



D7.3 Report on simulator test results and driver acceptance of PROSPECT functions

Niklas Strand, Bruno Augusto, Irene Gohl, Johann Stoll, Pablo Puente Guillen, Marie-Pierre Bruyas, Marie Jaussein, Evan Gallouin, Klaus Perlet, Mats Petersson, et al.

► To cite this version:

Niklas Strand, Bruno Augusto, Irene Gohl, Johann Stoll, Pablo Puente Guillen, et al.. D7.3 Report on simulator test results and driver acceptance of PROSPECT functions. [Research Report] IFSTTAR - Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux. 2018, 99 p. hal-02264013

HAL Id: hal-02264013

<https://hal.science/hal-02264013>

Submitted on 6 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



EUROPEAN COMMISSION
EIGHTH FRAMEWORK PROGRAMME
HORIZON 2020
GA No. 634149

Deliverable No.	D7.3		
Deliverable Title	Report on simulator test results and driver acceptance of PROSPECT functions		
Dissemination level	Public		Dd/mm/yyyy
Written by	Strand, Niklas	VTI	
	Augusto, Bruno	VTI	
	Gohl, Irene	AUDI AG	
	Stoll, Johann	AUDI AG	
	Puente Guillen, Pablo	TME	
	Bruyas, Marie-Pierre	IFSTTAR	
	Jaussein, Marie	IFSTTAR	
	Gallouin, Evan	IFSTTAR	
	Perlet, Klaus	BMW	
	Petersson, Mats	VCC	
	Johansson, Regina	VCC	
	Meltzer, Elin	VCC	
	Ljung Aust, Mikael	VCC	
	Braeutigam, Julia	BAST	
	Large, David	University of Nottingham	
Checked by	Last Name, Name	Organisation	Dd/mm/yyyy
	Bánfi Miklós Gábor	BME	
Approved by	Last Name, Name	Organisation	Dd/mm/yyyy
Issue date	Dd/mm/yyyy		



The research leading to the results of this work has received funding from the European Community's Eighth Framework Program (Horizon2020) under grant agreement n° 634149.

EXECUTIVE SUMMARY

The process of developing new automotive systems includes various testing cycles to assure a safe operation in traffic. Physical system testing on test tracks is very important for this purpose, but rather expensive and might only become possible in later stages of the development process. Using a virtual simulation environment offers a safe possibility of testing new systems in early stages of development. Additionally, driver-in-the-loop tests at test track and in a virtual simulator make it possible to evaluate driver reaction and potential acceptance by the future users of those systems. Within PROSPECT the new functions are investigated under various aspects in several simulator studies and test track studies.

This deliverable D7.3 gives detailed information of conduction and results of the each study. Three studies focus exclusively on the for Vulnerable Road Users (VRUs) specifically dangerous urban intersection scenarios. The first of those studies examines the driver behaviour in a turning situation when a bicyclist might be crossing. The described phenomena are looked-but-failed-to-see and failed-to-look. The second study, which provides an initial step in this line of research, analyzed the acceptance of issued forward collision warning times. The positioning of the potential accident opponent and the subjective feeling towards the criticality of the situation by the driver were key parameters taken into account. Last, but not least the acceptance of an intersection assist autonomous emergency braking systems was tested regarding the acceptance of potential buyers. The study was run for five days in a row for each participant to be able to judge the behaviour in a commuting situation.

Two studies focused on longitudinal scenarios. Both studies followed the same design, but one was conducted on a test track and the other one in a simulator. The main objective was to investigate drivers reactions to FCW warnings and Active Steering interventions in critical VRU scenarios in case of a distraction of the driver. Additionally, the test track study was used to validate the results from the simulator study.

The results of those studies are the basis for a wide acceptance evaluation of the systems. No system is an asset in increasing road safety if it is not accepted by the user and therefore turned off, if it is not required the system to be default on in consumer tests. Complemented by an additional acceptance study where the participants had to give their opinion of those systems after they watched videos of dangerous situations, the acceptance was analyzed based on questionnaires developed in PROSPECT and reported in Deliverable 7.2. This wholistic approach allows an expert discussion on the potentials of the PROSPECT functions in the future.

CONTENTS

Executive summary	2
1 Introduction	6
1.1 Background	6
1.1.1 Prospect	6
1.1.2 Deliverable outline	6
2 General methodology	7
2.1 Use Cases	7
2.2 Driving simulators	8
2.3 Acceptance methodology	8
3 Experiment #1 “Audi experiment t 7.2”	10
3.1 Theoretical background and focus of the study	10
3.2 method	13
3.2.1 Experimental design and driving tasks	13
3.2.2 Participants	15
3.2.3 Equipment and recorded data	15
3.2.4 Procedure	17
3.3 Results	18
3.3.1 Data analysis	18
3.3.2 Subjective results	18
3.3.3 Objective results	21
3.4 Conclusions	22
4 Experiment #2 “ Drivers’ Acceptance towards warnings based on driver models”	24
4.1 Methods	24
4.1.1 Methodology	24
4.1.2 Test setup	24
4.1.3 Research design and scenario	25
4.1.4 Procedure	26
4.1.5 Participants	27
4.2 Results	27
4.2.1 Brake onset after warning	27
4.2.2 Acceptance questionnaire	28
4.2.3 Correlation	31
4.3 Conclusions	31
5 Experiment #3 “University of Nottingham Acceptance testing”	33
5.1 Overview	33
5.1.1 Methodology	33
5.1.2 Demographics	35

5.2	Results.....	36
5.2.1	Defining Criticality of Situation	36
5.2.2	Assessment of System Acceptability	36
5.2.3	Assessment of System Trust.....	37
5.2.4	Intention to Activate/Use Warning Function	38
5.2.5	Intention to Activate/Use Braking Function.....	38
5.2.6	Willingness to Buy A PROSPECT-Like System	39
5.3	Analysis.....	40
5.4	Conclusions	40
6	EXPERIMENT #4 “DRIVER REACTION TO FCW WARNING AND AUTOMATIC STEERING INTERVENTION IN CRITICAL VRU SCENARIOS”	41
6.1	METHOD FOR SIMULATOR EXPERIMENT	41
6.1.1	Research design	41
6.1.2	Participants	41
6.1.3	Materials	42
6.1.4	Performance indicators.....	45
6.1.5	Procedure	45
6.1.6	Analysis	46
6.2	METHOD FOR THE TEST TRACK EXPERIMENT	46
6.2.1	Research Design.....	46
6.2.2	Participants	46
6.2.3	Equipment	47
6.2.4	Procedure	48
6.3	DATA ANALYSIS	49
6.4	DIFFERENCES IN METHOD BETWEEN SIMULATOR AND TEST TRACK EXPERIMENTS.....	50
6.4.1	Research Design.....	50
6.4.2	Methods	51
6.4.3	Procedure	51
6.5	RESULTS FROM PEDESTRIAN TESTS WITH FCW ONLY AT TEST TRACK AND IN SIMULATOR INCLUDING VALIDATION	51
6.5.1	Gaze Reaction Times from Test Track	52
6.5.2	Reaction Types	52
6.5.3	Steering Reaction Times from Test Track and Sim IV	52
6.5.4	Braking in Sim IV and Test Track.....	53
6.5.5	Distance to Conflict Object in Sim IV and on Test Track	55
6.5.6	Collisions in Sim IV and on Test Track	58
6.6	RESULTS FROM PEDESTRIAN TESTS WITH FCW W/VO AUTOMATIC STEERING INTERVENTION AT TRACK TESTING.....	58
6.6.1	Reaction Types	58
6.6.2	Steering Behaviour	59
6.6.3	Lateral Distance	60
6.7	RESULTS FROM CYCLIST TESTS WITH FCW W/VO AUTOMATIC STEERING INTERVENTION IN DRIVING SIMULATOR	63
6.7.1	Steering behaviour	63
6.7.2	Severity	67
6.7.3	Acceptance evaluation	67
6.8	CONCLUSIONS.....	71
7	Experiment #5 “Focus group and video-based experiment”	72

7.1	Global research design	72
7.2	Participants	72
7.3	Focus groups	73
7.3.1	Procedure	73
7.3.2	Materials	74
7.3.3	Results	76
7.4	Video-based experiment.....	80
7.4.1	Procedure	81
7.4.2	Materials	81
7.4.3	Results	83
7.5	Conclusions	91
8	Discussion and Comparison of results	94
8.1	Method related aspects	94
8.2	Overall Acceptability Evaluation.....	94
9	References	96
	Acknowledgments	98

1 INTRODUCTION

1.1 BACKGROUND

To test safety measures in real traffic may induce risks for drivers and therefore it is important to provide safe test environments with the human-in-the-loop, especially for early stage prototype systems. Such test environments can be set up in driving simulators or test tracks. A key to successful evaluation of active safety systems is to include the driver reaction and response to the system intervention to assure that the system itself does not impose any additional risks for the driver. Another important aspect for the evaluation is user acceptance. Without user acceptance the success of any active safety system is challenged. A low acceptance towards the system may lead to deactivating the system or not buying the system. Within PROSPECT, an acceptance methodology that can be used while evaluating active safety has been developed and used in the studies reported here.

1.1.1 Prospect

Recently active pedestrian safety systems with the capability of autonomous emergency braking (AEB-PED) in specific critical situations have been introduced to the market. Proactive Safety for Pedestrian and Cyclists (PROSPECT) is an EU funded project set out to improve the effectiveness of active safety systems targeting vulnerable road users (VRU). In PROSPECT there are two main targets to achieve this, namely: (1) by expanding the scope of the addressed scenarios, and (2) through improved overall system performance.

1.1.2 Deliverable outline

This deliverable addresses results from acceptance testing and testing of active safety measures on test tracks and in driving simulators. In the first chapter the background providing the rationale for conducting the tests are introduced followed by a brief introduction to the PROSPECT project and the objectives. In the second chapter the general methodology describing driving simulators, test tracks, and acceptance testing is described. Thereafter, each test and experiment is described in detail in their own chapters respectively. Following the chapters on each experiment, a chapter providing a discussion and a comparison of results is provided. In the last chapter of the deliverable the overall conclusions are provided.

2 GENERAL METHODOLOGY

2.1 USE CASES

Twelve use cases have been identified within WP2 and WP3 to be implemented in the demonstrators: 9 for cyclists and 3 for pedestrians (Figure 1). Even reduced, this number still addresses around 80% of all cyclist accidents investigated in WP3. Behaviours such as the velocity, distance and offset of the vehicle and VRUs (cyclists and pedestrians) are defined in Deliverable D7.2 (Report on methodology for balancing user acceptance, robustness and performance). Scenarios can then be realized on the test tracks or in simulator environments. For each use case, storyboards describe what the demonstrators intend to do based on the situation analysis and the risk assessment. To do so, each use case is divided into three distinct scenario groups: Safe Scenario (S), Critical Scenario (C) and Possible Critical Scenario (PC). Information is provided in Deliverable D7.2 (Report on methodology for balancing user acceptance, robustness and performance).

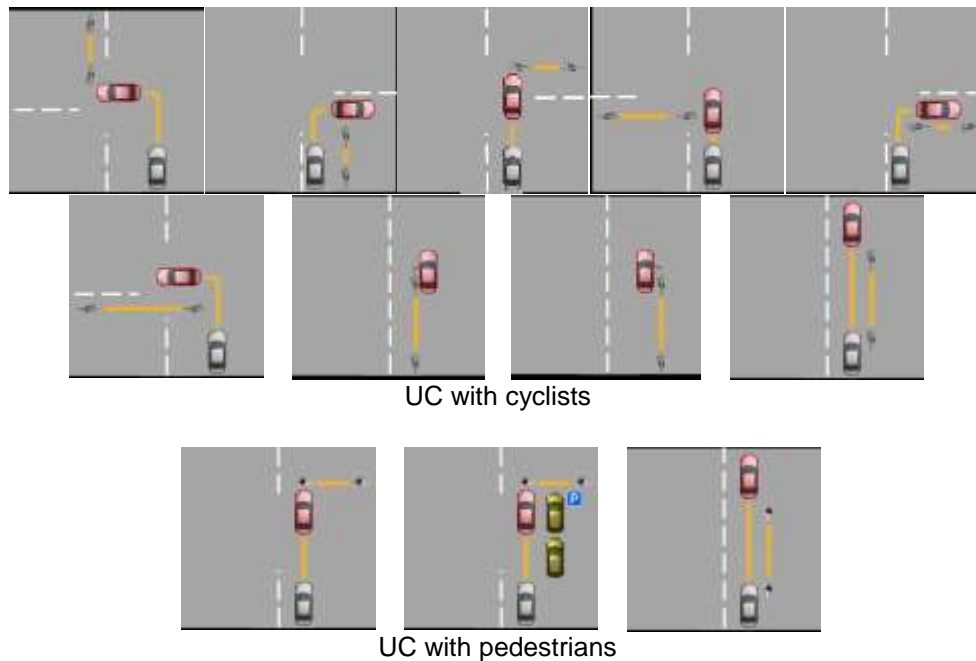












Figure 1: The 12 Demonstrator Use Cases defined within WP2 and WP3

A set of tests and experiments described in the current deliverable have been designed within T7.2 and T7.3 in order to evaluate the different PROSPECT systems. These tests were carried out in various contexts with the aim of covering most of the demonstrator use cases (Table 1) including pedestrians and cyclists and all functionalities of the PROSPECT systems.

Table 1: Experiments carried out within T7.2 and T7.3.

Partner	UC N°	VRU	Use case pictogram	Prospect function			Scenario		Activation		Added description
				Warning	Braking	Steering	Possible critical scenario	Critical scenario	True positive	False positive	
Audi /TME	5, 3	Cyclist		X			X	X	X	X	 VIL methodology & simulator expe.
UoN	2	Cyclist	 Reversed (driving on the left)	X	X		X	X	X	X	 Simulator experiment
VTI /Volvo	12	Cyclist/ Pedestrian		X		X		X	X		 Test track & Simulator expe.
IFSTTAR	2,4	Cyclist			X		X	X	X		 Lab. expe. Focus groups Video 
	10, 11	Pedestrian		X	X		X	X	X		

2.2 DRIVING SIMULATORS

In PROSPECT driving simulator studies were used to identify and tune the user acceptance of proactive safety systems for vulnerable road users, as well as to test active safety functions in a controlled and repeatable environment. Furthermore, driving simulators provided a testbed to provide replicable human-in-the-loop tests of active safety systems with active braking and steering capacity at varying stages of system development. The use and value of driving simulators as a tool for development of active safety functions was thoroughly described by for instance Fischer, et al. (2011). Simulated mock-ups of active safety systems—including the human machine interface—resembling PROSPECT features were therefore implemented in different driving simulators across partners throughout Europe.

2.3 ACCEPTANCE METHODOLOGY

Acceptance testing is an important part of the project since it provides knowledge on the user perception of the proactive systems developed within PROSPECT. It is crucial for the success of such active safety systems that they are acceptable for the drivers (e.g. judged useful and trusted). If not, they could be permanently turned off and would then have no effect on traffic safety. Moreover, interventions of active systems being rare, they may lead to unpredictable reactions from non-aware drivers being potentially frightened or startled when activated.

In this context, a specific acceptance methodology was developed within Task 7.3. This methodology is described in Deliverable D7.2 (Report on methodology for balancing user acceptance, robustness and performance). It integrates acceptance of false positive (warnings and/or interventions occurring at inappropriate times) and false negatives (no warning or activation when needed) and evaluates their influence on the drivers' acceptance.

Questionnaires are generally used to evaluate subjective components of constructs, such as acceptance and trust. The methodology is then based on existing questionnaires to be administered in tests and experiments that evaluate PROSPECT systems. By using common questionnaires, this work aims to enable an overall evaluation of the acceptance of the developed functions. Collecting them in such a way will ensure data is acquired in the same format and thus can easily be compared.

In the literature, acceptability is generally distinguished from acceptance. Acceptability is measured when the person has no experience of the system, as an "*a priori* measure of the extent to which a person thinks they will accept and use a system. Acceptance, on the other side, is determined after use and is a measure of how much a person actually uses a technology and the satisfaction with this experience".

Three questionnaires are completed at different times of the tests:

- **Before running the test/experiment:** questionnaire 1 (participant information) and 3 (global expected acceptance of the system or *a priori* acceptability, based on the "Perceived usefulness", and "Perceived satisfaction" questionnaire of Van der Laan et al. (1997) and the Trust questionnaire of Jian et al. (2000)).
- **During the test/experiment:** questionnaire 2 (feedback on each situation the participants are being faced with, based on the Risk Awareness measurement developed by Bellet and Banet, (2012)).
- **After the test/experiment:** questionnaire 3 (global acceptance of the system after having experienced it)

Before distributing the first questionnaire, the experimenter must give sufficient information about the functioning of the system to the participants. This information will allow the driver to gain some familiarity with the system and to create a mental model of what will be tested. The second acceptance questionnaire allows for measuring changes in acceptance as a result of the system experience.

The questionnaires have been developed with the objective to make acceptance evaluation not too invasive during the tests, and to disturb the test participants as little as possible. For that reason, the questionnaire completed during the experiment is very short. The two other questionnaires used before and after experiments could be longer as they do not interfere with the experiment.

3 EXPERIMENT #1 “AUDI EXPERIMENT T 7.2”

3.1 THEORETICAL BACKGROUND AND FOCUS OF THE STUDY

The output from task T2 defined the most relevant car-to-cyclist accident scenarios that should be addressed by next-generation VRU systems. Based on the results, the most relevant car-to-cyclist scenario is a driver approaching a non-signalized intersection or a T-junction and a cyclist crossing from a bicycle lane. In such situations the frequency to collide with a cyclist crossing from the right side is three times higher than with a cyclist crossing from the left side (Gohl, Schneider, Stoll, Wisch, & Nitsch, 2016). Furthermore, the frequency of collisions with a cyclist crossing from the right side as well as from the left side decreases when the driver’s maneuver intentions are taken into account. Figure 2 shows that left-turning drivers collided less frequently than right-turning drivers, irrespective of the orientation of the cyclist.

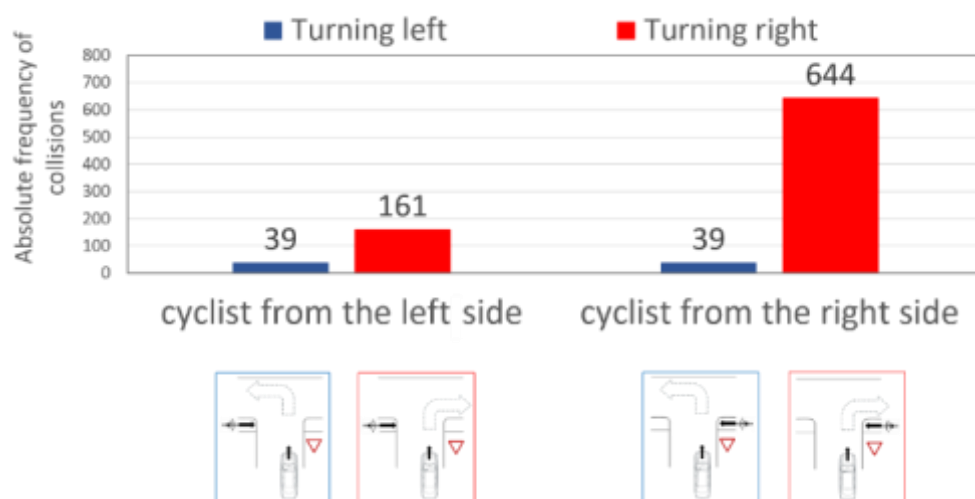


Figure 2: Absolute frequency of collisions depending on driver’s task and the crossing direction of the cyclist.

A general explanation why drivers turning right fail to manage situations with a cyclist from the right more often than drivers turning left is the drivers’ improper attention allocation strategies, as demonstrated by Summala et al. (1996). In unobtrusive field observations at T-junctions, Summala et al. (1996) studied drivers’ visual scanning strategy at left and right turns. The results showed that drivers’ visual scanning patterns differentiate according to their task goals: drivers turning left tended to look in both directions, whereas drivers turning right rather continued to look left. According to the authors, these different scanning patterns implied an attentional bias towards conflicting motor vehicles due to drivers’ erroneous expectations. These expectations are formed by drivers’ practice, integrating the perceived environmental cues in a certain context and task into hierarchically organized schemata including both what potential hazards may occur and where they may appear (Engström, 2011). This knowledge-driven processes, called “*top-down* processes”, may lead to critical situations if the drivers’ expectation does not match the actual situation. As a result, drivers turning right selectively looking for cars approaching from the left will

probably fail to notice the cyclist coming from the right. In addition, the probability to notice the cyclist coming from the right is decreased for right-turning drivers when the traffic density from the left is high, as demonstrated by Werneke and Vollrath (2012). In contrast, drivers turning left have to yield for cars from both sides. Even if the left-turning driver did not account for cyclists from the right, the chance of the cyclist being able to attract attentional focus through reflexive bottom-up processes is higher than for right-turning drivers since the cyclist appeared within the drivers' field of view (Summala & Räsänen, 2000).

Moreover, the latter provides an explanation why drivers turning right fail to manage situations with a cyclist from the right more often than with a cyclist from the left. As pointed out above, drivers turning right may fail to look in the direction of the cyclist crossing from the right, thus the cyclist appears entirely outside the drivers' field of view. If the cyclist appears within the field of view (e.g. cyclist crossing from the left), bottom-up selection may prevent drivers from colliding with that cyclist and thus reduce the overall frequency of collisions.

However, drivers turning right still collide quite frequently with cyclists from the left (see **¡Error! No se encuentra el origen de la referencia.**), though the chance of the cyclist being able to capture the driver's attention is increased due to the bottom-up selection. One contributing factor to the reduced ability of cyclists to capture the drivers' bottom-up processes within the drivers field of view is their poor sensory conspicuity, i.e. the degree of difficulty to perceive an object from its environment due to its physical characteristics such as size, brightness, illumination, color and movement (see e.g., Rogé, Ndiaye, Aillerie, Aillerie, Navarro, & Vienne, 2017; Tin Tin, Woodward, & Ameratunga, 2015). The results gained from task T2.1 support the hypothesis that poor sensory conspicuity may play a critical role in crash causation in such situations, as over a quarter of these accidents happened during nighttime or dawn (Gohl, Schneider, Stoll, Wisch, & Nitsch, 2016).

However, since most of these accidents happened during daytime, it may be assumed that drivers looked but failed to see the salient cyclist coming from the left appearing within their field of view. In the field of perceptual psychology this phenomenon is known as *inattention blindness* (e.g., Mack and Rock, 1998; Simons and Chabris, 1999) and "refers to the inability to detect salient stimuli appearing in the field of view if attention is allocated elsewhere" (Engström, 2011, p.14). For instance, Simons and Chabris (1999) demonstrated in an experiment that this phenomenon is truly existing. While participants had to count the number of passes between basketball teams in a video clip, a person dressed as a black gorilla walked through the mass of players. After the study, about 30-70% of the participants (White and Caird, 2010) reported that they did not notice this very salient object in their direct field of view. In recent years, many studies including in-depth accident analyses, self-reported near accidents and experimental studies conducted in simulators investigated the *looked-but-failed-to-see* error in the driving context (e.g., Brown, 2005; Herslund & Jorgensen, 2003; Koustani et al., 2008; Clarke et al., 2004; Mitsopoulos-Rubens and Lenné, 2012; Rogé, 2011; Rogé, 2017). The general consensus of these studies is that the looked-but-failed-to-see phenomenon may

arise from the observers erroneous expectations of what they are likely to see, and as a consequence they might unintentionally filter out the unexpected or infrequent objects without perceiving them. As a result, the looked-but-failed-to-see error could be less pronounced for drivers who cycle regularly themselves, compared to drivers who never cycle, since the former are more aware of the presence of cyclists. In fact, as demonstrated recently by Roge et al. (2017), cyclists' visibility depends on their cognitive conspicuity for car drivers rather than their sensory conspicuity. The authors concluded that attentional selection of a cyclist in the road environment during car driving depends mainly on top-down processing. But, as mentioned above, the results from T2.1 indicate that sensory conspicuity may be a contributing factor in such situations, in which a driver intends to turn right at a T-junction and a cyclist is crossing from the left. Top-down processes in terms of erroneous expectations of what kind of potential hazards may occur solely cannot explain the higher proportion of accidents that happened during nighttime or dawn.

Indeed, in natural driving situations, attention selection is typically the result of a dynamic interaction between both top-down and bottom-up selection. However, very little is known about this interaction and especially how and in which situations drivers bottom-up processes prevent them from colliding with cyclists when turning right at T-junctions even if drivers do not account for cyclists at all. Based on the literature above, it may be assumed that erroneous expectations will decrease drivers' ability to detect cyclists from both sides. If the visibility of cyclists depends on their cognitive conspicuity for car drivers, the car drivers who are aware of cyclists would avoid more collisions and would detect cyclists earlier than car drivers who do not account for cyclists. Since it was shown that drivers erroneous expectations lead to an attentional bias towards car objects (Summala, Pasanen, Räsänen, & Sievänen, 1996) resulting in a visual scanning behaviour biased to the left leg of the intersection, the probability to detect the cyclists from the left is higher than cyclists from the right, irrespective of the sensory conspicuity.

Moreover, if drivers' visual behaviour when turning right is biased towards car objects from the left, then cyclists who are coming from the right side will not be able to attract the drivers attention even if the cyclists sensory conspicuity is high. Therefore, drivers would show an equal detection performance of cyclists from the right side, regardless of the cyclists sensory conspicuity. In contrast, the results from T2.1 indicate that the ability to interrupt the drivers biased visual behaviour towards cyclists crossing from the left side depends on their sensory conspicuity. Therefore, it is hypothesized that drivers would show a better detection performance of cyclists from the left side with high sensory conspicuity compared to those with lower sensory conspicuity.

In addition, as demonstrated by Werneke et al. (2012), the extent of attentional bias towards the left side depends on traffic density from the left side. As a result, it may be assumed that in situations where no other objects (e.g. cars) from the left side appear, drivers would show a better detection performance of cyclists.

3.2 METHOD

3.2.1 Experimental design and driving tasks

In order to examine how and in which situations the drivers bottom-up processes prevent them from colliding with cyclists when turning right at T-junctions - especially if drivers do not account for cyclists - three situational factors were varied.

At first, two T-junction scenarios were randomized with either crossing traffic or without crossing traffic. Both T-junctions presented a yield sign indicating that the drivers had to give way. In the scenario with crossing traffic, one black and one red car were placed on the left side of the main road of the T-junction at a distance of 72 m and 30 m, respectively. When participants approached the T-junction, the red car started to cross the intersection with a mean velocity of 50 km/h from the left to the right once the relative temporal distance fell below 10 seconds. In contrast, the black car remained in its initial position. Both cars indicated to the participants that there would be some traffic at this T-junction. At the second T-junction there was no crossing or parked traffic at all.

Moreover, the crossing direction of the cyclist was varied between crossing from the left and crossing from the right side. Every cyclist was positioned at a distance of 53 m from the middle of the T-junction. When participants approached the T-junction, the cyclist started to cross the intersection with a mean velocity of 20 km/h once the participant reached a distance of 75 meters from the middle of the T-junction. To obtain comparable results, both T-junctions had the same geometry, so that the cyclist becomes visible at the same time, regardless of his crossing direction.

Since bottom-up processes are directly affected by the sensory conspicuity of an object, two different levels of cyclist's sensory conspicuity were selected as independent variable: low and high sensory conspicuity. In the high sensory conspicuity scenario, the cyclist was dressed in white and appeared against a colored background (see Figure 3 A). In contrast, in the low sensory conspicuity scenario the cyclist was also dressed in white but appeared against a similarly colored background (see Figure 3 B).

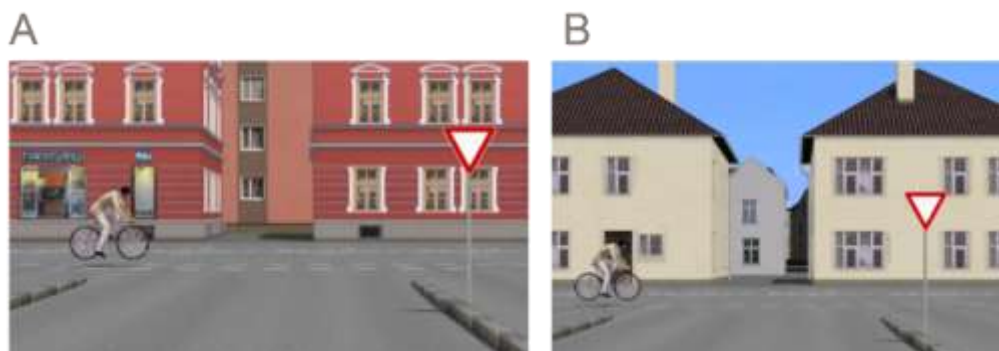


Figure 3: Levels of cyclist's sensory conspicuity varied within the study: high (A) and low (B) sensory conspicuity.

Since it is known that erroneous expectations will influence drivers' gaze and driving behaviour which decreases the drivers ability to detect cyclists, several rounds without crossing cyclists were needed to build up erroneous expectations of what potential hazards may occur, i.e. decrease the participants expectation of a cyclist crossing. In order to compare drivers' gaze behaviour with and without additional crossing traffic in non-critical situations (baseline-condition), each participant was confronted with both T-junctions, the one with the additional crossing traffic and the one without additional crossing traffic (within-subject design). Once participants completed two trials without crossing cyclists, they were confronted with a crossing cyclist at the next T-junction in a third trial. It may be assumed that this unexpected encounter with a cyclist will sensitize the participants to the presence of cyclists, resulting in a more alert gaze and driving behaviour. To ensure that participants are unaware of cyclists in all encounters, each participant can be tested only once in a critical incident. As a result, for all three situational factors varied within this experiment a between-subjects design in encounter situations was used (see Figure 4).



