Grant agreement no.: 815098

D4.2 – Guidelines and recommendations from simulations

Date of publication: 31-10-2021

Disclaimer

This report is part of a project that has received funding by the European Union's Horizon 2020 research and innovation programme under grant agreement number 815098.

D4.2 – Guidelines and recommendations from simulations			
Work package No.	4	**Work package Title**	Scenarios Design and Simulation Systems
Tasks involved in the reported results	T4.3 Results analysis, lessons learned and cross-fertilisation		
Deliverable owner	UNIVLEEDS		
Dissemination level	PU		
Due date	31/10/2021		
Delivery date	31/10/2021		

List of contributors		
Section	**Author(s)**	**Reviewer**
1	Haibo CHEN (UNIVLEEDS)	Patrick VAN EGMOND (LuxMobility), Thibaud LATOUR (LIST)
2	Evangelos PASCHALIDIS, Haibo CHEN (UNIVLEEDS)	
3	Nicolas BERT, Mohsen ZARE, Maxime LARIQUE (UBFC)	
4	Nicolas BERT, Mohsen ZARE, Maxime LARIQUE (UBFC)	
5	Roderick MCCALL, Alexandre BAUDET (LIST)	
6	Guillaume GRONIER, Lou SCHWARTZ (LIST)	
7	Mike JUMP, Sergio HENRIQUEZ HUECAS (LIV)	
8	Florian KUTZNER, Tobias VOGEL, Celina KACPERSKI (UMA)	

9	Haibo CHEN (UNIVLEEDS)	
10	All	

Version History			
Version	**Date**	**Main author**	**Summary of changes**
V0.1	31/08/2021	Haibo CHEN	Table of Content and responsibilities agreed
V0.2	13/09/2021	Evangelos PASCHALIDIS, Haibo CHEN	Chapter 2 added
V0.3	18/10/2021	Florian KUTZNER, Tobias VOGEL, Celina KACPERSKI	Chapter 8 added
V0.4	19/10/2021	Guillaume GRONIER, Lou SCHWARTZ	Chapter 6 added
V0.5	20/10/2021	Roderick MCCALL, Alexandre BAUDET	Chapter 5 added
V0.6	20/10/2021	Mike JUMP, Sergio HENRIQUEZ HUECAS	Chapter 7 added
V0.7	21/10/2021	Nicolas BERT, Mohsen ZARE, Maxime LARIQUE	Chapters 3 & 4 added
V0.8	25/10/2021	Haibo CHEN	Draft for peer-reviews
V0.9	29/10/2021	All	Review comments done
V1.0	31/10/2021	Haibo CHEN	Final check/submission

List of acronyms

Acronym	Meaning
ACC	Adaptive Cruise Control or Absolute Cardiac Cost
CAV	Connected and Autonomous Vehicle
CAVE	Cave Automatic Virtual Environment
CHAT	Check-Assess-Takeover
DA	Driver Assist
eHMI	external HMI
eVTOL	electric Vertical Take-Off and Landing
FAV	Fully Autonomous Vehicle
FoV	Field of View
GPS	Global Positioning System
GSR	Galvanic Skin Response
HMI	Human-Machine Interface
HO	Hand-Over
HOR	Hand-Over Request
HR	Heart Rate
L3, L4, L5	Level of automation
LC	Lane Centring
LED	Light-Emitting Diode
MR	Monitoring Request
ODD	Operational Design Domain
PAV	Personal Aerial Vehicle
PBQ	Pedestrian Behaviour Questionnaire
SAV	Shared Autonomous Vehicle

SoA	State-of-the-Art
TAM	Technology Acceptance Model
TO	Take-Over
TOR	Take-Over Request
TTC	Time To Collision
UAM	Urban Air Mobility or Unmanned Air Mobility
UMUX	Usability Metric for User Experience
UTAUT	Unified Theory of Acceptance and Use of Technology
VR	Virtual Reality
WM	Wrist Motility
WP	Work Package
XP1, XP2	Experiment 1 (driving simulator), Experiment 2 (the VR system)

Notice

This document complies with the European Blind Union's guidelines (http://www.euroblind.org/publications-and-resources/making-information-accessible-all) in order to be accessible to anyone, including blind and partially sighted people, and at the same time and at no additional cost.

Table of Contents

Table of Figures

List of Tables

Executive summary

PAsCAL is a user-centric research project aimed at accelerating the user-friendly evolution of connected, cooperative, and automated vehicles and transport systems, by addressing important issues relating to the role of humans in this evolution, in particular appropriate interactions of the autonomous vehicle with different road users including non-drivers. The difficulty to reproduce in reality safety-critical situations on the road, which involve highly automated vehicles, leads to the development of driving simulators to be used as an interactive virtual reality tool for the human factors studies in the project.

This deliverable reports the findings of five simulation experiments ranging from professional driving simulation to home study kits, from drivers to pedestrians, and from road to air.

While these experiments have different settings, targeted users and levels of automation as described, they carry out several common tasks including:

1. Correlate and analyse driver behaviour/reaction under different scenarios;
2. Assess the acceptance of new interfaces integrated in the simulators, including information feedback and entertainment systems;
3. Put forward recommendations describing ways to improve the CAVs design, so they will be useful and acceptable to future real users, and the future drivers' trainings;
4. Produce guidelines for WP6 pilot specifications (e.g., to design new use cases involving autonomous public transport and to define some variables which deserve to be tested in real conditions).

The main findings from these five experiments are summarised as follows.

Findings of "DRIVING SIMULATOR":

- It was observed that those participants who had some experience and knowledge of autonomous vehicles were able to get a more concrete idea of how an autonomous vehicle works, what could drive to an increased acceptability, more positive attitude and feelings towards autonomous vehicles.
- The results from the study of the effectiveness and acceptance of the different signals present in the CAV showed that audio signals

were preferred and considered the most effective by the participants. The voice signal was the most relevant signal for handover and taking over requests according to all participants in the experiment. While the experienced drivers were more responsive to the light signal, they agreed with the novices that it was more relevant as a confirmation of autonomous driving engagement, once it has been properly activated.

- The results from the analysis of the effect of the driving experience on the acceptability of the CAV showed that experienced drivers report higher trust than novice ones, with higher acceptability, more positive attitude, and lower perception of the risk associated with CAVs, which emphases the importance of knowledge transfer, training/education, and awareness of CAVs.
- This experiment also showed that although an information-rich HMI is better perceived in terms of usability, it does not lead to more trust for the driver. At times, the opposite is true. Some specific feedback about the car's level of perception can be perceived as a source of stress for the driver, for both experienced and novice drivers.

Findings of "VIRTUAL REALITY PLATFORM":

- Like the previous driving simulator, the Virtual Reality (VR) experiment also observed that the participants who had some experience and knowledge of CAVs declared a high level of trust during the VR experience.
- The VR simulation delivered to them a more concrete idea of how works a L5 vehicle and the services it could provide. Experimenting L5 CAV shuttles was a good surprise for most of them.
- The overall attitude and feelings of most of the participants, who were already positive before the experiment, increased when re-measuring after.
- The results also showed that vulnerable disabled participants preferred shuttles to conventional buses.
- Participants were in favour of premium L5 shuttles for the multimedia and infotainment services, combined with their superior design and comfort. Their willingness-to-pay, however, didn't increase while considering this option.
- Further research is needed to confirm these findings of acceptability by testing larger panels and real-life situations.

Findings of "HOME STUDY SIMULATOR":

- In this study only a single alert followed by a countdown were used, which resulted in participants often feeling stressed or hurried.
- Alerts often seen as annoying and interruptive had negative impacts on the participants feelings towards the CAV and increased mental load and feelings of control.
- Perceptions of the CAV change over time from feelings of fear (first visit) to issues to do with control and decision making (last visit).
- The ability to predict how and what kind of decisions the CAV will take was seen as positive. Uncertainty was perceived negatively and fear of an unexpected end to autonomous mode was present.
- As a level 4 vehicle still requires manual intervention, it places a responsibility and hence the need to be attentive at all times on the driver.
- Further work with more participants is required to obtain sufficient quantitative data, in combination with the rich and detailed qualitative data provided by the repertory grid analysis, to provide scientific evidence for assessment of real "driver" behaviours towards CAVs.

Findings of "IMMERSIVE ARENA":

This experiment has produced a number of observations including:

- Nationality, living country and having young kids seems to have an impact on CAV receptivity.
- When a CAV stops to let crossing a pedestrian, it is better to send a signal that the CAV will wait the pedestrian's crossing.
- A feedback is waited by pedestrians in all situation and particularly in dangerous ones.
- Presence or absence of a crosswalk already on the road does not play a significant role.
- When the CAV stops the use of a signal to show that the CAV is waiting that the pedestrian cross is needed. The projection on the road is well accepted in the cases of a pedestrian crossing is painted or not on the road.
- When no pedestrian crossing is painted on the road, the pedestrians mostly expect that the CAV doesn't stop. Thus, no signal seems

needed in this case. Or a discrete signal without honk can be used like a red light on the VAE or projected on the road.

- If a pedestrian crossing is painted on the road, pedestrians expect that the CAV stops.
- The CAV has to be easily identified in the traffic.
- Regulation and standardization of eHMIs are needed to ensure uniformity regardless of the manufacturer and improve predictivity, understanding and so acceptation of CAVs.
- The more promising eHMI in terms of UX and receptivity are text-based interfaces, but it raised some issues to be understood by everybody including visually impaired, illiterate, kids, persons not able to read the language used.
- The easier to understand and more elegant the eHMI is perceived to be, the better the acceptance of the CAVs equipped with it.

Findings of "HELIFLIGHT-R":

Immediate work was focusing on finalising setup of the testing environment, developing a series of briefing and de-brief questionnaires and obtaining approval from the University ethics committee. No experimental data has been collected. Once this approval has been granted, recruitment of volunteers will begin.

Other work carried out in WP4:

WP4 has carried out a State-of-the-Art review of **lessons learned** and **results found in other projects**, complementing the findings obtained from the aforementioned experiments. This task intends to explain how the aforementioned experiments are embedded in the overall research field, as well as where the simulations are placed in relation to the other research work. It focuses on several human-vehicle interactions such as:

- Interaction of the human driver with the autonomous vehicle, focusing on HMI designs for take-over request (TOR), their impact on behaviour and acceptance.
- Driver training: Given the novelty of the systems, drivers need to have accurate expectations and mental models. Studies in the existing literature have investigated the impact of training on driver behaviour and acceptance of autonomous vehicles.

- Interactions with pedestrians, focusing on the use, efficiency and acceptance of eHMI that aim in facilitating interactions of pedestrians with autonomous vehicles.
- Autonomous public transport, investigating the acceptance of autonomous shuttles after experiencing the system, and the needs of peoples with disabilities.
- Issues related to the acceptance of autonomous urban air mobility.

The results from these simulation experiments were also used to enrich the **multidimensional map of public acceptance** (see deliverables in WP3 for illustration). It was found that direct experience with CAV simulations, increases acceptance, including attitudes, affective reactions, intention to use and willingness to pay. It was also found that some degree of previous experience is necessary for furthermore immersive experience to yield positive effects. These findings have strategic implications. First, to increase acceptance simulators might offer a cost efficient and safe alternative to on-the-road prototypes. Second, "phasing-in" autonomous features stepwise, for example by exposure to partially automated vehicle features, seems more advisable than direct confrontation with L5 systems.

Scientific publications

Whilst the results of WP4 are described in this deliverable in detail, they are also presented in other formats of publication such as peer-reviewed scientific articles as listed in Table 0.1. When published, these publications will provide additional information about the WP4's work including specific and thorough literature reviews, further explanation of relevant studies and how the existing findings are used as well as where further work is needed.

Table 0.1: Potential Scientific publications from WP4

#	Title of Publication	Type of Publication	Title of the journal or equivalent	Status
1	Le véhicule autonome: un mode de transport urbain acceptable pour les personnes en situation de handicap moteur?	Conference	56eme congrès de la SELF[1]	Accepted

[1] https://ergonomie-self.org/congres-self/congres-2022/presentation-du-congres/

2	The effect of different types of signals on drivers' reactions during takeover in semi-autonomous cars	Journal	Safety Science	In preparation
3	Acceptability of the autonomous vehicle for disabled people	Journal	Ergonomics	In preparation
4	Acceptation des véhicules entièrement autonomes par les piétons lors de situation de traversée de route : expérimentation en environnement simulé	Conference	IHM 2022[2]	Submitted
5	French adaptation and validation of the Pedestrian Receptivity Questionnaire for Fully autonomous vehicles (PRQF)	Journal	TR-C[3]	In preparation
6	Comment les piétons interagissent-ils avec les véhicules sans conducteur?	Conference	SELF 2022[4]	Accepted
7	Receptivity and user experience of eHMIs: the results of a survey	Journal	Interacting with Computers	In preparation
8	Studying of Drivers in Automated vehicles Over Time	Journal	TR-A[5]	In preparation
9	Using Repertory Grids to Understand Driver Perception of Automated Vehicles	Conference	AutomotiveUI 2022[6]	In preparation
10	Vehicle Experience Concepts and Guidelines	Conference	AutomotiveUI 2022	In preparation

[2] https://ihm2022.afihm.org/fr/
[3] Transportation Research Part C: Emerging Technologies
[4] https://ergonomie-self.org/congres-self/congres-2022/presentation-du-congres
[5] Transportation Research Part A: Policy and Practice
[6] https://www.auto-ui.org/conference-series/conferences/

1 Introduction

1.1 Purpose and organization of the document

WP4 aims to collect attitudes, acceptances of participants exposed to CAV contexts (pedestrians, passengers) and user behaviours during the simulated use of CAVs. The simulated situations were based on relevant use scenarios for drivers and non-drivers, meeting their expectations and needs identified in WP3.

This deliverable (D4.2) presents the outcomes of five simulation experiments, focusing on the participants' experience of CAVs, ways to improve the CAVs design, and guidelines for the real-world pilots in WP6.

This deliverable starts with a State-of-the-Art review on what lessons learned and findings produced in other projects in Chapter 2, before describing the outcomes of the five simulation experiments, namely:

1. “DRIVING SIMULATOR” (Chapter 3),
2. “VIRTUAL REALITY PLATFORM” (Chapter 4),
3. “HOME STUDY SIMULATOR” (Chapter 5),
4. “IMMERSIVE ARENA” (Chapter 6) and
5. “HELIFLIGHT-R” (Chapter 7).

Chapter 8 presents the use of the outcomes of these simulation studies to enrich the multidimensional map of public acceptance developed in WP3 and verify if the attitudes and behaviours identified earlier are also observed during the WP4 tests in a simulated environment.

1.2 Intended audience of this document

The main audience for this document is twofold. The consortium members of the PAsCAL project first, especially partners responsible for other simulation experiments dedicated to training (WP5) and pilots (WP6), but also partners in charge of transversal analysis and cross fertilisation so they can have a clear idea of the material conditions the WP4 experiment were run.

A second but no less important audience is the wider research community for whom this document can serve as a basis for discussion of both the experimental protocols and the results that will emerge from the WP4 experiments.

1.3 Deviations

1.3.1 Covid-19 impact

The Covid-19 outbreak had a major impact on both organisation and people's mindset/attitudes towards some transportation means.

Regarding the organisation of the experiments, it impacted almost each step of it. Many of us had to organise themselves for working exclusively remotely, while we were still elaborating the research questions and designing the associated scenarios and experimental protocols. Ordering the necessary materials, developing the simulations, and setting up the simulators have been considerably retarded as soon as we weren't able to regularly access our facilities. Finally, most of the experiments need the subject to come at our respective facilities, what has been forbidden/made much more complicated for many months, depending on the sanitary rules of each country. For these reasons, the experiments that were supposed to start around September 2020 finally did from May to October 2021.

Regarding the people's mindset/attitudes among transportation, the Covid-19 was also a major changer. Most of the European citizens haven't been allowed to travel anymore for months and, when they were allowed again, we observed massive changes in the behaviours: some transport modes being avoided because associated to a higher risk of contamination, and some other ones becoming increasingly popular as they were seen as safe alternatives. So, we were forced to adapt some questionnaires for taking into account some recent changes in peoples' behaviour and, more generally, not to ignore such an event that made a big change.

1.3.2 Other factors

In addition to the deviations resulting from the Covid-19 pandemic, the work at Liverpool has suffered additional delays due to a number of factors, as follows.

First, the researcher hired to undertake the work was only available at 1/3 Full-Time Employment (due to the budget available) and so the work has automatically proceeded at a slower rate than initially planned for a full-time employee.

Second, ongoing maintenance and modernization works at the HELIFLIGHT – R simulator have suffered unexpected setbacks. Some of these upgrades were required to utilise X-Plane as the simulator outside world environment. The upgrades have, unfortunately, lead to delays related to integrating the new software and required hardware into the facility. This has meant that the experiments, as designed, cannot proceed until the upgrades are completed.

Third, the researcher has now left the University of Liverpool. A new researcher has joined the project as his replacement. However, formalizing the hiring and bringing the new researcher up to speed in the systems developed has made a further delay unavoidable.

2 Lessons learned from other projects

2.1 Introduction

Interactions of users with autonomous vehicles are expected to play a significant role in developing trust and acceptance. These interactions can occur both while interacting with the vehicle as its user or as an external road user (e.g., pedestrian). From an autonomous vehicle driver's perspective, it is necessary to continuously provide feedback to the user regarding the state of the vehicle with respect to the automation status. Moreover, for lower levels of automation (e.g., Level 3) an efficient design of Human-Machine Interfaces (HMI) is required to efficiently communicate takeover requests (TOR) to ensure safety and a positive experience for the user. On the other hand, drivers need to have well calibrated and realistic expectations (mental models) with respect to the automation limits and capabilities. Accurate mental models can minimise automation misuse and assist in developing trust and acceptance. Another aspect is the interaction of autonomous vehicles with vulnerable road users as pedestrians. In higher levels of automation, it may be the case that vehicles will be unoccupied, and pedestrians will have to interact with them in the absence of a human controller. Efficient communication in these situations has the potential to improve safety and enhance acceptance of autonomous vehicles by the other road users. Autonomous vehicles are also expected to have solutions in the form of public transport or flying vehicles. These can raise additional considerations with respect to safety and efficiency whole acceptance needs to be improved for all users including people with disabilities. The current chapter focuses on these aforementioned issues and presents findings from the existing literature, mainly from simulation studies. In particular, the main focus is on:

- *Interaction of the human driver with the autonomous vehicle*: This part covers HMI designs for TOR, their impact on behaviour and acceptance.
- *Driver training*: Given the novelty of the systems, drivers need to have accurate expectations and mental models. Studies in the existing literature have investigated the impact of training on driver behaviour and acceptance of autonomous vehicles.
- *Interactions with pedestrians*: This part focuses on the use, efficiency and acceptance of external HMI (eHMI) that aim in facilitating interactions of pedestrians with autonomous vehicles.

- *Autonomous public transport*: Findings from studies investigating the acceptance of autonomous shuttles after experiencing the system. The section also investigates the needs of peoples with disabilities.
- Issues related to the *acceptance of autonomous urban air mobility* (UAM).

2.2 Automated driving and human driver

2.2.1 Takeover requests, HMI and driver behaviour

In the highest degrees of automation (Level 4 or 5) the requirement for human intervention is expected to be trivial or unnecessary, as these vehicles will be able to perform a fallback and achieve a minimal risk condition. However, the currently available vehicles can only support features of partial automation (Level 2) with conditional automation (Level 3) to follow. The latter will be able to perform parts of the longitudinal and lateral components of the driving task requiring however a sufficient level of readiness from the driver to respond to TORs and resume manual control of the vehicle in emergency situations or when the Operational Design Domain (ODD) is no more supported. Existing literature has focused on several aspects of TORs as the communication interface with the driver, the performance of each system and their acceptance. As Level 3 vehicles have not been yet implemented, TORs and Human-Machine interfaces (HMI) have been mainly investigated in the context of driving simulator studies.

