roundabouts," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8715–8724, Jun. 2021.
- [36] A. Calvi and M. R. De Blasiis, "Driver Behavior on Acceleration Lanes," *Transp. Res. Rec.*, Jan. 2011.

- [37] M. Hunter, R. Machemehl, and A. Tsyganov, "Operational Evaluation of Freeway Ramp Design," *Transp. Res. Rec.*, vol. 1751, no. 1, pp. 90–100, Jan. 2001.
- [38] M. A. Brewer, K. Fitzpatrick, and J. Stanley, "Driver Behavior on Speed-Change Lanes at Freeway Ramp Terminals," *Transp. Res. Rec.*, Jan. 2011.
- [39] Steam. (2022) Steam with steamvr plugin. [Online]. Available: <https://store.steampowered.com/app/250820/SteamVR/>
- [40] Matlab. (2022) Roadrunner. [Online]. Available: <https://nl.mathworks.com/products/roadrunner.html>
- [41] Unreal. (2022) Unreal engine 4.27 documentation. [Online]. Available: <https://docs.unrealengine.com/en-US/index.html>
- [42] N. Beckers, O. Siebinga, Giltay, and V. der Kraan. (2021) Joan, a human-automated vehicle experiment framework. [Online]. Available: <https://github.com/tud-hri/joan>
- [43] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," 2017.
- [44] E. Chang, H.-T. Kim, and B. Yoo, "Virtual Reality Sickness: A Review of Causes and Measurements," *International Journal of Human-Computer Interaction*, vol. 36, pp. 1–25, Jul. 2020.
- [45] O. Siebinga, "A service by surf." [Online]. Available: <https://surfdrive.surf.nl/files/index.php/s/m55UYiSri8X4bvV>
- [46] Varjo. (2022) Varjo open source api. [Online]. Available: <https://developer.varjo.com/>
- [47] J. Peng, K. Lee, and G. Ingersoll, "An Introduction to Logistic Regression Analysis and Reporting," *Journal of Educational Research - J EDUC RES*, vol. 96, pp. 3–14, Sep. 2002.
- [48] E. Ostertagova and O. Ostertag, "Methodology and Application of One-way ANOVA," *American Journal of Mechanical Engineering*, vol. 1, pp. 256–261, Nov. 2013.
- [49] K. Kumari and S. Yadav, "Linear regression analysis study," *Journal of the Practice of Cardiovascular Sciences*, vol. 4, p. 33, Jan. 2018.
- [50] E. I. Obilor and E. Amadi, "Test for Significance of Pearson's Correlation Coefficient ()," *ResearchGate*, Mar. 2018. [Online]. Available: https://www.researchgate.net/publication/323522779_Test_for_Significance_of_Pearson's_Correlation_Coefficient
- [51] G. Kaya Uyanık and N. Güler, "A Study on Multiple Linear Regression Analysis," *Procedia - Social and Behavioral Sciences*, vol. 106, pp. 234–240, Dec. 2013.
- [52] T. K. Kim, "T test as a parametric statistic," *Korean Journal of Anesthesiology*, vol. 68, no. 6, p. 540, Dec. 2015.
- [53] S.-J. Kwon, J.-H. Chun, J.-H. Bae, and M.-W. Suh, "A study on the factors that improve the velocity perception of a virtual reality-based vehicle simulator," *International Journal of Human-Computer Interaction*, vol. 21, pp. 39–54, Sep. 2006.
- [54] C. Diels and A. Parkes, "Optimisation of speed perception in virtual environments by manipulation of the geometric field of view," *Perception*, vol. 39, pp. 281–282, Jan. 2010.
- [55] T. Melman, P. Visser, X. Mouton, and J. de Winter, "Creating the Illusion of Sportiness: Evaluating Modified Throttle Mapping and Artificial Engine Sound for Electric Vehicles," *J. Adv. Transp.*, vol. 2021, Oct. 2021.
- [56] S. S. Teja Lanka, S. Rajtmajer, J. Wu, and C. L. Giles, "Extraction and Evaluation of Statistical Information from Social and Behavioral Science Papers," in *WWW '21: Companion Proceedings of the Web Conference 2021*. New York, NY, USA: Association for Computing Machinery, Apr. 2021, pp. 426–430.
- [57] J. Koster and R. McElreath, "Multinomial analysis of behavior: statistical methods," *Behav. Ecol. Sociobiol.*, vol. 71, Aug. 2017.
- [58] A. B. de González and D. R. Cox, "Interpretation of interaction: A review," *aoas*, vol. 1, no. 2, pp. 371–385, Dec. 2007.
- [59] J. de Winter and D. Dodou, *Human subject research for engineers: A practical guide*, May 2017.
- [60] R. Armstrong, "When to use the Bonferroni correction," *Ophthalmic Physiol. Opt.*, vol. 34, Apr. 2014.