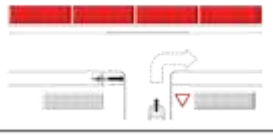





		ADDITIONAL CROSSING TRAFFIC FROM THE LEFT SIDE	CROSSING DIRECTION OF THE CYCLIST	
			from the left side	from the right side
CYCLIST'S SENSORY CONSPICUITY	high	with		
		without		
	low	with		
		without		

Figure 4: The resulting experimental design in encounter situations.

Although it is known that erroneous expectations will decrease the driver's ability to detect cyclists, the inverse conclusion is uncertain. Therefore, the hypothesis of whether or not expectations which take cyclists into account will increase the drivers ability to detect cyclists (regardless of the cyclists sensory conspicuity, crossing direction and the presence/absence of additional crossing traffic) will be tested. For this purpose, each participant was faced a second time with a crossing cyclist at the second T-junction.

Furthermore, the drivers' gaze behaviour was recorded with a head-mounted eye-tracking system (Eye Tracking Glasses 2 from SMI) in order to assess the drivers' detection performance. As a result of using a head-mounted eye-tracking system, participants are fully aware that their eye movements will be recorded. This knowledge may unconsciously bias the participants' gaze behaviour. To dissuade participants from thinking that they have to behave in an exemplary manner (as obtaining their driver's license), a cover story was constructed. The participants were instructed that the aim of the study was to examine which gaze parameters are suitable to predict the state of "mind wandering". In order to ensure that all participants understand the term in the same way, the following definition was introduced: "Mind wandering is a state where the thought processes that occupy the mind are on topics that are unrelated to the task(s) at hand" (Yanko & Spalek, 2013, p. 81) [and are often experienced by drivers that after these periods] "they can hardly remember any of specifics associated with the drive" (Yanko & Spalek, 2013, p. 81). Consequently, the main driving task of the participants was to drive a defined route several times and whenever they experienced mind wandering, they had to flash the vehicle headlight.

3.2.2 Participants

The study consisted of a sample of 92 participants, of whom 31 were female and 61 male. The participants' ages ranged from 19 to 56 years, with a mean age of $M=29.5$ years ($SD=8.5$ years). On average, participants obtained their driving license 12.2 years ago ($SD=10.8$ years). All participants requiring visual aid had to wear corrective contact lenses during the study in order to avoid interferences with the eyetracking system from SMI (Eye Tracking Glasses 2). The assignment of the participants to one out of eight groups was performed on the basis of their age, gender and average mileage per year.

3.2.3 Equipment and recorded data

The test was performed in the driving simulator described in section 2.1. The simulated environment was created using the tool-chain VTD for driving simulation applications. The road environment included houses, stores, signs, sidewalks, cycle paths and oncoming car traffic besides the two T-junctions of interest. Two T-junctions and five intersections are embedded within this environment and are located as shown in Figure 5

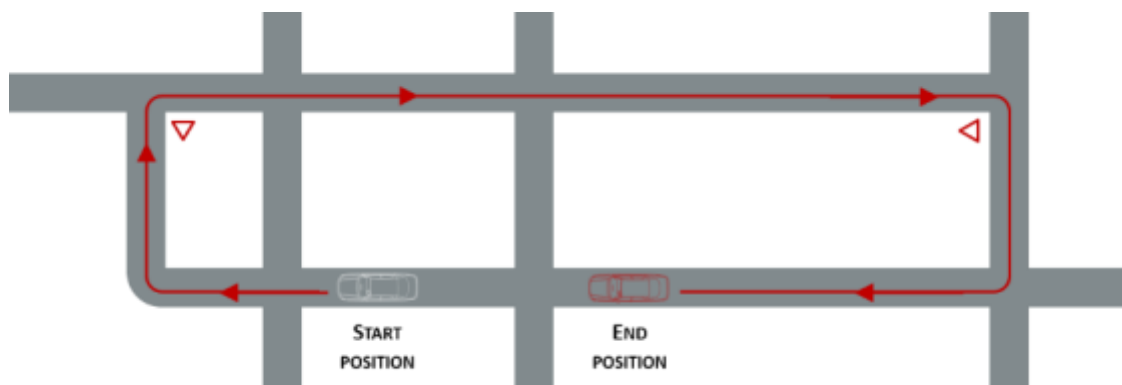


Figure 5: Location of the T-junctions within the simulated environment.

For each of the eight experimental design groups (see **¡Error! No se encuentra el origen de la referencia.**), three driving scenarios were created (24 driving scenarios in total). The first scenario was used to become familiar with the simulator and the route. During this scenario the investigator was seated in the rear seat and pointed the way. The second scenario was identical to the first and was used to build up erroneous expectations of potential hazards. In contrast to the first, the investigator was sitting outside the car this time. The third scenario included the crossing cyclist at both T-junctions. Each of the three driving scenarios took 8-10 minutes to complete.

In this study, driving data, gaze behaviour and subjective data were recorded. Driving data, speed, longitudinal acceleration, actuation of the turn indicator, remaining distance to both T-junctions and a binary variable, which displays whether or not the driver collided with the cyclist in the last driving scenario, were recorded by the ADTF software. Besides the driving data, eye-tracking information was collected with Eye Tracking Glasses 2 from SMI. The gaze information was recorded at a 60 Hz rate. Both datasets, vehicle and eye-tracking data, were synchronized based on a time marker (participant activating the turn indicator while directing the gaze in the same direction). Regarding subjective data, participants received three questionnaires within this study. The first was used to collect general information about the participants such as age, driving experience (mean mileage per year, years of possession of driver's license) and driving habits (frequency of car use, percentage distribution of driving time per road type including highway, rural and urban road). In the second questionnaire participants evaluated their perception of the driving situation after the first encounter with a crossing cyclist using the dimensions of the second questionnaire proposed in Deliverable D7.2. But instead of the continuous scales, participants had to indicate their level of agreement on 5-point Likert scales (see Figure 6).

With the following scales, how would you define the situation?						
Not critical at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Highly critical
Not frequent at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Highly frequent
Not foreseeable at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Highly foreseeable
Not controllable at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Highly controllable
Not frightening at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Highly frightening
Not stressing at all	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Highly stressing

Figure 6: Evaluation of participants' perception of the driving situation after the first encounter with a crossing cyclist.

Moreover, some questions were added concerning whether or not the driver has seen the cyclist in time and whether or not the driver has foreseen that a cyclist would cross at this T-junction. The third questionnaire was administered to measure changes in the drivers perception of the driving situation as a result of the increased awareness of cyclists. In addition, the participants were queried on a 5-point Likert scale whether they would like a driver assistance system in these situations. Finally, participants received some questions concerning their cycling habits (frequency of cycle use, usage of the cycle path against driving direction).

3.2.4 Procedure

After reading the brief instructions, the participants filled out the first questionnaire concerning their demographic characteristics and driving experience. Based on their age and mileages per year, participants were assigned to one of the eight experimental groups. Once the participants adjusted the seat and steering wheel of the car, the eye-tracking system was put on and the calibration was carried out. Subsequently, the participants were driving the first two driving scenarios without a crossing cyclist. There was a short break after both driving scenarios. After the second break, they drove the scenario a third time, but this time there was a crossing cyclist at both T-junctions. Between the two encounter situations the participants had to evaluate the situation with the second questionnaire. Shortly after the second encounter, they stopped a last time and filled out the third questionnaire. The procedure is summarized in Figure 7. The whole experiment for each participant lasted about 45 minutes.

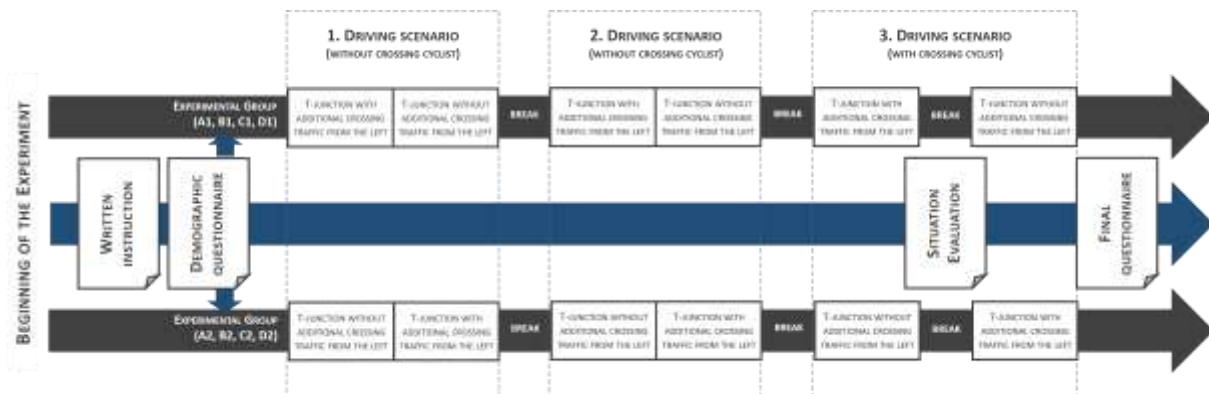


Figure 7: Experimental procedure in the driving simulator study.

3.3 RESULTS

3.3.1 Data analysis

The results presented here focus on the participants subjective perception of the driving situation as well as the participants objective detection performance in the first critical encounter with a crossing cyclist.

For data analysis the statistical software SPSS 24 was used. In order to describe the influence of the cyclists sensory conspicuity, the cyclists crossing direction and the presence or absence of additional crossing traffic on the evaluation of driver's perception of the driving situation with a crossing cyclist, a 2 (cyclist's sensory conspicuity) x 2 (cyclist's crossing direction) x 2 (additional crossing traffic) univariate ANOVA was conducted for each scale.

With regard to the analysis of driver's objective detection performance, a measurable parameter for the drivers detection performance needs to be defined. Therefore, in a first step for each driver, the moment at which s/he moves her/his gaze towards the cyclist was identified. This moment can be expressed in meters and represents the relative distance that is left before the bicycle path is reached. In order to account for different approaching velocities, the subject's relative distance was then divided by his/her velocity at this moment and represents the time left to react to the cyclist. Finally, a 2x2x2 univariate ANOVA was performed.

3.3.2 Subjective results

With regard to the criticality scale, the 2x2x2 ANOVA revealed a significant main effect only on the cyclists crossing direction ($F(1,87)=6.284$, $p=.014$, $\eta^2=.067$). The effect of cyclist's sensory conspicuity, additional crossing traffic and the interaction effects were not significant. As Figure 8 shows, the participants experienced the driving situation with a crossing cyclist from the right ($M=1.89$, $SE=.179$) as more critical than the driving situation with a crossing cyclist from the left ($M=2.52$, $SE=.176$).

With regard to driver's perception of the frequency of the experienced driving situation in real world, the ANOVA showed no significant main or interaction effects. Nevertheless, there is a marginal significant main effect of cyclist's crossing direction ($F(1,87)=3.481$, $p=.065$, $\eta^2=.038$). Participants reported that they experienced the driving situation with a crossing cyclist from the left ($M=2.69$, $SE=.144$) more

frequently than the driving situation with a crossing cyclist from the right ($M=2.31$, $SE=.142$) (see Figure 8).

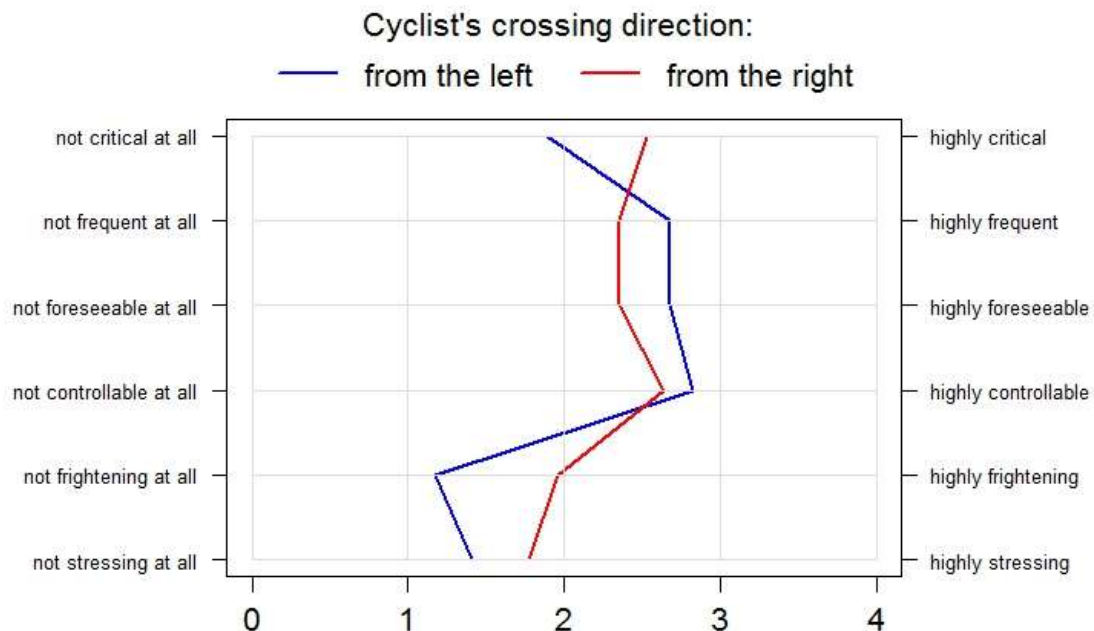


Figure 8: Participants' evaluation of the driving situation depending on cyclist's crossing direction.

With regard to the foreseeability of a driving situation, the ANOVA showed a significant main effect of the cyclists sensory conspicuity ($F(1,87)=11.036$, $p=.001$, $\eta^2=.113$). Interestingly, as shown in **¡Error! No se encuentra el origen de la referencia.**, participants reported that the driving situation with a crossing cyclist of low conspicuity ($M=2.77$, $SE=.127$) is more foreseeable than the driving situation with a crossing cyclist of high conspicuity ($M=2.16$, $SE=.132$). The remaining effects were not significant.

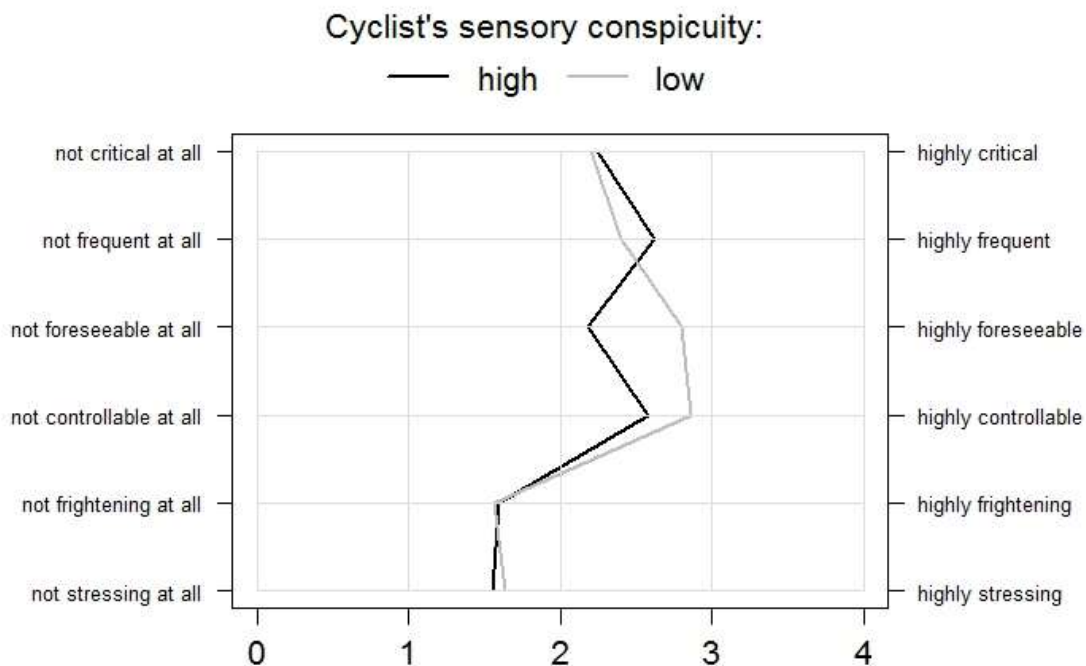


Figure 9: Participants' evaluation of the driving situation depending on cyclist's sensory conspicuity.

Regarding the controllability of a driving situation, the ANOVA revealed no significant main or interaction effect. However, there is a marginal interaction effect between the cyclists crossing direction and additional traffic ($F(1,87)=2.911$, $p=.092$, $\eta^2=.032$). As shown in Table 2, participants rated the driving situation with a crossing cyclist from the right more controllable when there was no additional traffic from the left side as compared to that with additional traffic. For driving situations with a crossing cyclist from the left, the opposite effect was found. Participants perceived this situation more controllable when there was additional traffic from the left side as compared to the driving situation without additional traffic.

Table 2: Perceived controllability of the driving situation depending on cyclist's crossing direction and additional traffic.

Cyclist's crossing direction			
		from the left	from the right
Additional traffic from the left side	yes	2.913s (SE=.216)	2.417s (SE=.211)
	no	2.689s (SE=.212)	2.917s (SE=.211)

With regard to the participants evaluation of how frightening the driving situation was, the ANOVA showed a significant main effect of the cyclists crossing direction ($F(1,87)=11.269$, $p=.001$, $\eta^2=.115$). With a mean value of 1.19 (SE=.172), participants reported a lower fright when the cyclist is crossing from the left side

compared to when the cyclist is crossing from the right side ($M=2.000$, $SE=.169$) (see Figure 8). The remaining effects were not significant.

Finally, regarding the stress scale, no significant main or interaction effects were found.

3.3.3 Objective results

With regard to the drivers detection performance, the 2x2x2 ANOVA showed two significant main effects (cyclist's crossing direction: $F(1,84)=11.023$, $p=.001$, $\eta^2=.116$; cyclist's sensory conspicuity: $F(1,84)=4.023$, $p=.048$, $\eta^2=.046$), but no significant influence of additional crossing traffic ($F(1,84)=2.060$, $p=.155$). Although there were no significant interaction effects found, the interaction between cyclist's crossing direction and cyclist's conspicuity was marginally significant ($F(1,84)=3.481$, $p=.066$, $\eta^2=.040$).

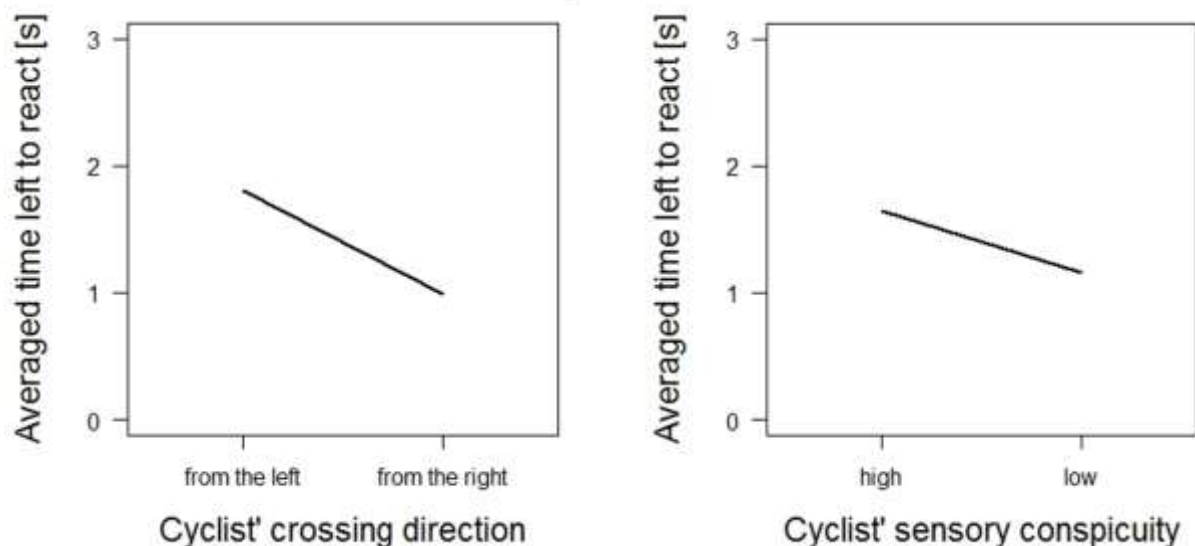


Figure 10: Drivers' averaged time left to react differed statistically significant depending on cyclist's crossing direction (left) and cyclist's sensory conspicuity (right).

¡Error! No se encuentra el origen de la referencia. (left) shows that drivers had significantly more time left to react to the cyclist coming from the left side ($M=1.810s$, $SE=.176s$) compared to the cyclist coming from the right side ($M=.993s$, $SE=.172s$). With regard to the main effect of the cyclists sensory conspicuity (see **¡Error! No se encuentra el origen de la referencia.**, right), drivers had significantly more time left to react on the cyclist with high sensory conspicuity ($M=1.648s$, $SE=.178s$) as compared to the cyclist with low sensory conspicuity ($M=1.154s$, $SE=.170s$). However, as indicated by the interaction between cyclist's crossing direction and cyclist's sensory conspicuity (see Figure 11), the effect of the cyclists crossing direction on driver's detection performance depends on the state of the cyclists sensory conspicuity. Whereas the average drivers' detection performance is almost the same for high and low cyclist's sensory conspicuity when the cyclist is crossing from the right side (low sensory conspicuity: $M=.975s$, $SE=.240s$; high sensory conspicuity: $M=1.010s$, $SE=.246s$), it is significantly different when the cyclist is

crossing from the left side (low sensory conspicuity: $M=1.333s$, $SE=.241s$; high sensory conspicuity: $M=2.287s$, $SE=.257s$).

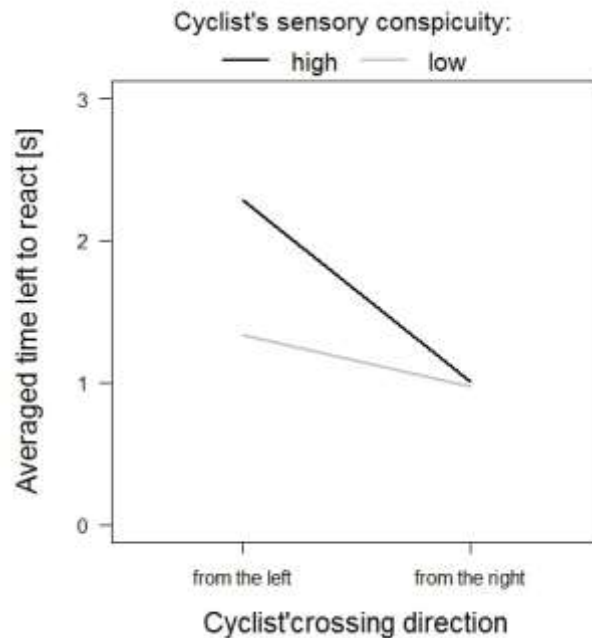


Figure 11: Interaction effect between cyclist's crossing direction and cyclist's sensory conspicuity.

3.4 CONCLUSIONS

The most relevant car-to-cyclist accident scenario that should be addressed by next-generation VRU systems is a driver approaching a non-signalized intersection or a T-junction and a cyclist crossing from a bicycle lane. Here it was found that the frequency to collide with a cyclist crossing from the right side is higher than with a cyclist crossing from the left side.

The literature review revealed that top-down processes (in terms of erroneous expectations of what potential hazards may occur) may explain both why drivers fail to manage situations with a cyclist from the right more often than with a cyclist from the left (failed-to-look) and why drivers still collide quite frequently with cyclists from the left (looked-but-failed-to-see). However, top-down processes solely cannot explain why the portion of accidents that happened during night time or dawn is higher for cyclist from the left than for cyclists from the right. An explanation could be that, in general, the drivers bottom-up processes may prevent drivers from colliding with a cyclist within the field of view (such as a crossing cyclist crossing from the left), but at night time or dawn the ability of cyclists to capture the drivers bottom-up processes may be reduced - especially if their sensory conspicuity is poor. As a result, the main goal of the study was to examine how and in which situations drivers bottom-up processes may prevent them from colliding with crossing cyclists when turning right at T-junctions.

Therefore, in a driving simulator study, participants had to turn repeatedly (in several rounds) right at two T-junctions. While in the first two rounds (out of three) the

participants experienced the turning situations without crossing cyclists, they were confronted with a crossing cyclist in the last round. Within the study three situational factors were varied: (1) the absence or presence of an additional crossing vehicle from the left side, (2) the cyclists crossing direction and (3) the cyclists sensory conspicuity.

The analyses of the objective detection performance showed two significant main effects: First, on average drivers had more time left to react to the cyclist coming from the left side compared to the cyclist from the right side resulting rather from an earlier detection than a lower approaching speed. This result confirms both (a) two baseline rounds are sufficient to build up erroneous expectations and (b) erroneous expectations bias the driver's attention allocation towards the left side and thus reduce the drivers' detection performance for a cyclist coming from the right. In general, this result confirms previous findings (Summala, Pasanen, Räsänen, & Sievänen, 1996). Secondly, the drivers had more time left to react on the cyclist with high sensory conspicuity as compared with a cyclist with low sensory conspicuity.

Although the interaction effect between the cyclist's sensory conspicuity and the cyclists crossing direction was only marginally significant, this result indicates that a higher cyclist's sensory conspicuity does not per se guarantee an earlier detection by drivers. Here it was found that when the cyclist is coming from the right side, then the cyclist's sensory conspicuity had only little influence on the driver's detection performance. In contrast, in driving situations with a cyclist coming from the left side drivers detected the cyclist with high sensory conspicuity much earlier than the cyclist with a low sensory conspicuity. As a result, future studies which examine the effectiveness of several measures to increase the cyclist's sensory conspicuity (e.g., fluorescent or retro-reflective clothing) have to take different driving situations into account.

Moreover, the enhanced detection performance of salient cyclists crossing from the left indicate that the drivers bottom-up processes may prevent the driver from overlooking the cyclist - especially if drivers do not account for cyclists at all. As a result, bottom-up processes can be considered as a natural defense system and are at least as important as top-down processes for traffic safety. However, future assistance systems should contribute to increase the cognitive conspicuity for cyclists as the results have shown that cyclists coming from the left side with low salience have not been detected considerably earlier as cyclists coming from the right side.

The initial hypothesis, that the additional crossing traffic could have an influence on the detection performance of drivers for crossing cyclists, could not be confirmed. One reason for this could be that drivers were expecting crossing traffic at one of the two junctions. Consequently, it could be assumed that drivers expected crossing traffic at each T-junction independently if they really experienced crossing traffic at this particular T-junction in the previous rounds or not.

4 EXPERIMENT #2 “DRIVERS’ ACCEPTANCE TOWARDS WARNINGS BASED ON DRIVER MODELS”

4.1 METHODS

4.1.1 Methodology

The aim of the TME study was to determine the acceptance of drivers to a Forward Collision Warning (FCW) system based on a comfort boundaries model. The comfort boundaries model was obtained from a previous study (DIV project) between TME, Autoliv and Chalmers [Boda et al. (2018)]. One of the aims of the DIV project was to analyse the brake onset of drivers (defined during the study as comfort boundary) when encountering a cyclist crossing the road in an intersection, as presented in Figure 14. Participants were instructed to drive and behave in the same way they would during normal traffic. The study was conducted in a driving simulator and in a test track using the same scenario. Boda et al. (2018) concluded that the moment in time when the cyclist first becomes visible to the driver (time to arrival visibility, $TTA_{visibility}$), had the biggest effect on the brake onset of the driver. The model that describes the brake onset dependant on the visible time of the cyclist. For the present study, the warning inside the comfort boundary was selected as the asymptotic value and the warning outside the comfort boundary was selected as the value below the lower 95% value.

The aim of the study conducted for PROSPECT was to determine the validity of DIV’s comfort model as a model to develop more acceptable warning times. The study was conducted using the AUDI Vehicle In the Loop (VIL) system.

4.1.2 Test setup

The VIL is a virtual reality simulator, which is coupled with a real car, i.e. the participant receives vestibular, kinesthetic and auditory feedback from driving a real car while he/she perceives the simulated environment from the traffic simulation (see Figure 12).



Figure 12: The Vehicle in the loop system used for TME study

The simulated environment was created using the tool-chain VTD (Virtual Test Drive) for driving simulation applications. The road environment included parking lots, houses, stores, signs, sidewalks, cycle paths and oncoming car traffic besides the intersection of interest (see Figure 13).



Figure 13. Simulated road environment realized for the TME study

Regarding the design of the warning, a visual-acoustic warning strategy was chosen. When participants approached the intersection of interest, both an acoustic signal and a visual icon within the instrument cluster are displayed once the participant reached a certain distance. To obtain comparable results, all participants were instructed to control the speed by activating the function Adaptive Cruise Control (ACC). The manipulation of the triggered signals was realized by the ADTF software. Simultaneously, the ADTF-software was used to record the driving data (i.e. actual vehicle speed, longitudinal acceleration etc.).

4.1.3 Research design and scenario

For the PROSPECT study, two warnings were tested: one inside the comfort boundaries and one outside the comfort boundaries of the drivers. The timings of the warning were calculated using the above-mentioned comfort boundary model obtained during DIV project. The scenario tested was the same scenario as in the DIV project: a cyclist crossing from the right (Figure 14). For the present study, the cyclist becomes visible at time to collision (TTC) of 4 seconds, i.e. 4 seconds before the car reaches the intersecting point between the vehicle and cyclist paths.