The TORs are communicated to the driver via three main HMI modalities (Morales-Alvarez et al., 2020):

- *Visual*: Visual HMI are usually presented in the form of arrows, icons or lights to inform the driver about a TOR. Visual HMI are considered as unobstructive however, they can be missed by drivers involved in non-driving related tasks.
- *Auditory*: Auditory HMI can have the form of sounds or speech messages to inform the driver about a TOR. Auditory HMI are clear to understand and do not require eyes-off-road time however, their message may not be clear to the driver while the more information provided, the longer the reaction time of a driver might be.

- *Haptic*: Haptic HMI attempt to attract driver's attention with vibrotactile systems. The message of a haptic HMI cannot be missed by a driver, but they cannot be used to warn for multiple alerts.

Bazilinskyy et al. (2018) conducted a survey to investigate receptivity of various HMI designs for TOR. Results in Figure 2.1 showed that among the auditory concepts, voice messages were preferred over beeps and other sounds. Regarding visual interfaces, icons and displays were more preferred compared to light or changes in the level of light. Finally, the most preferred vibrotactile approach was vibration of steering wheel.

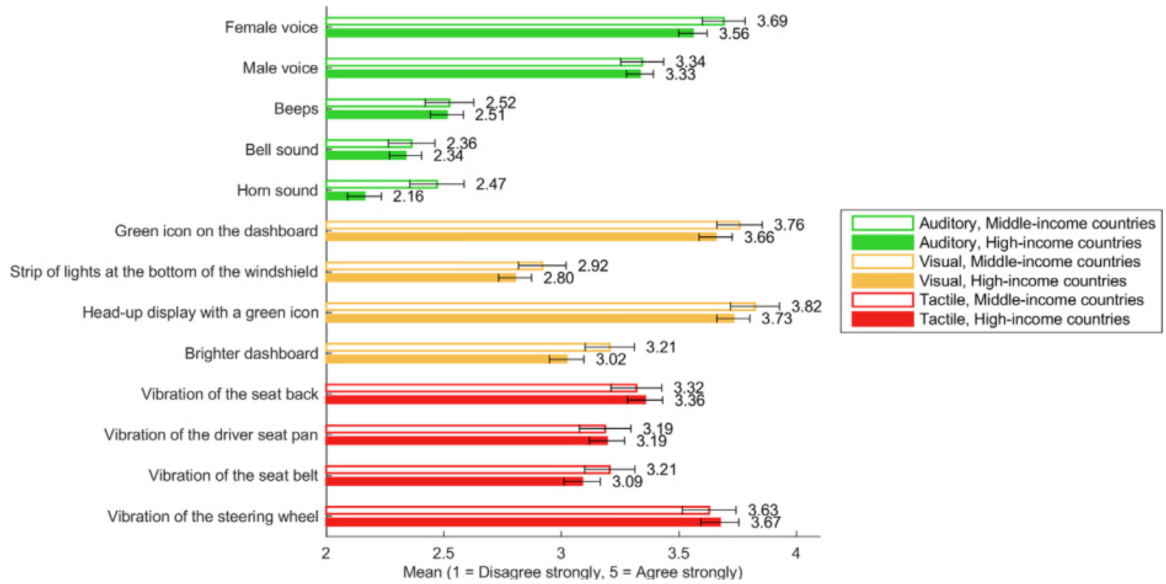
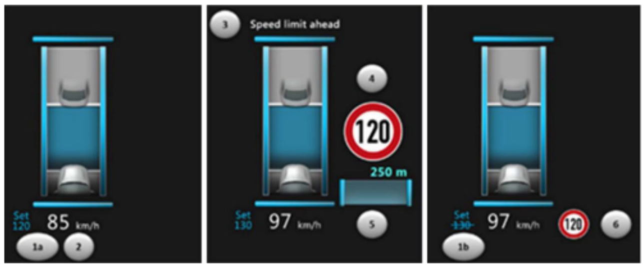
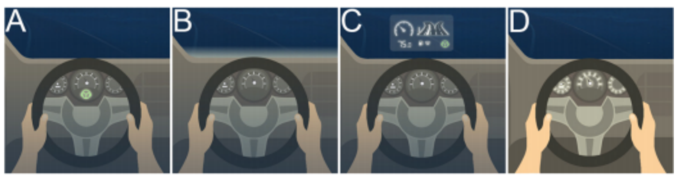
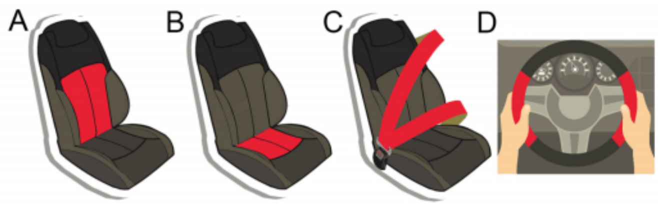
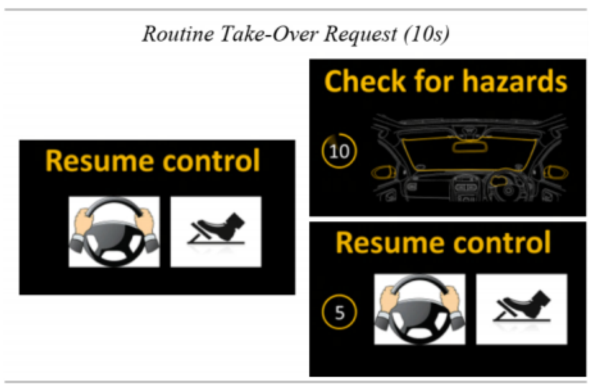
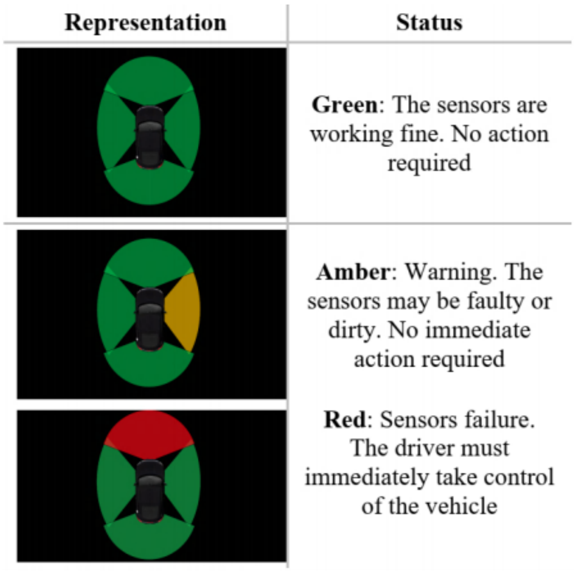
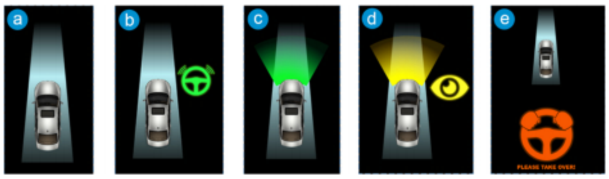


Figure 2.1: Perceptions about HMI concepts based on a survey of (Bazilinskyy et al., 2018)

The findings of Bazilinskyy et al. (2018) raise some considerations with respect to the implementation of the HMI. A first outcome is that people prefer clearly communicated messages that reduce ambiguousness. This is shown for instance in the preference of icons over lights or verbal messages over sounds. Another finding is that users may prefer less intrusive HMI. This is illustrated in the higher rankings of icons and displays over auditory messages or tactile solutions. This latter finding raises some concerns. As it will be presented later in this chapter, tactile and auditory HMI may have the potential to elicit faster reaction times in TOR. Hence, attention is needed in the implementation of HMI to reduce any annoyance and ensure acceptance. Some HMI concepts in the existing literature are presented in Table 2.1.

Table 2.1: Visual and Haptic HMI concepts

Icons indicating the automation status (gray = disabled; blue = available; green = active; red = take over) (Sebastiaan Petermeijer et al., 2017)	
Visual HMI presenting additional information to the driver as set and current speed, traffic events ahead etc. (Y. Forster et al., 2017)	
Concepts of takeover requests - (A) A green icon on the dashboard, (B) A strip of lights at the bottom of the windshield, (C) A head-up display with a green icon, (D) A brighter dashboard - (Bazilinskyy et al., 2018)	
Concepts of vibrotactile requests - (A) Vibrations in the seat back, (B) Vibrations in the seat pan, (C) Vibrations in the seat belt, (D) Vibrations in the steering wheel - (Bazilinskyy et al., 2018)	
Direct TOR vs "Check for hazards" advice before the TOR (Large et al., 2019)	

HMI providing information about sensors status (Large et al., 2019)	
Visual interface of the different vehicle system states. (a) Automation unavailable (b) automation available but not yet activated (c) automation activated (d) monitoring request (e) take-over request (Lu et al., 2019)	

The implementation of an HMI can vary within each of the main modalities with the potential impacts to have been investigated. Brandenburg and Chuang (2019) investigated the impact of an abstract (arrows indicating lateral or longitudinal control) versus a skeuomorphic (icons describing the type of control) visual design for the TOR. The authors found that a skeuomorphic representation led to faster and more accurate reactions of drivers especially at closer time headways. Moreover, participants in the study perceived the skeuomorphic representation as more desirable. Borojeni et al. (2016) compared the impact of different light emitting diode (LED) implementations and concluded that displays that inform drivers about the steering direction or moving lights can improve reaction time and time to collision. Yang et al. (2018) suggested a LED implemented on the vehicle's windshield that generates different light patterns depending on the conveyed message (automation status, danger detection, TOR etc.). Although no significant differences were found between the implementation suggested by the authors and a baseline visual HMI that was only activated in case of a TOR, participants reported higher levels of trust and acceptance for the LED display and also spent significantly longer time looking on the road. With respect to auditory approaches, Stojmenova et al. (2020) reported shorter reaction times when a directional sound from the (side of the danger) was applied compared to the non-directional sound display.

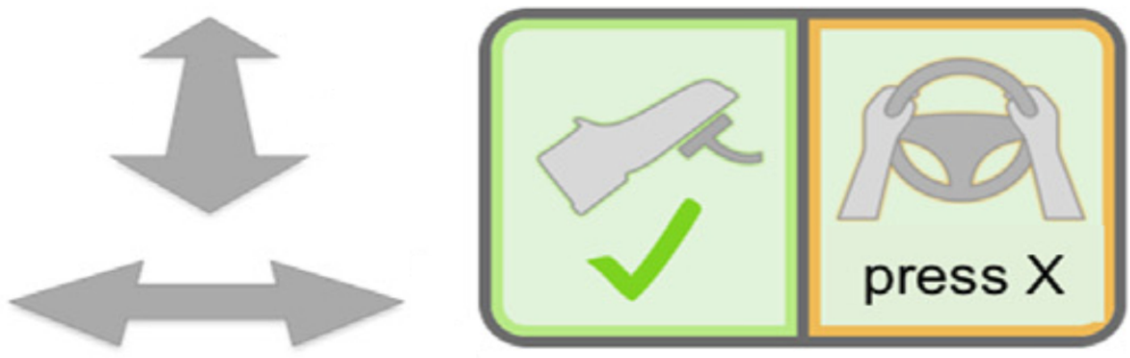


Figure 2.2: Abstract and skeuomorphic concepts, Brandenburg and Chuang (2019)

Several studies have focused on the impact of modality of an HMI on drivers' reaction time. The efficiency of warning systems was initially investigated outside the context of autonomous vehicles. Mohebbi et al. (2009) reported that vibrotactile and audio warning systems have the potential to improve reaction time in rear-end collision situations. Also, the authors concluded that tactile implementation led to better performance compared to auditory. Scott and Gray (2008) also found that a vibrotactile warning system resulted in better reaction times for the avoidance of rear-end collisions, compared to an auditory and a visual warning. Among the warning systems, higher reaction times were observed in the visual modality while all warning systems led to shorter reaction times compared to a no warning condition. For autonomous vehicles, Jeon (2019) reported faster reaction times with auditory HMI (either speech or a beep) compared to visual only. Sebastiaan Petermeijer et al. (2017) also concluded the same as in their study, regardless of the nature of a secondary non-driving related task (reading, calling or watching a video), auditory and vibrotactile TOR resulted in faster reaction times, compared to visual. Moreover, auditory and tactile requests were perceived as more useful by the participants of the study.

The main HMI approaches (or modalities) can be also combined to enhance their efficiency and improve driver performance after resuming control. In fact, multimodal HMI implementations have been found to reduce driver reaction time, mainly compared to a visual only interface, and be more well received by the drivers (Y. Forster et al., 2017; Yoon et al., 2019). Yun and Yang (2020) also found shorter reaction times in combined modalities while the slowest reaction time was observed in the visual feedback only condition. Despite the improvement in reaction time by multimodal implementations, driving behaviour does not necessary always improves after resuming manual control (S. Petermeijer et al., 2017).

On top of the typical HMI implementations, Lu et al. (2019) investigated the impact of a Monitoring Request (MR) before the TOR. Their study focused on driver behaviour before encountering a pedestrian at a zebra crossing. Participants that had the MR available had shorter reaction times and longer minimum time to collision hence, this is an additional design to be considered in the future. The MR+TOR design also led to lower perceived workload and higher levels of trust, compared to a TOR-only condition. However, when a MR was not followed up by a TOR resulted in later responses or no responses as the drivers might not be certain on how to react to the received warning. The presence of additional information before a TOR could be crucial in improving situational awareness. Large et al. (2019) found that when a warning was issued before the TOR led to significantly more mirror checks from the drivers, compared to a direct TOR.

Another important point to be considered is the driving behaviour following the issue of a TOR. For instance, Brandenburg and Chuang (2019) found that road geometry can affect the quality of a takeover (stronger breaking behaviour, higher lateral deviation, and increased response times) suggesting caution in curved roads by allowing more time to react. Moreover, the authors found that if a TOR is designed as a two-step process (lateral and longitudinal controls to be resumed separately), lateral control should be available first as it reduces the impact on lane deviation, especially on curved roads. Sadeghian Borojeni et al. (2018) also concluded that response to a TOR is faster on straight roads while drivers reacted slower in urgent events taking place on curved roads. Traffic density is another factor that has been reported to affect behaviour after a TOR request. In particular, increased density was linked to longer reaction time and shorter time to collision (TTC) (Gold et al., 2016; Radlmayr et al., 2014). However, So et al. (2021) reported that traffic density and road curvature did not affect reaction time in their study however, drivers took longer to stabilise the control of the vehicle after a TOR. Dogan et al. (2017) investigated the impact of an anticipated TOR (when the autonomous system was reaching its limit of maximum speed) with an unanticipated TOR (vehicle speed below the system boundary at the time of the TOR). While no significant differences were found in reaction time, in both cases, drivers took more time to regain the lateral control of the vehicle compared to a manual driving condition. This finding further highlights the importance of providing the human driver with sufficient time to regain control of the vehicle.

Despite the limitations or the potential of a TOR, initial levels of trust to autonomous systems increase with time and exposure. Large et al. (2019) conducted a multiple-day experiment and reported higher ratings of acceptance even after an emergency TOR on the fourth day. Moreover, Miller and Boyle (2019) concluded that as drivers become more familiar with a system their willingness to take more risks as engagement in non-driving related activities increases. Korber et al. (2018) also found that trust ratings of drivers increased after experiencing a series of TOR in a simulator. Providing details regarding the behaviour of the vehicle to some of the participants, did not affect the levels of trust or acceptance however, it improved their understanding of the system. An interesting finding of the study is that drivers' trust decreased when an explanation was provided for a TOR in a roadwork scenario. The authors assumed that a TOR request in this situation negatively affected the attitude of drivers' regarding the capabilities of the system, as they might have been under the impression of a more complex and competent system. This finding is raising a significant point of correctly "calibrating" the expectations of human drivers (else, their mental model) with respect to the capabilities of a specific autonomous system. A correctly calibrated mental model can adjust the perceptions and expectations of human drivers minimising issues of over-trust or distrust of autonomous vehicles and therefore their misuse. To that end, issues of driver training and education about autonomous vehicles are discussed in the next section.

Some main findings from the HMI and TOR literature can be summarised as follows:

- Current HMI approaches rely on visual, auditory and haptic solutions
- Visual HMI have been found less efficient in general in drawing drivers' attention
- Auditory and haptic HMI, although more efficient may be considered as disturbing
- HMI should deliver a clear message avoiding any uncertainty
- Multimodal HMI may be more efficient and well received by drivers
- Drivers have slower reaction times in curved roads, attention should be paid when a TOR is issued on these occasions
- Traffic density can affect reaction times in TOR

- Lateral control is taking longer to recover after a TOR. If resuming control takes place in two steps, drivers should resume lateral control first
- Overall, driving behaviour may be negatively affected after a TOR
- A warning prior a TOR has the potential to improve reaction times or situation awareness
- In general, the presence of HMI is preferred by drivers and perceived as useful

2.2.2 Driver training and education for autonomous vehicles

Drivers' responsibilities vary depending on the level of automation. In higher levels of automation as Level 4 or 5 human intervention is not of prime importance as these vehicles will be able to achieve a minimal risk condition however, this is not the case for the lower levels of automation. For instance, in Level 2 although some autonomous features are available, the driver should be monitoring at all times and intervene when system capabilities are reached, or an emergency situation occurs. The situation might be more complex in Level 3 as the driver is not required to monitor at all times and engage in a non-driving related task however, a TOR is likely to occur. It is essential for drivers to have a clear understanding with respect to the system capabilities (accurate mental model) while they have also developed additional procedural skills (navigation, hazard detection and vehicle control) to efficiently intervene when required (Merriman et al., 2021).

Several studies in the existing literature have focused on the development of training programs or approaches and investigated their impact on drivers' understanding of autonomous systems and the impact on their behaviour. A study by Mueller et al. (2020) highlighted the importance of intuitive training on how to monitor and evaluate interface indications for the correct interpretation of the system status in a Level 2 vehicle, for two different interface designs. The authors concluded that even trained participants failed to recognise whether an adaptive cruise control (ACC) had recognised the lead vehicle while they were not able to also state the reasons for the system being inactive. Moreover, participants relied on wrong sources of information to determine the status of the system. Significant differences however were observed regarding the lane centring (LC) systems, although correct scores were low both for trained and untrained participants. The authors concluded that driver education should extend from an interface-specific training approach and focus on the understanding of drivers on the system limitations and how to use the correct information to determine the status of a system. Another important benefit of training could be in preventing or minimising the negative impacts of first failure effects as it can provide the learner drivers with the opportunity to experience potential cases of failure. Hergeth et al. (2017) found that participants with prior experience reacted faster and more efficiently during a TOR. However, differences in behaviour between experienced and non-experienced participants were not significant during a second TOR which may imply that experience is another factor that improves interaction with an autonomous vehicle. It is interesting that trust levels of participants in automation increased after experiencing the system.