APPENDIX

A. Overview entire track

Figure 18 shows the overview of the mirrored track. Both sides have equal dimensions. The walls function as boundaries and as waypoint indicators. Another wall is placed between the two sides to make the other side invisible to participants.

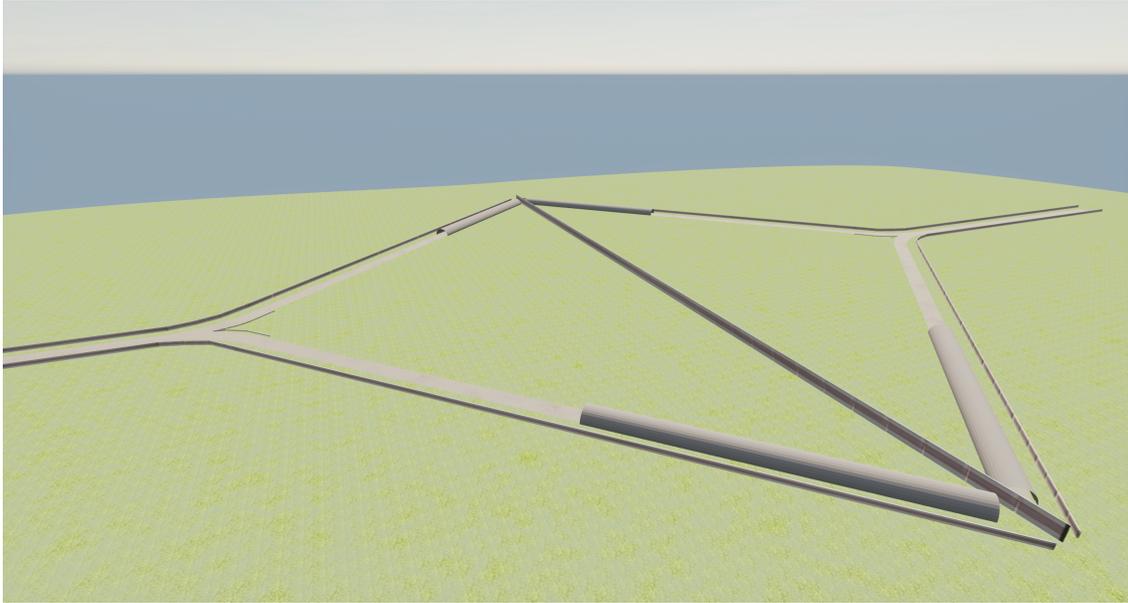


Fig. 18: Overview of the two-sided track.

B. Overview of the simplified curvatures at the merging point

Figure 19 shows the simplification made at the merging point.