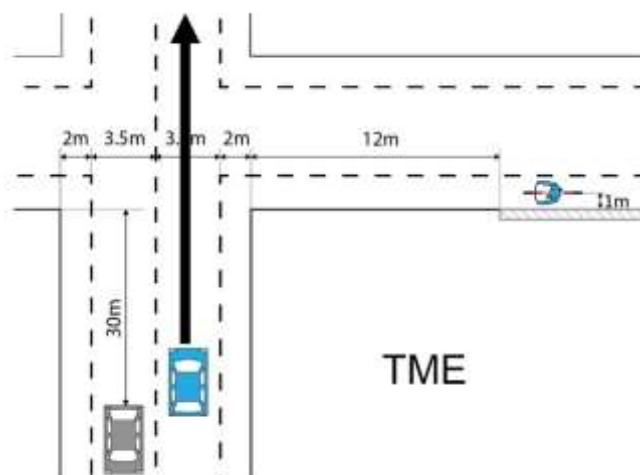


Figure 14: TME study scenario

In order to ensure that the cyclist does not become visible before a TTC of 4s, a blue wall (sight obstruction) on the right leg of the intersection was simulated (see Figure 15). Regarding the resulting warning times, a warning time inside the comfort boundaries of 2.6 s (asymptotic value of the model at $TTC=4$ seconds) and a warning time outside the comfort boundaries of 1.7 s (below the lower 95 % value of the model at $TTC=4$ seconds) was calculated. The hypothesis tested in the present study was that warnings outside the drivers comfort boundary would be better accepted as the situation would be critical in the perspective of the driver compared to the situation where the warning is inside their comfort boundary.



Figure 15: Simulated intersection according to the TME study scenario

4.1.4 Procedure

Participants were asked to drive three times around the test track to become familiar with the simulator and the route (Figure 16). After completing the familiarization phase, participants drove 9 times around the route and encountered 9 different conditions in the AUDI junction. Afterward participants had a break. Following the break, the same participants drove one more time around the route and encountered the cyclist crossing in the TME intersection (TME study). After this, the participant finalized the study and had to complete a survey regarding acceptance and willingness to buy the system experienced during the study.

For the TME study, the first objective was to validate that the model represents comfort boundaries of drivers, i.e. determine if the brake onset of the drivers is close to the warning. During the TME study, participants were asked to brake only after the warning was issued. The instruction given to the participant was the following: “During the second intersection, a warning will be signaled. After the warning is signaled, you are allowed to brake at anytime you find convenient (you decide when to brake depending on when you think the situation becomes critical)”. If a bigger gap was found between the warning time and the brake onset during the warning inside the comfort boundary compared to the gap during the outside boundary, it would indicate that drivers will not feel critical a situation when it is inside their comfort boundaries. The next step would be to determine if the activation time of the warning

would have an effect in acceptance (using the answers given during the survey completed at the end of the study). The hypothesis was that a warning outside the comfort boundary of the driver will result in better acceptance as it seen as a critical scenario.

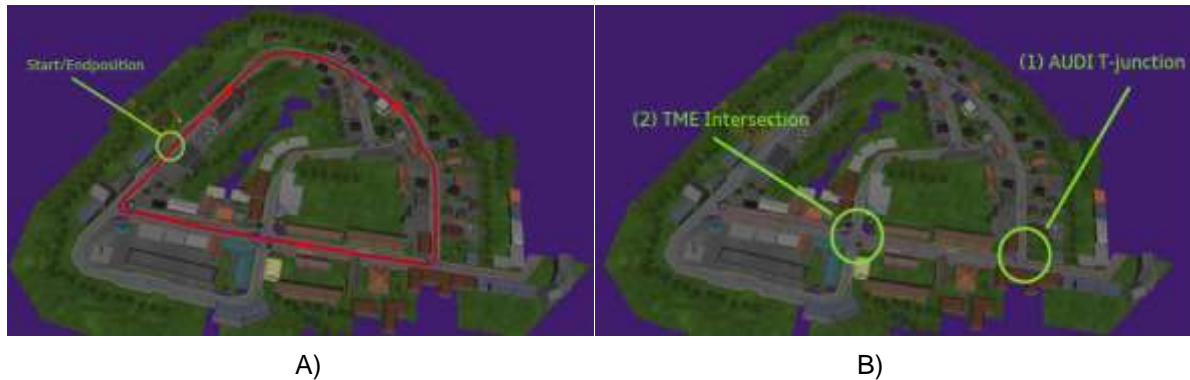


Figure 16: A) Route followed by each participants B) Intersection used for TME study

4.1.5 Participants

For the current study, 39 participants took part (32 males, 7 females). Participants were randomly allocated into one of the two groups (warning inside the comfort boundary or warning outside the comfort boundary). For the participants in the “warning inside the comfort boundary” group, the mean age was 23.65 years (SD=3.27) and for the “warning outside the comfort boundary” the mean age was 24.84 years (SD=5.93). The median of frequency of driving per week was “3-5 days” and the median of mileage per year was “20,000-30,000 km” for both groups. 7 participants from the “warning outside the comfort boundary” were removed due to problems with the recorded data.

4.2 RESULTS

4.2.1 Brake onset after warning

The first analysis done was to determine the time gap between the warning and the brake onset. This would determine how critical the situation was felt when the warning was issued. The hypothesis is that the warning inside the comfort boundary would not feel critical compared to the warning outside the comfort boundary. The mean gap time for the “inside comfort boundary” group was 1.32 seconds (SD = 0.99) and for the “outside comfort boundary” group was 0.65 seconds (SD = 0.38), as presented in Figure 17. A two sample t-test was done to compare both groups and it was found that there was a significant statistical difference between both groups ($t(29)=2.13$, $p=0.0421$). Behr et al. (2010) found that a participant braking reaction to an expected warning was 0.416 seconds (SD = 0.095). This means that the participants who received the warning outside the comfort boundary reacted almost immediately after the warning, i.e. they felt it was already a critical situation.

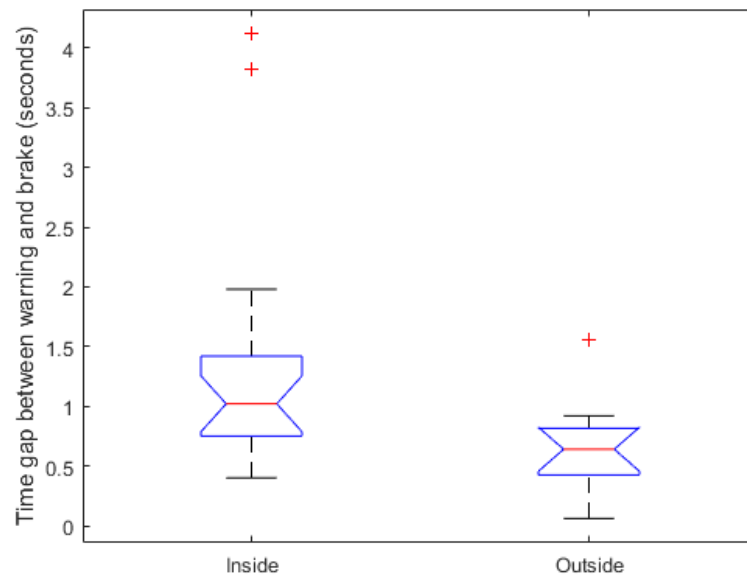


Figure 17: Time gap between warning time and brake onset time for both scenarios: inside and outside of the comfort boundaries

4.2.2 Acceptance questionnaire

Participants were asked to complete a survey at the end of the experiment. The survey consisted of 5 questions with a 5-point Likert-scale. The questions were the following:

- Do you think the warning was in time?
- How helpful was the warning to you?
- How disturbing was the warning to you?
- Do you think the warning was unnecessary?
- If it were available in the market, would you buy it?

A Wilcoxon rank sum test was done to compare the answers given by the participants in both groups. For the first (Was the warning on time?) and third question (Was the warning disturbing?), there was a statistical significant difference between groups. For the first question participants from the “inside comfort boundary” group score a median of 3 (“Just right”) compared to the “outside comfort boundary” group with a median of 5 (“Too late”). This indicates that participants felt that the “outside comfort boundary” warning was issued later than when they would have braked. For the third question, although both groups did not find it disturbing, the “inside comfort boundary” group found the warning more disturbing than the “outside comfort boundary” group.

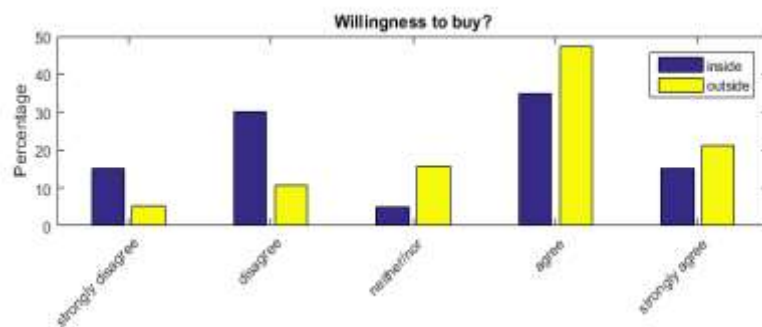
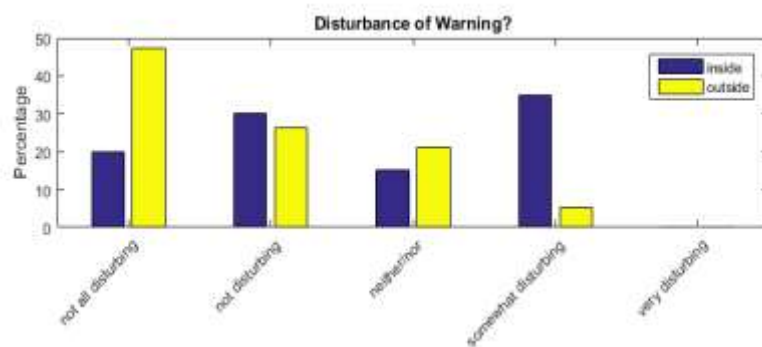
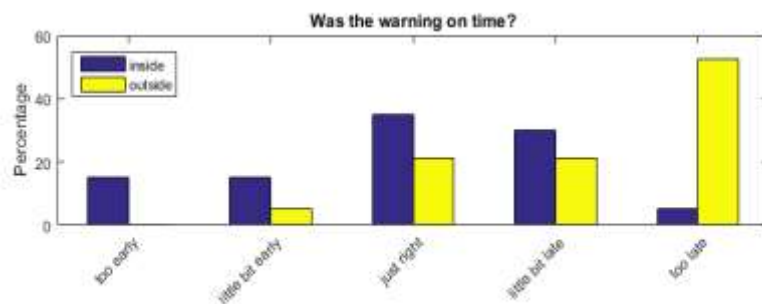
The second, fourth and fifth question did not show any significant difference between groups. For the second question (Was the warning helpful?) both groups found the warning helpful (both groups score a median of 4 which meant “Somewhat helpful”). For the fourth question, both groups found the warning necessary. And for the fifth question, both groups tended towards agreeing to buy the system, regardless of the

timing of the alarm. Table 3 shows the results of the analysis of the survey responses and Figure 18 shows the results of the survey by group.

Table 3: Responses per group to the five questions in the survey given to the participants after the drive

Question	Median “inside comfort boundary”	Median “outside comfort boundary”	p-value
Warning on time?	3 (“Just right”)	5 (“Too late”)	0.0013 *
Warning was helpful?	4 (“Somewhat helpful”)	4 (“Somewhat helpful”)	0.5206
Warning was disturbing?	2.5 (between “Not disturbing” and “Neither/Nor”)	2 (“Not disturbing”)	0.0308 *
Warning was unnecessary?	2 (“Not unnecessary”)	3 (“Neither/Nor”)	0.6622
Willingness to buy?	3.5 (between “Neither/Nor” and “Agree”)	4 (“Agree”)	0.1674

*Statistical significance at $\alpha < .05$



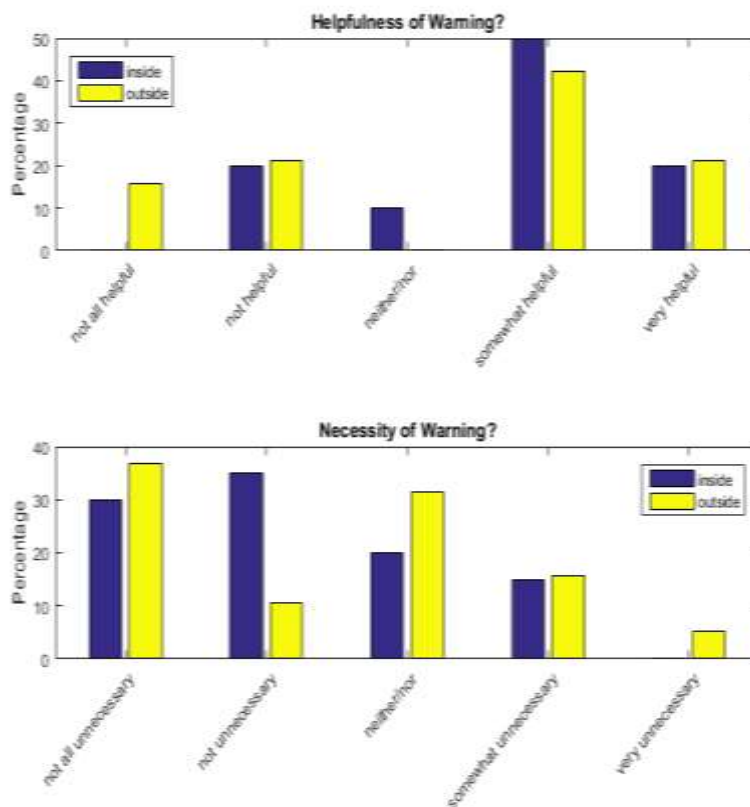


Figure 18: Survey results for both groups

4.2.3 Correlation

A spearman rank correlation test was performed to determine the effect of the gap between the warning and the braking time and the willingness to buy a PROSPECT-like system. It was found a slightly significant negative correlation between the time gap and willingness to buy ($r(31) = -0.3478$, $p = 0.0552$). This means that the closer the warning time is to the time when a driver would normally brake, i.e. when the driver feels a situation is critical, the more the driver will be willing to buy a PROSPECT-like system. It is therefore necessary to understand the comfort boundaries of people to avoid signaling warnings too early and reduce the acceptance of people towards the system.

4.3 CONCLUSIONS

The study demonstrated that it is possible to use comfort boundary models to determine more acceptable warning times for drivers. The braking reaction towards the warning outside the comfort boundary confirmed that the warning was issued when the scenario was already critical for the driver (outside the driver's comfort boundary). In similar way, the long gap between the braking onset and the warning inside the comfort boundary confirms that the driver was still in a "comfort zone" during the warning. This validates the comfort boundaries of the model. Secondly, through the surveys' answers, it was found that the warning outside the comfort boundary was less disturbing than the one inside the comfort boundary. It was also found that people would be more willing to buy the warning system that has the warning outside the comfort boundary. Although this was tested only in one scenario,

it is an important step to validate driver models as a method to develop better accepted warnings by the drivers.

5 EXPERIMENT #3 “UNIVERSITY OF NOTTINGHAM ACCEPTANCE TESTING”

5.1 OVERVIEW

The University of Nottingham (UoN) conducted a large-scale, longitudinal driving simulator study (N=48) to evaluate system functionality (T7.2) and issues of driver trust and acceptance (T7.3). Adopting the methodology developed by Large et al. (2017), the study took place in the Human Factors driving simulator at the University of Nottingham. The driving simulator is a medium-fidelity, fixed-based simulator comprising an Audi TT car located within a curved screen, affording ~270° forward and side image of the driving scene via three overhead HD projectors. A thrustmaster force-feedback steering wheel and pedal set are faithfully integrated with the original Audi steering wheel and pedals. STISIM Drive (version 3) was used to create an urban driving environment.

5.1.1 Methodology

Forty-eight experienced drivers were invited to attend at the same time on each of five consecutive days (Monday to Friday), and completed the same journey, which was presented to them as their daily commute.

The journey began on the outskirts of an urban environment and continued through the city. Towards the end of the drive, which lasted approximately 10-15 minutes, participants were asked to make a left turn. Shortly after this, drivers were advised that they had reached their destination and were asked to safely stop the vehicle.

The simulator was modified to replicate the PROSPECT functionality (utilising an audible warning and emergency braking intervention, as specified by D5.2). However, this was only triggered once during the week, when a cyclist was detected crossing the final road into which driver was turning. The intention was to replicate the ‘likely’ frequency of activation associated with a ‘real’ system. During the remainder of the study (i.e. all other visits to the simulator), the driving experience was routine, i.e. no cyclist hazard present during the manoeuvre. This approach improves upon other ‘single-visit’ simulator studies in that participants are not inundated with warnings and interventions in rapid succession (which can provide a false representation of the system, and is therefore likely to generate a poor assessment of acceptance), but rather experience the system only occasionally.

The study explored PROSPECT use-case 2 (see Figure 19), whereby a vehicle and a cyclist are approaching a crossing from the same direction. The cyclist wants to continue straight ahead while the vehicle intends to turn to the right. A collision risk occurs when the cyclist starts crossing the road at the instant when the car starts turning to the right.

Half of the participants (n=24) experienced a ‘true-positive’ system intervention, i.e. the system identified a critical incident – in this case, the cyclist crosses the road into which they were turning, and a warning is provided, followed by emergency braking (Figure 20).

The remainder of the participants (n=24) experienced a ‘false-positive’ intervention, i.e. a cyclist is detected as they approach the roadside and a warning is provided, followed by emergency braking. However, in this case, the cyclist actually stops before entering the roadway (Figure 21).

As such, it was expected that the latter intervention would be perceived as a ‘false-alarm’ as the cyclist did not actually enter the roadway. Nevertheless, the situation is still arguably ‘critical’ in so far as the system predicted that a collision would take place based on the current trajectories of both parties, and also likely to be highlighted by the PROSPECT system. Consequently, understanding the effect on driver acceptance and acceptability is still highly relevant.

During the study, acceptance was subsequently assessed utilising the approach developed as part of PROSPECT and documented within D7.2. In this case, only the ‘during’ and ‘after’ questionnaires (in addition to capturing demographic data) were employed to avoid biasing the results (i.e., to avoid creating expectations of behaviour from ‘the PROSPECT system’). It was also felt that this more accurately reflected a ‘real-world’ situation, whereby drivers would not necessarily be aware of the operational intricacies of active safety systems in their vehicle, and therefore an emergency intervention by the car would likely be unexpected.

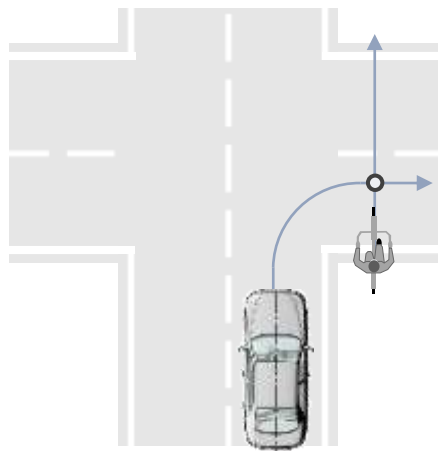


Figure 19- PROSPECT Use-Case 2 replicating in the study (Note: testing took place in the UK, and therefore the scenario was mirrored)



Figure 20: Cyclist crosses the road into which the driver is turning ('true positive'). Note: image is slightly distorted due to projection on an immersive, wrap-around screen; road turning is actually at 90° to main road.



Figure 21: Cyclist stops at road-side, but system still activates ('false positive'). Note: image is slightly distorted due to projection on an immersive, wrap-around screen; road turning is actually at 90° to main road.

5.1.2 Demographics

Forty-eight participants took part in the study, with 24 experiencing the true positive and 24 experiencing the false positive intervention. Participant demographics (age, gender and driving experience) are provided below in Table 4.

Table 4: Participant demographics

	Gender	Age	Years with licence	Mileage (Percentage on Highway)
True Positive (n=24)	12M, 12F	Mean 33.1, Range 21-64	Mean 12.9, Range 2-37	Mode 10-20k km (Mean 45%, Range 0-98%)
False Positive (n=24)	14M, 10F	Mean 32.5, Range 23-56	Mean 12.9, Range 2-38	Mode 10-20k km (Mean 41%, Range 0-90%)

5.2 RESULTS

5.2.1 Defining Criticality of Situation

Participants were presented with either a true positive or false positive ('false alarm') intervention. In both situations, the system provided an audible warning and emergency braking was applied. During the false alarm scenario, a cyclist was still present, but stopped at the roadside. As such, it was considered that the scenario may still be perceived as 'critical' as participants would be aware of the potential hazard. It is therefore noteworthy that participants rated both situations with similar levels of criticality (87.5 and 86.5, respectively) (Figure 22), using the CRITIC method (Common Risk awareness measurement meThod for Inter-population Comparisons) (Bellet and Banet, 2012).

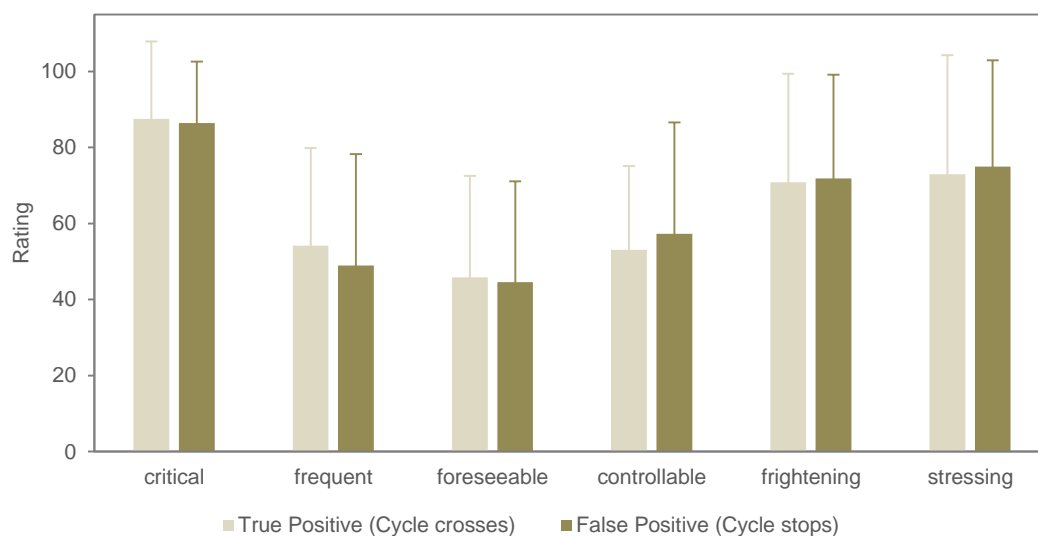


Figure 22: Ratings for Criticality

5.2.2 Assessment of System Acceptability

Using the approach proposed by Van Der Laan et al. (1997), participants rated items that collectively contribute to 'usefulness', 'ease of use' and 'satisfaction'. Combined, these can be interpreted as 'acceptability'. Participants generally rated the system highly on all scales (Figure 23). Overall, acceptability was equivalent across both scenarios (true positive and false positive), with ratings of 83.0 and 82.2 (out of 100), respectively.

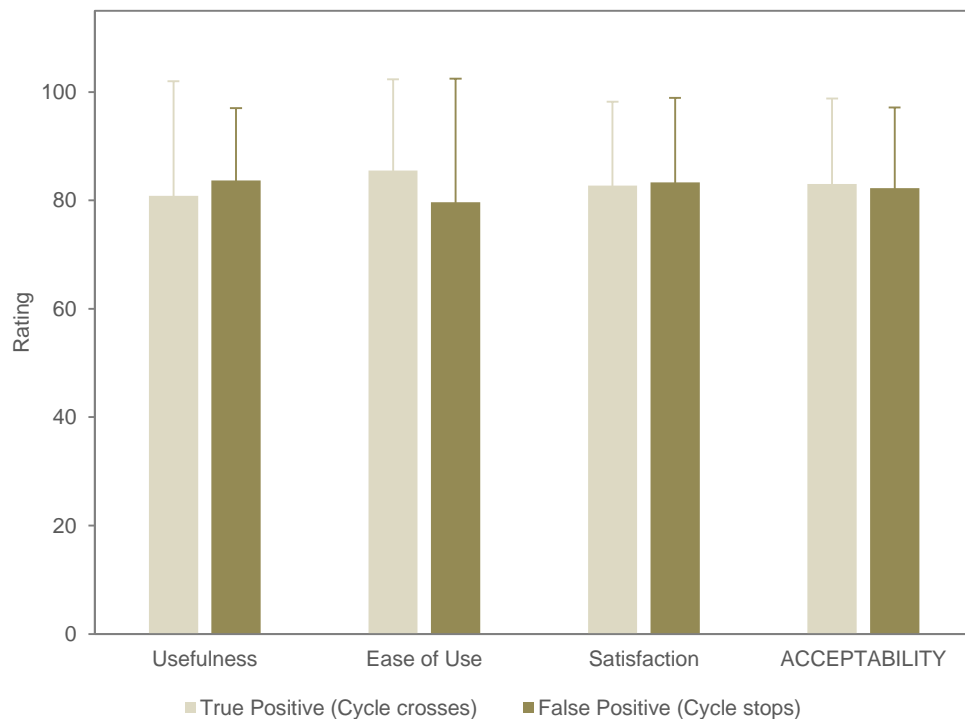


Figure 23: Acceptability Ratings

5.2.3 Assessment of System Trust

Using the Trust in Automation scale (Jian et al., 2000), participants rated their trust in the system. Trust is calculated by combining ratings for trust items and reverse-scored distrust items. Overall, trust was high in both scenarios (true positive and false positive), with ratings of 880 and 861 (out of 1200) respectively (Figure 24).

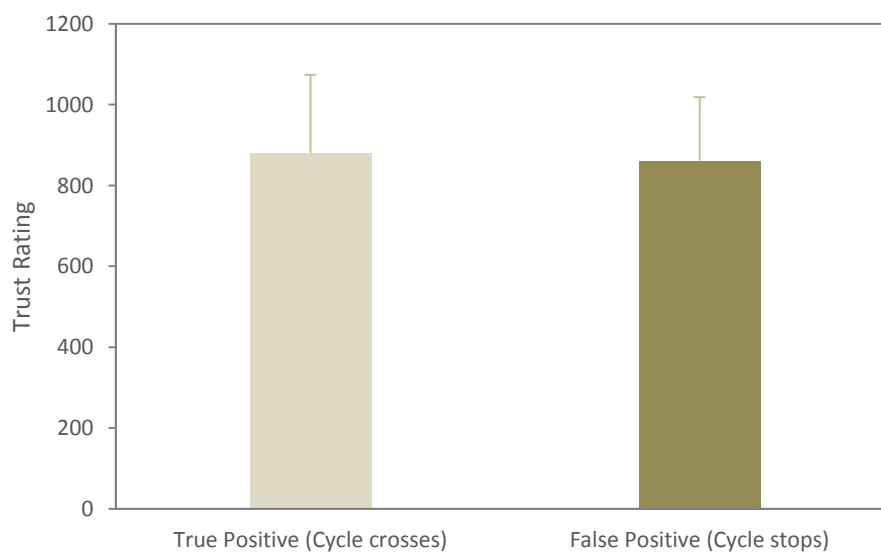


Figure 24: Trust Ratings

5.2.4 Intention to Activate/Use Warning Function

Participants reported a high likelihood of activating or using the **warning** function provided by the system, indicating that they would use or activate warnings in 79.1% of **all** driving situations based on their experience of a true positive activation, and 82.0%, when they were exposed to a false alarm (Figure 25). Nevertheless, it is worth highlighting that several participants reported that they were not aware of or did not hear the warning during the study.

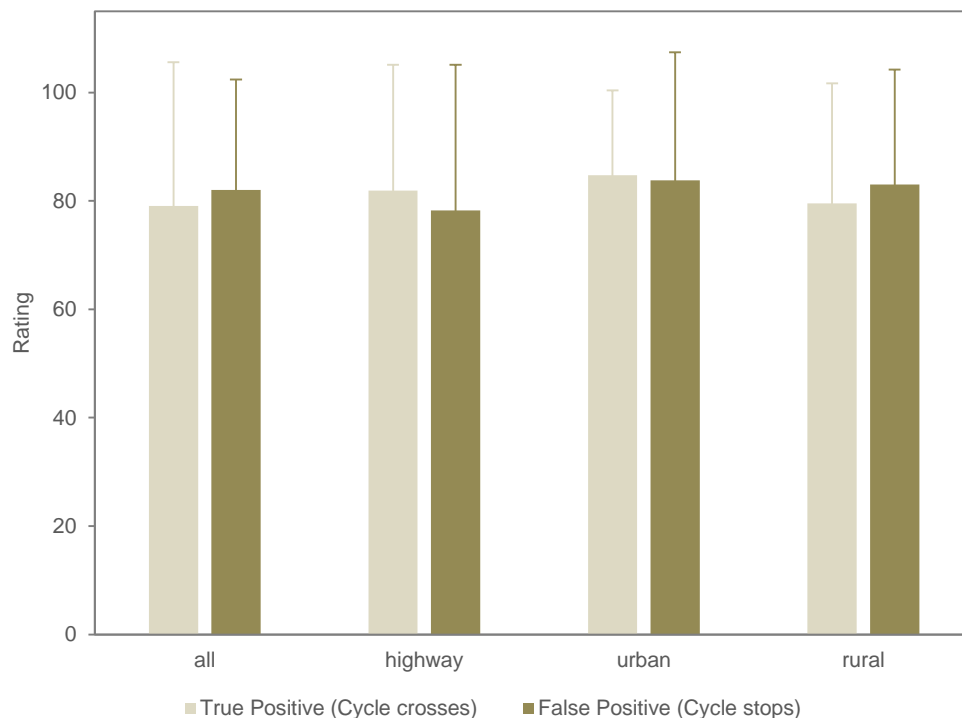


Figure 25: Intention to Use Warning Functionality

5.2.5 Intention to Activate/Use Braking Function

Participants were also positive about the **braking** function, indicating that they would activate or use this in 68.2% of **all** driving situations (based on their experience of a true positive intervention), and 73.1% of situations by those experiencing the false alarm (Figure 26).

The higher ratings made by those experiencing the false alarm is interesting. Combined with the high ratings for criticality (comparable to the true positive interventions), this may suggest that these drivers recognised the potential of the system, i.e. its capability to identify a potential cyclist hazard and predict its intention to enter the roadway, even when the cyclist ultimately stopped at the roadside. Drivers exposed to a more 'classic' false alarm, i.e. whereby the system activates but no cyclist is present (not tested here), may be less accepting of the technology.