Krampell et al. (2020) developed a training programme of 14 error-learning scenarios to highlight the limitations of a Driver Assist (DA) which corresponds to a Level 2 vehicle. A trained group of participants was more likely to understand critical situations where they had to take back the control of the vehicle, compared to a group that had only read the user manual (control group). The control group expressed difficulties in understanding changes in the road environment (such as absence of lane markings) that required human intervention, while they overall expressed more doubts regarding the capabilities of the system overall, compared to the trained group. In a similar approach, Yannick Forster et al. (2019) compared the impact of an interactive tutorial approach against the owner's manual. Their tutorial was based on MS Powerpoint and presented operating elements of the vehicle and interfaces. Moreover, a set of questions was provided to participants including explanations and details both for correct and wrong answers. The authors examined performance in correct transitions between manual, Level 2 and Level 3 levels but did not find significant differences between the two training approaches. However, participants who had some type of prior training formed a more accurate mental model compared to a control group that did not receive any training. Moreover, participants in both training conditions performed better in transitions across the different levels of automation. Payre et al. (2017) investigated the impact of training type compared with the involvement in a secondary non-driving task. As elaborated in the current study was considered a text explaining the features and use of the autonomous vehicle followed up by some questions. Participants in both elaborated and simple training also received some vehicle familiarisation during an in-vehicle practice session. During the study, participants experienced two system failures and higher reaction times observed in the first automation failure scenario. Moreover, participants who received elaborated training reacted faster and also interacted less with the pedals while regaining the manual control of the vehicle, while these participants also reported higher trust in the autonomous system. Sportillo et al. (2018) investigated the impact of more innovative forms of training such as driving simulator and virtual reality, compared to a simple user manual approach. The training covered the use of a Level 3 system regarding manual driving, automated mode and takeover requests for specific scenarios (road obstacle, road markings and system failure). For the first and the second TOR, the participants trained with a virtual reality approach and a simulator reacted faster with respect to the ones trained with the user manual however, no significant differences in time to collision were found regarding the third TOR. Participants in the user manual condition showed less initial confidence in the system which increased with time. Ebnali et al. (2019) followed a similar approach and investigated the impact of video and simulator training (versus no training) on takeover performance and takeover decision on whether a scenario is critical or not. Trained participants (both types of training) had quicker takeover responses however no differences were observed in speed behaviour after resuming control of the vehicle. Also, participants trained in the simulator had significantly better lateral control compared to not trained. Moreover, training significantly improved the accuracy rate of necessity for takeover especially for the critical situations. An interesting finding of this study was that training shifted trust of drivers towards moderate while higher levels were reported prior training. This outcome highlights an additional benefit of training towards the correct calibration of mental models, as these participants might have been prone to over trusting the capabilities of automation. Sahaï et al. (2021) evaluated the impact of three types of training (paper, video and driving) on TOR in a Wizard-of-Oz study. The authors found that in urgent TOR, participants in the driving training condition reacted faster but no impacts were found on visual behaviour.

One of the most elaborated approaches in driver training for autonomous vehicles is the CHAT (Check-Assess-Takeover) concept by Shaw et al. (2020). The CHAT training was developed for Level 3 automation and attempted to cover the following aspects:

- Formulation of well calibrated mental models
- Awareness about potential impacts of being out-of-the-loop
- Establishment of a procedure of tasks to be performed by human drivers for a safe transition from an out-of-the-loop state back to manual driving.

The main difference of the CHAT procedure is a sequence of checks a driver must perform from the time receiving a TOR to regaining the control of the vehicle. These actions refer to an initial check potential hazards in the surrounding environment and the assessment of the situation with respect to the vehicle's position and the road environment. The authors examined the impact of CHAT training versus a simple user manual approach. Although they did not find significant differences in terms of manual driving, significant differences were observed in the number of mirror-checks for the trained participants who also spend less time on a non-driving related task.

In brief, findings from driver training for autonomous vehicles can be summarised as:

- Any training programme needs to focus on the development of accurate mental models with respect to the capabilities of automation and hence calibration of trust in the system. This can be concluded by the benefits of some training in the identification of critical situations. Training has been found to either improve or calibrate trust towards automation and can be considered as a useful tool for the development of current mental models.
- Drivers need to be trained on how to use the HMI interfaces regarding the vehicle status rather than relying on abstract sources of information.
- Training programmes need to improve drivers' anticipation with respect to a potential TOR. Based on traffic characteristics and road geometry the drivers should be able to expect the occurrence of a potential TOR.
- Training programmes would potentially be more beneficial if exposed drivers to various system failures rather than educating them with manuals or other sources.
- The highest benefit that has been found from training is the takeover reaction time while less improvement has been found in behaviour after resuming control. However, this impact of training could be beneficial in emergency situations.
- Research on TOR and HMI has shown that drivers may need more time to properly control the vehicle when resuming manual control. Concepts like CHAT have focused on improving situation awareness before resuming control via a series of action to be performed by the driver after a TOR.

- Regardless training, drivers' responses in TOR improve with time. To that end, the most beneficial aspect of training could be in minimising the impact of first-time system failures.
- In most cases, any type of training (including user manual training) led to improvement in driver performance than no training. An issue to be considered is whether experienced drivers should also attend training courses before using any automated systems.

2.3 Interactions with other road users: pedestrians

Communication of autonomous vehicles with pedestrians is a major concern in the research community. Currently interactions of drivers with pedestrians are undertaken via visual contact and gestures however, in a hypothetical scenario of an unoccupied autonomous vehicle this approach will not be feasible. Several studies in the existing literature have suggested the use of eHMI that either communicate the intention of the vehicle or suggest an action for the pedestrian. The eHMI examined in present studies have the form of either visual or audio communication systems. The purpose of eHMI can vary and usually takes the form of (a) communicating vehicle's intention (status), (b) advise pedestrians, (c) indicate awareness of pedestrians' presence, while combinations of these implementations can also exist.

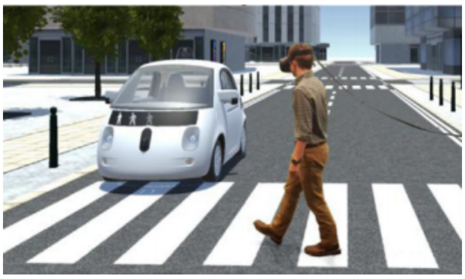
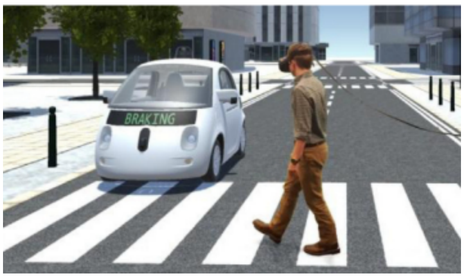
The efficiency of the various eHMI approaches has not been yet fully established in autonomous vehicles and pedestrians interactions however, it has been found that their presence is likely to increase receptivity of autonomous vehicles (Deb et al., 2018) or perceived predictability regarding their behaviour (Matthews et al., 2017). Stadler et al. (2019) concluded that the presence of eHMI can improve decision times of participants and reduce the error rates (e.g. in terms of right-of-way). Moreover, the authors found that eHMI can reduce the perceived task effort from a pedestrian viewpoint. Ferenchak and Shafique (2021) found that the presence of eHMI has the potential to improve trust and acceptance of interactions with autonomous vehicles. Similar benefits from the presence of eHMI regarding improvement of perceived safety and acceptance were also reported by P. Wang et al. (2021).

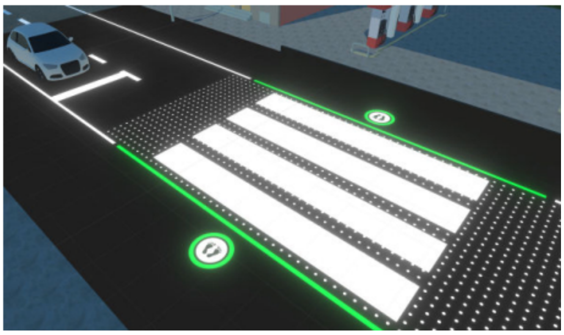
In existing studies, the modality eHMI is either visual or auditory. Regarding visual implementations, Löcken et al. (2019) mentioned that a good concept should use unambiguous signals and have high visibility. Colley et al. (2020) also mentioned unambiguousness as the most important aspect, with respect to auditory approaches. For instance, a honk can be perceived as greeting, warning, or a signal to let the pedestrian pass. Unclear messages should be avoided especially considering the presence of pedestrians with visual impairments that cannot perceive the motion cues of a vehicle. Colley et al. (2020) divided unambiguousness into three concepts namely, standardization, distinctiveness, and perceptibility. Deb et al. (2018) found that visual interfaces were perceived as safer, compared to audio. Table 2.2: Visual eHMI concepts presents some visual eHMI concepts found in literature.

Among visual designs, a walking silhouette or 'braking' in text were the most preferred in a study of Deb et al. (2018), compared to a flashing smile or no eHMI. In the same study, a verbal message was found to be the most preferred audible feature, compared to a beep signal or music. Verbal messages could also be a useful alternative for pedestrians with visual disabilities. Regarding behaviour, Deb et al. (2018) found that a verbal message and music elicited similar crossing behaviour while on the other hand, participants hesitated more with the beep sound. With reference to the authors, it could be the case that a beep sound might be also perceived as an alert and communicate a confusing message to the pedestrians. Participants in a Wizard-of-Oz study conducted by P. Wang et al. (2021) also perceived a text message as more efficient and easy to understand, followed by a symbol and then a virtual eyes interface.

Although differences in time to initiate crossing were not significant, a preference to start earlier in the presence of text interface was observed. In a study conducted by Ferenchak and Shafique (2021) participants reported higher preference for text eHMI on the grille had the highest average response followed by text on the roof, side mirror arrows, and LED windshield. Overall, text eHMIs outperformed non-textual interfaces in terms of preference. Bazilinskyy et al. (2021) investigated different combinations of eHMI in a video study. The authors found that red coloured eHMI reduced perceived safety to cross. Also, text messages led to higher perceived safety. Moreover, the authors concluded that for yielding vehicles, the placement of an eHMI on the windshield was a less efficient approach (participants less likely to cross).

Table 2.2: Visual eHMI concepts

A walking silhouette presented by Deb et al. (2018) and a similar to conventional pedestrian traffic light presented by Stadler et al. (2019)	 
A text message indicating advice ("After you") in Löcken et al. (2019) and a text message indicating intention in Deb et al. (2018)	 
A virtual eyes concept in Löcken et al. (2019)	

A virtual smile concept in Löcken et al. (2019)	
A crosswalk laser projection presented in Löcken et al. (2019) and alternative laser projections in Stadler et al. (2019)	 
Smart road infrastructure in Löcken et al. (2019)	

Löcken et al. (2019) investigated several visual designs, including both implementations communicating the message on the vehicle and on the road. Among the several approaches, a smart road indicating when to cross the street was perceived as safer and received higher trust while a virtual eyes concept was not preferred. A similar finding regarding the virtual eyes concept was also found by (Chang et al., 2017). Although this concept can improve perceived safety as it indicates vehicle's awareness regarding pedestrian's presence, it can also be confusing regarding the exact message communicated while it does not reflect basic information that influences crossing decisions such as, intention, speed or distance

(Chang et al., 2017). Li et al. (2018) investigated the impact of communicating the urgency of a situation by colour differences. The authors reported that green and red colours were interpreted more correctly however, a flashing amber indication might be perceived as safe or confuse pedestrians. This latter finding is another example of the impact of ambiguousness in an eHMI design and the confusing message it may deliver.

In their study, Löcken et al. (2019) concluded that information presented on the road (e.g. smart crosswalk or a laser crosswalk projected by the vehicle) were more preferred compared to information presented on the autonomous vehicle, as the latter was more difficult to perceive from distance. On the other hand, Stadler et al. (2019) found that on-vehicle display-based eHMI can improve decision time and error rates when crossing to a greater extent, compared to on-road projections. Moreover, display-based implementations are more well received by pedestrians. Further research towards those directions could provide further insights regarding the most suitable placement of visual eHMI.

The type of message that needs to be communicated by the autonomous vehicles has been also investigated in the existing research. She et al. (2021) reported that a commanding style significantly improved trust compared to advisory and informative, and advisory also significantly improved trust compared to informative. Moreover, when a cross decision was expected, participants in the advisory and commanding style conditions were more likely to make a cross decision compared to those in the informative style condition. Hochman et al. (2020) considered in a study a status message (“Slowing” or “Driving”) and an advice message (“Cross” or “Don’t Cross”). Text messages were also background coloured either green or red, depending on the type of message. Colour helped pedestrians understand the FAV (Fully Autonomous Vehicle) intention. Using a green background e-HMI, pedestrians had higher errors when they received status messages compared to advice messages.

Colley et al. (2020) investigated different types of auditory messages which types were categorised as intent, awareness or directive (advisory) also including participants with visual impairments in their study. The spoken text was the most preferred as sound messages, without the visual context, were always perceived as warnings. Moreover, in some cases sounds frightened the participants. Considering the presence of pedestrians with visual impairment, the concept of authority was considered as important by the authors. These pedestrians may rely on seeing people and follow their advice hence, in the absence of a driver the autonomous vehicle should have this role. On the other hand, this cannot be achieved via messages that communicate awareness or intention.

Although studies indicate that eHMI can influence crossing behaviour, this outcome has not been universally established. P. Wang et al. (2021) found that despite the presence of an eHMI, participants still primarily based their decision to cross on motion cues of vehicle as the approaching speed and distance. Moreover, Hochman et al. (2020) reported a learning effect over time, with participants in their study gradually improving their behaviour. Also, the authors reported that pedestrians' decision making depends on a combination of the e-HMI implementation and the car distance. Li et al. (2018) also reported that despite the presence of an eHMI that indicated a warning level, vehicle kinematics were still the most important factor for a crossing decision. These findings are an indication that eHMI can have a supplementary role on enhancing pedestrians' decisions based on vehicle kinematics. This is also supported by findings reported by Kaleefathullah et al. (2020) where it was found that the dependence on the information presented by the eHMI could have negative safety impacts in case of a failure to deliver the correct message. In future connected systems, an alternative approach could be the use of mobile phone warnings to pedestrians, to discourage crossing. Studies (Pooya Rahimian et al., 2016; P. Rahimian et al., 2018) have shown that mobile phone warnings can mitigate the negative impact of distraction (e.g. texting). However, this strategy could lead to overreliance to technology and negatively affect situational awareness of pedestrians.

Figure 2.3: Pedestrian receiving warning on mobile phone (P. Rahimian et al., 2018)

Finally, the efficiency of eHMI and change in pedestrian behaviour might be also subject to individual traits and characteristics Deb et al. (2018). Pedestrians with poor knowledge of traffic rules (e.g. assuming that pedestrian always has the right-of-way) were found to accept shorter gaps and take longer to cross the road. The use of eHMI might not be beneficial in these cases. On the other hand, pedestrians who intentionally violate traffic rules were found to take longer time before initiating a crossing. A potential reason could be that autonomous vehicles is a new traffic element hence they are uncertain of whether it is safe to brake rules and cross. With respect to these situations, the use of eHMI may improve road safety.

- It is still debatable whether eHMI should be the primary source of information in situations of interactions between autonomous vehicles and pedestrians. However, the presence of eHMI may improve pedestrians' trust and confidence as it shows awareness for their presence.
- The message delivered by eHMI should be clear and unambiguous. This could be one of the main reasons that visual or audio text messages are preferred over symbols or generic sounds.
- A simple indication of awareness for pedestrians' presence may be confusing, while when vehicle's intention or a suggestion towards the pedestrian are communicated, uncertainty may be reduced.
- When there is uncertainty regarding the readability of a message from far distance, the colour of the eHMI has been found to have an effect (red colour makes less likely a cross decision)
- Complete reliance on the message communicated by an eHMI may have negative safety implications in case of system failure. It is argued that eHMI should be a secondary/complementary source of information, together with the motion cues of the vehicles.

2.4 Use of driving simulators in CAV research

2.4.1 Driving simulators pros and cons

Driving simulators have been introduced in the driving behaviour and road safety research as an alternative data collection method, to the self-report attitudinal and behavioural scales (Helman & Reed, 2015). Some typical applications in the context of conventional driving cover speeding behaviour, lateral position behaviour, breaking response time, complex driving behaviours (e.g. secondary task), driving behaviour of a specific group of drivers, physiological measures observations, ecological validity etc. (Fisher et al., 2011). Driving simulators are a very useful tool in the research of autonomous vehicles as most of the technologies have not been yet implemented in actual vehicles. Applications cover the use and efficiency of HMI and TOR, reaction times, driving behaviour after a TOR, trust and acceptance. Moreover, driving simulator and virtual reality approaches are used to examine autonomous vehicles and pedestrians interactions.

Driving simulators provide a series of benefits, compared to field observations or questionnaire surveys. First, the driving environment is totally controlled and repeatable, and offers the opportunity for the representation of scenarios and situations that would be impossible to be implemented in real life (e.g. performance under alcohol influence, fatigue, new vehicle technologies etc.). Also, when studies focus on specific traffic situations, it is possible to examine all cases of interest which might not be observed in real traffic data. Moreover, there is not real risk involved for the participant drivers, since the environment is virtual. Another strong advantage of them is the capability of recording data with high frequency and precision. (De Winter et al., 2012; Helman & Reed, 2015). Also, apart from the observed behaviour, it is possible for the researcher to obtain information about the personal characteristics of the drivers (sociodemographic, attitudes, emotions etc.) and relate them to the observed actions and decisions.

On the other hand, driving simulators also suffer from a number of limitations. First, drivers are aware that they drive in a virtual rather than a real environment, therefore there is no real danger and other safety issues. Thus, the lack of real risk, which was already referred as an advantage, can be also a disadvantage (De Winter et al., 2012). Also, any type of reckless behaviour or traffic violations would not have any other real-life implications e.g., receiving a ticket for speeding. Another issue that derives from the use of driving simulators is simulator sickness (Helman & Reed, 2015). Simulator sickness may result in dropouts from studies or the adoption of a driving behaviour that reduces its effect but does not necessarily represent the actual driving behaviour of a driver. With reference to De Winter et al (2012), the effects of simulator sickness can be reduced by limiting the horizontal field of view, applying short driving scenarios (less than 10 minutes) with sufficient breaks for rest between the different runs.

2.4.2 Driving simulator validity

The level of simulators' validity (the extent that they represent actual driving behaviour or approximate the driving experience), when used in research, training or policy making is a crucial issue. The validity of driving simulators is mostly assessed in two levels, the physical validity (or fidelity) and the behavioural/predictive validity (Godley et al., 2002; Jamson, 1999).

The physical validity (or fidelity) of driving simulators is related to their physical characteristics, layout and vehicle dynamics. The closer a simulator approximates the real-life driving conditions the more fidelity is considered to offer (Godley et al., 2002; Triggs, 1996). Moving-based (dynamic) simulators are considered to provide more realism than fixed-base simulators (Godley et al., 2002). The fidelity of driving simulators is distinguished in four different levels, low, medium and high (or very high) (Rudin-Brown et al., 2009). Low fidelity simulators are often personal computers or workstations that also include steering wheel and pedals. Their graphical visual representation is limited and only a few features of vehicle dynamics can be captured. Medium fidelity simulators achieve an improved representation of the driving experience, they feature better graphic visualisation and more realistic vehicle dynamics. High fidelity (or advanced) simulators offer the most realistic driving experience. This type of driving simulators provides the best representation of real-life stimuli. They can also feature attributes as physical movement, related to the simulated driving environment, and they typically provide at least 180° field of view. Finally, very high-fidelity simulators provide close to 360° field of view, realistic graphical representation of the driving environment and a moving base that is able to simulate the physical forces related to medium levels of acceleration and deceleration.

Table 2.3: Driving simulators of different fidelity levels

A low fidelity simulator environment (Brandenburg & Chuang, 2019)	

A medium fidelity fixed based simulator environment (Almallah et al., 2021)	 a) Rural Environment b) Urban Environment
University of Leeds Driving Simulator (UoLDS) – high fidelity, dynamic (Paschalidis et al., 2020)	

The second type of driving simulator validity (behavioural/predictive validity) refers to the comparison of driving behaviour between real and simulated driving. It is often assumed that driving simulator validity is related to the behavioural validity. However, this assumption is not always true (Godley et al., 2002) and the same behavioural validity may be achieved with a lower fidelity and cheaper simulator. Rudin-Brown et al. (2009) also highlighted that, except for the issue of cost, lower fidelity simulators may offer more advantages e.g., easiness in programming and extraction of data and they stress a potential need for the development of a driving simulator that will provide a specific level of fidelity but at the same time will provide a sufficient level of behavioural validity. Behavioural validity is further distinguished in absolute and relative validity (Blaauw, 1982; Fisher et al., 2011). The absolute validity is achieved when the numerical values, related to driving behaviour, between real and simulated driving are similar, while in the case of relative validity the numerical values in the two cases are different but they follow similar magnitudes and patterns.