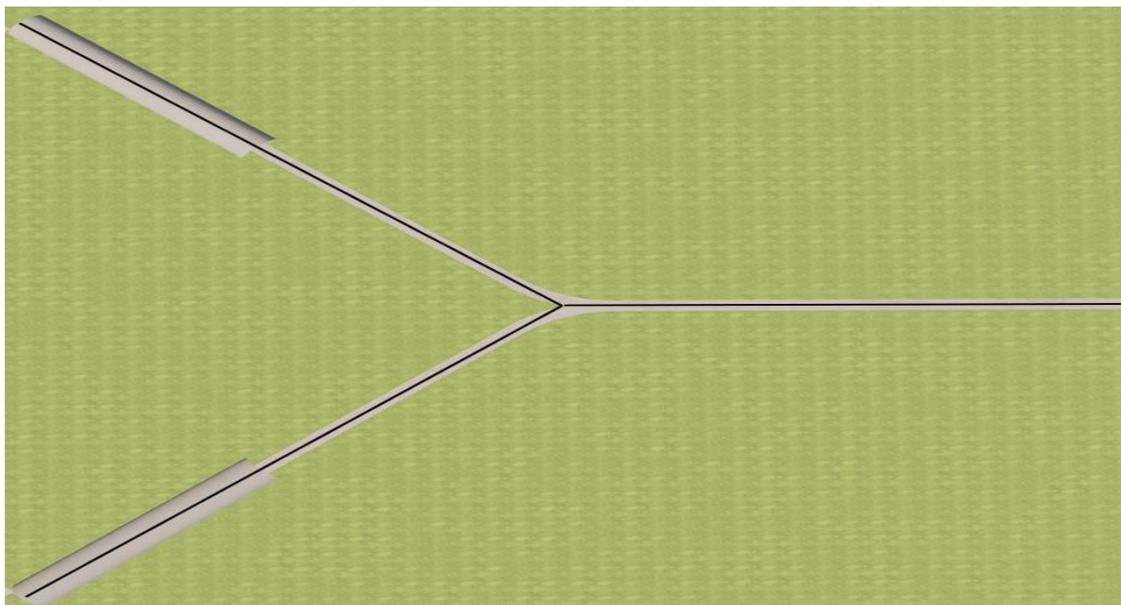


Fig. 19: One side of the track with the curvatures at the merging point simplified.

C. Analysis of trails with multiple conflict resolution times

Figure 20 visualizes the general pair-wise behavior for all conditions. During analysis, it is discovered that the CRT can be found multiple times. The average amount of CRTs found during a trial is 4.8, 6.2, and 5.5 for conditions 1, 2, and 3 respectively. After analyzing the trials it is concluded that the final CRT is the most reliable one in this work.