It is also interesting to note that the ratings associated with the intention to employ the braking function on a highway were lower, particularly for those experiencing the false alarm. This suggests that participants would be less accepting of emergency autonomous braking on higher speed roads, particularly where this might be deemed unnecessary (as may be the case for a false alarm).

Nevertheless, it is worth highlighting that because of the nature of the experimental scenario (and the simulated environment), drivers were already slowing down or braking when they approached the junction to turn (as defined by the use-case). Moreover, as a fixed-base simulator, there is an absence of the physical forces associated with braking. As such, some participants reported that they were unclear whether the vehicle had braked or not.

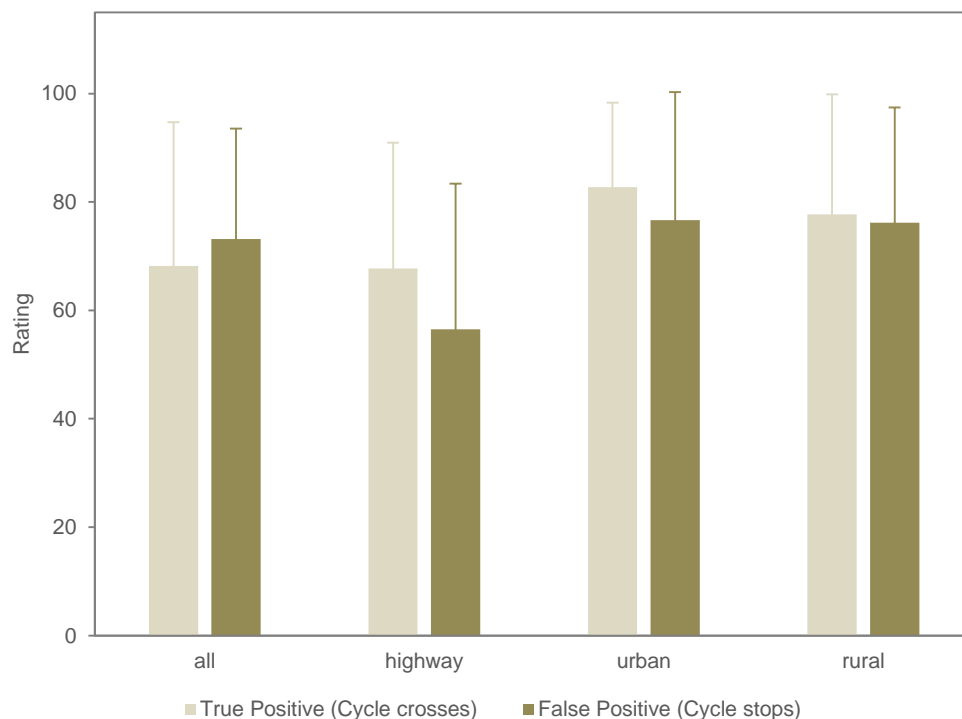


Figure 26: Intention to Use Braking Functionality

5.2.6 Willingness to Buy A PROSPECT-Like System

Overall, participants indicated a high likelihood of buying the PROSPECT system based on their experience during the simulator study. For those who were presented with a true positive, this was 78.3%, whereas for participants experiencing the false positive alarms, the likelihood of buying the system was slightly lower at 68.3% (Figure 27). Even so, it is worth noting that factors such as price and reliability were also highlighted (in written comments) by several participants as important considerations (in addition to performance) when making this decision.

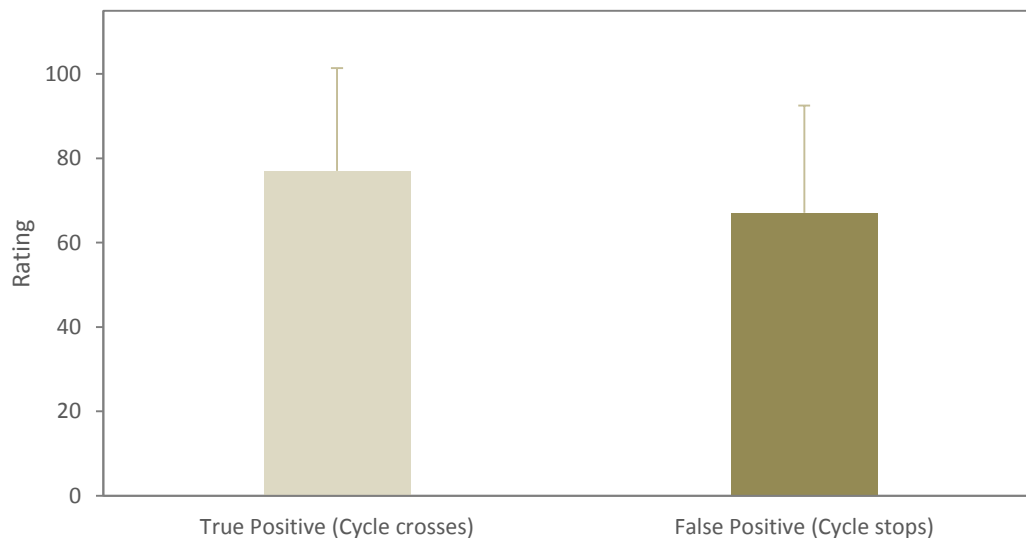


Figure 27: Willingness to Buy a PROSPECT-like System

5.3 ANALYSIS

Independent samples T tests were conducted for all measures, using 'Intervention' (true-positive versus false-positive) as the grouping variable. There were no significant differences for any of the reported measures between participants experiencing either the true positive or false positive alarm (all $p > 0.05$), suggesting high acceptance of the PROSPECT system even in situations where drivers' only experience was a false alarm.

5.4 CONCLUSIONS

The study aimed to evaluate system functionality (T7.2) and issues of driver trust and acceptance (T7.3) associated with the PROSPECT system, by employing a between-subjects, longitudinal simulator with 48 participants. Using this approach, with a single activation on day 5, provided a more ecologically-valid experience to determine acceptance/acceptability than a single-visit simulator study, in which participants experienced multiple activations.

Overall, results suggest high acceptance of and high willingness to buy the PROSPECT system, even in situations where drivers' only experience was a false alarm, although this varied slightly depending on the road situation (e.g. ratings suggest that participants would be less accepting of emergency braking on higher speed roads, particularly if this was associated with a false alarm).

6 EXPERIMENT #4 “DRIVER REACTION TO FCW WARNING AND AUTOMATIC STEERING INTERVENTION IN CRITICAL VRU SCENARIOS”

6.1 METHOD FOR SIMULATOR EXPERIMENT

The first part of the fourth experiment was performed as a driving simulator study run at VTI in Gothenburg, Sweden. The study had the main objective to assess driver responses to interventions and warnings of a PROSPECT like system. A secondary objective was to validate results from Sim IV through a comparison with test track results derived from T7.1 and reported in D7.1. The vehicle control strategies as well as the HMI strategies are based on T5.3 and T5.4.

The study applied for approval to the regional ethics review board in Linköping, Sweden. As of the judgement of the ethics review board there were no ethical concerns in relation to Swedish law on human subject research. Thus, no objections were made by the board. Although there were no objections towards the study by the ethics review board, the research was planned and conducted in accordance with the ethical guidelines of the American Psychological Association (APA), e.g. regarding informed consent and debriefing after study participation.

6.1.1 Research design

In the driving simulator study a short scenario (warning-TTC 1.7 s) and a long scenario (warning-TTC 2.2 s) was used. The target varied between a doll representing a pedestrian and a moving bicycle target, and the HMI varied between providing a warning or combining the warning with a steering intervention. See table 5 for an overview of the research design. The methodology used for acceptance evaluation is described in the Deliverable D7.2 (Report on methodology for balancing user acceptance, robustness and performance).

Table 5: Description of experimental conditions in the driving simulator

Condition	Participants	Target	Warning TTC (s)	Intervention TTC (s)
A	16	Doll	1.7	-
B	18	Bike	1.7	1.4
C	18	Bike	1.7	-
D	15	Bike	2.2	-
E	18	Doll	2.2	-

Note. For acceptance evaluation all conditions are included, while for test track comparisons only doll conditions are used, and for the simulator study the bike conditions are used.

6.1.2 Participants

In total 85 participants participated in the study. In the acceptance evaluation all 85 were included. In the results comparing simulator data and test track data 26 participants were included. All participants were recruited through advertisement on social media including sponsored material on Facebook and LinkedIn. All participants were rewarded with either two cinema tickets or gift cards valid at a flower shop.

6.1.3 Materials

6.1.3.1 Driving simulator

The VTI Driving Simulator IV (Sim IV), located in Gothenburg, Sweden, was used in the experiment. Sim IV was designed and specified with high level requirements to be able to provide a realistic simulation of real driving impressions (Jansson, et al. 2014). See Figure 28 for an illustration of Sim IV.



Figure 28: VTI Driving Simulator IV, on location Gothenburg, Sweden.

The motion system is capable of substantial linear movement over both x and y axes of a sledge, plus translation and rotation of the hexapod for a total of 8 Degrees of Freedom (DOF). The visual system uses nine projectors to provide the driver with about 180-degree forward field of vision. Furthermore, the sound system with a 6.1 surround set-up allows directional sound from objects outside of the cabin (Jansson, et al. 2014). For an overview of the technical specifications of Sim IV, see Table 6.

Table 6: Technical specifications of VTI Driving Simulator IV.

<i>Motion system</i>	
Pitch (degrees)	± 16.5
Roll (degrees)	± 16.5
<i>Linear system</i>	
Amplitude (m)	
Surge	$\pm 2.5 / \pm 0.31$
Sway	$\pm 2.3 / \pm 0.32$
Velocity (m/s)	
Surge	$\pm 2.0 / \pm 0.8$
Sway	$\pm 3.0 / \pm 0.8$
Acceleration (m/s^2)	$\pm 5.0 / \pm 6.5$
<i>Visual system</i>	
Forward view (degrees)	>180
Rear-view mirrors (LCD screens)	3
Average resolution on screen* (arc minute per line pair)	
Horizontal	5.0**
Vertical	2.5**
<i>Exchangeable cabin</i>	
Passenger car	Volvo XC60
Heavy truck	Volvo FH16

*The human eye has 0.59 arc minute per line pair, ** ± 0.5 .

To produce realistic simulations the simulation software used in Sim IV is based on open standards and in-house developed software (Jansson, et al. 2014). The simulator software has three main components: (1) the ViP Core, (2) VISIR, and (3) SIREN. The ViP core is used to run simulations and contains for example the scenario and vehicle dynamics. VISIR is used to render computer graphics and includes a scripts tool to generate roads. SIREN is used to produce sounds from the sound model (engine and tire noise, other vehicles) as well as from recorded sounds such as warnings used by the human-machine-interface.

During the experiments Sim IV was set-up with a small SUV passenger car cabin (Volvo XC60) and two out of the three rear-view mirrors were activated (side mirrors). Moreover, an eye-tracking system was used to record the test participants eye gaze. A secondary task system was also made available in the simulator, given the need to create artificial distractions under which critical conditions could be generated.

6.1.3.2 Warning strategy

The warning strategy was based on a human-machine-interface including a warning with a brake pulse, audio warning, and a red blinking LED light. The HMI was then administered at either a short (1.7 [s]) or slightly longer (2.2 [s]) TTC to form the warning strategy. In one of the conditions the warning strategy was also paired with an automatic steering intervention.

6.1.3.3 Scenarios

The simulator drive consisted of two parts, training, and the actual experiment. Under training, the subjects drove on a straight road where they could test drive the vehicle for common driving manoeuvres, i.e. braking, accelerating and lane changes. Before the training was completed, the drivers were exposed to the distraction task 5 times.

The accustomization period was followed by the experiment run on a generic rural road with a lane width of 3.3m. Very light traffic was present on the oncoming lane and no traffic on the test vehicle direction. For the duration of the drive, the test subjects were instructed to use the cruise control as much as possible. The cruise control had a fixed speed of 71km/h which could not be adjusted.

The test subjects drove for roughly 11 km while experiencing a total of seven distraction tasks. These consisted of an auditory cue which prompted the drivers to read six numbers on a secondary screen close to the passenger seat. Each number was presented for 200ms and followed by a blank screen that lasted 200ms, before the next number was rendered. The subject was required to press the distraction screen while reading the numbers, otherwise they would not be presented. All distraction tasks were triggered on straight sections of the road and oncoming traffic was deactivated before each distraction instance, so as to not hinder the driver from completing the task.

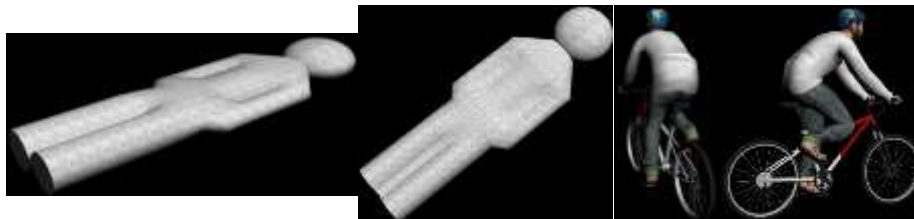


Figure 29: 3D models of critical objects.

After the seventh and final distraction task the subjects were presented with the critical scenario. In this situation a stationary doll, or a moving bicycle (see **Figure 28**), was placed in front of the simulated vehicle while the test subject was busy with the distraction task, and, consequently, its gaze was away from the road. Choice of doll or bicycle depended on the test condition, and the bicycle had a constant speed of 21km/h moving in the same direction as the simulator vehicle, while the pedestrian represented by the doll did not move. Regardless of which critical object was used, it was always spawned in the same position in the world, roughly 2.3m from the middle of the road. The road section where the drivers met the critical object was modelled after the test track ASTAZERO, in line with test track experiments running the same study in the real world, see **¡Error! No se encuentra el origen de la referencia.** 30.



Figure 30: Test Track and Virtual world, critical object spawn position.

predefined
The actions
were taken in the following order: (1) distraction task, (2) spawn of critical object, (3) warning, and (4) intervention. Each set of TTC triggers contains a reference TTC for each action, and two such sets were used in the simulator experiment (Table 7.)

Table 7: TTC trigger sets for the SIM IV experiment.

	Set A (TTC in sec)	Set B (TTC in sec)
Distraction	3.6	4.11
Spawn of Target Object	1.78	2.28
Warning	1.7	2.2
Intervention	1.4	-

In total, five different experiment conditions were tested in the simulator. TTC sets A and B with and without intervention and the bicycle as the critical object, and TTC set A with no intervention and the doll as the critical object.

6.1.4 Performance indicators

In the simulator study two performance indicators are analysed, namely:

- Distance bike right shoulder at time-to-collision = 0 [s]
- Lateral displacement relatively to doll
- Crashes (hitting the bike target)

The first indicator shows the criticality of the steering occurring in the scenario, the second shows more detail on the steering, while the third one is an indicator showing the severity of the situation.

6.1.5 Procedure

Upon arrival at the driving simulator facilities the participants were first welcomed and provided information about their participation to be able to provide an informed consent towards participation in the study. Next, they were introduced to the driving simulator and the secondary task. After calibration of the eye-tracking system they drove a practice scenario to familiarize with the driving simulator and the secondary task. After the familiarization they drove the experimental route including one of the experimental conditions at the end of the route. When the simulator driving was finished the participants answered all questionnaires. Before receiving the reimbursement, a debriefing session was held where the experimental leader answered any posed questions and disclosed more detail about the project and the experiment.

6.1.6 Analysis

The performance indicator distance bike right shoulder at time-to-collision = 0 [s] was analysed using a one-way uncorrelated analysis of variance, while collisions were analysed using a descriptive approach which was deemed sufficient due to the amount of crashes recorded.

In the results section covering the simulator data only conditions B, C, and D are included. These conditions are using the bicycle target with a relative speed differentiating it from the doll target (pedestrian). The conditions using a doll target were used for the comparison of driving simulator and test track results and the results are therefore used in the section covering the results from the comparison.

6.2 METHOD FOR THE TEST TRACK EXPERIMENT

The second part of the fourth experiment, performed by Volvo Car Corporation, was similar in setup to the first part by VTI, but took place on a test track. The main purpose was to evaluate driver reactions to warnings and interventions.

The test track experiment ended with a critical pedestrian scenario which occurred at a time when the driver was visually distracted by a secondary task. After a pedestrian dummy appeared on the road in front of the subject vehicle, the vehicle warned the driver of the impending threat through a forward collision warning (FCW). In addition to the warning, half of the participants also received an automatic steering intervention. The study procedure was approved by the Regional Ethics Review Board in Gothenburg, Sweden.

6.2.1 Research Design

The experiment, having a between subjects design, consisted of four conditions, differing only in the timing of the warning and the presence or absence of the steering intervention at the critical scenario. The time to collision (TTC) at which the warning and steering interventions were activated are shown in table 8 below.

Table 8: TTC of the warning and steering intervention for the four different experiment groups.

Group	Participants	Warning TTC (s)	Steering Intervention TTC (s)
F	16	1.7	1.0
G	15	1.7	-
H	19	2.2	-
I	15	1.95	1.0

6.2.2 Participants

A total of 108 Volvo Cars employees participated in the experiment, of whom 65 were included in the results. They were recruited via a participation request email and a participation questionnaire. The general inclusion criteria were:

- Having a valid driver's license and having driven a minimum of 5,000 km the previous year
- No whiplash, neck or heart problems
- No test drivers
- Not having previously participated in similar tests with surprise elements

For several reasons, not all data was relevant for inclusion in the analysis. Data from 43 participants had to be excluded because of warnings or interventions starting too early or too late, the participant looking up at the road before getting the warning (rendering it impossible to analyse their reaction to the warning), or technical issues with the dummy.

6.2.3 Equipment

6.2.3.1 Test Track

The experiment was performed on the rural road track at AstaZero in Sandhult, Sweden (figure 31). One lap of the rural road track is approximately 5.7 km long, with one lane of width 3.3 m in each direction. The straight stretch of road in the area highlighted in red in figure 31 is where the critical pedestrian scenario took place. This road segment has a slight upward incline.

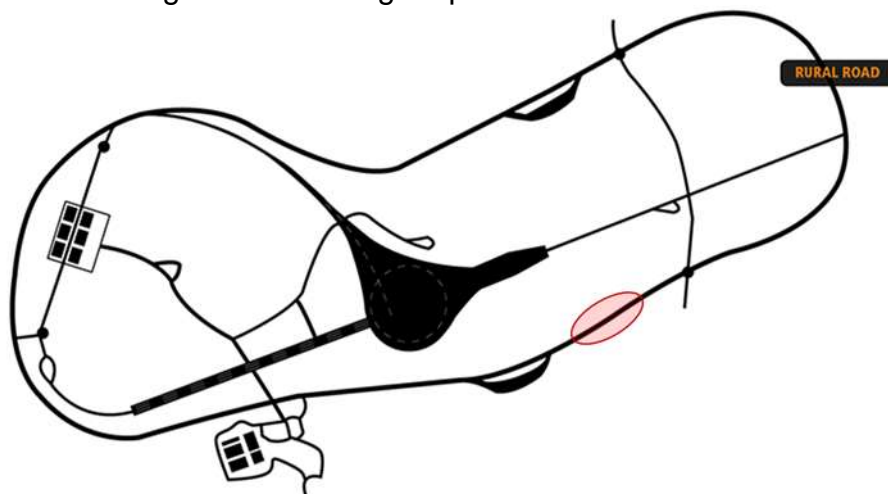


Figure 31: Rural Road test track overview. The red ellipse points out the area where the dummy was inflated.

6.2.3.2 Test Vehicle

The test vehicle was a Volvo V60 MY2018 (SPA platform), equipped with test software to enable warning and steering requests. The test vehicle was driven with cruise control (CC) set to a speed of 75 km/h, giving an actual velocity of approximately 71 km/h.

The forward collision warning (FCW) consisted of a sound, a brake pulse, a blinking icon in the graphical head up display (HUD), and an icon in the instrument cluster.

The steering intervention requests a certain pattern of pinion steer angle movement from the system. If the driver should try to override the steering intervention, they will be able to, although the system will still strive to follow the steering wheel movement pattern, and the resulting vehicle path might be influenced by both system and driver input.

6.2.3.3 Data Logging Equipment

The vehicle was equipped with Dewesoft measurement equipment to allow for the logging of vehicle signals, video data, and GPS data. Dewesoft and GPS-equipment were also used to trigger the secondary tasks, the warning and the steering intervention at the pedestrian event.

6.2.4 Procedure

Upon arriving at the test track, the participant received both written and verbal information about the study procedure, to be able to give informed consent for participation. After that, the test leader and the participant entered the test vehicle where the participant received instructions for the distraction task, and got to practice completing the task until he/she was accustomed to it. After the participant was told how to activate the cruise control, they were informed that there would be no other traffic on the track, and that the test leader would not converse with them during the drive. A test leader was present in the back seat of the car during the entire drive to start the secondary tasks and make sure that the drive proceeded safely while the participant performed the tasks. If needed, the test leader would warn the driver to ensure that they resumed control of the vehicle.

The participant was instructed to start the task as soon as possible after being prompted by an auditory cue. Upon hearing the sound, they were to immediately touch the display with a finger and read the displayed numbers out loud. The numbers would not show up unless activated by the finger press, and thus the participant had to hold his/her finger on the display for the duration of the task. Each task consisted of a sequence of six random one-digit numbers, with each number appearing for 200 ms followed by a blank screen for 200 ms.

The drive lasted for three laps on the track, amounting to approximately 15 minutes. The secondary tasks were initiated on straight sections of the road and were completed 15 times in total. The purpose of the secondary task was for participants to be visually distracted enough that their reactions to the warning and intervention would be comparable to that of unsuspecting drivers in real world critical situations. The critical pedestrian scenario occurred while the driver was occupied with the final secondary task.

After the scenario, the experimental drive ended, and the test leader informed the participant that the obstacle in the road was a balloon test dummy. The participant was asked to fill out two questionnaires about their experiences of the critical scenario and their opinions about the warning and steering intervention. They were then debriefed about the actual purpose of the study and allowed to ask questions.

6.2.4.1 Critical Scenario

The critical pedestrian scenario was triggered to start when the participant was busy performing the secondary task, and hence looking away from the road. The task was triggered by a GPS signal, and approximately 20 meters later, the test vehicle ran over a cable laying across the path, triggering the inflation of the pedestrian dummy. By the time the participant was in the middle of the number sequence, the warning was activated to alert the driver of the obstacle in the lane. The dummy measured 179 cm in length and was 60 cm at its widest, and the left side of the dummy was positioned approximately 2.15 m from the middle of the road.



Figure 32: Test track pedestrian dummy.

6.3 DATA ANALYSIS

The data were processed in Matlab, where the dependent variable results were calculated. Most variables, such as vehicle velocity and distances, were taken from vehicle signals or GPS data.

Figure 33 shows the progression of a critical event and the definition of the measured dependent variables, for example gaze and steering reaction times. Gaze data for the test track experiment was determined by manually coding video data of the driver to establish a point in time where the participant had fixated his or her gaze on the road ahead.

Steering reaction time is the time between the FCW and steering onset time. Steering onset time was calculated using two different methods: one for the comparison data, and one for only the analysis of the test track data. For the comparison between the simulator and the test track data, steering onset time was determined by first differentiating the steering wheel angle signal three seconds prior to the start of the distraction task, and finding the standard deviation of this derivative over the three second span. Then, the point of the steering start was determined as the first time after the warning where the derivative of the steering wheel angle signal reached 18 times the previously determined standard deviation.

For the analysis of only the test track data, the steering onset time was manually coded by determining the point in time where the steering wheel angle signal started to increase from its minimum value.

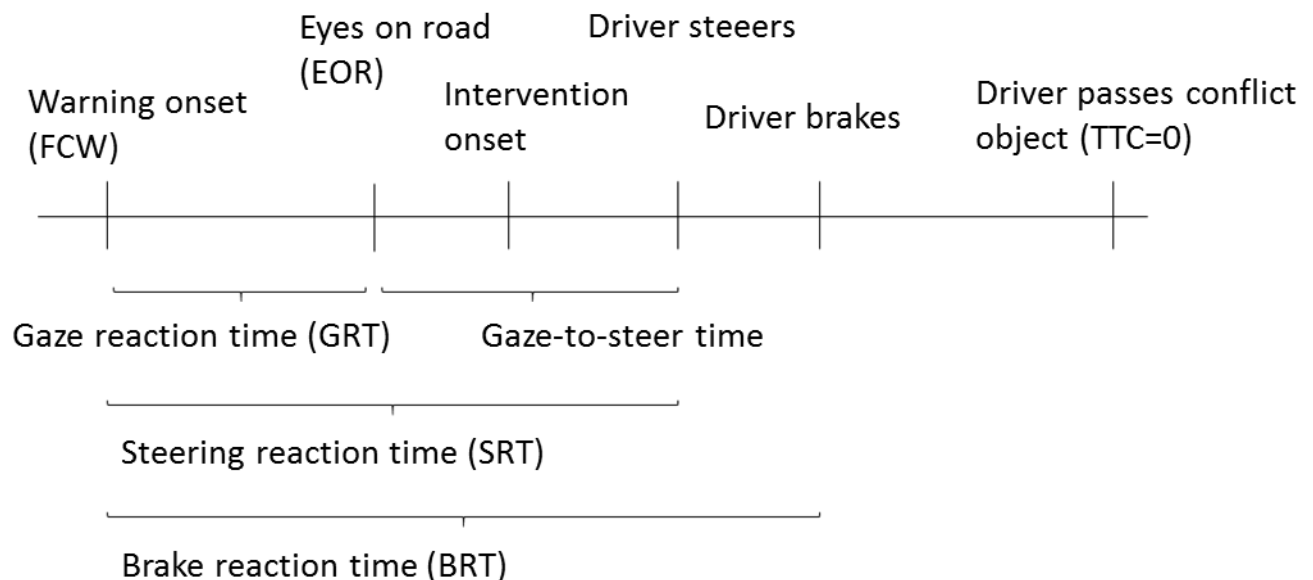


Figure 33: Figure describing the different dependent variables and how events unfold over time

The results were then analysed statistically in Minitab, with ANOVAs (one-way and two-way), T-tests and Chi-Square tests being performed according to the situation's requirements. For the comparison between simulator and test track data, a 2 (test environment) x 2 (warning-TTC) ANOVA was performed in order to describe the influence of test environment and warning-TTC on the various dependent variables.

6.4 DIFFERENCES IN METHOD BETWEEN SIMULATOR AND TEST TRACK EXPERIMENTS

The simulator and test track experiments were designed with the intention that comparing data between the experiments would allow for validation of the simulator results and thus create validity for experimental conditions that were only run in Sim IV.

6.4.1 Research Design

Four conditions were compared: two from Sim IV with the dummy as the critical object and only warning. The independent variables were the test environment and the timing of the warning. No groups with a steering intervention were included in the comparison.

Table 9: Test environment and TTC of the warning for the four different experiment groups.

Group	Test Environment	Warning TTC (s)	HMI
A	Simulator	1.7	Warning only
E	Simulator	2.2	Warning only
G	Test track	1.7	Warning only
H	Test track	2.2	Warning only

6.4.2 Methods

The methods used in the two experiments needed to be the same to be able to conclude that any differences in results were because of the test environment and not due to confounding factors.

Although careful consideration was given during the planning phase to ensure the similarity of methods, some details differ between the simulator and the test track:

- 1) The simulator warning featured a blinking LED in the windshield while the test track warning instead had a blinking graphical icon in the HUD.
- 2) The simulator scenarios featured oncoming traffic throughout the drive (although not during the critical scenario or any secondary task), while the test track was clear of any other traffic
- 3) The cruise control in the simulator was deactivated by the warning, whereas in the test track experiment it was still active after the warning
- 4) The middle of the dummy in the simulator was positioned 2.3 m from the centre line, while the middle of the test track dummy was positioned 2.45 m from the centre line, resulting in a bit more overlap for the simulator scenario

These factors are important to consider in the analysis, as they might have impacted the results.

6.4.3 Procedure

The procedures between the simulator and test track studies differed slightly:

- 1) While the simulator participants were given a practice session to try out the simulator and the secondary task, the test track participants were not able to practice driving on the test track before the experimental sessions started. They were, however allowed to practice the secondary task while the vehicle was standing still until they felt comfortable with it.
- 2) In the simulator, the secondary task was repeated seven times during the experimental session, while on the test track, it occurred about 15 times.
- 3) The Dewesoft and GPS systems used to trigger the start of the FCW on the test track were not completely reliable, and as such there was a slight variation in the times at which the FCW started. In the simulator, the triggering of these events was much more reliable and better replicable.

6.5 RESULTS FROM PEDESTRIAN TESTS WITH FCW ONLY AT TEST TRACK AND IN SIMULATOR INCLUDING VALIDATION

6.5.1 Gaze Reaction Times from Test Track

Gaze reaction times were measured from the onset of the warning to the point where the driver's gaze was directed towards the road and aimed to evaluate the effect of the warning. These were obtained in the test track experiment but not for the Sim IV experiment. Between conditions in the test track study, a two sample T-test showed no statistically significant difference in gaze reaction time for warnings at TTC=1.7 s ($M=0.70$, $SD=0.08$) and warnings at TTC=2.2 s ($M=0.75$, $SD=0.30$); $t(21)=-0.70$, $p=.49$. This is expected since it is the same warning only perceived at different distances from the conflict object.

6.5.2 Reaction Types

Few participants reacted by braking alone: one of 26 in the simulator, and two of 34 on the test track. In the simulator, 24 of 26 participants (92%) braked as well as steered in their attempt to avoid a collision, whereas on the test track, only 13 of 34 participants (38%) reacted by both braking and steering. Moreover, of those who reacted in both ways, 12 of 13 test track participants steered first and then braked, with 10 of 24 simulator participants reacting the same way. The remaining 14 simulator participants braked first and then steered, a pattern that was only observed for one test track participant.