In literature, there is a series of studies investigating behavioural validity comparing driving behaviour in real and simulated environments. Two of the most examined aspects of driving behaviour, regarding behavioural validity, are the speeding and lateral position behaviours. Godley et al. (2002), investigated behavioural validity in terms of speed. The authors confirmed relative but not absolute validity in terms of speed, comparing the same road layouts in a simulator and field environments. Towards the same direction, Yan et al. (2008) studied the validity of driving simulator in terms of speeding and safety. Their results revealed absolute validity between simulated and real driving. Bella (2008), tested behavioural validity for two-lane rural areas. The results of this study indicated the existence of relative validity but also of absolute validity for most of examined cases. Y. Wang et al. (2010) used a medium fidelity driving simulator, in order to investigate behavioural validity and confirmed relative but not absolute validity. Risto and Martens (2014) compared the differences in headway choice between an instrumented vehicle and driving simulator without however finding significant deviations. McGehee et al. (2000) compared drivers' reaction times in real and simulated environment and found statistical equivalence between the two cases. As a general conclusion, driving simulators have been found to provide relative validity (patterns in behaviour between simulation and real life are in the same direction) however, absolute validity is study dependent.

In the context of autonomous driving simulator fidelity and validity has received less attention. Eriksson et al. (2017) investigated reaction time to handover and takeover requests conducting simulator and on-road experiments. Drivers were in general faster in their responses in the on-road situation. Moreover, the authors did not find significant differences regarding workload, perceived usefulness and satisfaction between the two contexts. Yeo et al. (2020) examined different setups and found that VR and motion can improve feelings of presence and fidelity and reduce simulator sickness. Moreover, the authors suggested a combination of VR and Wizard-of-Oz approach as the most preferrable. Finally, Bellem et al. (2017) examined the use of driving simulators for investigating passenger comfort. Participants experienced a lane change and a deceleration manoeuvre in a dynamic simulator and a test track. The study confirmed overall relative validity in terms of comfort ratings and the authors explained the differences as a result of speed underestimation in simulator environments.

2.5 Autonomous public and shared transport systems

2.5.1 Research on the acceptance of autonomous shuttles

Autonomous vehicles are also expected to have the form of shared or public transport modes. Shared solutions have been investigated in the existing literature in the form of autonomous shuttles, allowing participants to directly experience a ride with the vehicle. Some examples of autonomous shuttles used in existing studies are illustrated in Figure 2.4. Direct experience is likely to have a positive impact on acceptance. For instance, Nordhoff et al. (2021) conclude that the majority of participants in their study formed a more positive opinion after experiencing (either rather positive or very positive) followed by no change. Moreover, Paddeu et al. (2020) investigated perceived trust and comfort prior and after experiencing a trip with an autonomous shuttle and reported significant differences, while A. Salonen and Haavisto (2019) also suggested that perceived safety can be improved via direct experience. Bernhard et al. (2020) reported that the perceived experience in an autonomous minibus trial affected the acceptance of the system. The impact of experience can be also reflected on other aspects. For instance, Paddeu et al. (2021) found that willingness to pay for autonomous taxi may increase after experiencing an autonomous ride. This section presents some findings from existing studies with respect to passengers' perception and acceptance after experiencing autonomous shuttles.

Figure 2.4: Examples of autonomous shuttles used in studies, Nordhoff et al. (2019)- left and Madigan et al. (2017) - right

Perceived safety is one of the main issues towards the acceptance and use of autonomous public transport. Paddeu et al. (2020) found that participants reported higher levels of trust when Shared Autonomous Vehicles (SAV) had lower speed (8 km/h) versus higher speed (16 km/h) and participants used seats facing forward while the same findings were also concluded in terms of comfort. Hilgarter and Granig (2020) also concluded from interviews conducted after experiencing a Level-3 autonomous shuttle that technological advancement and lower speeds can increase perceived safety. This outcome is also supported by the findings of Bernhard et al. (2020). However, not all studies share the same findings with respect to speed. In a study of Chen (2019), participants evaluated positively five attributes of an autonomous shuttle service in terms of overall satisfaction, namely, speed, stability and comfort, safety, convenience, and information clarity. However, the lowest score was observed with respect to speed (15 km/h). Speed was evaluated as low and hence, that implementation of an autonomous shuttle was considered unsuitable as part of the overall road transport system. Molina et al. (2021) collected responses after experiencing an autonomous shuttle and reported that passengers were comfortable at higher speeds (up to 50 km/h). In a study of Lundgren et al. (2020) participants reported lower scores for comfort and speed compared to safety, usefulness and satisfaction (with an operating speed of 18 km/h); the service was not regarded as competitive at lower speeds. Similarly, speed of 15 km/h was evaluated as very low in a study conducted by Bernhard et al. (2020), while the majority of participants perceived the service as safe. However, higher perceived safety may be misleading as it could be driven by the lower speeds that participants experience in field studies (Paddeu et al., 2021). Lower speeds can be also undesired from other road users as they can have a negative impact on traffic flow and increase congestion (Rombaut et al., 2020).

Aside the issue of speed, there are still remaining issues regarding safety, security and vehicle performance. For instance, Molina et al. (2021) mentioned that although the vast majority of their participants in this study felt confidence regarding the performance of the system there were still some concerns regarding the absence of a human monitoring the system in future fully autonomous shuttle or loss of GPS signal. Moreover, deviations from expected behaviour (e.g. compared to a human driver) can demotivate the use of the system as it may be perceived as an indication of its limited capabilities. Nordhoff et al. (2019) reported that limitations in capabilities such as obstacle avoidance or requirement for human intervention in case of deviations from the pre-planned trajectory (especially at Level 3 of automation) can negatively impact perception. Given that technology is still immature, experiencing lower levels could enhance acceptance. Hilgarter and Granig (2020) also suggested that due to limited experience of passengers with the autonomous systems, the presence of a human controller who can intervene in case of emergency can increase trust. Nordhoff et al. (2021) also reported that supervision is an important factor for the acceptance of autonomous shuttles. The authors found that participants preferred remote supervision versus onboard supervision. Perceived level of control could be a factor affecting acceptance, hence with reference to this study solutions as buttons or apps to communicate with external control rooms could enhance positive attitudes. Ramseyer et al. (2018) conducted a shuttle experiment and reported that risk perception declines when there is a driver on board. Moreover, the authors found that shuttle passengers tended to humanise the shuttle (comparing its behaviour to a human driver) and notice aspects of manoeuvring that usually are not observed as turning or braking behaviour. A. O. Salonen (2018) concluded that although an autonomous shuttle may be perceived safer compared to a conventional bus, sense of lack of security (e.g., for cases of harassment or violence) in the absence of a human driver remains high. Another aspect that can raise concerns is the management of emergency situations. Nordhoff et al. (2021) also suggested that willingness to share with others can affect intention to use. Finally, Piatkowski (2021) suggested that some additional features have been as security cameras, arrival clock and route map have been reported as desirable.

Apart from the technical and safety issues that may exist owing to the novelty of this technology, autonomous public transport will be ultimately evaluated by potential passengers as part of the transport system based on its efficiency and reliability. Participants in a study of Nordhoff et al. (2021) found autonomous shuttles easy to use, however their interaction with the system was within a simple context. Upon their implementation, new issues will arise as identifying, booking, ticketing, entering, getting seated, and leaving the shuttle. A. Salonen and Haavisto (2019) conducted a series of interviews to passengers who experienced an autonomous shuttle. The authors concluded that the service was perceived as safe while the absence of a human driver was not perceived to have negative safety implications. On the other hand, available routes and flexibility were the most important factors to use the service. Nordhoff et al. (2021) found that compatibility with current mobility patterns is an important factor for intention to use autonomous public transport while Molina et al. (2021) reported that autonomous shuttles (or public transport in general) are perceived as a useful element in an integrated multi-modal transport system. To that end, lower speeds may reduce the perceived usefulness as they could lead to longer travel times and hence less efficiency. Nordhoff et al. (2019) also found a positive attitude towards using autonomous shuttles, especially if used as feeders towards other public transport systems. The complimentary role of autonomous shuttles was also illustrated in a study of Hilgarter and Granig (2020) as participants from rural areas were more positive towards autonomous solutions mainly due to lack access to public transport. The idea of autonomous shuttles as an integral part of a multi-modal public transport system was also found in a study of Piatkowski (2021). Finally, factors as environmental friendliness can also promote the use of autonomous shuttles (Bernhard et al., 2020).

Using the Unified Theory of Acceptance and Use of Technology (UTAUT), Madigan et al. (2017) found that users' enjoyment of the system, had a strong impact on the intention to use autonomous shuttles. Other significant factors were performance expectancy (a concept similar to perceived usefulness of the technology acceptance model), social influence and facilitating conditions (perceived capability to use autonomous shuttle). On the other hand, effort expectancy (similar to perceived ease of use) was not found to have a significant effect. Adapting some similar concepts with the previous study, Bernhard et al. (2020) found that performance expectancy followed by effort expectancy were the most influential factors for acceptance, together with socio-demographics as age and gender. Lundgren et al. (2020) also found that performance expectancy (including route reasons) and effort expectancy motivate not to use the services. Chen (2019) applied the technology acceptance model (TAM) in the context of autonomous shuttle acceptance after experiencing a trip. Results indicated that attitudes and perceived enjoyment are related to intention to use however, the author did not find a significant impact of trust.

A summary of some main findings regarding autonomous shuttles is presented below:

- Although perceived as safe in general, some studies have reported concerns related to higher speeds
- Technical aspects of the technology as loss of signal or obstacle avoidance need to be resolved to enhance acceptance
- The onboard presence of a human controller may be necessary at the early stages of deployment
- The presence of human staff can improve perceived security
- Monitoring (either onboard or remote) is a very important element. Passengers' perception of control needs to be ensured for instance with buttons that allow communication with a control centre
- Passengers may tend to compare the behaviour of an autonomous shuttle with a human driver and deviations from the latter may be perceived as negative
- Passengers mostly perceive autonomous shuttles as a complementary element of a multi-modal transport system. Their speed levels and overall performance needs to be competitive to increase behavioural intention to use.
- Autonomous shuttles (and public transport) should require the minimum possible effort from the passengers to use.

2.5.2 Users with disabilities and autonomous transportation

The introduction of fully autonomous vehicle transportation is expected to improve the mobility needs of people with disabilities independent travel as it is expected to provide a cheaper and faster transportation alternative (Hwang et al., 2020b) however safety issues still remain as reported for instance by Bennett et al. (2020) regarding people with visual impairments. With respect to blind people, Bennett et al. (2020) identified four main areas of interest namely, desire and hope for independence, scepticism over the AV meeting the needs of blind people, safety concerns and affordability. In a study of Brinkley et al. (2020), visually impaired respondents reported expectations for fewer crashes of lower severity after the introduction of autonomous vehicles. Other areas of improvement regarded emergency response, less traffic congestion, shorter travel times and fewer emissions. On the other hand, equipment failure, interactions with other road users (conventional vehicles, pedestrians), hacking and legal liabilities were areas of main concern. Although participants believed that their needs as blind users were considered in the development of autonomous vehicles, a strong majority was concerned about the implementation of laws that will discourage them from using this technology. Hwang et al. (2020a) reported some similar expected benefits as freedom of travel, cost savings, safety and improved accessibility. Hwang et al. (2020b) surveyed participants with disabilities and found that opinions were split with respect to overall safety while only a minority of participants perceived autonomous vehicles as safe under severe weather conditions.

When it comes to public transport and the existing systems available, people with disabilities have reported challenges both with respect to the public transport and the built environment. Regarding the former, issues related to the efficiency of services (connectivity, frequency etc.), poor design (lack of shelters, steep ramps, lack of lifts), driver attitude (e.g. unawareness of drivers for the additional needs, staff assistance in general), poor presentation of information, and in-vehicle facilities (narrow spaces, buses steep to get on/off etc.) (Low et al., 2020; Park & Chowdhury, 2018). Also, visually impaired users of current public transport systems rely in several cases on mobile technologies and applications that provide detailed audio information about their trip. Future systems need to ensure compatibility and enhancement of with these systems while they should also address potential limitation (e.g. loss of GPS signal).

The needs of disabled users remain similar also with respect to autonomous public transport systems. Patel et al. (2021) found in the context of incorporating AV in paratransit that these services need to provide accessibility to a wide range of destinations. Moreover, safety occurred as a main factor including human assistance to get on/off board. Also, it was reported that any apps related to the use of the system (e.g. booking) need to account for the needs of blind people and finally, the built environment around the service should consider users with mobility impairments. Kassens-Noor et al. (2021) investigated attitudes of people with disabilities towards future autonomous public transport systems and found that people with disabilities tend to be more dependent on public transport. In the same study, it was found that people with visual impairment were more likely to report willingness to use autonomous public transport, compared to people with mobility disabilities. Apart from typical issues shared among all users as safety and security, people with disabilities may have additional accommodation needs. Absence of human staff or assistance could be a factor to discourage people with disabilities from using autonomous public transport modes (Kassens-Noor et al., 2021). Similarly, Hwang et al. (2020b) reported that half the respondents of their study still preferred the presence of on-board staff on autonomous vehicles. This outcome is an additional indication that people with disabilities are concerned regarding accessibility and safety issues in the absence of a human in charge.

Figure 2.5: Examples of accessibility testing (Riggs & Pande, 2021)

Despite concerns, Hwang et al. (2020b) found in a stated preference study that people with disabilities favour the choice of autonomous vehicle solutions . An interesting aspect to be considered was that single-ride autonomous vehicles were more preferred to the shared vehicle alternative. Time savings but also willingness to share are aspects to be considered regarding this study. Also, respondents who had a negative opinion about the public transport system were more likely to favour autonomous solutions. This finding suggests that competitive autonomous transport solutions could attract a share of road users that do not favour existing public transportation alternatives. Cordts et al. (2021) also shared some similar findings with respect to autonomous vehicles in general, which were perceived as safe and well received by individuals with disabilities.

The main findings with respect to disabled people can be summarised as follows:

- Disabled people have the same safety and security concerns about autonomous vehicles which may be magnified depending on the type or level of their disability.
- The presence of human on-board is still very important as it does not only offer a sense of safety for someone to monitor the system but can also have an assistive role for disabled people.
- Shared autonomous vehicle solutions generate expectations of improved accessibility and flexibility. Competitive and account for the additional needs of disabled people can be more attractive.
- The design of the vehicles (space, ramps etc.) needs to cover the needs of people with disabilities.
- System information need to be provided in an inclusive way.
- Some people with disabilities rely on technological devices. Future systems need to ensure compatibility and offer solutions that address existing limitations.

2.6 Autonomous air mobility

The introduction of autonomous vehicles also extends to air transportation. A particular concept that has gained the interest of the research community is the Urban Air Mobility in the form of Personal Aerial Vehicles (PAV) or as usually mentioned (electric) vertical take-off and landing (eVTOL) vehicles. This technology could be the solution to the continuously increasing traffic volumes and passenger demand for reliable travel times and improved safety and security (Ahmed et al., 2021). Their acceptance has been mostly investigated in the form of questionnaire surveys and shared several common aspects with road transportation as cost, travel time, reliability of services and safety (Al Haddad et al., 2020). Goyal (2018) concluded in that in the context of unmanned air mobility (UAM) services, participants in a survey stated considerably higher preference for a human operated vehicle compared to the remotely operated and partial or fully automated options. The presence of a human pilot may increase passenger confidence due to the ability to intervene in case of an emergency situation. Moreover, people very likely perceive staff as a figure of authority that could prevent situations of harassment or violence. To that end, Goyal (2018) reported that sharing an UAM with strangers is conditionally accepted and additional assurances as a security screening process may need to be considered. The issue of safety in UAM can be more complex, compared to the road modes of transport, and factors as weather, distance, population density and type of terrain are likely to affect acceptance (Ragbir et al., 2020).

Figure 2.6: The use of VR (Janotta & Hogreve, 2021) and simulator (Perfect et al., 2017) in flying vehicle research

Moreover, due to their specific characteristics flying (urban) vehicles raise some further considerations. For instance, extensive presence of such vehicles creates concerns of visual pollution in the skies of the cities (Al Haddad et al., 2020) or increase in noise levels (Edwards & Price, 2020). Privacy is another perceived problem as people may feel exposed when being watched from above (Goyal, 2018). Privacy concerns raise also in the case of drones not only for passenger transport purposes, but in general. (Aydin, 2019; Clothier et al., 2015). Public may seek answers as to how, why and by whom UAVs will be operated (PytlikZillig et al., 2018). Cybersecurity is one issue reported with respect to the UAM (Goyal, 2018).

Challenges in the acceptance of UAM and autonomous flying vehicles or airplanes may be a result of the lack of knowledge regarding these systems or the lack of adequate education and sources of information. For instance, in a study of Reddy and DeLaurentis (2016), general public stated that movies and mainstream media are used as sources to learn about unmanned aircrafts. On the other hand, the knowledge of stakeholders comes from trade literature or personal experience. More recently, Aydin (2019) derived the same conclusion with respect to the sources used by the general public to obtain information about drones. This may create a distorted impression and negatively affect the acceptance of drones, UAVs and autonomous vehicles in general. Lack of knowledge and understanding of how autonomous systems operate can negatively impact acceptance. For instance, Al Haddad et al. (2020) reported that uncertainty about urban air taxis could affect the perceived time horizon for the adoption of this technology by the end users while Goyal (2018) mentioned that a part of the public is not willing to be an early adopter till these autonomous vehicle technologies prove safe. To that end, Rice et al. (2019) mentioned that familiarity has a positive correlation in the willingness to use autonomous flights. People are expected to be more acceptable towards these technologies as their familiarity increases.

Unlike the positive impact on acceptance after experiencing an autonomous vehicle (Liu & Xu, 2020) the same effect was not found in a study conducted by Astfalk et al. (2021). In particular, the authors did not find significant differences in the perception about an air taxi after participants observing a flight demonstration of it. The authors concluded that alternative approaches of experiencing an air taxi as virtual reality (VR) could potentially help in improving attitudes and acceptance. The concept of VR or simulator testing of UAM and eVTOL has received less attention compared to road autonomous vehicles. Marayong et al. (2020) presented a simulator approach based on a Cave Automatic Virtual Environment (CAVE) environment without however reporting results of a particular case study. Janotta and Hogreve (2021) conducted a VR experiment to investigate attitudes towards air taxis. Results showed overall a positive experience although safety concerns were also reported. Persson (2019) simulated and tested the dynamics of different types of aerial vehicles in a VR environment combined with a motion system. The authors collected physiological signals related to stress as heart rate or electrodermal activity, without however providing the results of a comparison across the tested cases. This study consists of an initial indication with respect to the positive impact of VR on acceptance, although further research is required. Moreover, the aforementioned study was desktop computer based hence and did not capture impacts as vehicle dynamics that would require more advanced facilities. Perfect et al. (2017) presented a training program for PAV while investigating the potential for mass adoption of these vehicles. The authors organised their training around 24 key skills required to pilot a flying vehicle which were distributed in four lessons. The authors tested their training approach in a study using the Heliflight-R full motion flight simulator at the University of Liverpool. Participants perceived the training as useful and straightforward but challenging. The authors compared performance with a previous study and found benefits from training especially regarding the skills of landing and deceleration descent. Although the scenario of PAV may be distant, direct experience of its operation could assist in the acceptance of these vehicles or other services as air taxis.

Some main findings with respect to autonomous air mobility are summarised below:

- Autonomous air systems are perceived as any other transport systems in terms of cost, travel time and efficiency hence, they need to be competitive compared to road alternatives.
- The absence of a human pilot may have additional safety impacts given the nature of the system (not ground based).
- Similar to other road alternatives, the absence of human staff can reduce perceived security (e.g. violence, harassment).
- Negative impact of UAM can extend to other aspects as visual or noise pollution and privacy (watched by users of UAM when on the ground).
- There is lack of accurate information and education regarding flying autonomous vehicles which can reduce their acceptance.
- Direct experience with an autonomous system has been found to increase acceptance of road alternatives. This should be incorporated in future research of autonomous air systems as the majority of existing studies relies on questionnaire surveys.