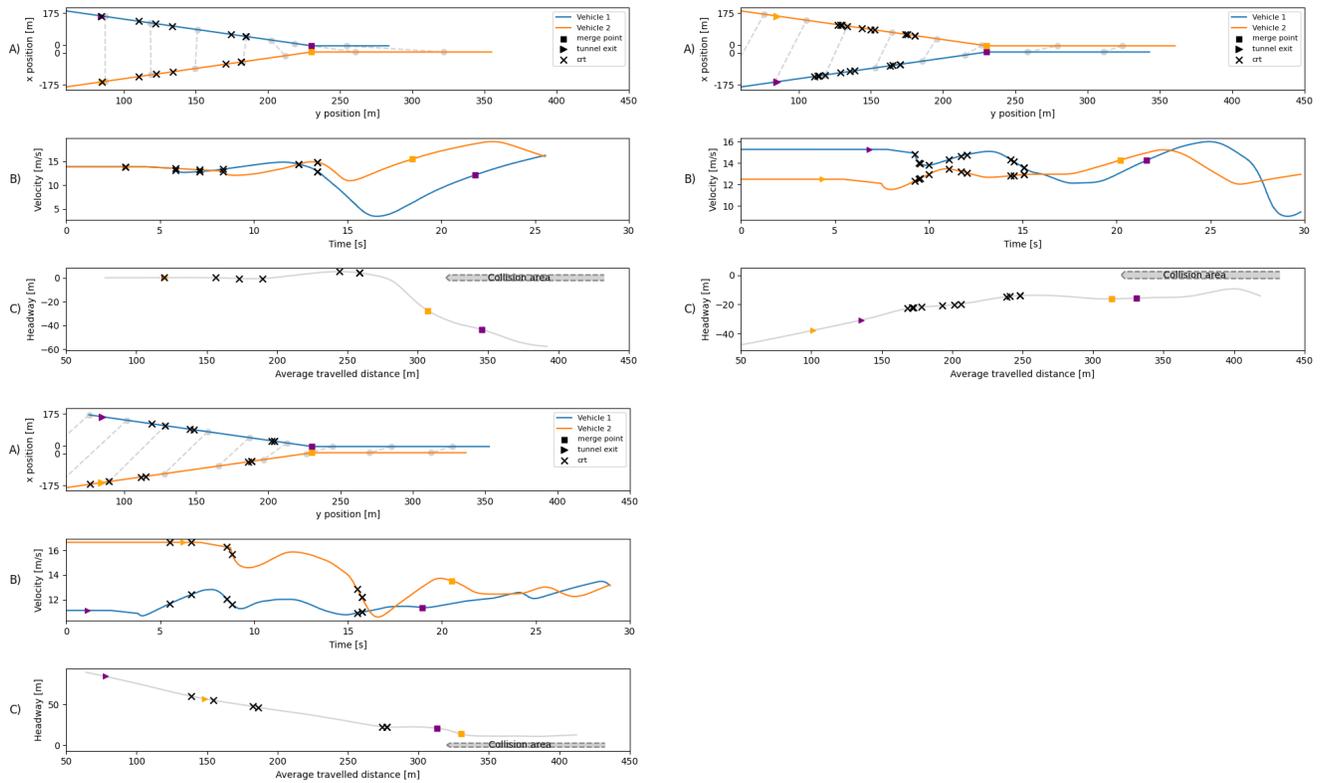


Fig. 20: Visualization of the pair-wise behavior for each condition. Results show that in each condition multiple CRTs are found.

D. Head mount rotation analysis to determine the threshold

Figure 21 shows two examples for each condition of the head mount rotation against time during the interactive section. The analysis shows that if participants look straight, the output is near 1 and each time their attention switches to the other vehicle, a peak arises. This peak is steeper near the merging point compared to peaks at the beginning of the interactive section. After analyzing all trails, a threshold of 0.95 is chosen. Values > 0.95 indicate that a participant is focused on the road, whereas values < 0.95 imply one is focused on the other vehicle. The figures for all trails can be found in the Data Availability section.