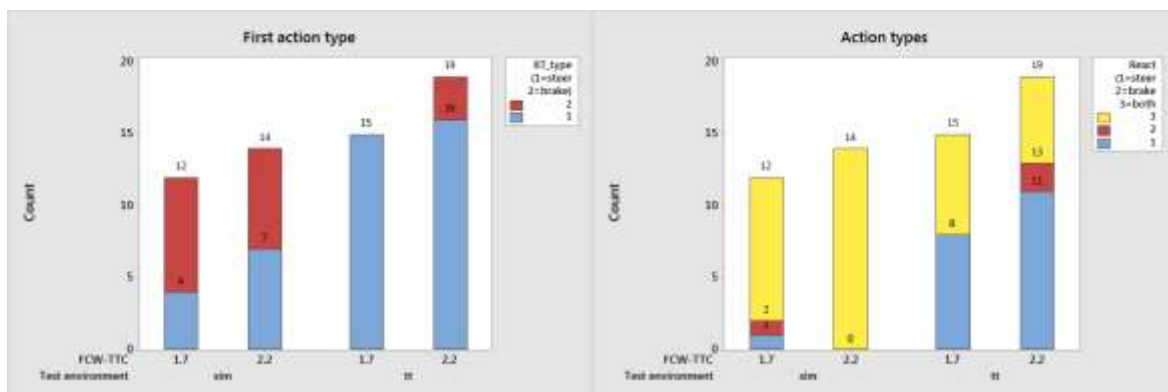


Figure 34: The left figure shows the first avoidance action and the right figure shows all avoidance actions

The initial reaction is the one to occur first of either steering or braking. In the majority of cases for the simulator experiment, the initial reaction is braking (15/26 participants), while on the test track, steering accounts for over 90% of initial reactions (31/34).

6.5.3 Steering Reaction Times from Test Track and Sim IV

The time from the warning onset until the participant started to steer was analysed from test track and Sim IV data. A two sample T-test from the test track data showed no statistically significant difference in steering reaction time between the warning-TTC=1.7 s ($M=0.95$, $SD=0.14$) and warning-TTC=2.2 s ($M=1.20$, $SD=0.60$) conditions; $t(17)=-1.67$, $p=.11$. This indicates that participants react to the warning in the same amount of time regardless of the distance to the conflict object in these two cases.

The time it takes participants to start steering from the time when they looked up at the road was examined for the test track experiment. A two sample T-test showed no statistically significant difference in eyes on road to steering started time for the warning-TTC=1.7 s ($M=0.25$, $SD=0.10$) and warning-TTC=2.2 s ($M=0.42$, $SD=0.62$) conditions; $t(16)=-1.10$, $p=.29$.

To validate the simulator, steering reaction times were compared between the Sim IV and test track results. There is a slight difference in warning HMI between the two test environments that could have impacted the reaction times, where the Sim IV has a blinking LED rather than a HUD warning. However, there was no statistically significant main or interaction effect in steering reaction time between the Sim IV and the test track as determined by the 2x2 ANOVA. There is no gaze data from the simulator; however, since steering reaction times do not differ significantly between test environments, it is possible that gaze reaction times do not either.

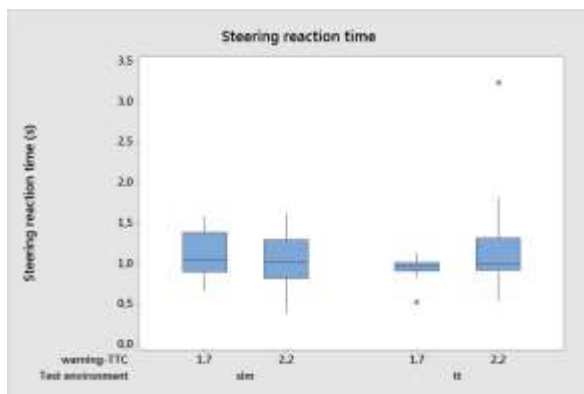


Figure 35: Steering reaction times were measured from the onset of the FCW until the start of the steering manoeuvre

6.5.4 Braking in Sim IV and Test Track

An analysis was done of the velocity maintained when the car passes the conflict object (TTC=0 s) or when its velocity reaches 0 km/h, whichever comes first. Sim IV participants had a significantly lower velocity than test track participants as shown by a two-way ANOVA, with a mean of 42.7 km/h ($SE=3.75$) compared to the test track's 60.6 km/h ($SE=3.29$) ($F(1,56) = 12.88$, $p<.005$). Furthermore, the same ANOVA demonstrated a lower velocity for those with an early warning ($M=45.5$ km/h, $SE=3.36$) than those with a later warning ($M=57.8$ km/h, $SE=3.69$) ($F(1,56) = 6.05$, $p<.05$). Being given an earlier warning allows participants more time to decrease their speed.

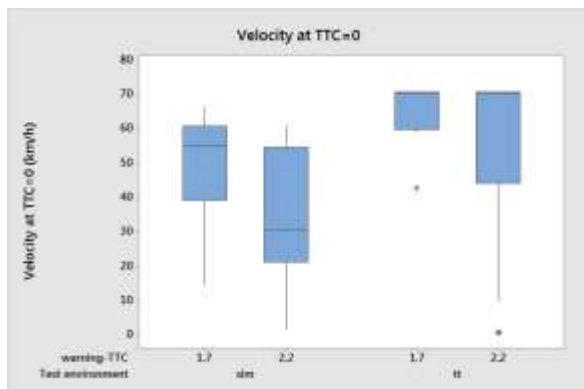


Figure 36: The velocity was measured at the point where the vehicle passed the conflict object

To explain the difference in velocity for the different test environments, one might consider the fact that the cruise control is deactivated by the warning in the simulator. However, the one test participant who did not brake had a velocity of 66.5 km/h (when the warning was given at the later time of $TTC=1.7$ s), a difference of only 4.5 km/h. This suggests that, at least for participants in the group with the late warning, the difference in speed decrease in the simulator is not affected remarkably by the cruise control deactivation.

No simulator participants had come to a standstill before arriving at the conflict object, suggesting that, had they not veered away from it, they would have collided with the conflict object. Thus, steering is the principal reason that collisions are avoided. Yet, simulator participants seem to possess an instinct to brake that is not observed to the same extent on the test track.

There is no conclusive evidence that the reason for the difference in braking propensity lies in the test environment. It is possible that method differences might have been confounding factors to the difference in braking tendency between the two test environments. First, the fact that the simulator experiment featured oncoming traffic earlier during the drive might have discouraged drivers from initially veering into the opposite lane, especially since the critical situation occurs on the way up a hill, blocking participants' views of the situation beyond it.

Second, the warning HMI differed between the simulator and the vehicle on the test track. In the simulator, the warning had a blinking LED display that was not present in the test vehicle, but was replaced instead with a less salient blinking graphical HUD image. If the brake pulse in the simulator was slightly more salient in Sim IV, it might have primed drivers to brake rather than steer.

Third, in the simulator, the onset of the warning deactivated the cruise control that all drivers were using, a phenomenon that did not occur in the test track experiment. It is unlikely that such an occurrence could have primed the drivers to brake since the speed difference over time between the deactivation of the cruise control and the brake reaction by the driver ($M=0.97$ s) is quite low and might not have been felt by the driver.

Although the result differences in braking propensity could be due to method differences, it might also be possible that the results are due to environmental factors after all. Further studies must be done, diminishing the differences in method, to support or oppose this possibility.

6.5.5 Distance to Conflict Object in Sim IV and on Test Track

The 2x2 ANOVA performed on data from all test participants except those who avoided collision by only braking showed a statistically significant main effect of warning-TTC on the lateral distance to the conflict object; $F(1,54)=26.97$, $p=.00$. Thus, those who received the warning 2.2 seconds before collision passed the conflict object with more lateral margin ($M=1.11$, $SE=0.10$) than those receiving the warning 1.7 seconds before the collision ($M=0.33$, $SE=0.11$). The effect of test environment on lateral distance to the conflict object was close to statistically significant ($F(1,54)=3.99$, $p=.051$), showing a tendency towards larger lateral distances in the simulator compared to on test track.

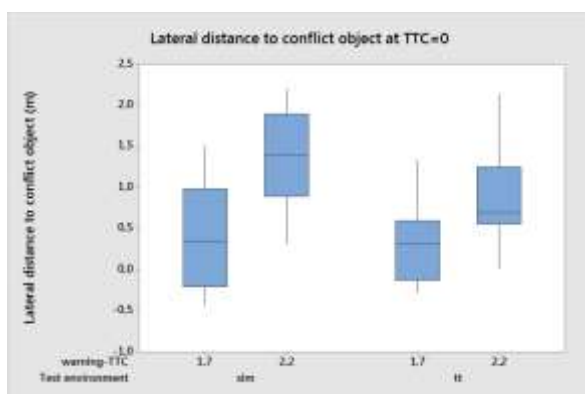


Figure 37: Comparison of the lateral distance between the vehicle's right side and the conflict object

Although the numbers confirm that there is a slight lateral difference between the two test environments, the following diagrams plotting the trajectory of each participants' vehicle over time show that the vehicles in Sim IV follow a path quite similar to that seen on the test track.

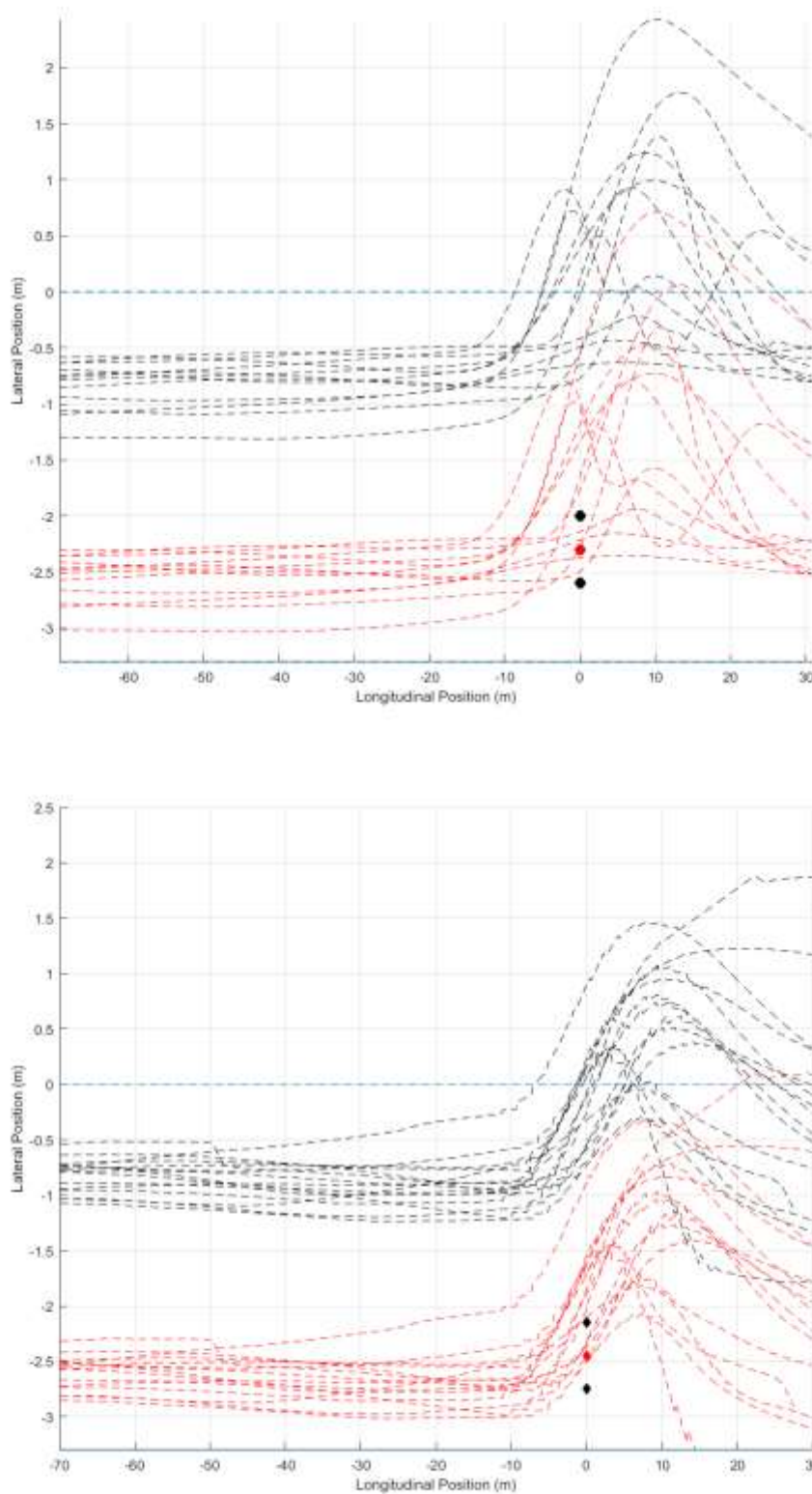


Figure 38: Trajectories for late warning (TTC=1.7 s) in Sim IV (above) and on test track (below). The diamonds mark the location of the conflict object. The black line is the left side of the vehicle and the red line is the right side of the vehicle.

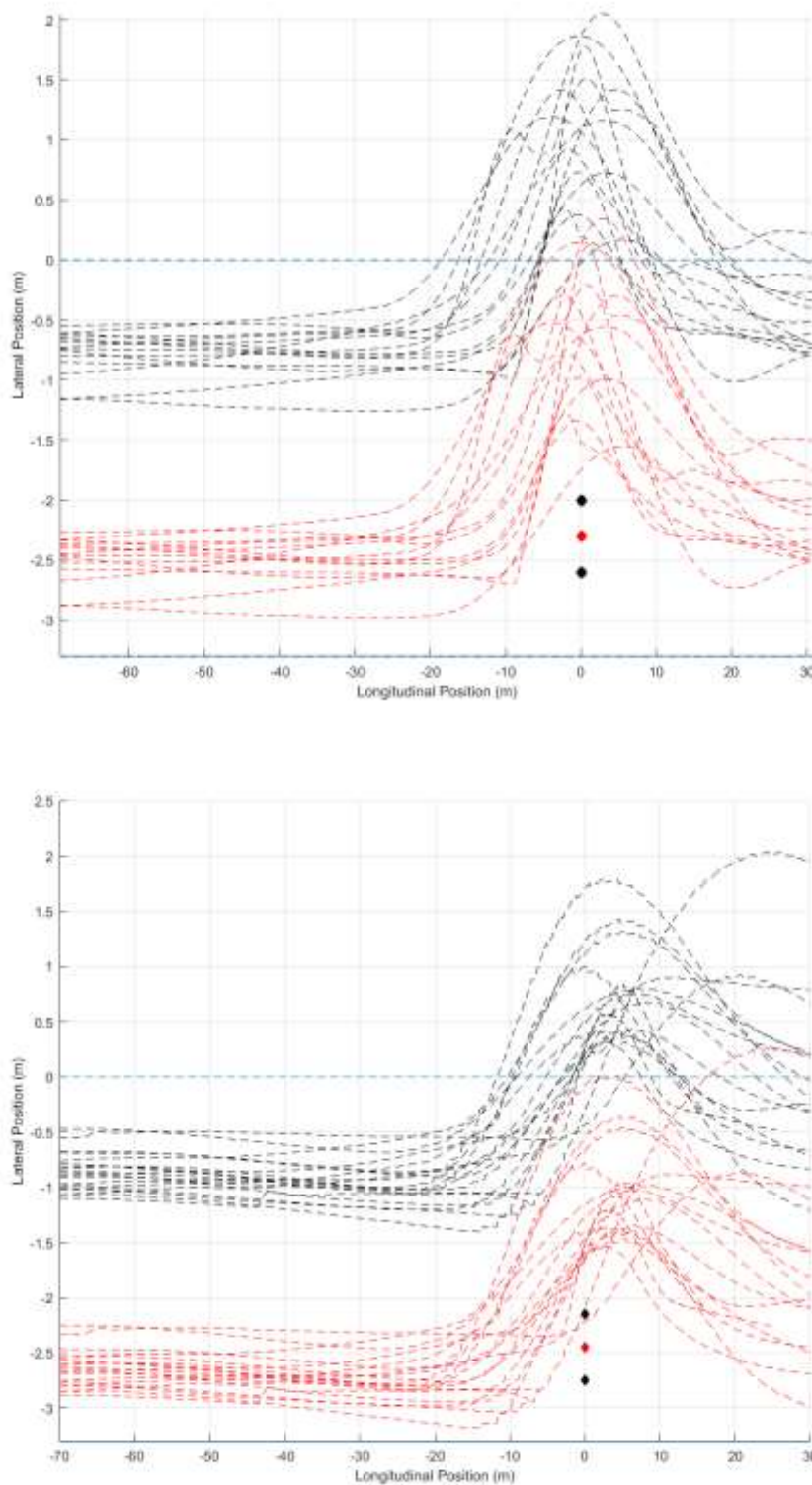


Figure 39: Trajectories for early warning (TTC=2.2 s) in Sim IV (above) and on test track (below). The diamonds mark the location of the conflict object. The black line is the left side of the vehicle and the red line is the right side of the vehicle.

6.5.6 Collisions in Sim IV and on Test Track

In the Sim IV experiments, there were no collisions in the group with the early warning (TTC=2.2s), while five out of 12 participants collided when given a late warning (TTC=1.7s). In the test track experiments, results similarly showed that only one of 19 participants collided in the group with an early warning, and in the group with the later warning, 5 of 15 participants collided.

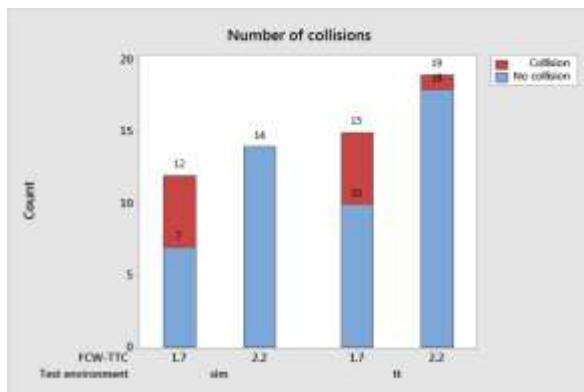


Figure 40: The number of collisions observed in the different conditions

6.6 RESULTS FROM PEDESTRIAN TESTS WITH FCW W/VO AUTOMATIC STEERING INTERVENTION AT TRACK TESTING

6.6.1 Reaction Types

In the scenario where the warning was given 1.7 seconds before the vehicle reached the conflict object, all 15 participants steered first. This accords with other studies (Eckert, 2011) showing that shorter TTC leads to steering reactions rather than braking reactions. In the group with the warning at a TTC of 2.2 seconds, three participants braked before steering, as was also the case in the group with the warning at a TTC of 1.95 seconds combined with a steering intervention at TTC 1.0 second.

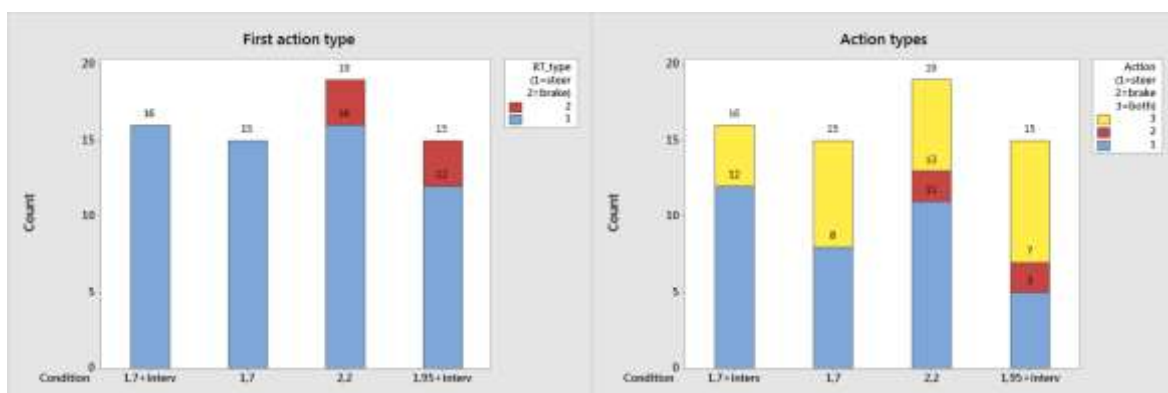


Figure 41: The left figure shows the first avoidance action and the right figure shows all avoidance actions.

6.6.2 Steering Behaviour

Observation of the steering wheel angle signal shows that, after beginning their steering manoeuvre to the left (steering counter clockwise), drivers start turning the steering wheel back (clockwise) before reaching the conflict object. The following graphs show how this is the case for two participants in different conditions.

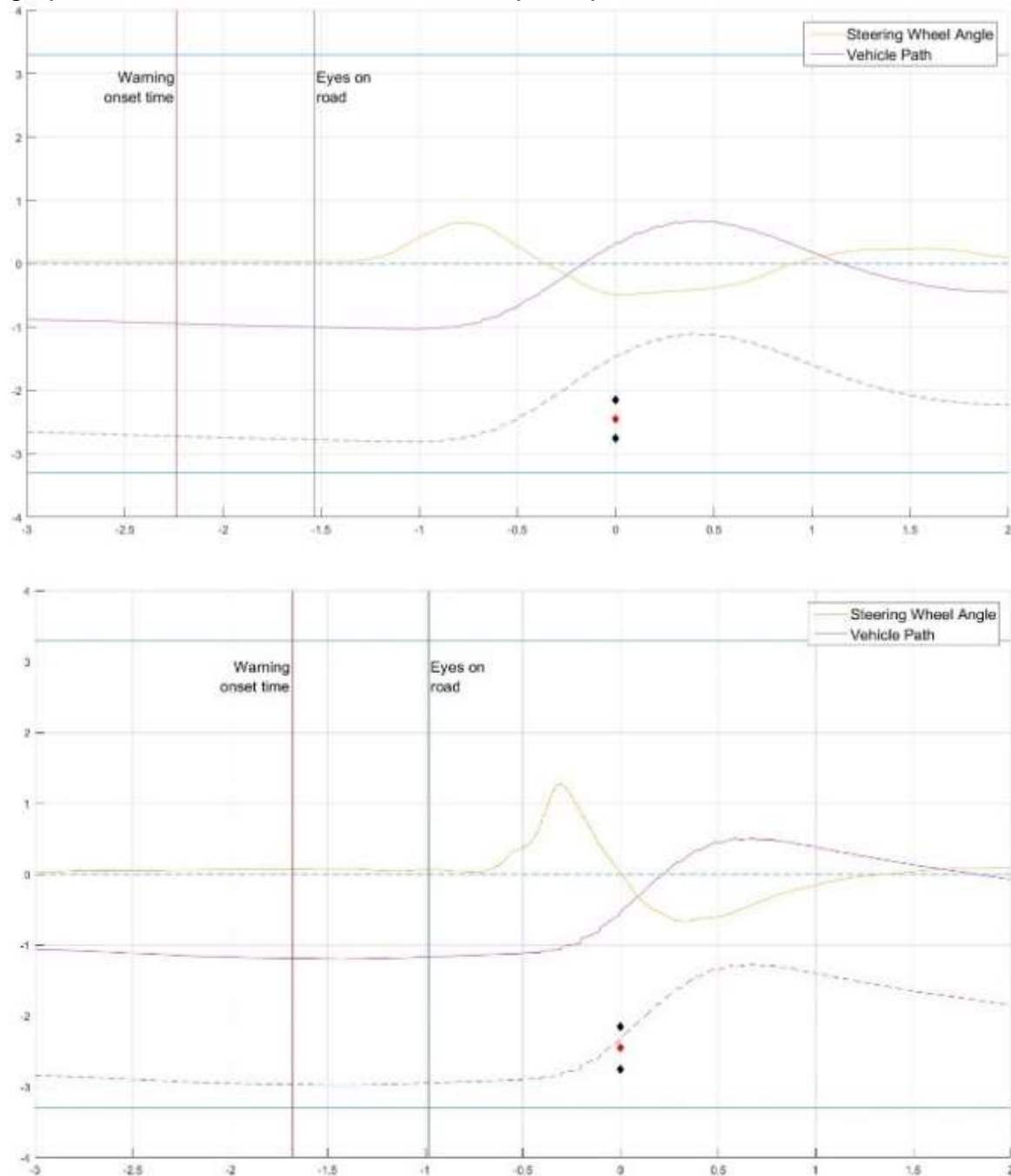


Figure 42: Above: Steering wheel angle and vehicle path of one driver with warning at $TTC=2.2$ s. Below: Steering wheel angle and vehicle path of one driver with warning at $TTC=1.7$ s.

A one way ANOVA showed no statistically significant difference in steering reaction time (time from warning to the start of the steering manoeuvre) between the four groups, indicating that the participants in all four groups started to steer as soon as possible after getting the warning. Probably even in the group where the participants received the warning “early” (TTC=2.2 s), the situation was perceived as critical and required a quick steering action.

6.6.3 Lateral Distance

A one-way ANOVA shows that the group with warning-TTC=2.2s (without automatic steering) has a significantly larger lateral distance to the conflict object at TTC=0 s compared to the groups with warning TTC=1.7 s (without automatic steering) and TTC=1.95 s (with automatic steering) ($F(3,59)=6.22$, $p<.05$), indicating that the longer time and distance you have before the conflict object, the bigger possibility you have to steer away further from it.

It was expected that the group with warning at TTC=1.95 s (with automatic steering) would have a bigger lateral distance to the conflict object than the group with warning-TTC=1.7 s (also with automatic steering) since they received the warning earlier. One explanation that this was not the case could be that the participants in the former group were driving around 25 cm further to the right in lane during the secondary task compared to the latter group. This difference corresponds quite well to the difference in lateral distance to the conflict object between warning-TTC=1.95 s ($M=0.23$) and warning-TTC=1.7 s ($M=0.54$). Another difference between these two groups was that all participants in the warning-TTC=1.95 s group had had time to get their eyes on the road before the automatic steering intervention started, while around half of the participants in the group with warning-TTC=1.7 s still had their eyes off the road. Further analysis is necessary to determine if and how this has an effect on the lateral distance to the conflict object.

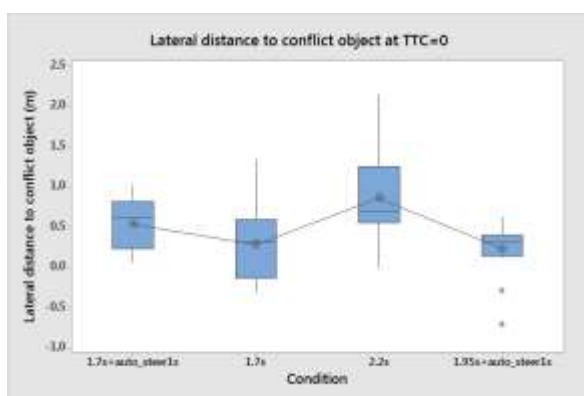


Figure 43: Lateral distance to conflict object at TTC=0. In this graph participants that only braked to avoid a collision were removed since they do not reach TTC=0.

A Pearson product-moment correlation coefficient was computed and showed that the position in the lane before the conflict object does not affect the total lateral deviation. Hence, participants move a similar distance laterally, regardless of their initial lateral position in lane, which means that the initial position in the lane

correlates to the lateral distance to the conflict object $r=0.31$, $n=63$, $p<.05$ and the maximal lateral deviation in lane, $r=0.29$, $n=63$, $p<.05$, as shown in Figure 44.

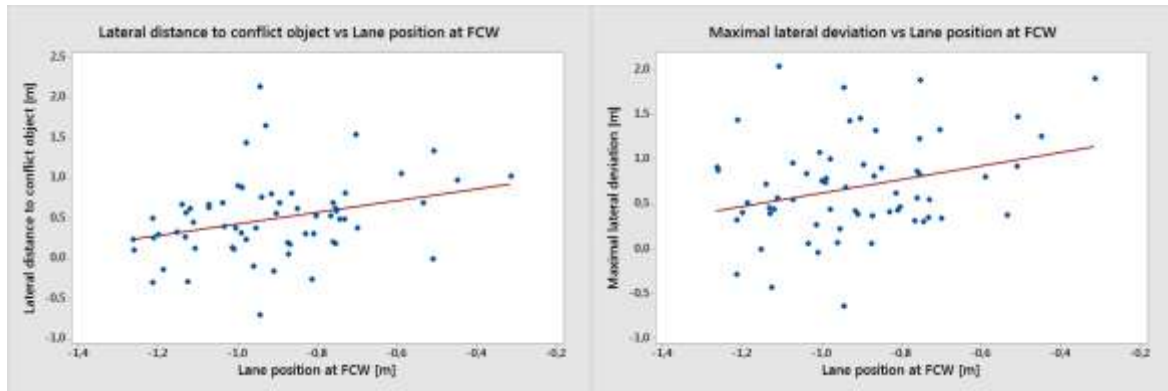


Figure 44: Scatterplots showing all participants except those to come to a standstill before the conflict object. Lane position is measured from the left front corner of the vehicle to the centre line (positive values are to the left)

A two sample T-test from the test track data showed no statistically significant difference in maximum lateral deviation (the amount of lateral shift between the time of the warning and the point in the interval 70 m before the conflict object to 30 m after it where the vehicle is at its leftmost position) between the group with warning-TTC=1.7 s + automatic steering at TTC=1 s ($M=0.86$, $SD=0.50$) and the group with warning-TTC=1.7 s without automatic steering ($M=0.67$, $SD=0.59$); $t(27)=0.98$, $p=0.34$. It thus seems that the presence or absence of steering assistance does not affect how much the vehicle moves laterally.

Greater lateral distance from the conflict object, which was observed in the group with warning-TTC=2.2 s (without automatic steering), also results in taking up more space in the oncoming lane, which can be an issue if there should be oncoming traffic. However, looking at the maximum position in the oncoming lane reveals that the other conditions are as far in the oncoming lane as this group, but at other points. A one-way ANOVA shows a statistically significant difference only between the group with warning-TTC=1.95 (with automatic steering) ($M=0.29$, $SD=0.42$) compared to the group with warning-TTC=1.7 s (with automatic steering) ($M=0.86$, $SD=0.50$) and the group with warning-TTC=2.2 s (without automatic steering) ($M=0.88$, $SD=0.51$) ($F(3,59)=5.62$, $p<.05$). This also shows that the group with warning-TTC=1.7 s with an automatic steering intervention does not steer further out into the oncoming lane compared to the group with the same warning but without the intervention ($M=0.67$, $SD=0.59$).