3 Findings from Driving simulator

3.1 Overview

3.1.1 Context

The adoption of CAVs is targeted at various societal benefits such as reducing pollution (Bansal et al., 2016; Anderson et al., 2014), traffic accidents caused by driver error (NHTSA, 2008) and increasing mobility and personal safety (Anderson et al., 2014). However, some concerns include travel safety (Bansal & Kockelman,2018), personal comfort while driving (Kyriakidis, Happee & DeWinter, 2015), vehicle hacking (Kennedy, 2016; Tennant et al., 2017) and data privacy (Collingwood, 2017; Howard & Dai, 2014). While the public generally agrees that CAVs are more secure than conventional modes of transport (Liu, Yang & Xu, 2019; Becker & Axhausen, 2017), there are also concerns about possible equipment failures (SeapineSoftware, 2014; Bansal et al, 2016) and lack of traffic control. In summary, the extensive data on people's perceptions of the consequences of mass adoption of CAV and personal uses show that opinions are mitigated. Therefore, we decided to conduct a simulated autonomous driving experiment with the aim of confronting people with a level 3 semi-autonomous individual car and finding out their acceptability, attitude, and behaviour towards L3 CAVs.

3.1.2 Purpose of the study

The main objective of the study was to evaluate the behaviours, attitudes, and acceptance of the drivers regarding level 3 autonomous vehicles. The specific objectives were defined as:

- Research Question 1.1:
  How the HMI design affects efficiency and acceptance of taking over requests?
- Research Question 1.2:
  How the HMI design regarding the feedback about the vehicle perception of its environment affects passengers' trust?
- Research Question 1.3:
  Does driving experience affects the driver's likelihood to accept and use a L3 CAV?

To answer these different research questions, an experimental protocol and variables were defined (i.e., D4.1). The main variables established are:

1. The study population was divided into novice drivers (experience of less than 40,000 kilometres) and experienced drivers (experience of more than 40,000 kilometres); 40 000 km of driving experience was decided as enough for being classified as an experienced driver, because after 40 000 km, drivers are able to use their peripheral vision effectively in maintenance of lane-position (Summala et al., 1996; Lehtonen et al., 2014).
2. All participants have been exposed to two different combinations of signals for handover requests (HOR) and taking over requests. (TOR): graphic and sound signals were identical, but combination A added a light signal while combination B added a vocal signal.

Table 3.1: Combinations of signals

	Combination A	**Combination B**
Graphic signal (touchscreen)	X	X
Sound signal (speaker)	X	X
Vocal signal (speaker)		X
Light signal (LED strip)	X	

3. Half of the participants were faced with a touchscreen displaying, in addition to the autonomous driving-related information, dynamic feedback about the vehicle's perception of its environment, while the other half were faced with a simpler touchscreen, keeping all autonomous driving-related information but displaying no feedback.

   Feedback information include:

   - Current legal speed
   - Type of road (urban, countryside or motorway)
   - Eventual event ahead (traffic jam, roadwork, parked vehicles)
   - 3-level quality of perception

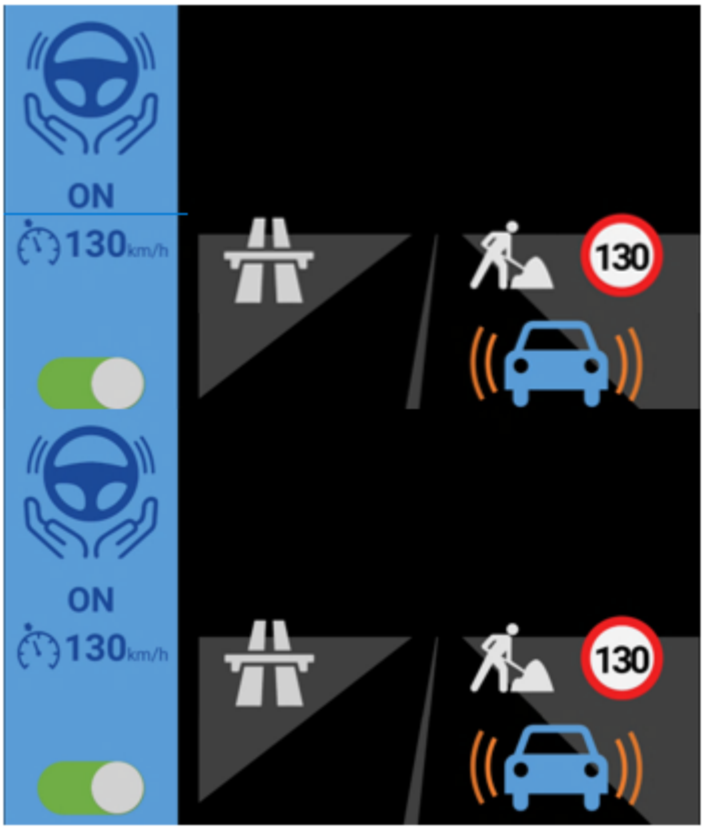


Figure 3.1: Screenshot of the touchscreen displaying feedback

3.1.3 Study population

42 volunteers consented freely to participate to this study. They were invited to the laboratory facilities for the experiments, which lasted a total of about two hours (considering the completion of the questionnaire and the interview). 3 individuals were excluded due to either trouble with completing the experiment or data acquisition issues. Finally, data from 39 subjects, including 21 females and 18 males, were studied.

The average age of participants was 33 (±16) years old, (36 ±14 for females; and 28 ±19 for males). The participants were diverse in terms of gender, age and socio-professional category (see Table 3.1). 24 participants (6 males and 18 females) with driving experience more than 40000 km were considered experienced drivers and 15 participants (12 males and 3 females) with experience less than 40000 km were classified as novice drivers. Table 3.2 presents the demographic characteristics of study population.

Table 3.2: Demographic characteristics of the study population

Demographic Variables	Number of participants	Percentage
Gender		
Male	18	(46%)
Female	21	(54%)
Age		
18 – 29	21	(54%)
30 - 39	6	(15%)
40 - 49	6	(15%)
50 - 59	2	(5%)
> 60	4	(10%)
Socio-professional class		
Unemployed	14	(36%)
Senior managers and intellectual professions	12	(31%)
Employees	8	(21%)
Intermediate professions and occupations	2	(5%)
Retired	2	(5%)
Tradesmen and business owners	1	(3%)
Educational level		
Primary education	3	(8%)
Secondary education	6	(15%)
Post-secondary education	9	(23%)
Bachelor's or equivalent	9	(23%)
Master’s Degree or higher	12	(31%)
Experience		
Driving experience <40000 km	15	(38%)
Driving experience >40000 km	24	(62%)
Receiving Feedback during simulation		
Yes	22	(56%)
No	17	(44%)

Participants' experience and knowledge about autonomous vehicles has been checked prior incorporation to avoid potential biases in the experiment. Figure 3.2 presents the past experience and knowledge of 39 participants regarding autonomous vehicles.

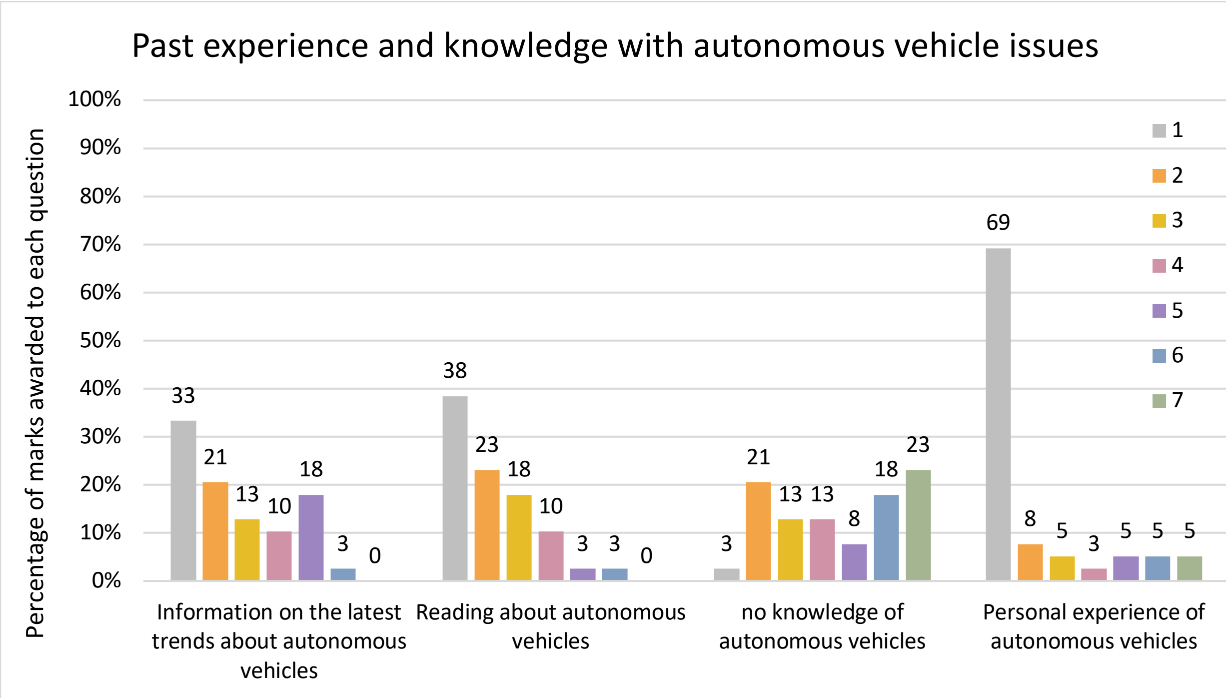


Figure 3.2: Past experience and knowledge regarding CAVs

All participants have a limited experience with autonomous vehicles (M= 2/7; SD= 1.9). Furthermore, they reported not being well informed about the latest trends concerning autonomous transportation (M= 2.7/7; SD= 1.6). More than 60% of our participants have never used autonomous functionalities and only 15.4% had already travelled on an autonomous shuttle. 38.5% of participants declared having autonomous functionalities in their car. It appears that a relatively naive population regarding knowledge/experience with the autonomous vehicle participated in the study.

3.1.4 Simulation system

This experiment was carried out on UTBM’s automotive driving simulator.

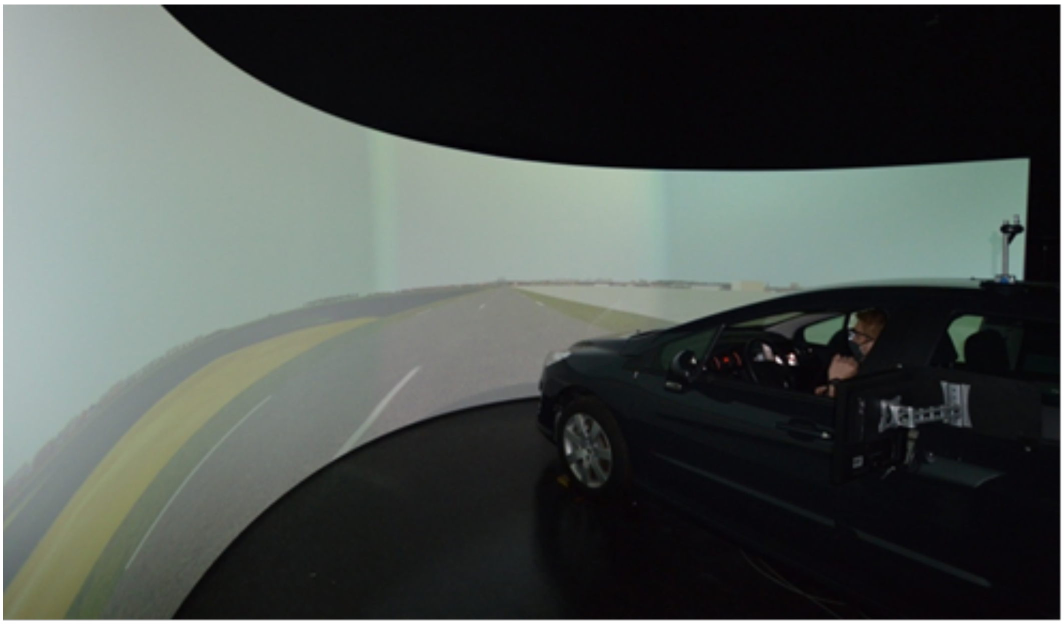


Figure 3.3: UTBM's driving simulator: cabin and front screen

The participants had to drive in a simulated level 3 autonomous car. An HMI delivering the autonomous driving-related signals and providing the de(activation) command has been co-design and co-developed by UBFC and Inetum. This is a multichannel HMI, able to deliver graphic, light, sound, and vocal signals.

Figure 3.4: UTBM's driving simulator: embedded multichannel HMI

A detailed description of the simulation system is available in the deliverable D4.1 Scenarios and experimental protocols.

3.1.5 Simulation scenario

After a tutorial and a resting time outside the car, the simulation scenarios chains 7 periods of manual driving alternated with 7 periods of autonomous driving, along a 15 to 20 minutes ride mainly composed of countryside roads and a motorway segment. The driver encounters crossroads, other vehicles and some more or less unusual events like roadworks, traffic jams, vehicles parked on the road etc.

A detailed description of the scenario is available in the deliverable D4.1 Scenarios and experimental protocols.

3.1.6 Metrics

Physiological and subjective parameters were measured in addition to video recording to characterise the subject's behaviour and to determine their acceptance regarding of the autonomous vehicle.

3.1.6.1 Subjective measures

Acceptability was measured through the questionnaires and interviews concerning attitude, trust, perceived risk, willingness to pay, change in mobility, ease of use, etc.

Acceptability was also studied in relation to the type of feedback provided by the graphical tablet HMI (second research question) and in relation to the driving experience of the participants (third research question).

In addition to descriptive statistics (mean, standard deviation, percentage), a statistical analysis was carried out to specify the significant differences observed between the different variables tested.

3.1.6.2 Objective measures

Several physiological data were also measured: Heart Rate (HR), Galvanic Skin Response (GSR),Wrist Motility (WM), and also pupils' diameters, eye movement and sight directions thanks to eye tracking glasses. We also calculated the reaction time: the time that subject perceives the handover/takeover signal and react to delegate or take back the control of driving (pushing the button on the tablet).

The Absolute Cardiac Cost (ACC) directly derivates from Heart Rate: after the tutorial and before starting the simulation, the participants were asked to remain seated for 5 minutes so we could measure the heart rate at rest. The following equation is used to determine ACC:

$$ACC_i = HR_i - HR_{rest}$$

Only ACC and GSR results are part of this deliverable. Eye tracking results need a longer exploitation and will be exploited in future publications.

The sample sizes for the objective variables are slightly smaller than for the subjective variables because we were forced to withdraw some participants data because of technical issues.

A detailed description of the metrics is available in the deliverable D4.1 Scenarios and experimental protocols.

3.2 Acceptability of L3 CAVs

3.2.1 Trust and attitudes towards CAVs

Figure 3.5 and Figure 3.6 present the results of level of trust measurement for 39 participants in the driving simulator.

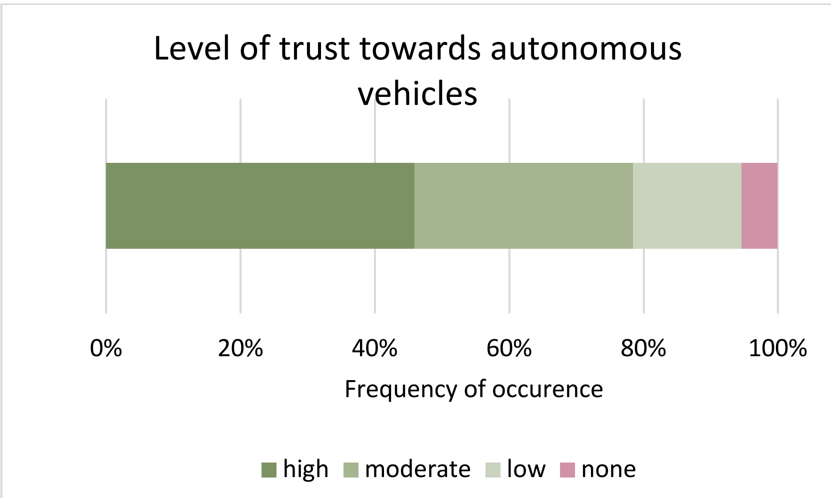


Figure 3.5: Level of trust towards CAVs

All participants reported feeling confident throughout the experience in the driving simulator. During the interview, 44% of participants reported a high general level of trust in the technology (some even reported "total confidence in the system"), 31% a moderate level, and 15% a low level of such trust. Following the experiment in the simulation, only 5% of participants declared that they did not trust the autonomous vehicle.

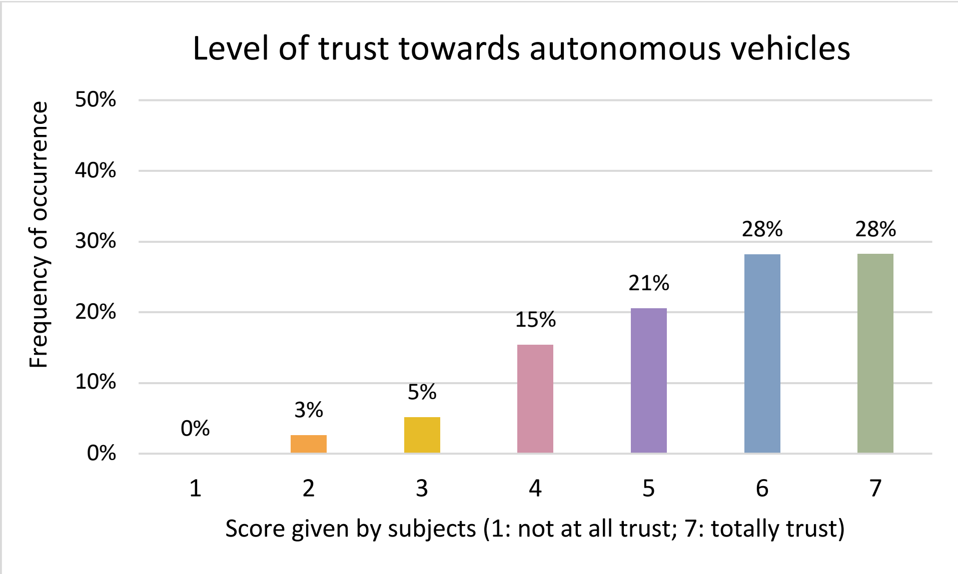


Figure 3.6: Level of trust towards CAVs

Similarly, they declared on a scale of 1 to 7 an average trust towards CAV of 5.5/7 (standard deviation = 1.3; Min = 2; Max = 7). Some statements confirm this effect such as: the subjects felt "super comfortable" and "didn't want to drive again" when the autonomous mode was available.

A large proportion of our participants found as safe the simulated autonomous driving system. It is worth to mention that 15.4% of the participants believed this to be because they were in a simulation situation (and therefore not really in danger). Indeed, 43.6% of the participants were ready to engage in some activity (such as checking their cell phone, making a phone call, reading a book, etc.) while in the autonomous mode and 35.9% intend to do so, which supports this idea of confidence and feeling safe in the autonomous mode. More than 20% of the respondents did not feel ready to shift their attention completely away from the road/vehicle. They said they "need time to adapt" or "to get to know how the autonomous car works".