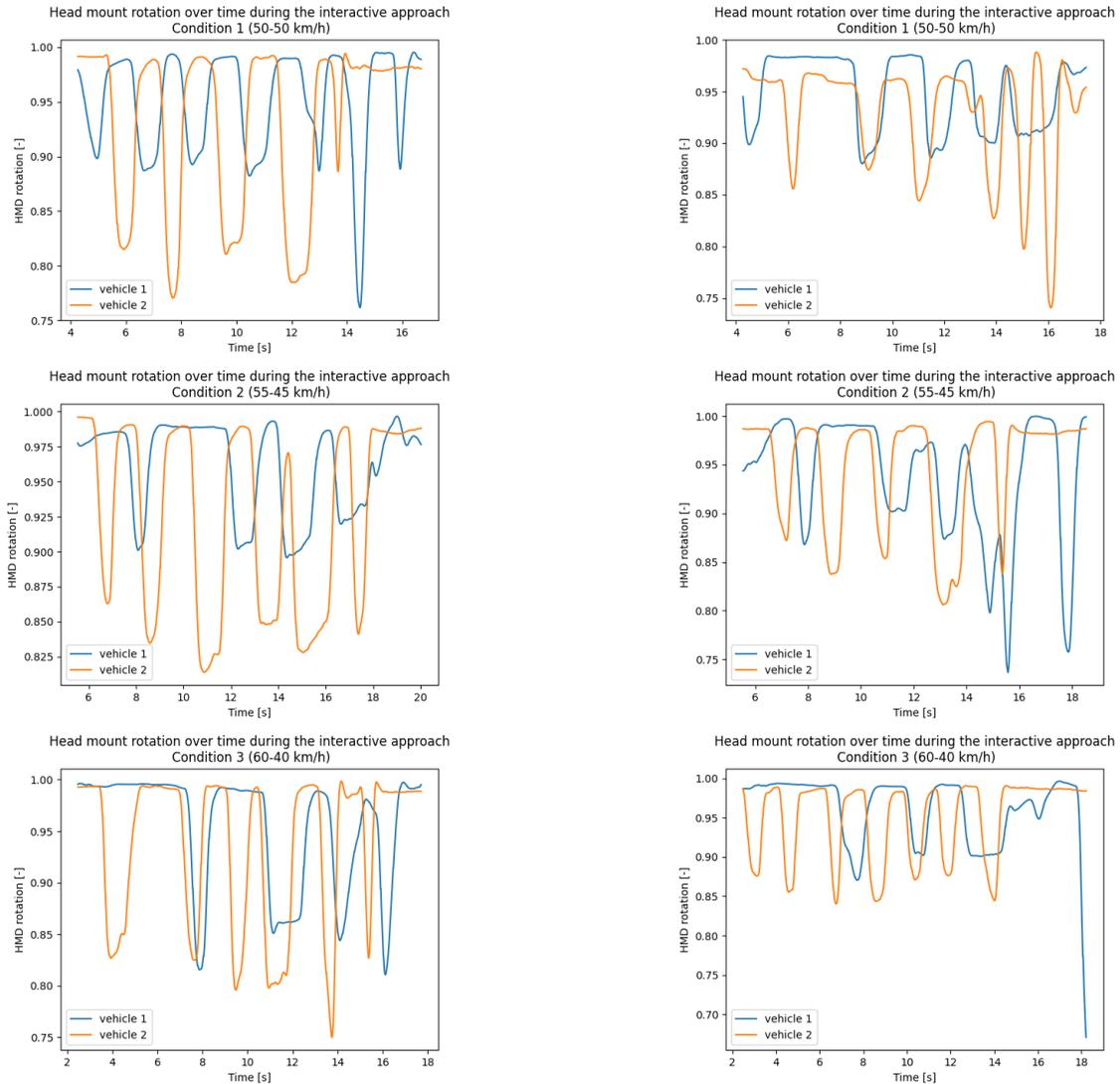


Fig. 21: Head mount rotations against time to determine the threshold when one switches between areas of interest. Peaks arise when one fixates on the the opponent.

E. Histograms for resolved conflict at traveled distance for each condition

Figures 22, 23, and 24 show the histograms and kernel density estimations of each condition. The histograms show the number of occurring traveled distances for specific CRTs captured in bins.

CRT distribution on average traveled distance condition 1 (50-50 km/h)

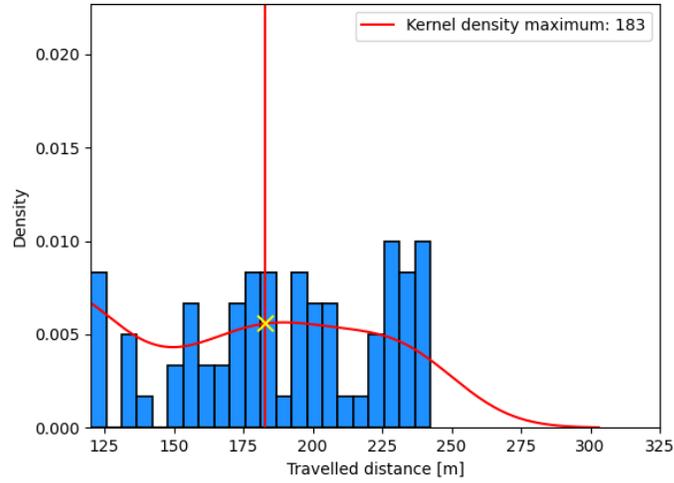


Fig. 22: Histogram and kernel density estimation condition 1

CRT distribution on average traveled distance condition 2 (55-45 km/h)