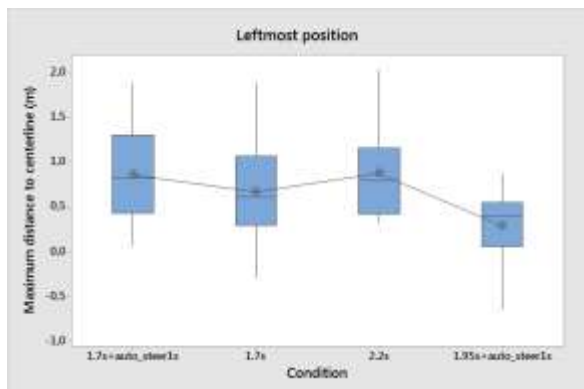


Figure 45: The furthest left position shows how far in the oncoming lane the vehicle is. It is measured from the vehicle's left side to the centre line. Positive values are to the left.

A two sample T-test showed no statistically significant difference in lateral distance to the conflict object (at TTC=0) between the two groups with the same warning-TTC (1.7 s). The group with automatic steering had slightly bigger distance to the conflict object (M=0.54, SD=0.32) compared to the one without automatic steering (M=0.28, SD=0.44). This result indicates that the automatic steering does not affect the lateral distance to the conflict object significantly but gives somewhat larger margins to the object. It seems the timing of the warning has larger effect: a two sample T-test showed statistically significant larger lateral distance to the conflict object (at TTC=0) the warning-TTC=2.2 s (M=0.86, SD=0.55) compared to the warning-TTC=1.7 s (M=0.28, SD=0.44); $t(29)=-3.82$, $p<0.005$. The group that received the warning earlier end up around half a meter further away from the conflict object in lateral distance when they pass it.

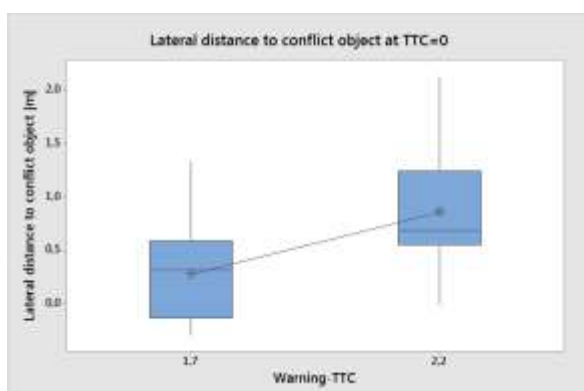


Figure 46: Comparison of lateral distance to the conflict object for the late and early warnings without steering assistance

Comparing the two groups that received only a warning and no intervention shows more collisions for a later warning (at TTC=1.7s) than for an early one (TTC=2.2s); only one crash in 19 participants was observed for the early warning group, while one third of participants collided in the late warning group (5/15), although this difference is not statistically significant.

Two groups differed only in the presence or absence of a steering intervention, and had the same late warning at a time to collision of 1.7 seconds. Comparing these two groups with a Chi-Square Goodness-of-Fit Test illustrates that there were significantly fewer collisions when the steering intervention was present ($\chi^2(1, N=5) = 5, p < .05$). There thus seems to be a trend towards a steering manoeuvre from the system being able to help in collision avoidance. However, when including lateral distance to the conflict object in the analysis, the steering intervention does not contribute as strong of an improvement as can be seen when comparing the warning times. As such, this suggests that the best thing in helping the driver to control a critical situation is to allow them more time to handle it.

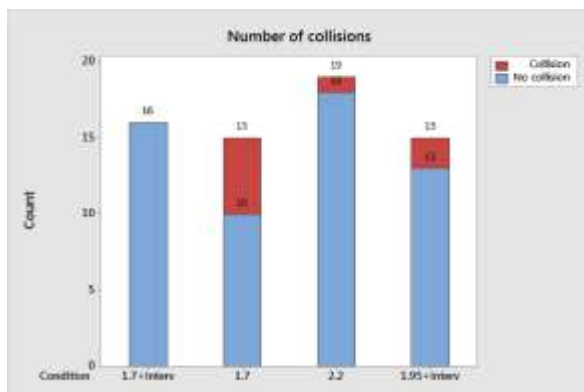


Figure 47: The number of collisions observed in the different conditions on the test track

6.7 RESULTS FROM CYCLIST TESTS WITH FCW W/WO AUTOMATIC STEERING INTERVENTION IN DRIVING SIMULATOR

6.7.1 Steering behaviour

The descriptive statistics, as shown in Table 10, shows that the two conditions encompassing only a warning are similar while the condition using the full intervention has slightly shorter distance to the target. All values could be considered within a safe passing distance to the bike target.

Table 10: Descriptive statistics for distance bike right shoulder at time-to-collision = 0[s]

Condition	<i>n(md)</i>	<i>M</i>	<i>SD</i>
B	18(0)	1.123	0.429
C	16(1)	1.433	0.712
D	13(2)	1.429	0.805

A one-way uncorrelated analysis of variance showed no significant effect of the warning strategy (headway and HMI) on the measurement of distance bike right shoulder at time-to-collision = 0 [s]: $F_{2,44} = 1.249, p = 0.297$.

As gathered from the analysis in the previous section, there was a difference in the average distance from the doll at $TTC = 0$ sec, between intervention and non-intervention trials for warning at $TTC = 1.7$ sec. For the given sample this effect is deemed non-significant, however the difference in maximum lateral displacement as seen in Table 11 indicate that condition B has a smaller value compared to the others.

Table 11: Descriptive statistics for maximum lateral displacement relatively to bike

Condition	<i>n</i>	<i>M</i>	<i>SD</i>
B	18	1.123	0.429
C	17	1.433	0.713
D	15	1.428	0.805

A one-way uncorrelated ANOVA show that there is a significant difference in lateral displacement: $F_{2,47} = 6.999$, $p = 0.002$. A Bonferroni multiple comparison show that the maximum lateral displacement is significantly smaller in condition B compared to C ($p = 0.003$). This can possibly be explained by how early the vehicle starts steering away from the threat. Figure 48 and figure 49 depict the steering angle in relation to the longitudinal position of the vehicle in the road. It is possible to see that when the intervention is on, figure 48, steering is initiated roughly 10m before the intervention off condition, figure 49. This period (0.5 to 0.7 sec) is characterized by a plateau on the steering angle which follows an initial positive slope. This is explained by a driver counter action which tries to mitigate the steering intervention torque.

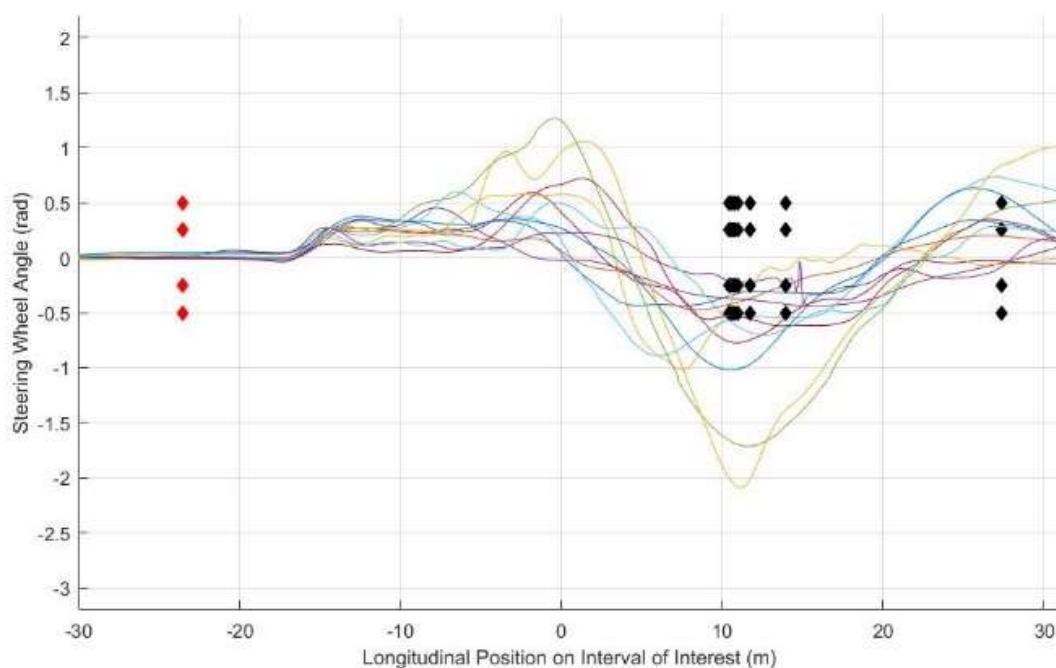


Figure 48: Steering angle in relation to the longitudinal position of the vehicle in the road with the intervention activated. The red diamonds mark FCW at $TTC=1.7$ s and the black diamonds mark the location of the bike at $TTC=0$ s.

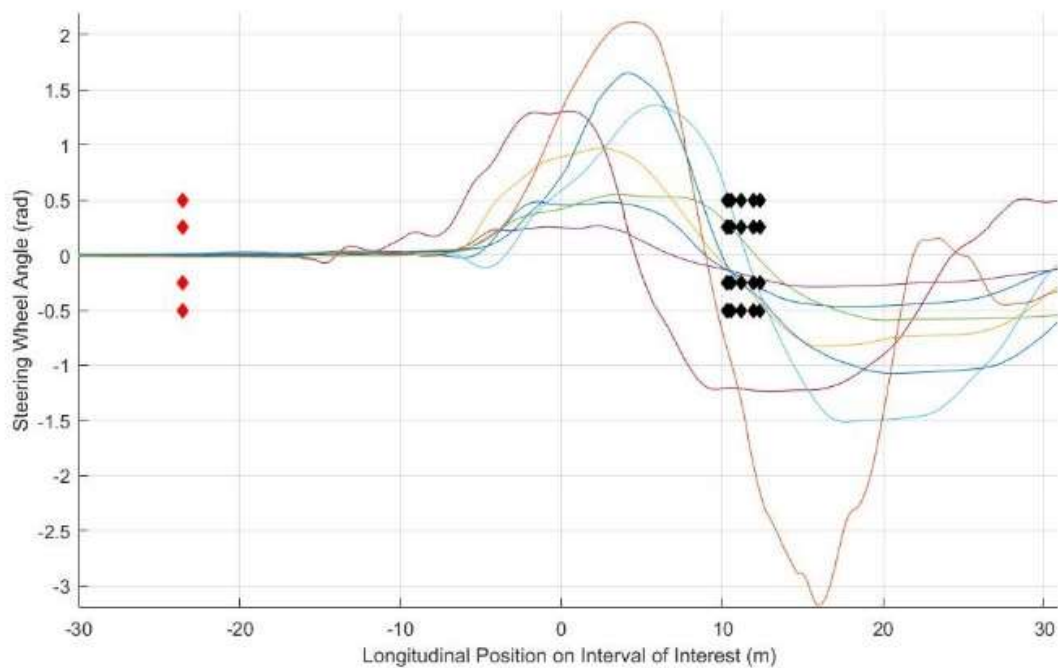


Figure 49: Steering angle in relation to the longitudinal position of the vehicle in the road with the intervention not activated.

While the plateau is active the vehicle is building a yaw angle relatively to the bike, figures 50 and 51, which places it in a more favorable trajectory when the driver takes over with a conscious steering action. This, we believe, explains why the maximum steering angle and maximum lateral displacement are higher in the no intervention case than in the intervention case where a more favorable initial trajectory required less correction efforts.

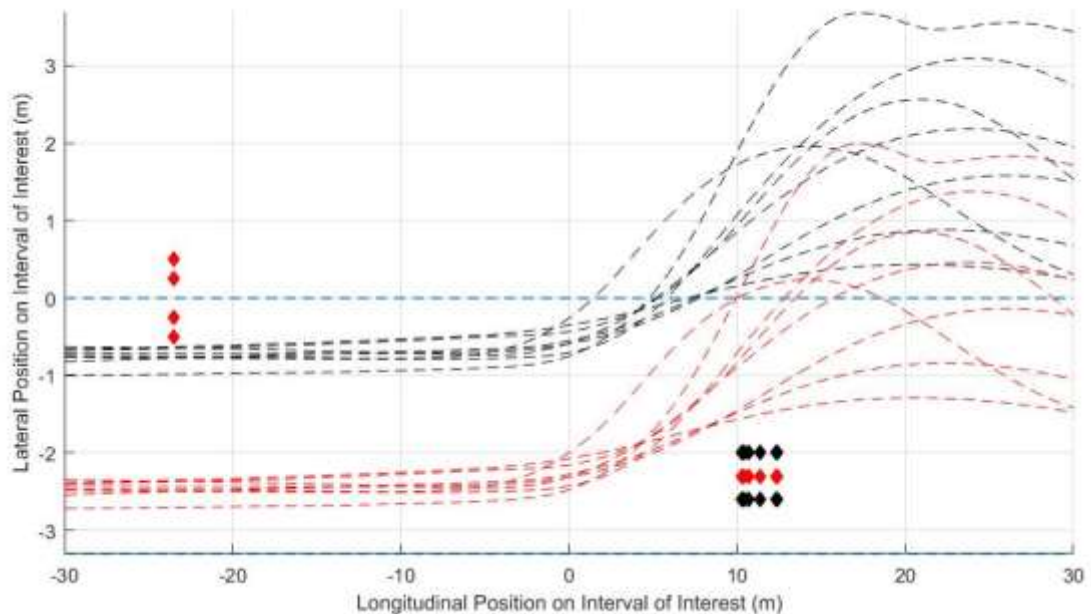


Figure 50: Lateral position relative to the bike with the intervention not activated

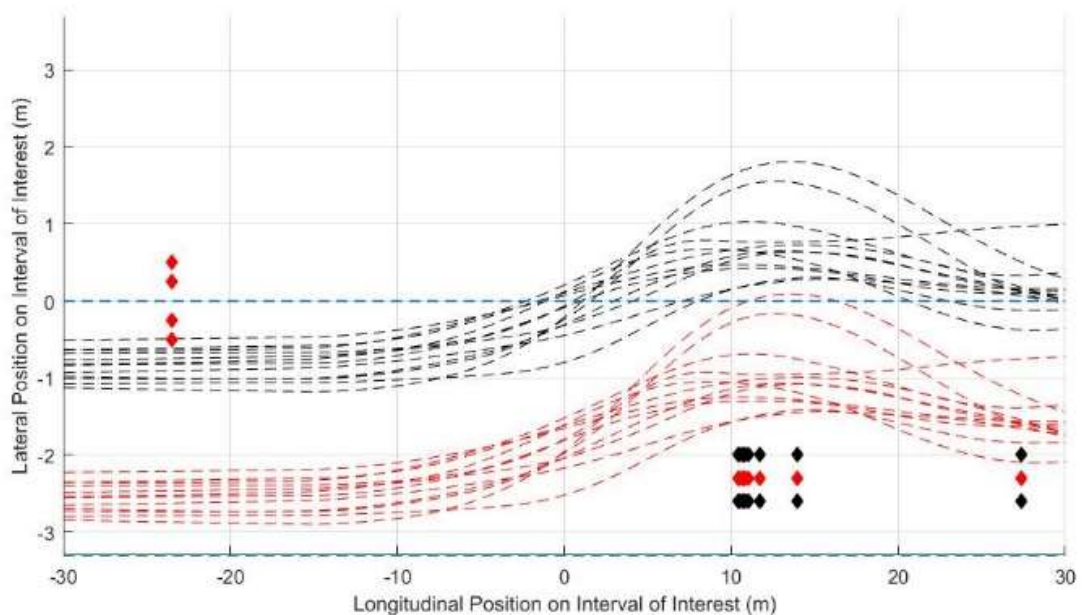


Figure 51: Lateral position relative to the bike with the intervention activated

In a bigger sample, if statistical significance could be attached to the difference between vehicle and doll at $TTC = 0$ sec, the same explanation could be used to motivate that difference; more control from the driver perspective may imply smaller margins.

6.7.2 Severity

The analysis of severity in terms of the amount of crashes per conditions revealed that there was no difference between the three groups as no crashes occurred in any of the conditions.

6.7.3 Acceptance evaluation

6.7.3.1 Criticality of the conflicts

The criticality of the scenarios was assessed with the CRITIC method (Common Risk awareness measurement method for Inter-population Comparisons) from Bellet and Banet (2012). Six dimensions, that allow assessing the criticality of the driving situation are rated from 1 to 5 (with 1 = not critical at all and 5 = highly critical).

The situations were judged significantly more critical when the participants met the stationary doll on their course than when it was the moving bike (Kruskal-Wallis, $H=6.025$, $p<.05$). Fear (Kruskal-Wallis, $H=4.257$, $p<.05$) and stress (Kruskal-Wallis, $H=6.407$, $p=.011$) were also judged significantly higher with the doll, while frequency of the event (Kruskal-Wallis, $H=4.405$, $p<.05$) and controllability of the situation (Kruskal-Wallis, $H=6.064$, $p<.05$) were judged significantly higher with the bike (Figure 52).

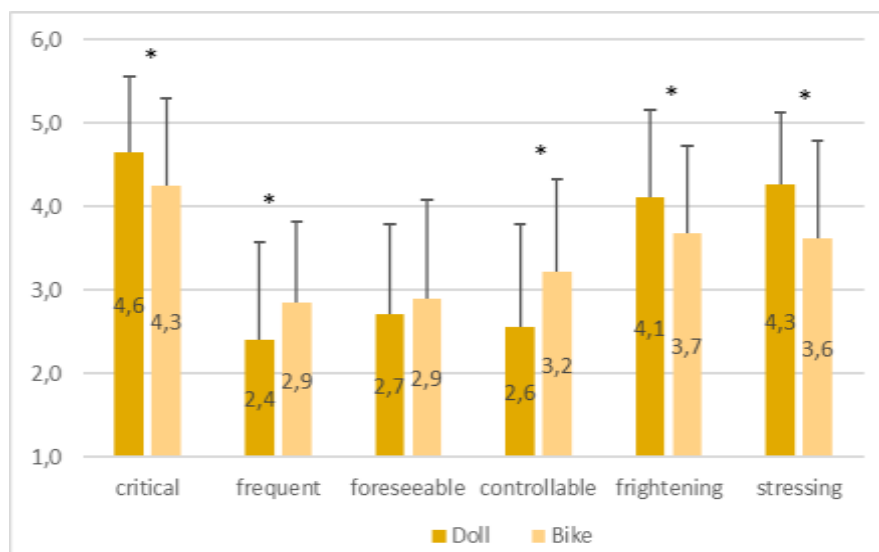


Figure 52: Criticality rating according to the object (doll versus bike).

The same dimensions were also compared according to the Time To Collision (1,7 s versus 2,2 s). No significant difference was found between the modalities except in terms of foreseeability. As expected, when the driver met the critical object with a TTC of 2,2 seconds, the situation was judged more foreseeable than with a TTC of 1.7 seconds (Kruskal-Wallis, $H=3,805$, $p=0,05$) (Figure 53).

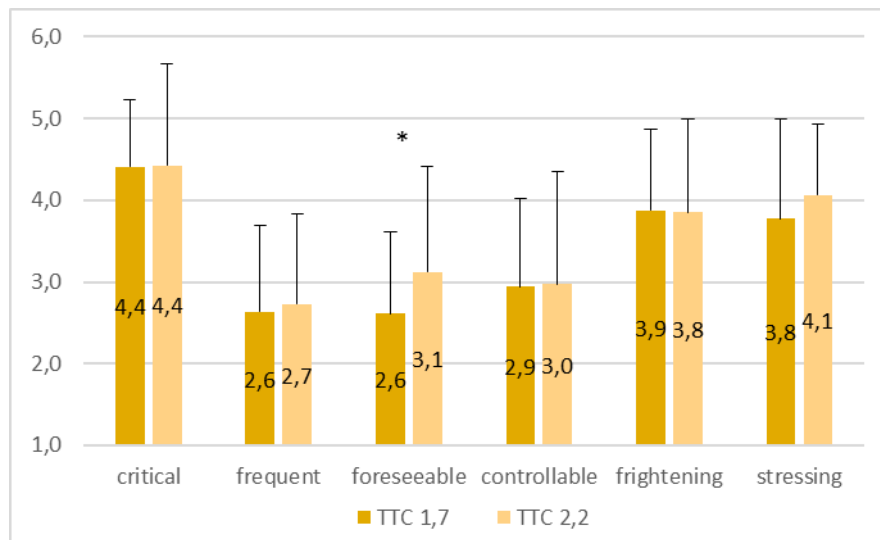


Figure 53: Criticality rating according to TTC.

No significant difference was found between warning and full steering intervention on the 6 investigated dimensions.

6.7.3.2 Acceptability

As described in the Deliverable D7.2 (Report on methodology for balancing user acceptance, robustness and performance) acceptability of the PROSPECT-like systems was measured after the experiment with the questionnaire from Van der Laan et al. (1997). Participants rated the system very positively and acceptability value is high (mean = 4,3 on a scale from 1 to 5) whatever the groups. (Figure 54).

No significant difference was observed depending on the object met (bike or doll), on the TTC (1,7 s or 2.2 s) or regarding the PROSPECT functionality (steering or warning).

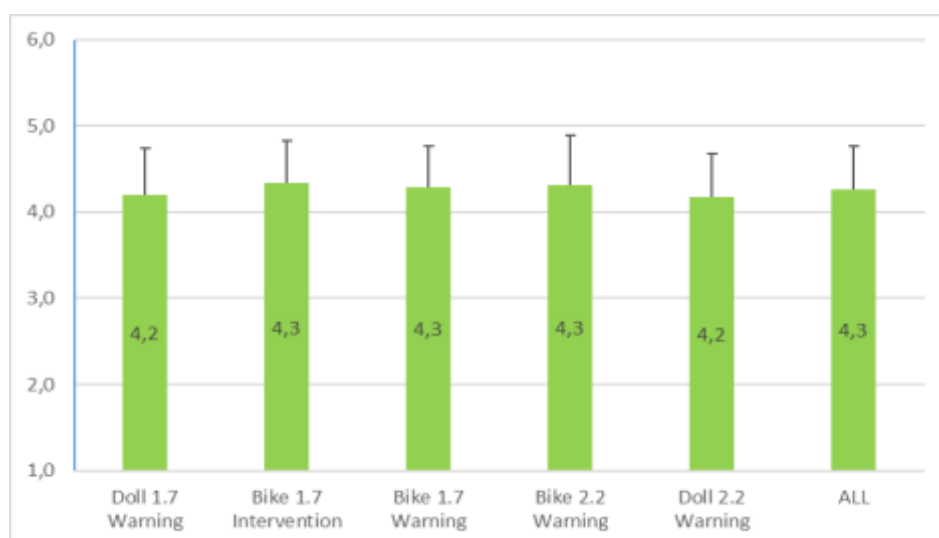


Figure 54: Acceptability of a PROSPECT-Like system according to the groups.

However, the acceptability value varied significantly according to the drivers' opinion on the help provided by the system. Only few drivers ($n=11$, 13% of the sample) considered that the system did not behave in a right manner while giving a warning or a steering (1 to 3 in each of the 5 groups). The results show that these drivers gave a significantly lower acceptability value than those who found that the system behaved in a right manner [Mann-Whitney, $U=176$, $p<.01$] (Figure 55).

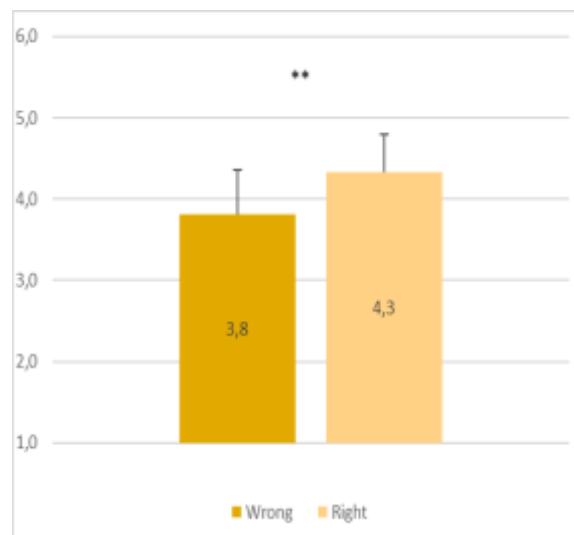


Figure 55: Acceptability of a PROSPECT-Like system according to system behaviour

6.7.3.3 Intention to use the system

Overall, participants reported a quite high likelihood of using the PROSPECT functionalities, which nevertheless differs significantly across the systems [$\chi^2(2)=43.659$, $p<0.001$]. They were significantly more positive towards the warning than for the braking [$t(72)=1.952$, $p=0.055$] or the steering [$t(72)=4.768$, $p<0.001$] and for the braking than for the steering [$t(69)=3.834$, $p<0.001$] (Figure 56).

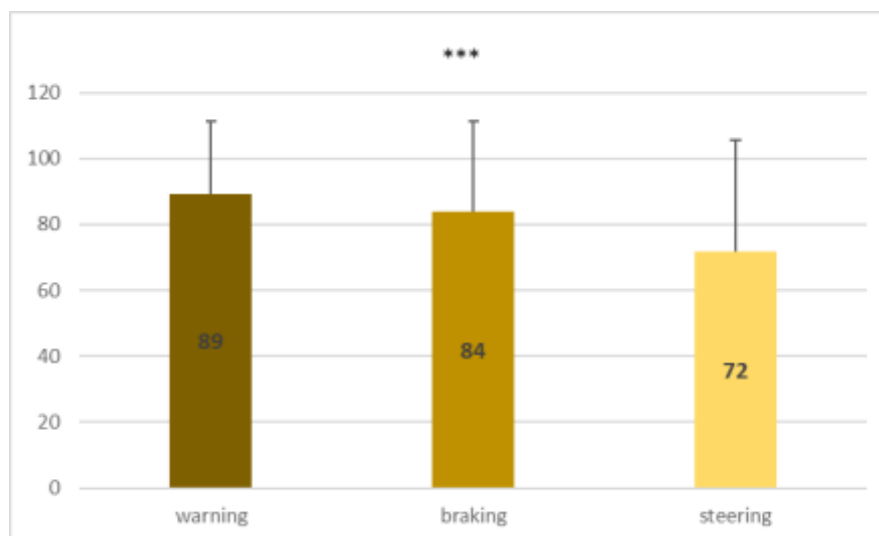


Figure 56: Intention to use the different functionalities.

6.7.3.4 Willingness to buy a PROSPECT-like system

Finally, the willingness to buy a PROSPECT-Like system was investigated. The willingness values are quite high and no significant difference is observed according to the groups (Figure 57).

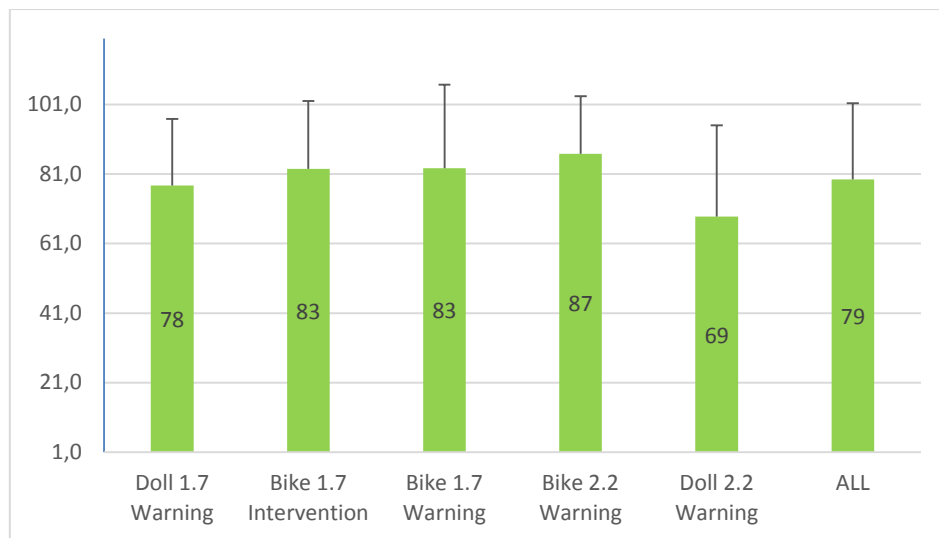


Figure 57: Willingness to buy a PROSPECT-Like system according to the groups.

No significant difference is found depending on the TTC (1,7 s or 2.2 s) or regarding the PROSPECT functionality (steering or warning). However, the willingness to buy differed according to the object met during the experiment (bike or doll). Those who met a bike gave significantly higher willingness to buy values than those who met a doll [Kruskal-Wallis, $H=4.904$, $p<.05$] (Figure 58).

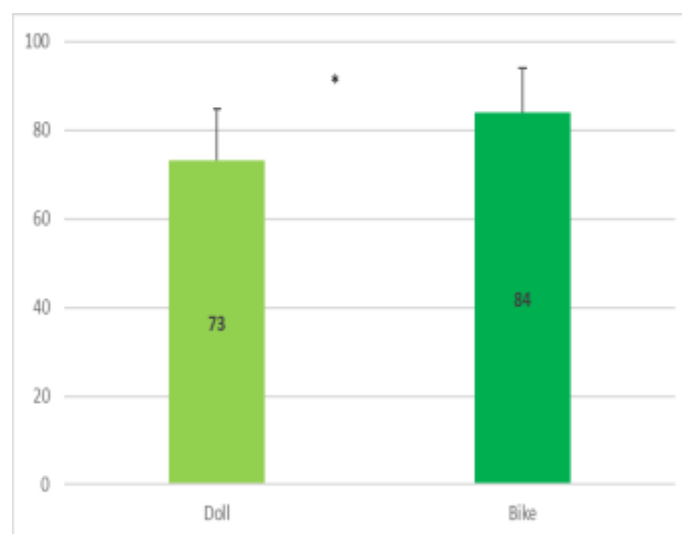


Figure 58: Willingness to buy a PROSPECT-Like system according to the objects met.