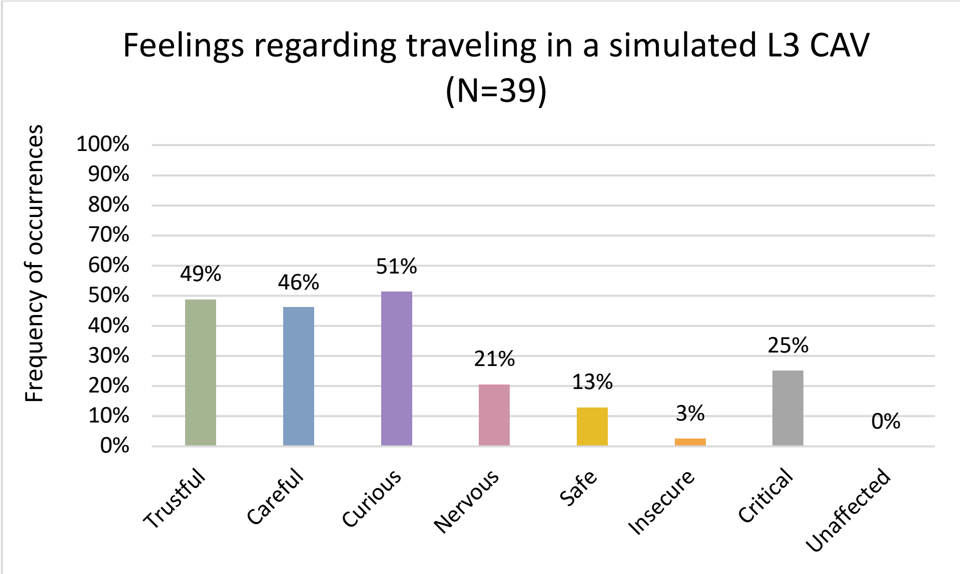


Figure 3.7: Participants' feelings during the experiment in CAV

Figure 3.7 shows the feelings of the participants during the experiment on CAV. 49% of participants reported feeling trustful, while 51% reported as curious, and 46% as careful. Only 3% of participants felt unsafe. The figure 4 shows more precisely the participants' level of trust during handover/takeover transitional periods.

Participants feel confident about switching autonomous driving on and off, as does the general level of trust in the functioning of the autonomous system (Figure 3.8). 74.4% of participants reported being very confident when interacting with the autonomous driving interface. Participants reported being "totally confident", " immediately confident", and having "clear information". Some of them described the system as "confident, natural, intuitive and easy to use". No subject challenged the interface of autonomous system or expressed any distrust.

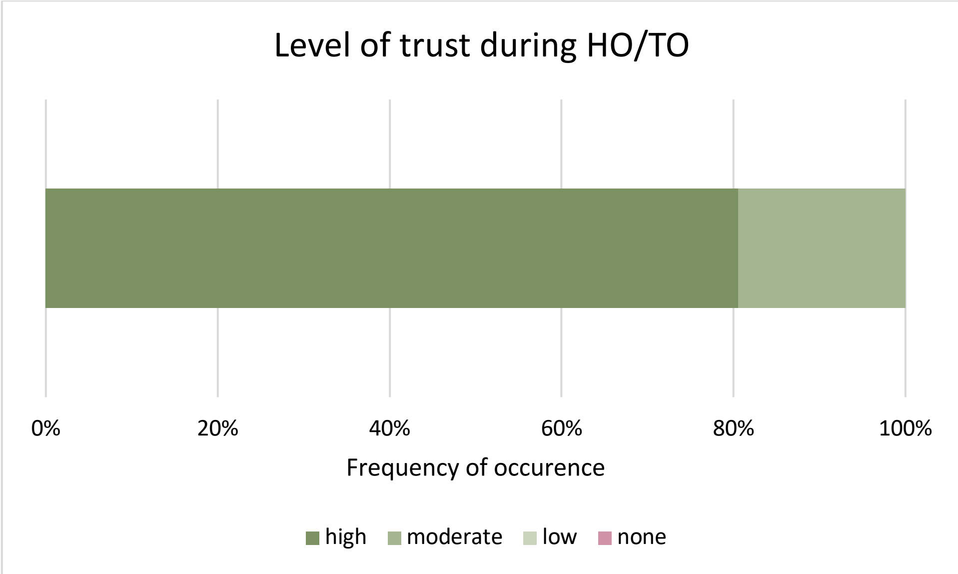


Figure 3.8: Level of trust towards CAVs during HO/TO

The measured trust indicators were positive. The participants felt generally confident during the autonomous driving simulation. It seems that the simulation experience has a bias to giving a realistic view of the safe running of an autonomous vehicle. This bias will be further investigated in the analysis of the questionnaires about the attitude and acceptability of the autonomous vehicle.

Table 3.3: ACC, GSR & reaction time of HO/TO

	Mean	SD	p-value
ACC (N=32)			
Handover	3.22	4.73	0.001***
Total Manual trials	-0.7	3.95	
Takeover	0.28	4.52	0.001***
Total autonomous trials	-2.51	3.61	
Handover	3.22	4.73	< 0.001***
Takeover	0.28	4.52	
GSR (N=31)			
Handover	6.64	6.00	0.915

Total Manual trials	6.68	6.03	
Takeover	6.56	6.30	0.388
Total autonomous trials	6.30	6.01	
Handover	6.64	6.00	0.013**
Takeover	6.56	6.30	
Reaction time (N=36)			
Handover	3.07	1.12	0.451
Takeover	3.03	0.69	

p < .05* p < .01** p < .001***

Physiological measurements at the time of handover and takeover are shown in Table 3.3. The heart rate (HR) of the 32 subjects was lower when required to take over control of the system than when required to release it (0.28 vs 3.22; p < 0.001). This could be explained because the heart rate was higher in the manual mode than in the autonomous mode; therefore, it remains higher at the time of handover than at the time of taking control of the system. Heart rate was higher during handover compared to that measured during the manual mode trials before receiving the handover signals and the results was statistically significant (p=0.001). The same results were observed at takeover, meaning that heart rate was higher compared to all autonomous driving trials (p=0.001). The increase in heart rate during handover or takeover could be due to the stress or cognitive load required to manage these situations.

The GSR that detect the different conductance of the skin when for example a person is under stress were significantly higher for the Handover than the Takeover (6.64 vs 6.56 µS; p = 0.013). The average of GSR at handover conditions was not significantly higher than in all autonomous driving trials (p=0.915). The average of GSR at takeover conditions was not significantly higher than in all autonomous driving trials (p=0.388). This difference was not significant between the average GSRs for all manual driving trials and handover (p=0.915). The GSR results imply that in the autonomous mode, subjects feel less under pressure and the stress increases at the moment of takeover. This pressure continues to increase in the manual mode and slightly decreases at the time of the hand-over.

Reaction time during handover and takeover are similar and show no significant difference (p=0.451).

3.2.2 Effect of the simulation experiment

Figure 3.9 shows the comparison between several categories of questions asked online before experiment and the questionnaire post simulation.

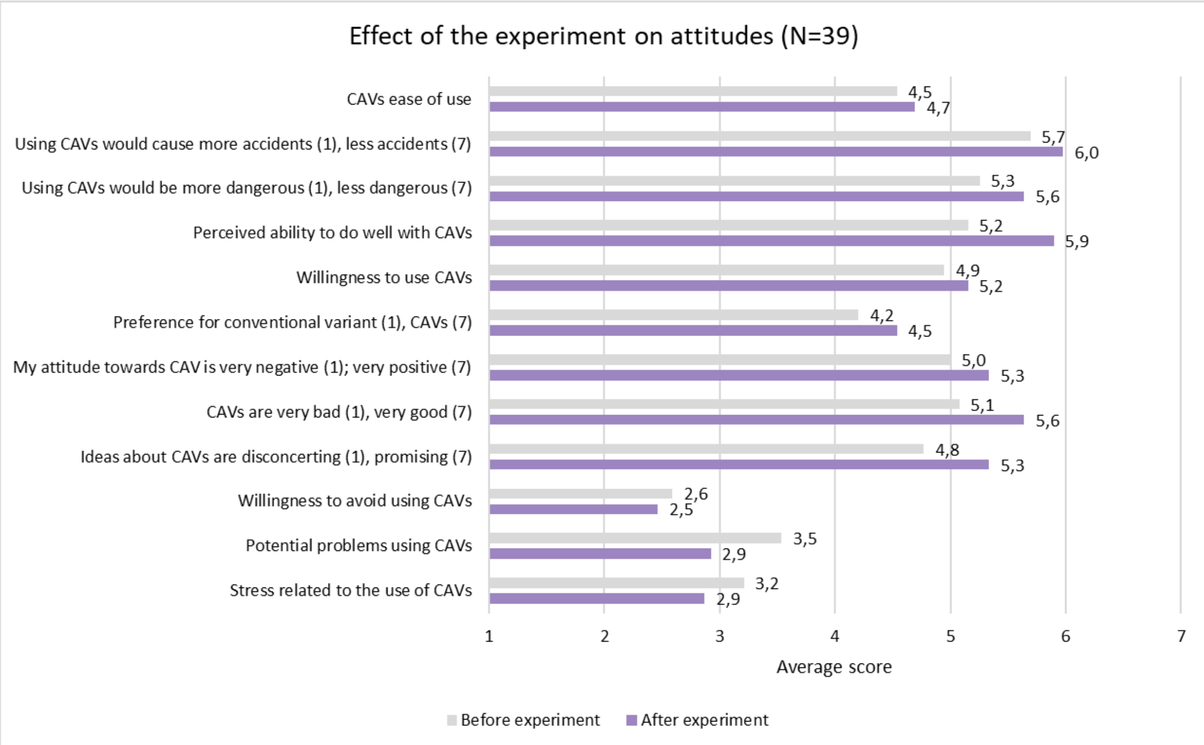


Figure 3.9: Effect of the experiment on attitudes

The participants' opinion on autonomous vehicles was generally more positive after the experiment compared to 15 days before the experiment. They felt less stressful using of autonomous vehicles after experiment compared to their earlier responses (3.2/7 vs. 2.9/7; p = 0.317). Although this difference was not statistically significant, it seems their opinion has changed after experiment: they reported being more willing to use autonomous vehicles (3.5/7 vs. 2.9/7; p = 0.084) and would avoid these vehicles to a smaller extent (2.6/7 vs. 2.5/7; p = 0.525). Possible reasons for non-significant results could be due to the small number of subjects surveyed.

General attitude of the participants regarding autonomous vehicles was more promising (4.8/7 vs. 5.3/7; p= 0.197), and they perceived better (5.1/7 vs. 5.6/7; p = 0.002) autonomous vehicle after experiment. We observed the spontaneous attitude of subject also increased after the experience (5.0/7 vs. 5.3/7; p = 0.056). Participants reported in interviews that they could "do whatever one wants in autonomous mode, while keeping an eye on the road, one enjoys the landscape." They are expecting an improvement of some aspects of life: " Perfect for people with reduced mobility or who need assistance." Participants would also prefer the autonomous vehicles compared to their conventional ones after the experiment and these results were statistically significant (4.2/7 vs. 4.5/7; p = 0.070).

Furthermore, all participants felt more motivated to use autonomous vehicles (4.9/7 vs. 5.2/7; p = 0.385) but also more capable of using them (5.2/7 vs. 5.9/7; p = 0.001) after the experiment. The words expressed in interviews support these data: "general ease of use", "intuitive car and tablet".

Regarding Perceived Risk, if a large proportion of the population used autonomous vehicles, participants would consider the trip less dangerous (5.3/7 vs. 5.6/7; p = 0.058) and the number of accidents lower (5.7/7 vs. 6.0/7; p = 0.043) after the experiment.

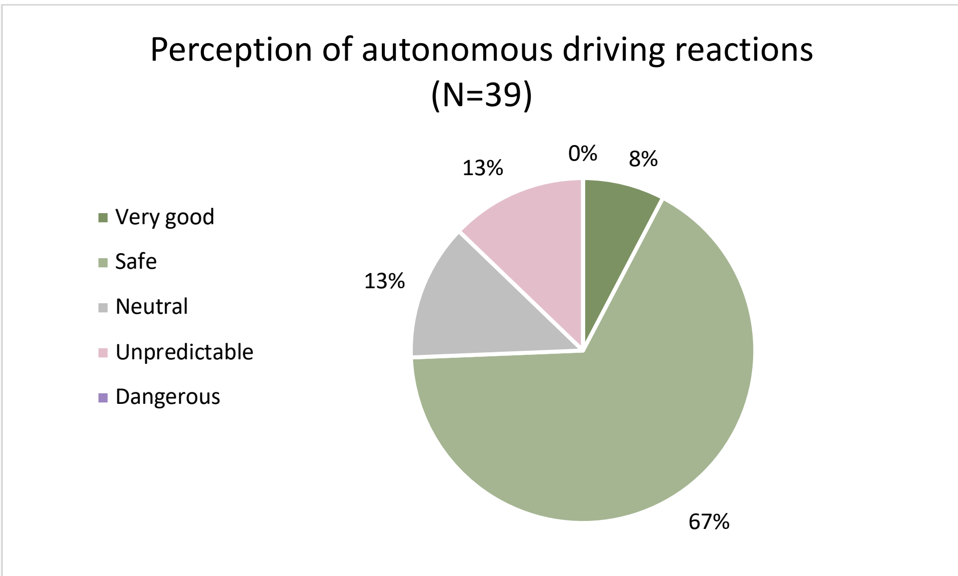


Figure 3.10: Risk perception of autonomous driving reactions

The participants were asked to rate the reactions of the autonomous vehicle (Figure 3.10): 13% found it unpredictable, 67% mentioned the safe reactions ("pleasant car management"), 8% qualified the reaction of the vehicle as very good ("great autonomous mode, speed/road are controlled and requires less attention"). One of the reasons that some subjects found CAV as unpredictable might be due to the driving simulator, as some of interviewees reported it "Too sensitive when accelerating", "the car turns at the last moment in autonomous mode", "the brake was too hard".

While the participants seemed to enjoy the simulation experience to learn about the autonomous vehicle and to get a feel for how such a mode of transport works, it is interesting to see if they would be willing to pay to have autonomous technology in their vehicle.

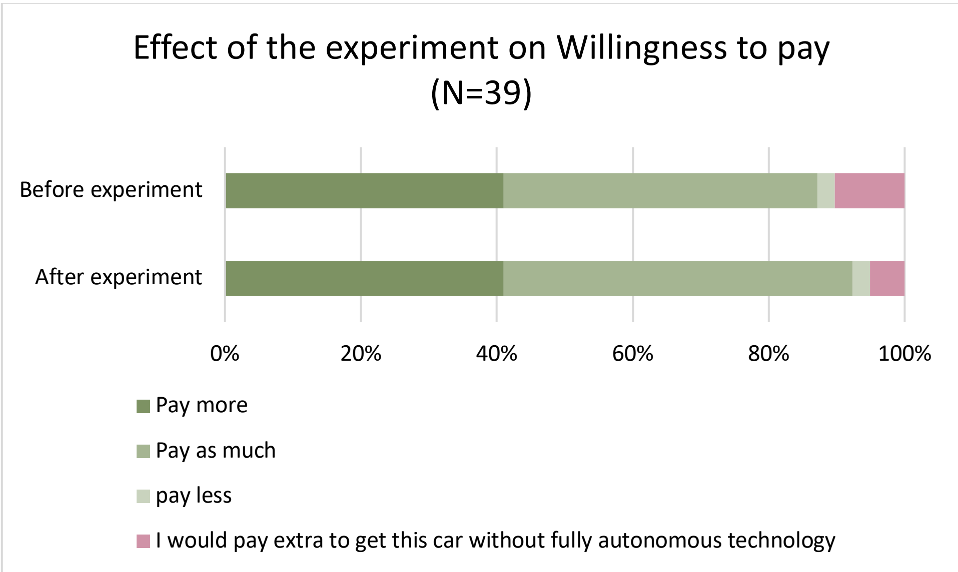


Figure 3.11: Willingness to pay for autonomous technology

The autonomous driving simulator experiment had no effect on participants' willingness to pay. Although they expressed confidence and a more positive attitude towards the autonomous vehicle, participants were no more willing to pay more (46% vs. 51%) to integrate autonomous technologies into their vehicle or when buying a future vehicle.

We could therefore see that the participants were generally confident during the experience in the driving simulator. The participants, who had limited experience and knowledge of autonomous vehicles, were able to get a more realistic idea of how an autonomous vehicle work. It seems that they have a positive acceptability towards the autonomous vehicle. Even though it was a simulated experience, their attitudes and feelings seem to have evolved positively after trying the simulated autonomous vehicle.

3.2.3 Effect of autonomous driving

Table 3.4 presents the physiological measurement of the subjects for 14 trials in manual and autonomous modes in the driving simulators.

Table 3.4: ACC & GSR during manual and autonomous driving

	Trials	Manual M (SD)	Autonomous M (SD)	P-value
ACC (N=32)	1	1.84 (4.45)	-3.25 (3.70)	
	2	-0.79 (4.64)	-2.55 (3.56)	
	3	-0.18 (4.23)	-2.22 (3.73)	
	4	-1.35 (4.50)	-2.07 (3.93)	
	5	0.99 (4.17)	-2.15 (4.67)	
	6	-1.67 (4.57)	-2.34 (4.74)	
	7	-2.31 (4.54)	-2.34 (3.80)	
	Total	-0.70 (3.95)	-2.51 (3.61)	0.005**
GSR (N=31)	1	6.40 (6.03)	5.87 (5.77)	
	2	6.15 (5.61)	5.83 (5.65)	
	3	6.26 (5.63)	6.15 (5.71)	
	4	6.23 (5.50)	5.85 (5.72)	
	5	6.54 (6.29)	6.50 (6.52)	
	6	6.42 (5.98)	5.95 (5.69)	
	7	6.25 (5.72)	6.09 (5.55)	
	Total	6.68 (6.03)	6.30 (6.01)	0.002**

$p < .05^{*}$ $p < .01^{**}$ $p<.001^{***}$

The heart rate (HR) was significancy lower during the autonomous driving phases than during the manual driving phases (-2.51 vs -0.70; p=0.005). These results indicate that during the different trials in manual driving mode, the HR of the subjects was statistically different (non-parametric Friedman's test between the ACC means p<0.001). This Heart rate was lower at the two last trials. However, the subjects' HR was not statistically different throughout the driving experience in the autonomous trials (p=0.1).

The GSR is significancy higher when driving in manual mode than when driving in autonomous mode (6.31 vs 5.98 µS; p=0.002). The results physiological analysis suggest that subjects feel more relaxed during autonomous driving phases than during manual driving phases. It can imply that that the subjects had less cognitive/physical workload when the vehicle is in autonomous mode than in manual mode.

However, we observed contradictory results, as the HR was not statistically different between trials in autonomous mode but did show a significant difference in manual mode. In contrast, the GSR was significantly different in the various autonomous driving trials but not in the manual driving trials. The possible reason for this result could be the impact of other study variables (duration of trials, the feedback from the on-board tablet, the signal configurations or different events faced such as road works, vehicle breakdown...) on the subjects which differed between the trial. The simulators might also impact on our results. Therefore, we propose to study these results in a pilot study.

3.2.4 Effect of driving experience

We want to study the possible differences between two types of participants: experienced drivers and novice ones regarding trust towards CAVs.

Experienced drivers reported a higher level of trust than novices. 50% of experienced drivers has a high level of trust towards the system compared to 33% of novices. Similarly, 27% of novices report a low level of trust towards the autonomous vehicle compared to only 8% of experienced drivers. It seems that the level of experience may influence trust towards the autonomous vehicle.

Furthermore, experienced drivers rated their trust towards the autonomous system as 5.8/7 while this score was 5.1/7 for novice drivers. The interviews also confirm these results as some experienced drivers speaking of "total confidence", " an intuitive system that people can have confidence in", while some novices were "more cautious", "waiting to see how the technology will develop" and want "time to have confidence".