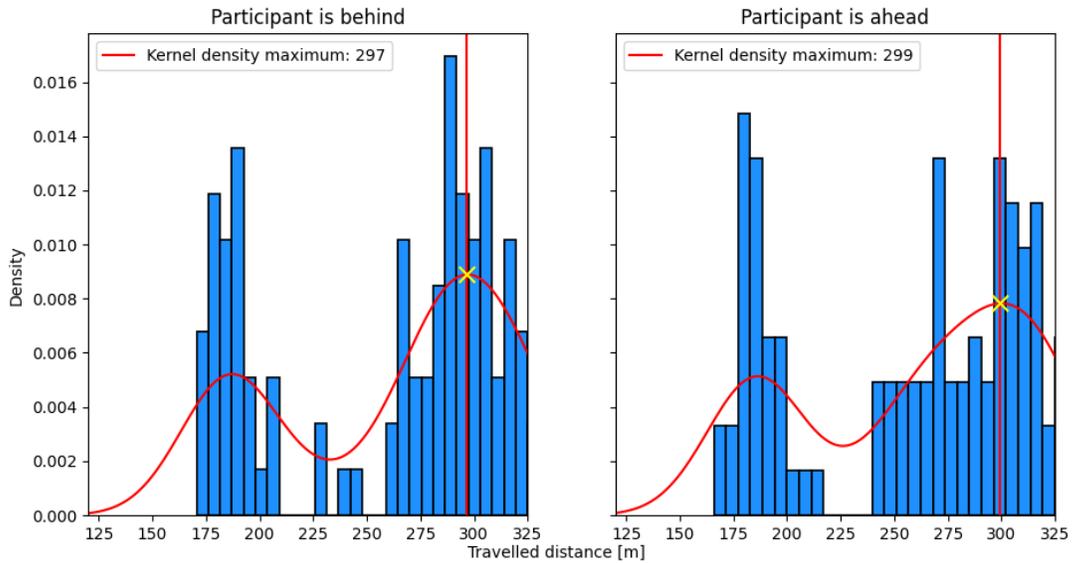


Fig. 23: Histogram and kernel density estimation condition 2

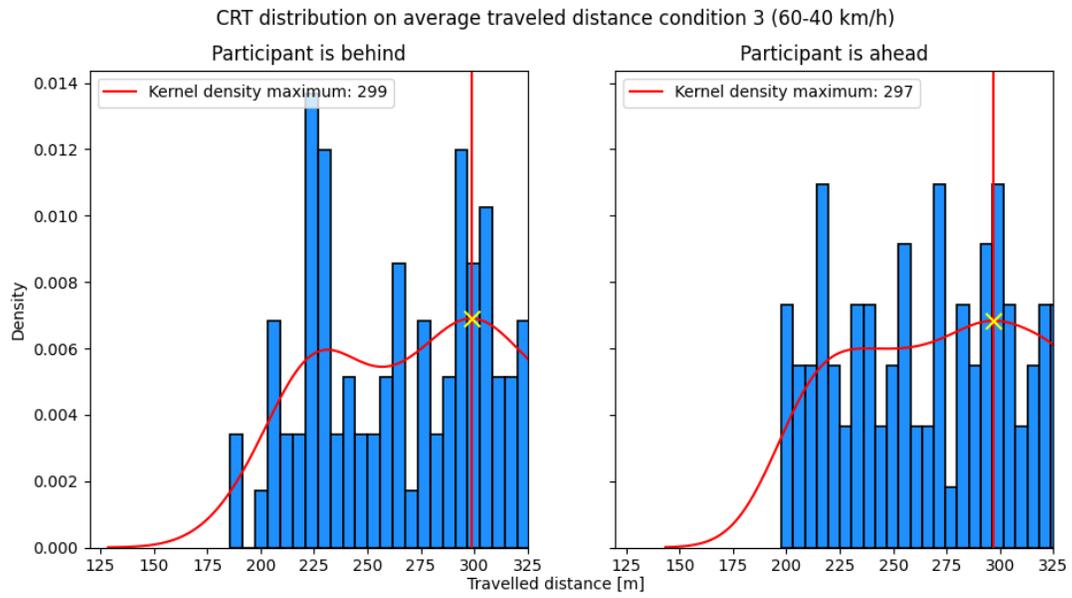


Fig. 24: Histogram and kernel density estimation condition 3

F. Limitation on velocity bumps

Figure 25 illustrates the negative velocity bumps with red squares. These bumps always occur when participants take control. This is dependent on the initial velocity. Greater initial velocities result in greater negative velocity bumps.