6.8 CONCLUSIONS

The analysis that was done to compare the data from test track and simulator showed that the steering reaction times did not significantly differ between the two different test environments, indicating that the scenario was perceived as similarly critical. However, the actions drivers performed to avoid a collision differed. On the test track, the majority of the drivers steered as first action and then almost half of the participants also braked after the first steering action. In the simulator, the majority of the drivers braked as a first action, and then steered. Only one participant steered as an only action. The speed when passing the conflict object was significantly lower in the simulator than on test track. Further studies are needed to conclude what the difference in collision avoidance actions depends on.

There was no statistically significant difference in number of collisions between the test environments. There were more collisions among the participants that got the late warning. Looking at the lateral distance to the conflict object, the distance was significantly larger when receiving the warning earlier (TTC=2.2 s) compared to later (TTC=1.7 s), showing the importance of earlier warnings. There was also a tendency towards larger distances in the simulator compared to test track. The earlier warning also gives the driver the possibility to brake to avoid a collision.

Pedestrian tests with automatic steering intervention at TTC=1.0 s in combination with a FCW at TTC=1.7 s resulted in fewer collisions than with only FCW (at TTC=1.7 s) on the test track. The results from the test track also show a tendency towards larger lateral distance to the conflict object in the group that got an automatic steering intervention compared to the group with the same warning time but without the intervention.

From the results of the driving simulator data comparing the three bicycle conditions (B,C,D) it can be concluded that there is no statistically significant difference between the warning and intervention strategy used when comparing the dependent variables of crashes and distance bike right should at TTC=0 s. However, equally important is that the added steering intervention did not impose any new safety hazards to the driver in the situation. It is therefore concluded that although a strategy using only FCW may seem sufficient a strategy also using an automatic steering intervention as well could add additional safety effects as they are done with less variation and in a more controlled way compared to a human driver without such support. The earlier automatic steering intervention at TTC=1.35 s in the cyclist scenario changed the direction and lateral position of the vehicle until the drivers own steering intervention, which causes the driver needs to make a smaller lateral movement to avoid a collision in the cyclist scenario.

7 EXPERIMENT #5 “FOCUS GROUP AND VIDEO-BASED EXPERIMENT”

7.1 GLOBAL RESEARCH DESIGN

IFSTTAR conducted two experiments to assess “*a priori*” acceptability of the PROSPECT systems. The Acceptance Methodology described in chapter 2.3, which is based on existing theories and on scientifically validated acceptance model items, was applied with the three questionnaires (Before, During and After) for the video-based experiment and the questionnaire After only for the focus groups.

The main objective of the video-based experiment was to assess the risk awareness of the drivers and their acceptability of the PROSPECT functionalities. To do so, the participants were immersed with video recordings from the point of view of the driver showing conflicts between a car and VRUs and the activation of the PROSPECT systems.

In addition, focus group experiments were carried out to obtain in-depth and richer information than the questionnaire items. Indeed, the main purpose of focus groups is to collect the participants’ opinion on a given subject. They allow participants to express and compare their points of view with other members of the group and are a means that has proven to be efficient and effective in assessing the acceptability of the new technologies. According to Horberry et al. (2014), “abstract concepts such as the perceived effectiveness and usefulness might be better served by focus groups”; such concepts being important components of the acceptance evaluation.

7.2 PARTICIPANTS

The protocols have been submitted to the IFSTTAR internal ethics committee, which stated that this research does not fall under the French public health code and does not require consulting a committee for the protection of persons. The information on the experiments was given in advance to each participant and all signed a consent form.

The participants were recruited by a recruitment agency and received 60 € in compensation for the video experiment and 80 € for the focus groups that lasted longer. Both experiments were carried out at IFSTTAR in Lyon, France in April-May 2018.

The sample consisted of drivers who regularly drive. Three specific age groups were considered, intermediate aged-drivers being excluded. The driving experience was also taken into account:

- young novice drivers (18 to 25 years), who are allowed to drive but do not have yet their full driving license. French drivers generally obtain their full license after 3 years of driving (2 under certain circumstances),
- middle-aged drivers (30 to 50 years), who have been driving for more than 10 years
- elderly drivers (over 70 years), who still drive regularly

Additionally, middle-aged drivers were divided into 2 groups according to their car equipment in terms of Advanced Driver-Assistance Systems (ADAS) and their habit of using them (drivers who often use ADAS and drivers with little or no ADAS). The full sample of our population is shown in Table 5.

63 people participated in the video-based experiment aged from 19 to 85 years, including 22 men and 41 women. The sample consisted of 15 young drivers, 17 middle-aged drivers familiar with ADAS, 16 middle-aged drivers not familiar with ADAS and 15 elderly drivers.

Starting from a goal of 24 people, the focus groups were made of 21 participants aged from 18 to 72 years old, including 9 women and 12 men. They were separated in 3 groups homogeneous by their driving experiences and age: 8 young drivers, 8 middle-aged drivers and 5 elderly drivers. Due to the difficulty of recruiting drivers over the age of 70, only 5 elderly drivers participated.

Table 5: Demographic details.

Categories	Video-based experiment	Focus groups
Young drivers	15 participants 10 F, 5 M, mean = 21 years old	8 participants 2 F, 6 M, mean = 21 years old
Middle-aged drivers without ADAS	16 participants 13 F, 3 M, mean = 42 years old	4 participants 2 F, 2 M, mean = 45 years old
Middle-aged drivers with ADAS	17 participants 9 F, 8 M, mean = 39 years old	4 participants 2 F, 2 M, mean = 42 years old
Elderly drivers	15 participants 9 F, 6 M, mean = 75 years old	5 participants 1 F, 4 M, mean = 76 years old
Total	63 participants 41 F, 22 M	21 participants 7 F, 14 M

7.3 FOCUS GROUPS

7.3.1 Procedure

Each focus group lasted around 2.5 hours and was structured in five parts, which were always conducted in the same order:

1. **General Information:** in order to start the discussion and break the ice, the participants introduced themselves and explained the features of their vehicle and their daily path. They also specified if they use ADAS during their drive often and regularly.

2. **Interaction VRU – Drivers:** the participants explained how they feel towards VRU interactions on the road. Videos showing conflict situations were presented, so that participants could give more in-depth opinions.
3. **PROSPECT Systems Presentation:** PROSPECT systems were presented and explained by the moderator, illustrated by videos. In order to get their first impressions, the participants discussed then the advantages and disadvantages of such systems. Keywords were noted directly on a board in plain view with the aim to bounce back to the terms they used throughout the focus group.
4. **Acceptability:** in this part, participants gave their opinion on their acceptability of a PROSPECT-like system, what do they think about it and if they would use them and buy them. Questions from the UTAUT model (“Unified theory of acceptance and use of technology” of Venkatesh et al., 2003) were also added to point out social acceptance. In this model, user intentions to use a system is explained by four key constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions.
5. **Performances:** finally, a discussion was engaged about the performance such system should meet and how they would accept false alarms.

The discussions were transcribed from the videos using the "Sonal" software, which allowed us to segment the video clips, assign them themes, and identify the participants. The qualitative analyses were made with a thematic content analysis. This analysis made it possible to identify the strong ideas that are shared but also those which oppose the participants. The main lines of thought that emerge around the acceptability of the PROSPECT system were then extracted.

7.3.2 Materials

The focus groups were held in a dedicated room and were audio and video recorded. A video projector and a laptop showed some videos to the participants in order to explain the PROSPECT system functioning and also to provide examples of situations where the system would be useful. Questionnaires were distributed to the participants to retrieve their general information at the end of the sessions, and to collect data about the PROSPECT system acceptability.



Figure 28: Focus group room.

Videos recorded at IDIADA test track within T7.1 have been used to explain how the PROSPECT systems work. Participants could see real test situations and thus better immersed themselves in the emergency braking and the emergency steering functions. One video shows a crossing scenario, where a cyclist-dummy crosses from the left. The car performs an evasive steering manoeuvre to avoid the accident (Figure 29).



Figure 29 a & b: Cyclist crossing and steering maneuver.

Other videos were also provided to explain the PROSPECT functionalities (alarm, emergency braking and emergency steering) showing the different phases from the VRU detection to the activation of the system (Figure 30). In the following situation, the system a) detects a VRU that could collide with the vehicle and triggers an alarm b) looks for available space for automatic steering while anticipating VRU movement, c) engages an automatic emergency braking and d) engages an automatic steering as enough space is available (from: <https://www.youtube.com/watch?v=zAeEnLr3WYk>).



b)



a)

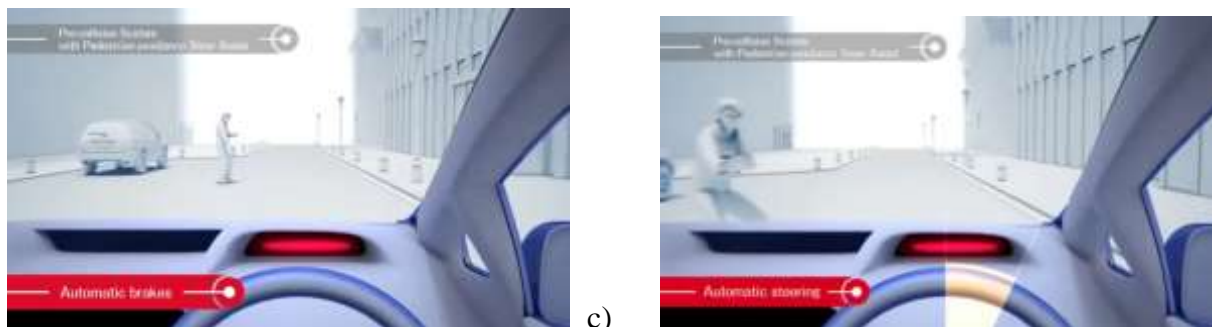


Figure 30: System functioning from VRU detection (a) to steering activation (d).

7.3.3 Results

7.3.3.1 Interactions Drivers – VRUs

The strongest idea stated by the participants is that they have difficulties perceiving VRUs and anticipating their behaviour when driving. Each category of participants recalled situations where pedestrians hidden by a car crossed without paying attention to the traffic (*“Behind a car when they pass, you do not see them”*). Some participants also spoke about pedestrians who get off the bus and cross without looking around. They also evoked difficulty perceiving cyclists in blind spots.

Another strong idea was that vulnerable road users would not respect the traffic rules: “they are *“dangerous”*, *“unruly”* and *“unconscious”*. Cyclists are described as being even more dangerous and difficult to anticipate than pedestrians because of their higher speed when riding. On the other hand, pedestrians were considered as road users that behave unpredictably and are less visible to the drivers. In general, the participants thought that they have to adapt themselves to vulnerable users to avoid accidents and that they must be *“wary”* of them. Participants of all ages declared themselves very attentive to children by the roadside, when they pass near a school or simply when they see them.

Regarding the development of bike paths, some drivers considered that they are well designed for cyclists, but that cyclists do not respect them. On the other hand, others (drivers who are also regular cyclists) found that the road improvements are not sufficiently adapted, and that this forces them to disregard the rules of the road, and thus to create conflict situations.

7.3.3.2 Advantages and disadvantages of PROSPECT-like systems

After the PROSPECT functionalities have been explained, the participants gave some first impressions:

Advantages:

- Safety: all participants first mentioned the safety provided by such a system. Middle-aged drivers pointed out that it would reduce accidents, improve their protection and especially those of others, while elderly drivers said that the system would allow an earlier and faster reaction for the driver.

- Performances: young drivers declared that the system would provide a better “*vision*” of the danger. For the middle-aged and elderly drivers, the detection of danger would be faster. Furthermore, elderly drivers added that such systems could be more effective in case of visibility problems (sun, rain, fog, etc.).
- Peace of mind: all participants agreed that a PROSPECT-like system adds peace of mind to the driving, because they would be less afraid of overrunning a VRU. Middle-aged drivers evoked the psychological impact of an accident that could be avoided.
- Assurance: elderly drivers imagined that such a system could lower insurance premiums.
- The elderly: middle-aged drivers thought it could be very helpful for their parents.

Disadvantages:

- Price: all the participants imagined that a PROSPECT-like system would greatly increase the price of their car.
- Concentration: young drivers declared that having such a system could lead to reduce their vigilance and attention to the road. Middle-aged drivers also stated that they would be less attentive. For the elderly drivers there would even be a risk of drowsiness. This was considered as a major drawback of the systems.
- Car change: middle-aged and elderly drivers raised the problem of car change. Indeed, some people drive several cars and could forget that the car driven is not equipped by the system.
- Electronic addition: elderly drivers raised the problem of adding electronics that could lead to more maintenance and increases the risk of malfunction.

However, despite the many disadvantages evoked by the participants, they generally believed that the benefits would be more important than these disadvantages.

7.3.3.3 Acceptability

The discussion guide of focus groups was based on the UTAUT model developed by Venkatesh & al. (2003) and adapted by Adell & al. (2014) to driver support systems. Expected effort, which corresponds to the degree of facility associated with the use of a system was not discussed, because the drivers did not fully experience the systems. A section about trust has been added instead.

The four following dimensions were investigated:

Expected performance (perceived usefulness): “degree to which an individual believes that the use of a system can help him achieve a profit in his performance” (Venkatesh & al., 2003).

All participants were of the same opinion, i.e. a PROSPECT-like system would be very useful because it improves the road safety and reduces the number of accidents. For some young drivers, once in the market, the system will become indispensable. For others, the system does not seem 100% useful since it is possible nowadays to drive without. The majority of middle-aged and elderly drivers shared the opinion that the system would increase the road safety and provide “*peace of*”

mind". Middle-aged drivers possessing ADAS added that they would feel "safer" and that the system would bring "*fewer accidents and fewer injuries*". Elderly drivers recognized making sometimes driving errors ("*It's an asset, because after all you can always be ineffective*"). For these drivers, everyone should feel safer with such a system.

On the other hand, the participants mentioned some limitations. For example, the interventions of the system could generate other accidents while avoiding a vulnerable road user. Participants mentioned that emergency braking could result in a rear-end collision. Middle-aged drivers were worried that the system triggers too often with too many alarms, too many emergency brakings or too many automatic avoidance manoeuvres ("*I would be afraid the system anticipates too much!*"). Audible or visual alarms could also be annoying for the middle-aged drivers due to information overload, while the young drivers feared to be surprised by alarms. Elderly drivers did not worry about the alarm function and found it more "*reliable*" than the other functions (braking and steering).

Trust is defined "as an antecedent to reliant behaviour, a willingness to accept vulnerability in expectation of a positive outcome" (Miller & al., 2016).

Overall, the participants declared that they would trust this system. More specifically, the middle-aged drivers possessing ADAS linked their opinion about the PROSPECT system to their confidence in their car equipment and their experience of it ("*I trust my equipment because for three years I have never had a problem, so for me it would be 100% reliable.*"). However, drivers who do not use ADAS raised the problem of a dysfunctional cruise control, and declared that this could also happen with PROSPECT-like systems. For the majority of middle-aged drivers, a good experience of use would increase their trust in the system ("*After if we see in the long run that it works great, obviously it will be widely accepted*"). Young drivers shared the same opinion ("*Because as long as I have not tested these systems and I have not seen the results from my own experience, I think I will always be very careful*"). Young drivers stated that they would need visual evidence or statistics. One elderly driver, more pragmatic, argued that no technology is 100% reliable ("*We don't need to be defeatist, but it is reasonable to say that it will not be 100% reliable.*"). Another one recalled his long experience with technology and declared he could trust this technology ("*We saw the birth of computers, see where they are today, so we trust.*").

However, some fears were expressed regarding the risk that too much trust in the system could generate. In each group, drivers declared that they could become overconfident in the system (Young drivers: "*we rely on our safety*", "*I'll be less careful*"; middle-aged: "*we become less attentive because we know that the machine can all manage*"; and elderly drivers "*a loss of concentration. I am less focused, the system is there*"). Some drivers (young and middle-aged) added that owning and using these systems could make them less engaged in the driving activity and lead to less attentive and more passive driving, which could be a major drawback of the systems ("*the goal of the system is to help, not to make us less vigilant.*"). This feeling is not shared by the elderly drivers who do not think that such systems would have an impact on their driving behaviour, evoking their long driving experience.

Social influence: “degree to which an individual thinks that the use of the system will enable him to satisfy the people and the significant norms of his home group while giving him a positive image” (Venkatesh & al., 2003).

Participants declared that their entourage would recommend using the systems. Advice from their groups seemed to be mandatory because if no one talks about it, the participants will not be willing to have it. Middle-aged drivers declared needing “*to inquire people who have already used it*”. They stated that elderly people like their grandparents or people with disabilities would need the system more. The young drivers also declared they need other opinions from their entourage like family or “*word of mouth*”. One of them specified however that the experience of use is more important and added that he would not recommend the system without having experienced it. Elderly drivers thought that their children would encourage them to buy a PROSPECT-like system due to their age-related difficulties in the driving.

Intention to use: “degree to which an individual thinks that he intends to use the system” (Davis, 1989).

Overall, participants preferred the alarm function. Young drivers were not homogeneous in the intention to use alarms, with some of them thinking that it is an important function and others that this is an additional noise nuisance because of all the alarms already present in vehicles. For middle-aged drivers not familiar with ADAS, the alarm function should be sufficient (“*For example, simply activate a signal without emergency braking*”), while middle-aged drivers with ADAS would be ready to activate all PROSPECT functionalities. Elderly drivers also preferred the alarm function only because they cannot replace it, while they said being able to brake or steer by themselves.

Doubts were evoked on the use of the emergency braking function. Elderly drivers who trust themselves would rather not activate and stated that the alarm function should be sufficient by giving time to brake by themselves. Middle-aged drivers possessing ADAS were more confident in the technology and considered that emergency braking is useful, especially when aware of being not always attentive to the road.

Even more than for braking function, participants were generally very suspicious of the steering function. Only middle-aged drivers possessing ADAS declared that they would intent to use it. Most of the participants stated that the emergency steering could lead to more accidents than it could save.

Regarding usage context, participants were generally more ready to use all functionalities in the city. One middle-aged driver possessing ADAS recognized that the system would be essential in this environment. Elderly drivers were generally of this opinion, pointing out that the higher number of vulnerable road users is encountered in the city. Young drivers declared that they would not like to be able to activate / deactivate the different functions, because they would be afraid of forgetting to activate the system back when necessary. They would prefer using this system in city only, unlike other drivers with longer driving experience (middle-aged and elderly drivers), who found the systems equally useful in cities and rural areas (“*There are rural roads where there is activity*”).

7.3.3.4 Willingness to buy a PROSPECT-like system

Participants stated generally they would be likely to buy PROSPECT systems; the main limitations being trust and price. Participants imagined that such systems would significantly increase the price of the car. This problem was especially raised among the young drivers still students. When price is not considered, the willingness to buy increases significantly (*“Without taking into account the price, there is no problem”*). However, one single middle-aged driver using ADAS said being ready to buy such systems even though the price was so high that it would be necessary to cancel another comfort option (*“I think I might not take a leather seat, I'll take off an option but that I think I would take it, because it's a driving aid of quality, which actually reassures”*).

Trust also played an important role in the willingness to buy PROSPECT-like systems, especially for some young drivers. For them, being confident in the systems is mandatory (*“For me it's not because of the price, I do not have complete confidence in it”*).

7.3.3.5 Acceptable performance of the system

For most participants any system failure would be totally unacceptable. The false positives (activation in a non-critical situation) were especially pointed out as unacceptable situations. For the participants, they would render the systems useless and even dangerous: too many alarms could become annoying, while emergency braking or steering could lead to other accidents. Some participants also pointed out that false positives could decrease the drivers' confidence in the system.

False negatives (no activation in a situation where a collision is imminent) were accepted even less. However, a margin of error seemed to be tolerated by some drivers who considered that no technology is 100% reliable.

The fact that an activation could be different from that foreseen by the driver appeared to be more acceptable than the two previous situations to the condition that the collision is avoided. In the case of imminent collision situations, some participants recognized that they could be not as good as the system.

7.4 VIDEO-BASED EXPERIMENT

The video-based experiment was run in order to investigate the driver's acceptability of the PROSPECT functions. Drivers were faced with a series of 20 videos from a driver point of view covering various use cases derived from accident scenarios. These videos were selected from the naturalistic observations conducted within WP2 in Barcelona by IDIADA. Each video showed a conflict between a car and a cyclist or a pedestrian, and froze when the situation becomes critical. A message informed then the driver about the action the PROSPECT system would have taken (warns the driver or brakes).

The three questionnaires “Before”, “During” and “After” described in the Deliverable D7.2 (Report on methodology for balancing user acceptance, robustness and performance) were used at different times of the experiment to assess the PROSPECT function acceptability.

7.4.1 Procedure

Each session, lasting around 1 hour, was organized with the following protocol:

1. **Reception and welcome of participants:** participants were provided with a pen and the questionnaires prepared in the order of the session.
2. **General information:** participants completed the consent form and started filling in the first part of the "before" questionnaire (participants' information).
3. **Procedure and instructions:** Before the experiment itself, the same videos as for the focus groups were shown to explain the PROSPECT system functioning. Then two other videos were displayed as examples to explain how to fill the "during" questionnaire.
4. **Questionnaire "Before":** the participants completed the questionnaire "Before" and evaluated their prior acceptability of the PROSPECT functions (*a priori* acceptability).
5. **Watching videos and Questionnaire "During":** participants completed the questionnaire "During" after each video shown.
6. **Questionnaire "After":** after viewing the 20 videos, the participants completed the questionnaire "After" (acceptability of the system).
7. **End of the experiment and data recovery:** participants were thanked for their participation and all completed questionnaires were collected.

7.4.2 Materials

The participants sat at around 2 meters in front of a screen (1.8 m wide * 1 m high) where the videos were shown. A paper-board was installed in the room to explain the course of the session. Up to 8 participants could attend one session at the same time.



Figure 31: Positioning of the experimental room.

20 videos were selected among the large dataset collected by IDIADA during their naturalistic observations in Barcelona. Each video showed a conflict from the driver's

point of view between the car driven and a cyclist (5 cases) or a pedestrian (15 cases), and froze when the situation became critical (Figure 32).

- In 8 cases the flashing word 'ALERT' at the bottom of the screen informed the participant that the PROSPECT system would give a warning in this situation (Figure 32). All of them were showing a conflict with a pedestrian.
- In the 12 other scenarios, the message informed the participant that the system would activate an emergency braking (Figure 33), 7 of which were showing a conflict with a pedestrian while the remaining 5 showed a conflict with a cyclist.

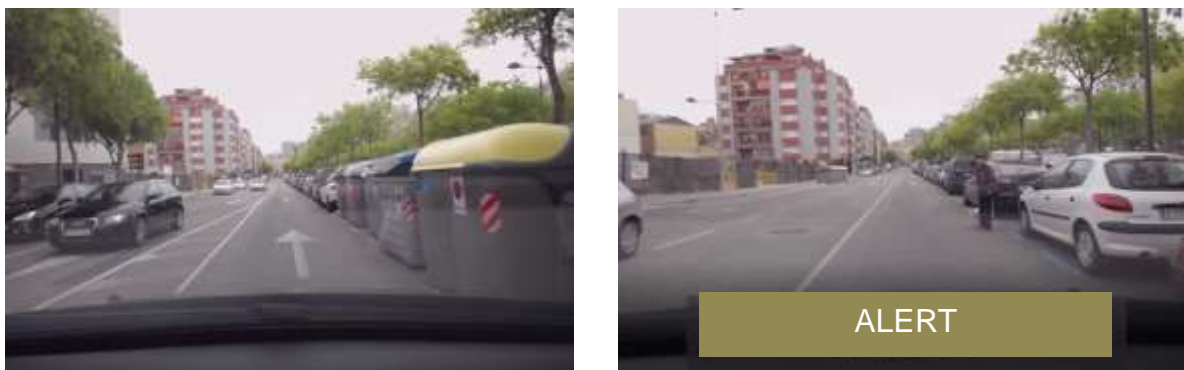


Figure 32: Driver view and warning given.

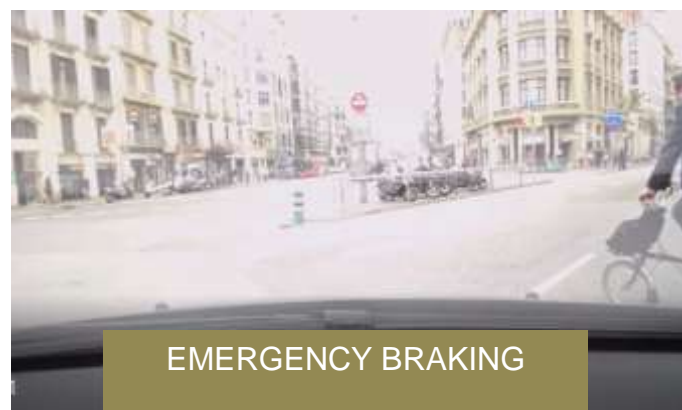


Figure 33: Activation of emergency braking.

Among these videos, different use cases to be implemented in the demonstrators (deliverable D3.2: Specification of the PROSPECT demonstrators) were investigated:

- UC_DEM_2: vehicle turns right and traffic in same direction (2 videos),
- UC_DEM_4: crossing situation from left (1 video),
- UC_DEM_10: pedestrian crosses from right (3 videos),
- UC_DEM_11: pedestrian crosses from right, obscured (2 videos).

The other videos showed scenarios that were identified as critical scenarios that could lead to accidents. These scenarios came from the deliverable D3.1 – The addressed VRU scenarios within PROSPECT and associated test catalogue:

- UC_CY_T1_U: vehicle turns to left and traffic is in same direction (2 videos),
- UC_PD_2: crossing a straight road from off-side (2 videos),
- UC_PD_3a: crossing at a junction from the near-side; vehicle turning across traffic (1 video),
- UC_PD_4a: crossing at a junction from the near-side; vehicle not turning across traffic (1 video),
- UC_PD_6: crossing a straight road from off-side; with obstruction (3 videos),
- UC_PD_7a: along the carriageway on a straight road away from vehicle (2 videos),
- UC_PD_7b: along the carriageway on a straight road towards vehicle (1 video).

The order of the videos was randomized and different for each test session.

7.4.3 Results

7.4.3.1 Criticality of the conflicts

To evaluate the criticality of the conflicts shown among the 20 videos, the CRITIC method (Common Risk awareness measurement meThod for Inter-population Comparisons) from Bellet and Banet, (2012) is used. This method investigates the feeling of the drivers on the situations they have been faced with regarding various dimensions. Six of these dimensions are selected in this study. The criticality scale allows for assessing the criticality of the driving situation. This is a subjective assessment of the situational risk, which requires the perception of the critical event and the evaluation of the danger of this threat. The predictability of the situation is assessed in terms of frequency (the type of situation is frequent or not) and its foreseeability (such an event is foreseeable or unexpected). The situational controllability refers to the participants' assessment of their abilities to adequately manage such a situation. Finally, the participants' emotional feeling allows for taking into account the effect of the situation on the driver in terms of fear and stress.

The criticality of the conflicts is judged significantly higher in situations where an emergency braking is shown than for those where a warning is shown [$F(1,18)=9.78$, $p=.006$]. These situations are also considered significantly less controllable [$F(1,18)=13.696$, $p=.002$], more frightening [$F(1,18)=7.198$, $p=.015$] and stressful [$F(1,18)=10.142$, $p=.005$] (Figure 34).

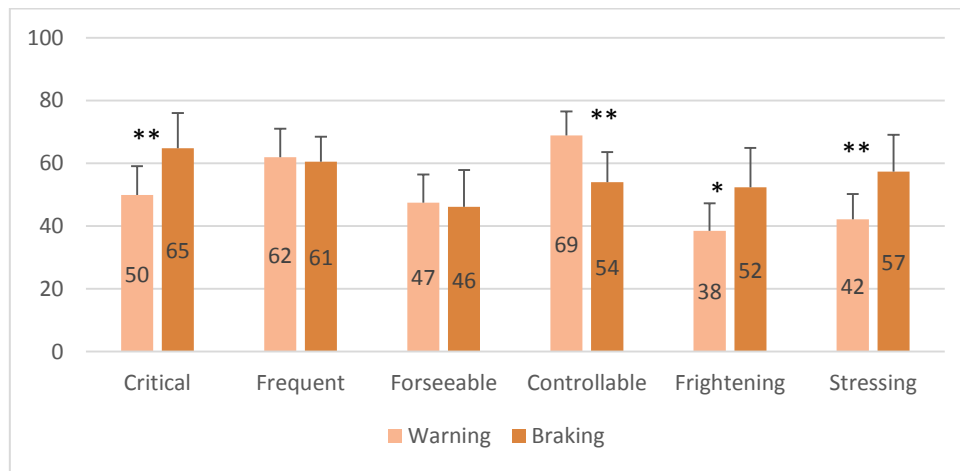


Figure 34: Video rating according to the PROSPECT system action¹.

No significant difference is found according to the type of VRU: the risk assessment does not vary whatever the VRU is a pedestrian or a cyclist. However, when the VRU is coming from the right the situation tends to be judged more critical [$F(1,18)=3.424$, $p=.081$] and is considered significantly less frequent [$F(1,18)=5.412$, $p=.032$] than when they come from the left (Figure 35).