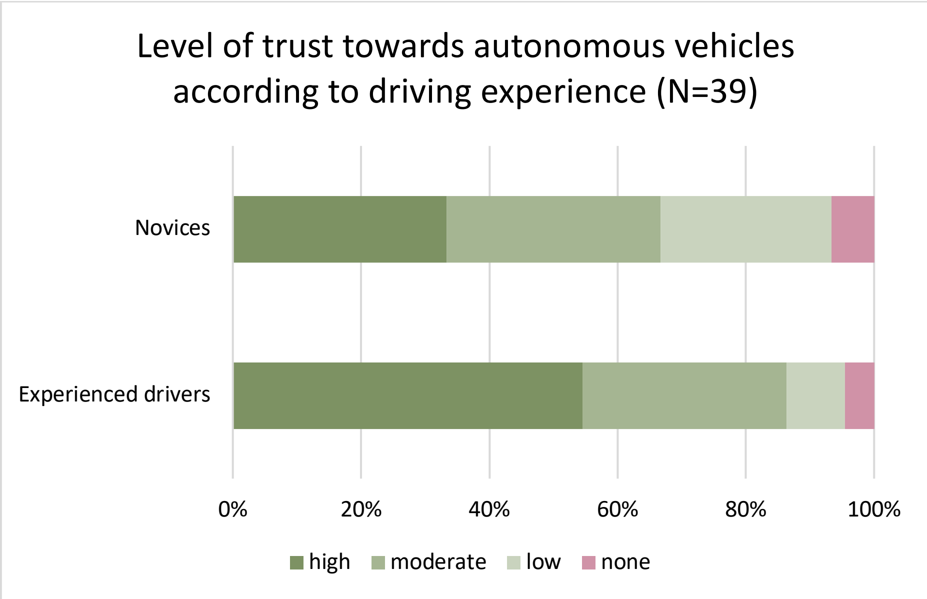


Figure 3.12: Level of trust towards CAVs according to driving experience

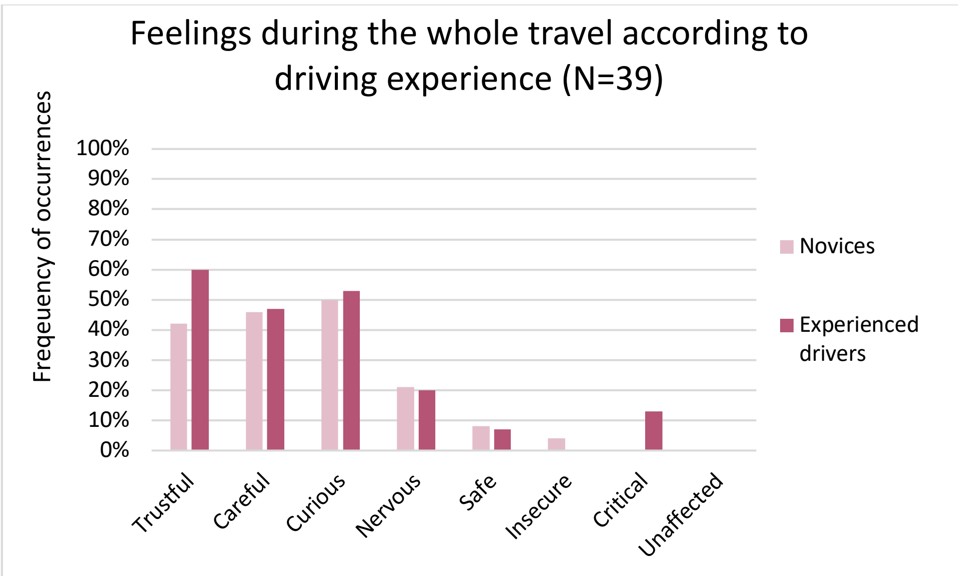


Figure 3.13: Feelings according to driving experience

This finding was confirmed by the questionnaire item on trust levels in which experienced drivers reported higher trust in the system than novice ones (60% versus 40%). The trust indicators studied show that experienced drivers had higher confidence to the autonomous vehicle. It seems that driving experience could influence attitude of people regarding autonomous vehicles.

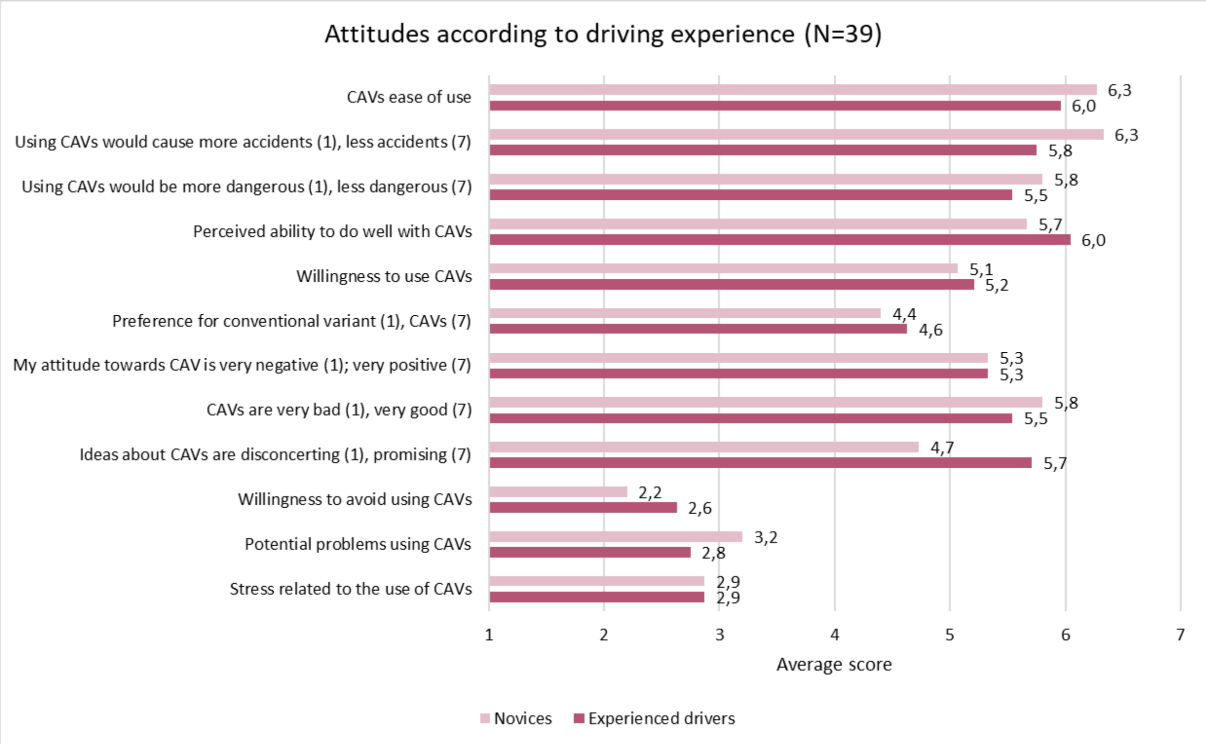


Figure 3.14: Attitude according to driving experience

The results concerning the general attitude of the participants seem to confirm a trend in favour of the experienced drivers over the novices. Some items show that experienced drivers are more positive towards the autonomous vehicle than novice participants. While the items related to the attitude of the participants do not seem to show any difference, it can be noted that the novices feel that they have less difficulties in using autonomous vehicles than the experienced drivers (3.2/7 vs. 2.8/7). The trend observed previously is confirmed since experienced drivers significantly consider autonomous vehicles as more promising (4.7/7 vs. 5.7/7; $p = 0.042$) and would prefer to use this type of vehicle for their journey (4.4/7 vs. 4.6/7) unlike novices. In the category "Modification of mobility behaviour", experienced drivers believe that they would use autonomous vehicles more readily (5.2/7 vs. 5.1/7) and do better than novices (6.0/7 vs. 5.7/7). In addition, novice drivers perceive more risks in relation to autonomous vehicles than experienced ones. Novices perceive the use of autonomous vehicles by a large part of the population as more dangerous (5.8/7 vs. 5.5/7), but also the number of accidents in this situation as higher (6.3/7 vs. 5.7/7; $p = 0.048$) than experienced drivers. The following figure confirms the low perceived risk measured.

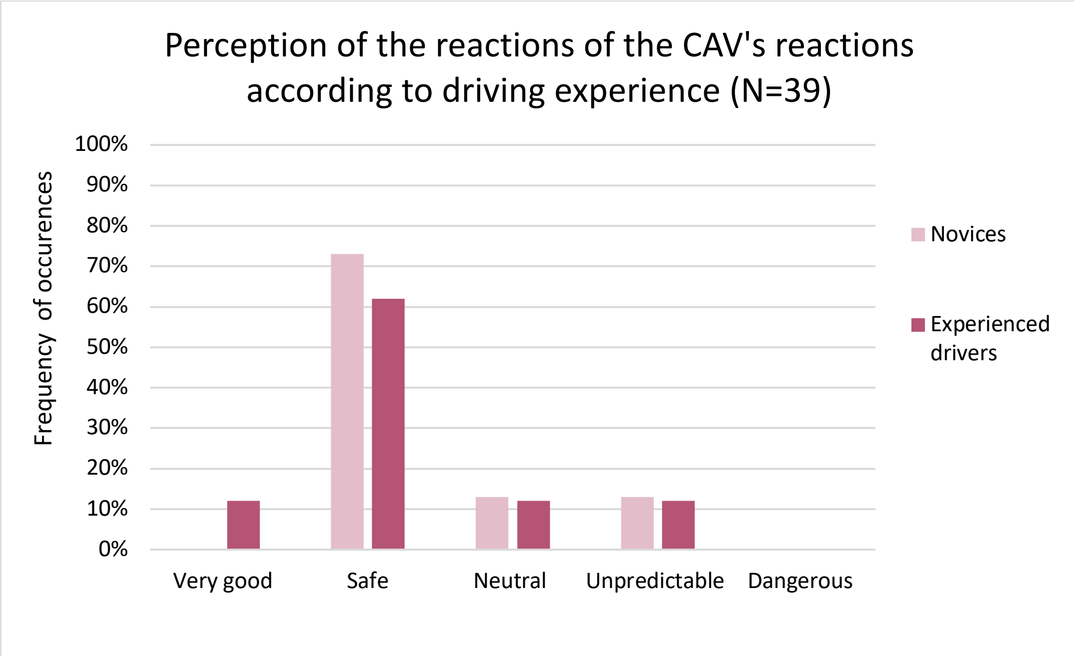


Figure 3.15: Perception of the CAV's reactions according to driving experience

The 39 participants involved in the experiment were asked about the reactions of the autonomous vehicle. Both experienced drivers and novices agreed that some of the CAV's reactions were unpredictable. We have already mentioned the limitations of the driving simulator. On the other hand, they also agree on the idea that the autonomous vehicle is safe (73% for novices and 62% for experienced drivers). It is interesting to note, however, that 12% of the experienced drivers declared that the autonomous system was very good, compared with 0% of the novices. It is interesting to see if this trend is observed in terms of behavioural intentions.

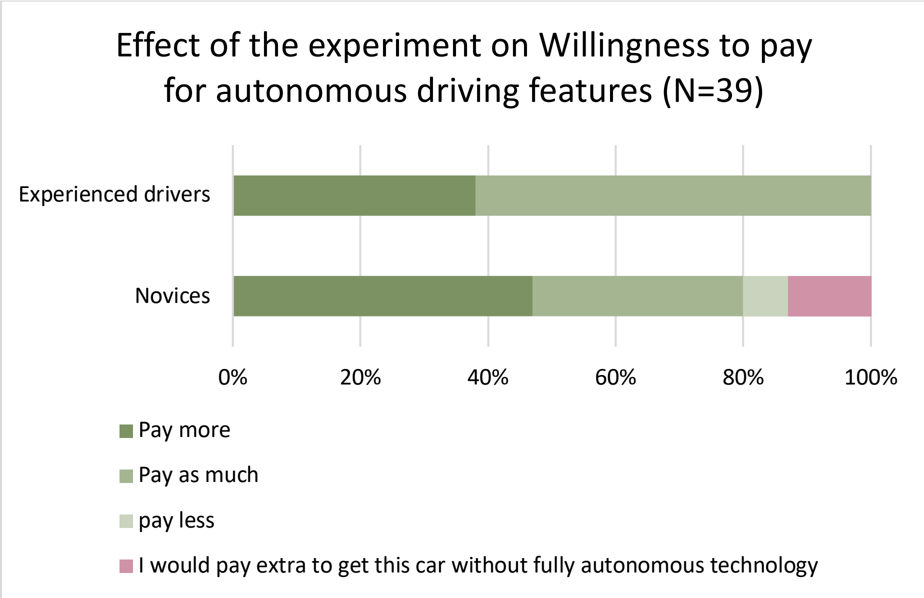


Figure 3.16: Effect of the experiment on willingness to pay

Participants, whether experienced or novice drivers, do not differ in their willingness to pay for autonomous technologies. Of note, 7% of novice drivers would be willing to pay less for autonomous technology and 13% would be willing to pay more for their car not to be equipped with autonomous technology. On the other hand, all experienced drivers say they are ready to pay as much or more to acquire autonomous technology for their personal vehicle.

To answer research question 3, we can say that the experienced drivers declare themselves more confident than the novice participants. Their acceptability of the autonomous vehicle seems to be better. Their attitude is more positive and their perception of the risk linked to autonomous vehicles is lower. So, if there is an overall positive effect of experimenting through a simulation on the acceptability of the autonomous vehicle, it seems to be higher among experienced drivers.

Table 3.5: ACC, GSR & reaction time of HO/TO, driving experience

		Mean	SD	P-value
ACC (N=32)				
Handover	Novices	4.53	4.87	0.116
	Experienced drivers	2.20	4.49	
Takeover	Novices	-0.09	4.26	0.750

	Experienced drivers	0.56	4.81	
GSR (N=31)				
Handover	Novices	8.83	6.91	0.068
	Experienced drivers	5.06	4.85	
Takeover	Novices	8.66	7.38	0.075
	Experienced drivers	5.05	5.09	
Reaction time (N=36)				
Handover	Novices	2.64	0.54	0.102
	Experienced drivers	3.38	1.33	
Takeover	Novices	3.07	0.76	0.874
	Experienced drivers	3.00	0.65	

p < .05* p < .01** p<.001***

Table 3.5 presents the physiological variables of the subjects during handover and takeover transitions according to their driving experience. The novices have a higher HR than experienced drivers during handover despite it was not statistically significant (p=0.116). On the other hand, we do not notice any difference in HR between novices and experienced drivers for taking over (p=0.750), maybe because HR during takeover was fairly lower than during handover for both novices and experimented drivers, after several minutes without active driving.

The GSR results confirmed that the experienced drivers feel less stress compared to the novice ones. Their mean GSR values increased less for both handover and takeover (5.06 vs 8.83 µS; p=0.068 and 5.05 vs 8.66 µS; p=0.075).

The reaction times between novice and experienced drivers for handover (p=0.102) and takeover (p=0.874) were statistically insignificant. However, it seems that novice drivers delegate driving faster (2.64 vs. 3.38 second). The physiological measurements confirm that experienced drivers probably feel more comfortable and relaxed to enjoy the autonomous driving than novices. The results of the interviews also indicated that novice drivers feel less trust than experienced ones during the experiment.

3.2.5 Effect of a feedback about the vehicle's perception of its environment

The reported level of trust is globally positive for the whole population, this positive effect is even more pronounced among experienced drivers. We were also interested in the effect of the type of information shared with the driver. Half of the participants completed the trip with a basic interface and the other half with a more comprehensive interface in terms of information related to the car and the environment. As a reminder, the subjects in the condition "without feedback" had the following information represented on the graphic tablet: a pictogram related to autonomous driving, a pushbutton to activate/deactivate autonomous driving, the vehicle represented. The subjects in the condition "with feedback" had all of this same information represented and in addition: the authorised speed, the speed of the vehicle, a pictogram representing the type of environment (countryside, city), and a coloured dynamic signal informing the quality of perception by the car of the environment. Research question 2 looks at the possible differences between the two conditions ("with feedback" vs. "without feedback").

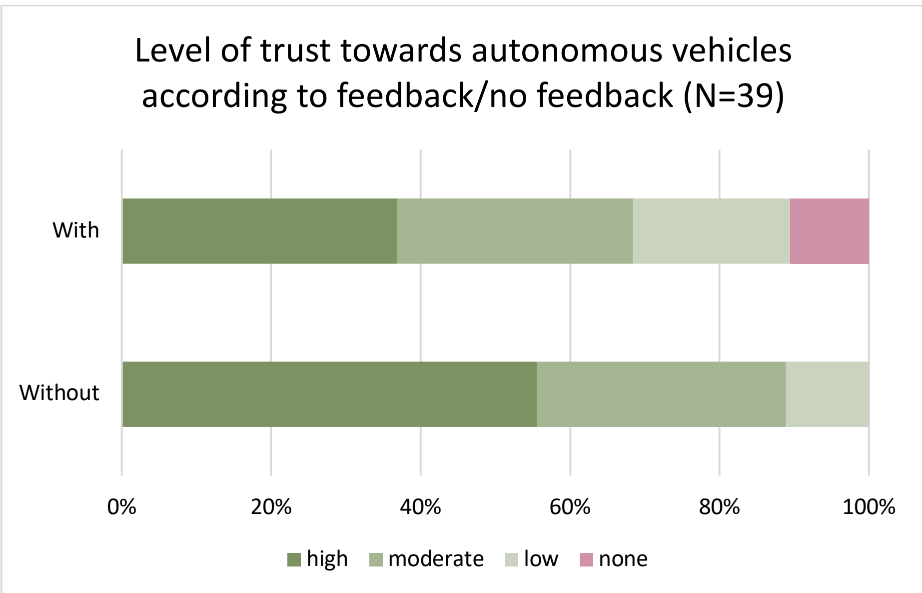


Figure 3.17: Level of trust towards CAVs according to feedback/no feedback

As seen previously, the whole population reported feeling quite trusting during the experiment. The scores on a Likert scale indicate no difference between the participants in the “with feedbacks” condition and the participants in the “without feedbacks” condition (5.5 vs. 5.5/7). Regarding the level of confidence reported, more participants who had the basic interface reported a high level of trust compared to participants who had the full interface (59% vs. 32%). Similarly, 9% of participants in the “with feedback” condition expressed that they did not trust the autonomous system. To explain this trend, we can speculate that an indicator specifying the good or bad perception of the environment by the car present on the full interface could have disturbed the drivers' trust. Indeed, participants in the “with feedback” condition declared that they "do not necessarily want to know that the car does not perceive its environment well" because "it is stressful to know that the car does not perceive the environment well". Some declared that they "do not want to have this information" and rather "let the car manage this information without sharing it with the driver".

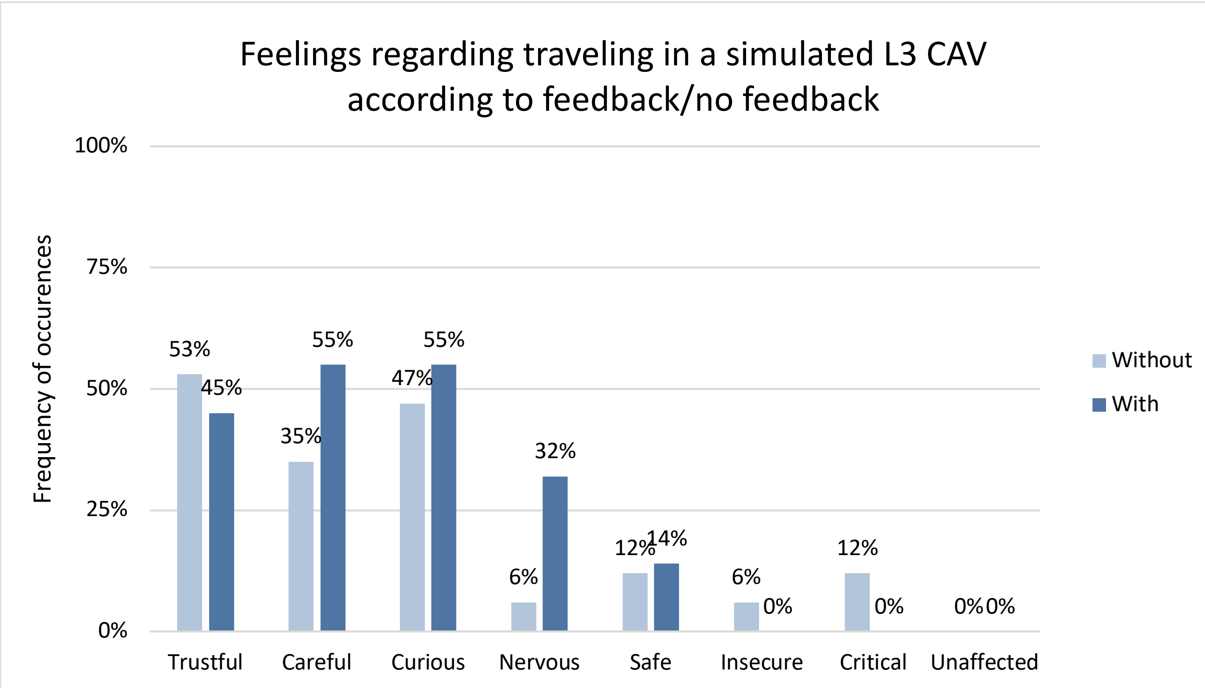


Figure 3.18: Feelings regarding traveling in a simulated L3 CAV according to feedback/no feedback.