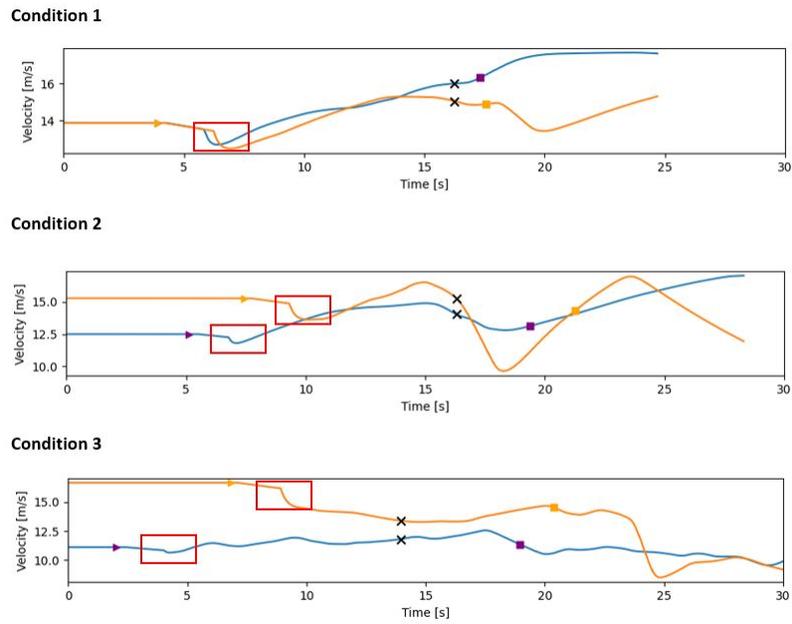


Fig. 25: Visualization of the negative velocity bumps for condition 1, 2, and 3.

G. Velocity traces ahead vehicle until behind vehicle leaves the tunnel

Figure 26 shows the velocity traces of the ahead vehicle exiting the tunnel until the behind vehicle exits the tunnel. What can be seen is that the ahead participants start speeding after exiting the tunnel. The time between the ahead and behind vehicle leaving the tunnel is 2.25 seconds in condition 2 and 4.67 seconds in condition 3. This suggests that speeding has the most effect on condition 3. Furthermore, it indicates that the interactive time for condition 3 is much longer compared to condition 2, which influences the outcome of the CRTs.

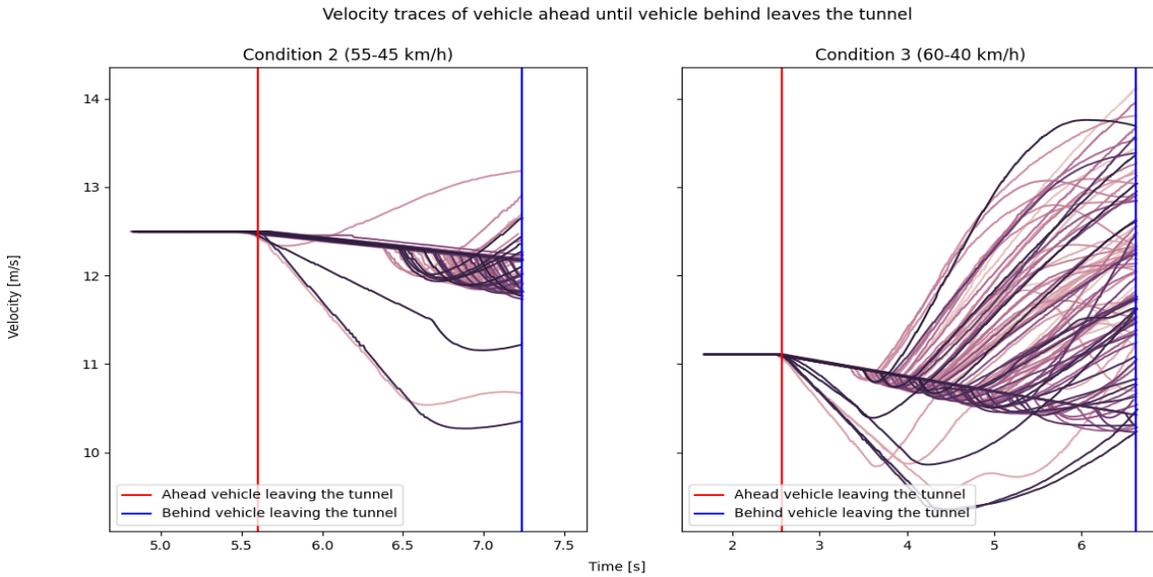


Fig. 26: Visualization of the velocity traces of the ahead vehicle in conditions 2 and 3.