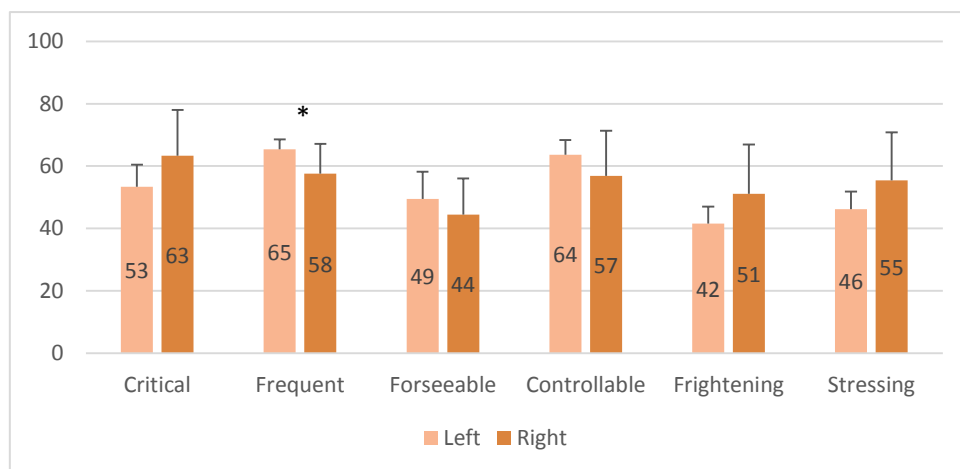


Figure 35: Video rating according to the VRU localization².

These dimensions do not vary according to the age and mileage of the participants. However a gender effect is shown, women rating significantly higher the fear [$F(1,61)=8.489$, $p=.005$] and stress [$F(1,61)=8.341$, $p=.005$] generated by the situations on the videos than the men; while the men tend to rate more controllable the situations than the women [$F(1,61)=3.820$, $p=.055$] (Figure 36).

¹ Mean values given by the 63 participants to the videos showing a warning or a braking (rated from 0 to 100).

² Mean values given by the 63 participants to the videos showing a VRU coming from the left or from the right (rated from 0 to 100)

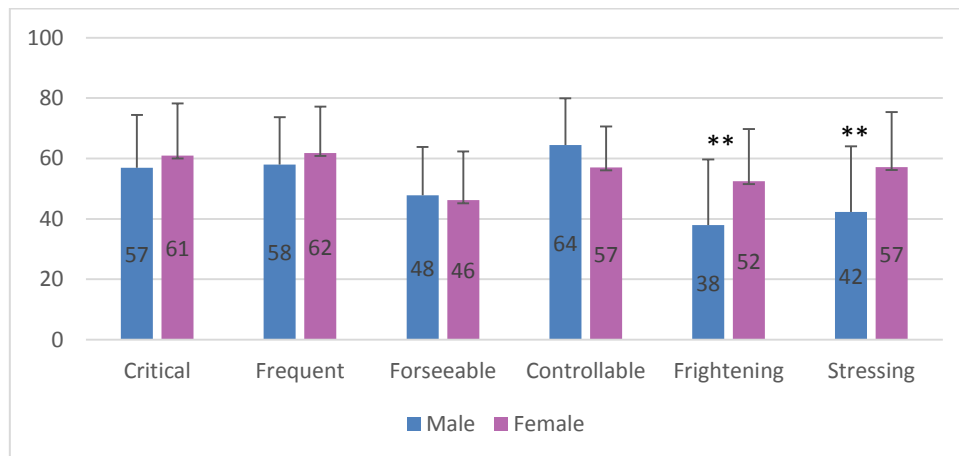


Figure 36: Criticality rating according to the gender³

7.4.3.2 Acceptability and trust

Acceptability is measured through different types of items and questions derived from Van der Laan et al. (1997) questionnaire. The main dimensions examined here are the “perceived usefulness”, the “perceived ease of use” and the “perceived satisfaction” which, when combined, provide an overall measure of acceptability.

Each dimension was examined before the experiment, i.e. before having seen the videos simulating an action of a PROSPECT-like system and at the end of the experiment, i.e. after having somehow experienced the system. Participants rated the system very positively and acceptability values are high (81 & 82 on a scale from 0 to 100) in both cases. No significant difference is shown between the investigated dimensions of acceptability before and after the experiment (Figure 37).

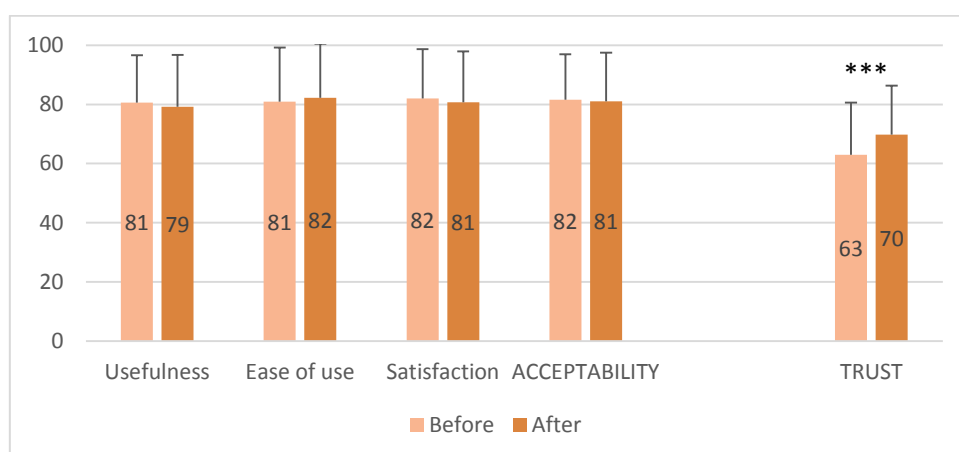


Figure 37: Acceptability and Trust before and after the experiment⁴.

³ Mean values given to the 20 videos by the males and the females (rated from 0 to 100)

⁴ Acceptability and trust measured on a scale from 0 to 100

Trust in the system is also an important aspect of the acceptance evaluation, which is evaluated using the 12 items of the Jian et al. (2000) scale. Trust is calculated by combining trust and distrust items in reversing distrust scores where relevant (Spain & al., 2008). It is important to note that the trust value is much lower than the acceptability values. However, trust significantly increases [$t(62)=-4.137$, $p<.001$] when participants have a better knowledge on the PROSPECT functionalities and especially after experiencing situations where the systems could help them to manage conflicts with VRUs.

No significant difference is found according to the age and gender of the participants. However, participants who drive less than 10 000 km/year give significantly lower acceptance values than those who drive more (mean = 77 vs 85) [Mann-Whitney, $U=347$, $p=.04$].

Strong correlations are also shown between the acceptability and trust values (Pearson correlations) and the rating of the participants to the 20 conflicts visualized on the videos. Participants who gave the highest criticality values to the videos are also those who gave the highest acceptance values to the system (see Table 6

Table 6: Factors that correlate with acceptability and trust.

).

Finally, attitude of the drivers towards automation was investigated and correlations are shown between the acceptability and trust values and the readiness to drive a highly automated vehicle⁵ (Table 6). No correlation is found between acceptability and trust values and the pleasure declared by the participants for manual driving⁶.

Table 6: Factors that correlate with acceptability and trust.

Correlations		"Dirving a highly automated car"	Video rating: criticality
Before	Acceptability	$r=.328^{**}$	$r=.462^{**}$
	Trust	$r=.330^{**}$	NS
After	Acceptability	$r=.287^{*}$	$r=.530^{**}$
	Trust	$r=.513^{**}$	NS

** $p < .01$ * $p < .05$

7.4.3.3 Intention to use the system

Data being not normally distributed, non-parametric statistics are executed to measure the intention to use. Comparison of the repeated measures are performed

⁵ Participants were asked how they would agree (from 0 to 100) to the following sentence "I would enjoy driving a highly automated car"

⁶ Participants were asked how they would agree (from 0 to 100) to the following sentence "I enjoy manual driving"

using Friedman's test showing significant differences in the intention to use according to the PROSPECT functionalities and to the contexts of use (Figure 38).

Participants reported a higher intention to use the warning than the emergency braking and the steering [$\chi^2(2)=17.871$, $p<0.001$], the lowest intention to use being always regarding the steering function.

The context also has a significant overall effect on the intention to use the systems. The participants were most likely to use the systems in urban areas [$\chi^2(2)=30.696$, $p<0.001$]. Reporting to use it on a highway or an express road is lower. The same differences are shown for each functionality: for the warning [$\chi^2(3)=18.635$, $p<0.001$], for the emergency braking [$\chi^2(3)=54.908$, $p<0.001$] and for the steering [$\chi^2(3)=24.009$, $p<0.001$]. It is important to note that the braking and steering functionalities obtain quite low values of intention to use on high speed roads. Only a warning seems to be acceptable in this context.

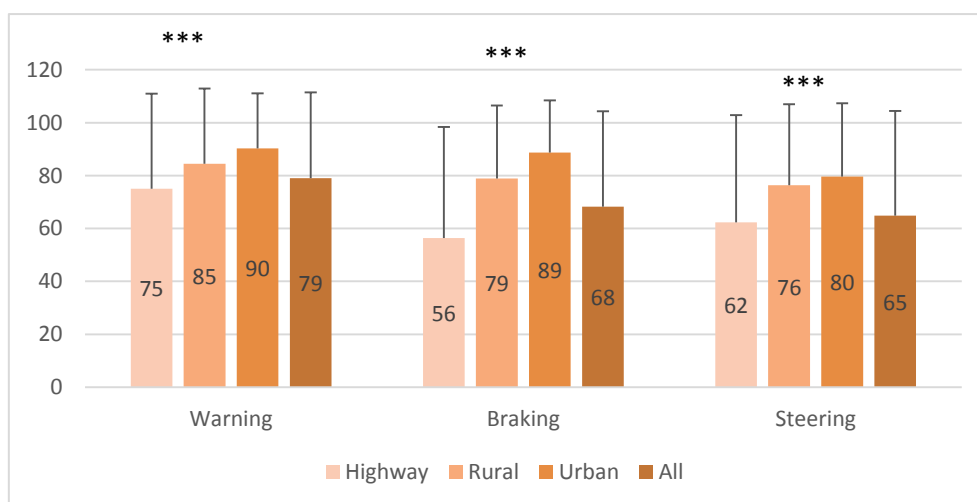


Figure 38: Intention to use according to PROSPECT functionalities and contexts.

The intention to use the braking and the steering functions are influenced by the age of the participants (Figure 39). Middle-aged participants are significantly more likely to use these functionalities than elderly, the lowest intention being expressed by the younger participants [for the braking function: Kruskal-Wallis, $H=5.912$, $p=.052$ & for the steering function Kruskal-Wallis, $H=9.790$, $p=.007$]. No difference in terms of age is observed regarding the intention to use the warning.

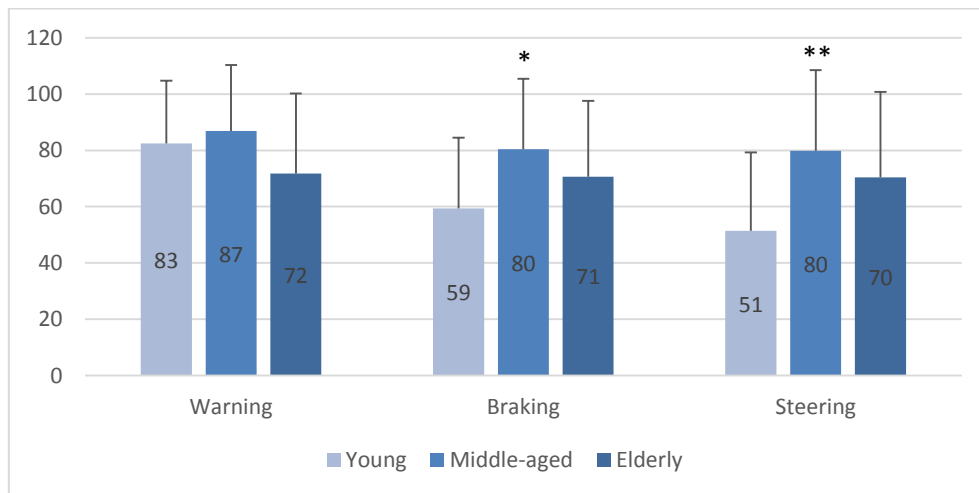


Figure 39: Intention to use according to the age.

Some differences are also observed between participants who drive less than 10 000 km/year and those who drive more (Figure 40). The latter being much more ready to use the systems than the former regarding the warning [Mann-Whitney, $U=367$, $p=.055$], the braking [Mann-Whitney, $U=293.5$, $p=.004$] and the steering [Mann-Whitney, $U=334$, $p=.022$].

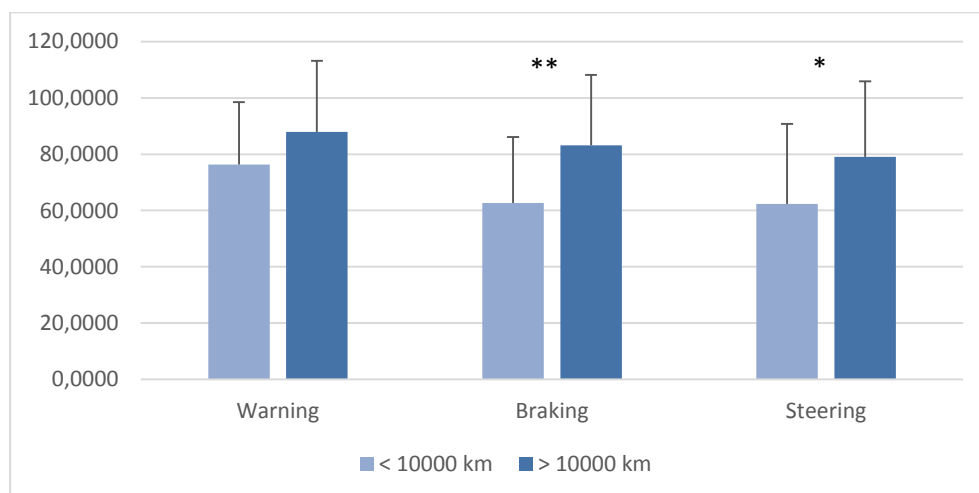


Figure 40: Intention to use according to mileage.

Finally, strong correlations (Table 7) are found between the intention to use the PROSPECT functionalities, acceptability, and trust (before and after the experiment). The readiness to drive a highly automated vehicle declared by the participants⁷ also correlates with the intention to use the systems, and the criticality rating as well.

Table 7: Factors that correlate with intention to use.

⁷ Participants were asked how they would agree (from 0 to 100) to the following sentence “I would enjoy driving a highly automated car”

Correlations		Intention to use		
		Warning	Braking	Steering
Before	Acceptability	,328**	,353**	,345**
	Trust	,410**	,424**	,456**
After	Acceptability	,588**	,525**	,390**
	Trust	,485**	,454**	,389**
"Driving highly automated car"		,375**	,302*	,353**
Video rating	Criticality	,253*	,346**	,337**
	Fear	-	-	,349**
	Stress	-	-	,332**

** p <.01 * p <.05

7.4.3.4 Willingness to buy

The willingness to buy is not very high before the experiment but significantly increases after the experiment when the participants have experienced on the videos situations where Prospect systems can be helpful [$t(62)=-2.229$, $p=.029$] (Figure 41).

The willingness to buy also differs significantly from the acceptability before [$t(62)=-3.686$, $p<.001$] and after [$t(62)=-2.774$, $p=.007$].

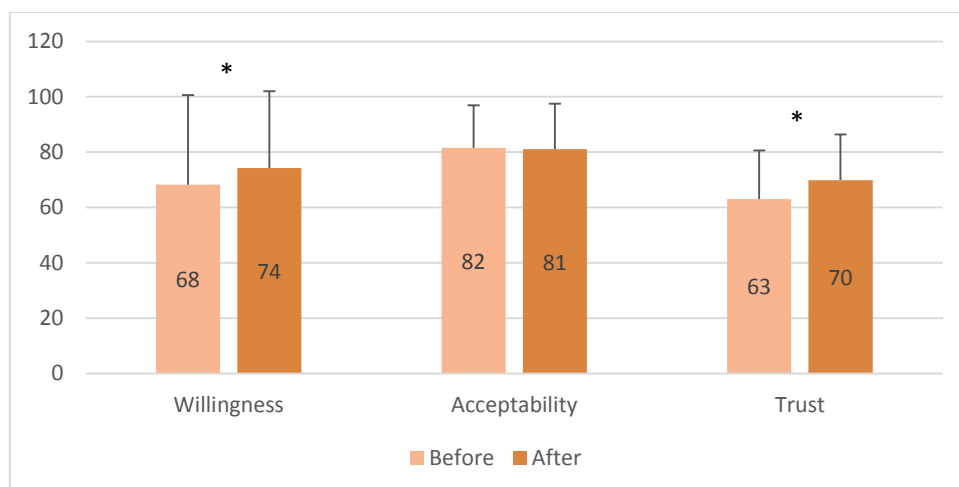


Figure 41: Willingness to buy before and after the experiment.

No age or gender effect is found. However, participants who drive less than 10 000 km/year declare themselves less likely to buy PROSPECT systems than participants who drive more, may this be before the experiment [$U=-347.5$, $p=.037$] or after the experiment [$U=329.5$, $p=.019$].

Finally, strong correlations (Table 8) are found between the willingness to buy the PROSPECT systems, acceptability, and trust (before and after the experiment). The readiness to drive a highly automated vehicle also correlates with the willingness to

buy. Regarding the video rating, criticality rating and fear rating turn to be linked with the willingness to buy as well.

Table 8: Factors that correlate with willingness to drive.

Correlations		Willingness to buy	
		Before	After
Before	Acceptability	,471**	,430**
	Trust	,536**	,527**
After	Acceptability	,532**	,721**
	Trust	,526**	,670**
"Driving highly automated car"		-	,287*
Video rating	Criticality	,310*	,516**
	Fear	-	,277*

7.4.3.5 Acceptance of a system failure

At the end of the experiment, the participants were asked how they would accept a system failure such as a false positive (activation in a non-critical situation) or a false negative (no activation in a critical situation). The low values given by the participants indicate that globally the system failures are judged quite unacceptable. However, the results show that false positives could be better accepted than false negatives [$t(62)=3.565$, $p=.001$]. The drivers seem to be better disposed to accept a system that does too many false positive activations than a system that misses to indicate critical situations (Figure 42).

Acceptability of false positive varies significantly for the warning [$t(62)=5.998$, $p<.001$] and for the braking [$t(62)=2.275$, $p=.026$], but no significant difference is found for the steering. The values, which are especially low for the steering intervention, suggest that a system failure would be hardly accepted in this case.

Acceptance of false positives or false negatives do not differ according to the driver profile in terms of age, mileage and gender.

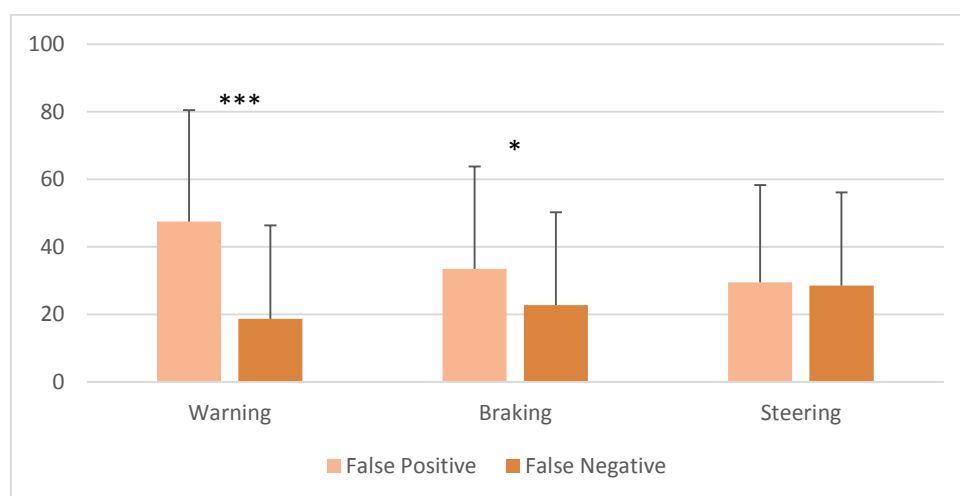


Figure 42: Acceptance of system failures according to PROSPECT functionalities.

Negative correlations are found between acceptance of a system failure and acceptability and trust in the system (Table 9). Drivers who declared the highest values of trust and intent to use PROSPECT-like systems are those who accept the less a system failure. This could indicate that these drivers only imagine such systems as very reliable and as systems in which they can have strong confidence.

Table 9: Factors that correlate with acceptance of system failure.

Correlations		Acceptance of system failure	
		False positive	False negative
Before	Acceptability	-	-,262 [*]
	Trust	-	-,302 [*]
After	Acceptability	-	-,348 ^{**}
	Trust	-	-,452 ^{**}
Driving automated car		-	-,249 [*]

7.5 CONCLUSIONS

The study conducted at IFSTTAR aimed at investigating the acceptability a priori of PROSPECT-like systems. The participants were asked about their opinion on the global functioning of ADAS close to the system developed in the PROSPECT project. General opinions were gathered through focus groups and questionnaires during a video-based experiment.

Results tend to establish that most of the participants would accept a PROSPECT-like system with a high score of 80% in the video-based experiment. Such a result is strongly linked with the criticality rating of the videos shown. The more the participants have estimated that the situations were risky, the more they have declared that they would accept a PROSPECT-like system. Likewise, in the focus groups, participants the most aware of the risks associated with VRUs also declared strongest acceptability of these systems, which substantiates Choi and Ji (2015) findings on the influence of the perceived risk on acceptability. Additionally, although none of the participants in the focus groups reported any accident with pedestrians or cyclists, they declared that this could happen and were aware of the consequences. Participants considered then that such systems would bring comfort and reassurance, and were convinced that they could become indispensable to the drivers, as it happened in the past with power steering or ABS.

On average, participants expressed confidence in the described systems, which tends towards 70% at the end of the video-experiment. Indeed, trust in the systems increases after having experienced situations where PROSPECT functionalities could help to avoid accidents. An interesting point is the influence of the attitude of the drivers towards in-car technologies in their confidence in the systems. As a result, drivers ready to drive highly automated cars would rely more on PROSPECT-like systems. Nevertheless, some limitations were expressed during the focus groups, with participants exposing concerns about potential consequences of an overconfidence in the systems. They stated that this could make them less engaged in the driving activity and lead to less attentive and more passive driving, which was considered as a major drawback of the developed systems.

If the intention to use the system obtained globally quite high values in the video experiment, which was consistent with the acceptability level, it is interesting to note some differences in terms of functionalities and driving environment. Participants of the video experiment showed a more pronounced intention to use the warning functionality than the braking and the steering, the lowest intention to use being the steering function. Focus group participants also raised concerns about the steering function, far more than for the warning and the braking. They were worried that a steering intervention could potentially lead to collateral accidents and then produce more harm than benefit. These results reveal different levels of trust according to the functionalities of the system. Participants of the experiments showed then a greater acceptance of the warning and braking functionalities and seemed much more suspicious towards the steering function. The intention to use also varied according to the context. Participants of the video experiment were more ready to use all functionalities in urban areas, which is consistent with the opinion collected during the focus groups. Most of the problems reported by the participants regarding interactions with cyclists and pedestrians concerned urban areas and a fewer occurred in rural areas. This could be linked to the complexity of the environment, higher in the city than in the countryside, which may distort the perception of the risks. Regarding highways and express roads, only warnings seem to be acceptable, since braking and steering functionalities obtain quite low values of intention of use.

A significant correlation was found between the willingness to buy a car equipped with a PROSPECT-like system and acceptability and trust in the systems. Participants who had a high level of acceptability were those who declared being the more inclined to buy the system, which is consistent with acceptability models. However, some differences between the driver groups were found during the focus groups. While the drivers familiar with ADAS appeared to be highly convinced by the utility of a PROSPECT-like system (as long as the performance is at a high level), the other drivers were less enthusiastic. Some differences between age groups were also observed. The young drivers expressed little confidence in the systems and preferred waiting for feedback from their close surroundings before considering using or buying it, while elderly drivers, although they were aware that their entourage would be reassured if they had PROSPECT-like systems, thought that they do not really need them.

Finally, the acceptance of system failures was investigated. The low values given by the participants indicated that they were judged quite unacceptable in general. However, differences were made between false positives and false negatives. It is worth to note the stronger rejection of a non-activation of the system rather than an unnecessary activation, a result that was supported by the focus groups. An interesting negative correlation was also found between acceptability of the system and acceptability of false positives. Indeed, the drivers the most convinced by the systems were also those who reject the most system failures, which show the link between acceptability and reliability of the system. It has to be noted that such a result was not observed within the experiment carried out in Nottingham. In their experiment, it seems that the drivers recognised the potential of the system, i.e. its

capability to identify a potential cyclist hazard and predict its intention to enter the roadway, even when the cyclist ultimately stopped at the roadside. Drivers exposed to a more 'classic' false alarm, i.e. whereby the system activates but no cyclist is present (which was not tested there), may be less accepting of the technology.

8 DISCUSSION AND COMPARISON OF RESULTS

8.1 METHOD RELATED ASPECTS

The use of driving simulators and test track for driver-in-the-loop tests of active safety functions provided a safe and repeatable set-up with drivers in the loop. These tests are valuable for the evaluation of safety critical functions that are designed to work while the driver is still in the loop and capable of interfering with the intervention. The simulator studies covered in this deliverable also covers a broad range of use-cases and variations of PROSPECT-like system functionality and are thus complementing each other.

8.2 OVERALL ACCEPTABILITY EVALUATION

Acceptance testing is an important part of the PROSPECT project. It provides knowledge on users' perception of the systems developed within the project, and an indication of their likelihood to purchase such a system. Fundamentally, it is crucial for the success of such active safety systems that they are acceptable to the drivers (e.g. judged to be useful and trusted).

Acceptability was assessed in 4 experiments and in different experimental environments. Eight Use Cases for Demonstrators and the 3 functionalities of the PROSPECT systems (warning, braking and steering) were investigated.

- Audi/TME used the VIL methodology in a simulator experiment
- University of Nottingham made a simulator experiment
- VTI/Volvo carried out a test track and a simulator experiment
- IFSTTAR performed 2 laboratory experiments: video-based experiment and focus groups

A common 'acceptance' methodology based on existing questionnaires was used, which is described in D7.2 (Report on methodology for balancing user acceptance, robustness and performance).

Results show a high likelihood of acceptance of PROSPECT systems, whatever the experimental conditions or the system investigated (VTI, UoN, IFSTTAR). However, ISFTTAR experiment showed that the more participants are aware of the risks associated with VRUs the more they give high acceptability values, which substantiates Choi and Ji (2015) findings on the influence of the perceived risk on acceptability.

The participants also expressed high confidence in the systems (UoN, IFSTTAR). Indeed, trust in the systems increased after having experienced situations where PROSPECT functionalities could help to avoid accidents. An interesting point from IFSTTAR experiment was the influence of drivers' attitude towards in-car technologies in their confidence in the systems. As a result, drivers ready to drive highly automated car would rely more on PROSPECT systems.

Participants were most positive towards the warning function, but nevertheless indicated also a high likelihood of using the braking and steering functions (VTI, UoN,

IFSTTAR. It is interesting to note the influence of the driving environment. Participants declared being more ready to use all functionalities in urban areas. Regarding highway and express road, only warnings seem to be acceptable, since braking and steering functionalities obtain quite low values of intention of use.

Willingness to buy was influenced by various factors, such as: the situation experienced (VTI: dummy versus bike), and the time at which the warning occurred (TME). Participants' willingness to buy increased after they were presented with 'critical' situations (IFSTTAR). A significant correlation was also found between the willingness to buy a car equipped with a PROSPECT system and acceptability and trust in the systems. Participants who expressed a high level of acceptability were those who declared being the more inclined to buy the system, which is consistent with acceptability models.

9 REFERENCES

- Adell, E., A. Várhelyi, and L. Nilsson. How Is Acceptance Measured? Overview of Measurement Issues, Methods and Tools. In *Driver Acceptance of New Technology: Theory, Measurement, and Optimization* (M. A. Regan, T. Horberry, and A. Stevens, eds.), ARRB Group Limited, Vermont South, Victoria, Australia, 2014, pp. 73–88.
- Bellet, T. and Banet, A., 2012. Towards a conceptual model of motorcyclists' Risk Awareness: A comparative study of riding experience effect on hazard detection and situational criticality assessment. *Accident Analysis & Prevention*, 49, pp.154-164.
- Choi, J. K. and Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle, *International Journal of Human–Computer Interaction*, 31: 692–702.
- Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly* 13 (3): 319–341. doi: 10.2307/249008
- Eckert, A., Hartmann, B., Sevenich, M., & Rieth, P.E. (2011). Emergency Steer & Brake Assist – A Systematic Approach for System Integration of Two Complementary Driver Assistance Systems.
- Fischer, M., Sehammar, H., Ljung Aust, M., Nilsson, M., Lazic, N., & Weiefors, H. (2011) Advanced driving simulators as a tool in early development phases of new active safety functions. Proceedings of the 3rd international conference on road safety and simulation, Indianapolis, USA, September 2011
- Jansson, J., Sandin, J., Augusto, B., Fischer, M., Blissling, B., & Källgren, L. (2014) Design and performance of the VTI Sim IV. Proceedings of the Driving Simulation Conference 2014, Paris, France, September 2014.
- Jian, J.Y., Bisantz, A.M. and Drury, C.G., 2000. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1), pp.53-71.
- Large, D.R., Burnett, G., Morris, A., Muthumani, A., Matthias, R. 2017. A Longitudinal Simulator Study to Explore Drivers' Behaviour During Highly-Automated Driving. 8th International Conference on Applied Human Factors and Ergonomics (AHFE 2017)
- Miller, D., Johns, M., Mok, B., Gowda, N., Sirkin, D., Lee, K., & Ju, W. (2016, September). Behavioural Measurement of Trust in Automation: The Trust Fall. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, No. 1, pp. 1849-1853). Sage CA: Los Angeles, CA: SAGE Publications.

Van Der Laan, J.D., Heino, A. and De Waard, D., 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation research. Part C, Emerging technologies*, 5(1), pp.1-10.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

ACKNOWLEDGMENTS



The research leading to the results of this work has received funding from the European Community's Eighth Framework Program (Horizon2020) under grant agreement n° 634149.