Again, participants facing the “without feedback” interface have more higher trust scores than participants facing the “with feedback” one. They reported higher trust during the experiment (53% vs. 45%). Also, to confirm this trend, drivers who had the full interface (with feedback) were more careful (55% vs. 35%) and nervous (32% vs. 6%) than participants who had the basic interface (without feedback). The trend seems to be confirmed in terms of trust. The interviews confirm this. The “without feedback” interface seemed "very intuitive", "simple to understand", "well done" whereas the full interface, even if it was "well thought out and well done", people "did not understand what some of the pictograms" displayed on the touchscreen were for.

Table 3.6: ACC, GSR & reaction time, feedback/no feedback

	With feedback M(SD)	**Without feedback** M(SD)	**p-Value**
ACC (N=32)			
Handover	2.64 (3.58)	3.97 (5.96)	1
Takeover	0.46 (4.57)	0.04 (4.61)	0.837
Total Manual	-0.89 (3.71)	-0.46 (4.37)	0.955
Total Autonomous	-2.66 (3.92)	-2.31 (3.31)	0.925
GSR (N=31)			
Handover	6.74 (7.08)	6.49 (4.34)	0.679
Takeover	6.77 (7.54)	6.28 (4.32)	0.650
Total Manual	6.74 (7.08)	6.59 (4.45)	0.594
Total Autonomous	6.53 (7.23)	5.98 (4.00)	0.622

Reaction time (N=36)			
Handover	2.84 (0.73)	3.41 (1.47)	0.252
Takeover	3.11 (0.73)	2.92 (0.64)	0.427

p < .05* p < .01** p<.001***

We cannot see any effect of presence or absence of a feedback, informing about the vehicle's perception of its environment, on the physiological variables, nor on the reaction times. So, the eventual stress induced by concerning information about a low perception level of the vehicle shows no physiological effect.

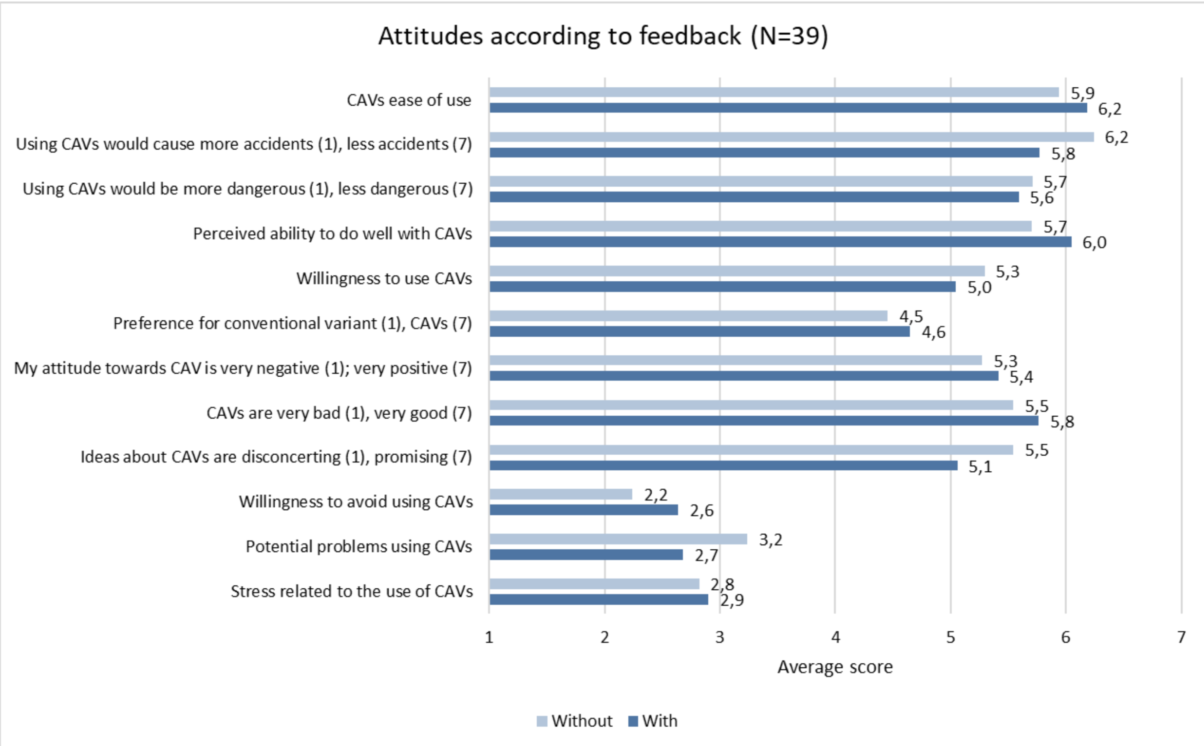


Figure 3.19: Attitudes according to feedback/no feedback

Figure 3.19 representing the answers to the post-simulation questionnaire according to the feedback (with or without) seems to indicate that subjects in "without feedback" condition have a better perception of autonomous vehicles than subjects in "with feedback" condition. The "Attitude" category shows that subjects in condition “with feedback” consider autonomous vehicles to be slightly more stressful than participants in condition “without feedback” (2.9 / 7 vs 2.8 / 7). Moreover, the participants exposed to the full interface wish to avoid more autonomous vehicles on the road (3.2 / 7 vs 2.7 / 7). The results obtained for the “General attitude” category indicate that subjects in condition “without feedback” perceive autonomous vehicles as more promising (5.5 / 7 vs. 5.1 / 7) than subjects in condition “with feedback”. These results show that subjects in no return condition perceive autonomous vehicles more positively. Also, these subjects judge autonomous vehicles to be less dangerous (5.7 / 7 vs 5.6 / 7) and capable of reducing the number of accidents (6.2 / 7 vs 5.7 / 7).

Some of the results obtained in the “Modified mobility behaviour” category seem to be in line with the previous results and indicate that participants in the condition of no return are slightly more likely to use autonomous vehicles if they are available compared to the subjects. in the condition with return (5.3 / 7 against 5.0 / 7). However, the results also indicate that subjects in the condition “with feedback” felt that they could cope more easily with autonomous vehicles than subjects in the condition “without feedback” (6.0 / 7 vs. 5.7 / 7).

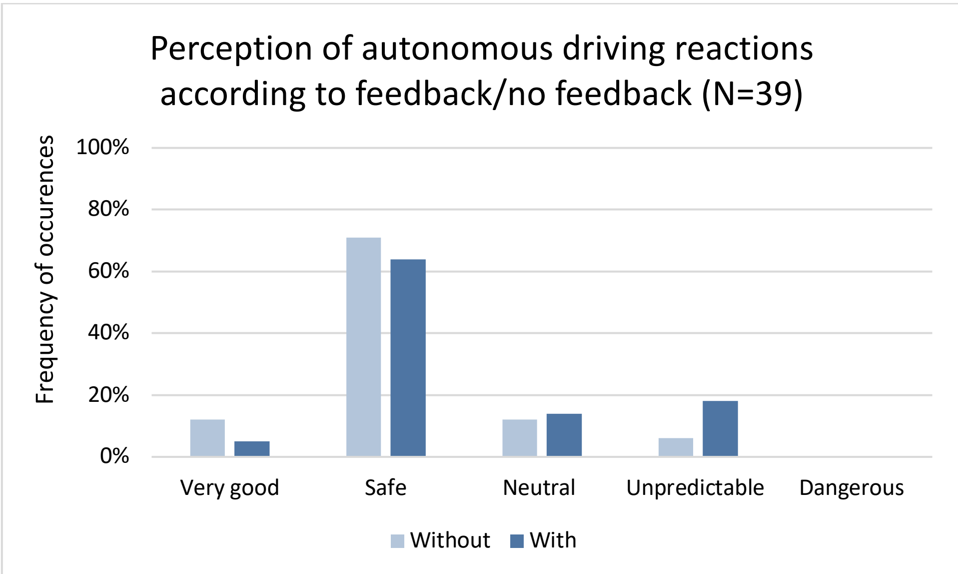


Figure 3.20: Perception of autonomous driving reactions according to feedback/no feedback

When participants are asked about perceived risks in terms of autonomous mobility, again, participants informed by the basic interface seem more favourable than participants exposed to the full interface. For example, there are more drivers in the no-feedback condition who rated the autonomous system as very good (12% vs. 5%) and safe (71% vs. 64%). Also, more of the subjects exposed to the full interface mentioned the unpredictable autonomous system than the participants exposed to the basic interface (18% vs. 6%). Once again, it seems that the trend is confirmed. Drivers who were informed by basic information have a more favourable acceptability of autonomous technology than participants exposed to more and more complete information. It is interesting to see if this trend is observed in terms of willingness to pay.

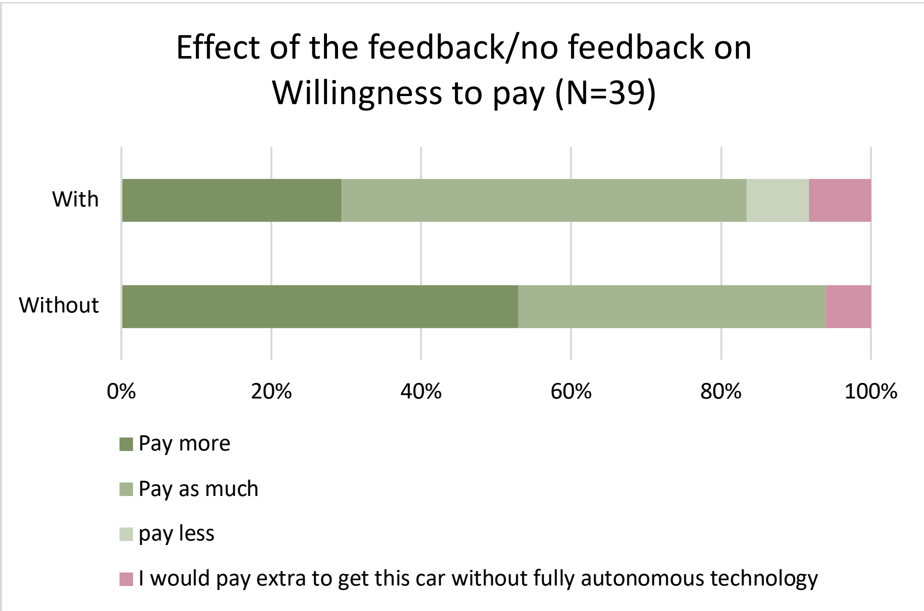


Figure 3.21: Willingness to pay for autonomous driving features according to feedback/no feedback

Here again, the participants of the “without feedback” condition are more likely to declare that they are ready to pay more in order to equip their personal vehicle with autonomous technology for a future purchase (53% vs. 32%).

3.3 Acceptance, effectiveness, and usability of HMIs

We have seen through the “with vs. without feedback” conditions that the CAV HMI had an impact on the acceptability and trust of the participants. We will now look at the overall impact of the different signals present in the HMI (graphical, audio and visual) on the participants' trust and acceptability, particularly during the TO and HO phases.

3.3.1 Comparison of the different signals

During the interviews, immediately after the experiment, the participants were asked what types of signals they had perceived during the TOR and HOR phases while driving in the simulator. Figure 3.22 showcases the frequency of their spontaneous answers.

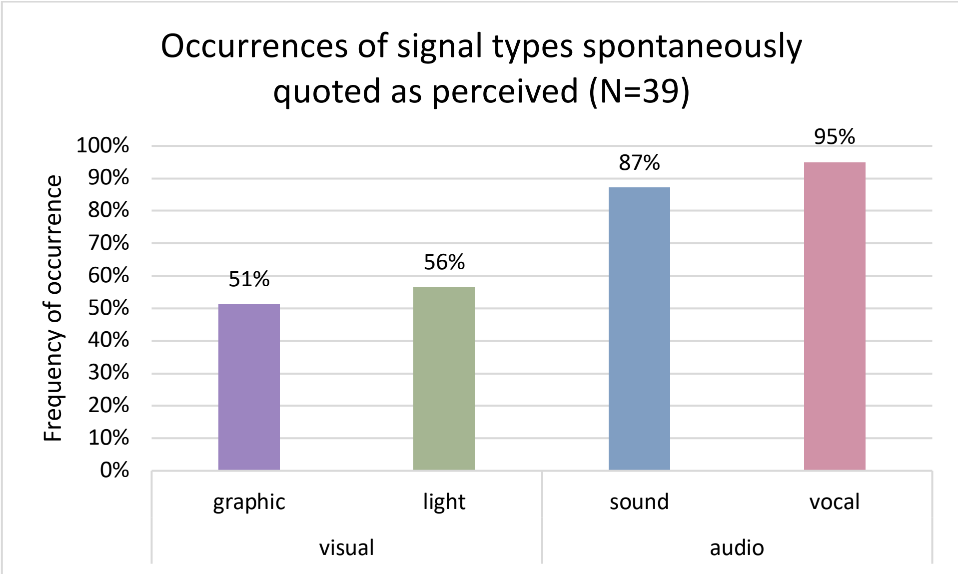


Figure 3.22: Signals perception

Spontaneously, most of the participants mentioned and remembered audio signals: 95% of the participants declared that they had heard the vocal signal and 87% the sound signal. In parallel, 56% of them perceived the light signal at the time of driving transitions and 51% said they saw the graphic signal on the touchscreen.

For most participants, the audio signal was "perceived first each time" and then "the graphic message was seen after". They were "more attentive to sound rather than to image".

Then, the participants were asked to specify effectiveness and satisfaction for each of the signals and for the 2 combinations of signals.

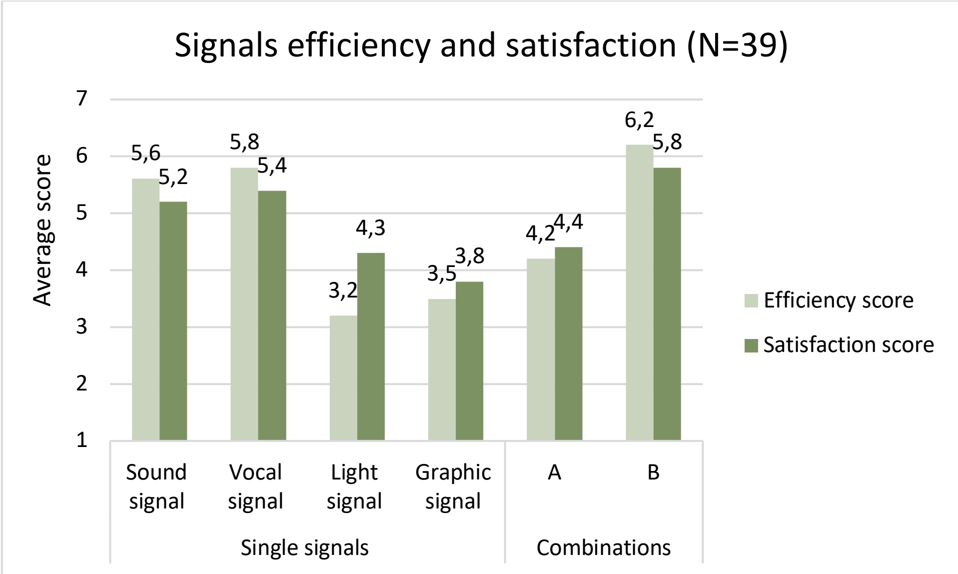


Figure 3.23: Efficiency and satisfaction of different signals

For the whole population sample, it clearly appears that audio signals are the most effective and the most satisfactory. The vocal signal is evaluated at 5.8/7 in terms of effectiveness: it is judged "the most relevant", "more relevant than the beep (sound signal)", "the most effective". The vocal signal is also the most satisfactory and pleasant (5.4/7): it is "the most precise" and "the clearest", even if it can be perceived by some subjects as "intrusive". The sound signal is also considered very effective (5.6/7): "relevant", "very effective". The pleasant sound emitted (5.2/7) is "non-intrusive", "soft to the ear", even though "there should be no noise in the car, people talking, music".

Graphic signal on the touchscreen is "less effective than the sound", "not the most effective" with a score of 3.5/7 on average. The subjects confirm that it is "a good complement to the audio". The light signal was also judged "less effective", or even "ineffective" during the TOR and HOR phases. On the other hand, 33% of participants said that it was "more relevant for confirming autonomous driving" is engaged. The subjects are divided on the satisfaction aspect of the light signal (4.3/7), some find it " aesthetic " and that it " does not disturb " and others consider it " disturbing " or even " blinding " in a dark simulation environment.

The trend is confirmed for the combinations of signals. Combination B (sound and vocal) is more effective (6.2/7) and more pleasant (5.8/7) than combination A (sound and light) (4.2 and 4.4/7) in the TOR and HOR phases.

The acceptability of the HMI is also studied by the usability and attractiveness of the interface.

Table 3.7: ACC, GSR & reaction time of HO/TO, signal combinations

		Mean	SD	P-value
ACC (N=32)				
Handover	A	3.89	4.46	0.011*
	B	2.38	5.86	
Takeover	A	0.52	4.63	0.178
	B	-0.23	5.09	
GSR (N=31)				
Handover	A	6.61	5.89	0.468
	B	6.69	6.25	
Takeover	A	6.44	6.20	0.542
	B	6.78	6.55	
Reaction time (N=36)				
Handover	A	3.24	1.66	0.362
	B	2.95	1.01	
Takeover	A	3.06	0.69	0.519
	B	3.04	0.81	

p < .05* p < .01** p<.001***

Table 3.7 indicates that the subjects' HR are lower when exposed to B signals combination, i.e. when they receive a vocal signal informing that they have to take over, than when they are exposed to A combination, i.e. when they receive a light signal informing, they have to take over control of the vehicle (3.89 vs 2.38; p = 0.011). We notice that HR results confirm interviews' results showing that the participants preferred the vocal signals compared to light signals, at least for handover/takeover requests.

In contrast to the HR results, the GSR results show no explicit trend according to the signals combinations.

Reaction times do not show a noticeable difference either.

Usability of the touchscreen was measured using the Usability Metric for User Experience (UMUX) questionnaire.

UMUX items were scaled between 1-point for "Strongly Disagree" to 7-point for "Strongly Agree." Participant scores were recoded to maintain a score from 0 to 6, using the method described by Finstad (2010): "Odd items are scored as [score - 1], and even items are scored as [7 - score]" (p. 326; Brackets in original). This method of subtraction is borrowed from the SUS and eliminates the negative/positive keying of the items. The sum of the four UMUX items was divided by 24, and then multiplied by 100 to achieve parity with the SUS score.

The calculated UMUX score is 76.2/100. This is a good usability score. It shows that the participants appreciated the usability of the graphical interface. In the interviews, participants referred to an "intuitive", "easy to use", "functional", "user-friendly" interface with "simple graphical elements".

Attractiveness of the interface was studied via the Attrackdiff questionnaire. Four categories are calculated:

- Pragmatic quality,
- Hedonic quality - stimulation,
- Hedonic quality - identity,
- Overall attractiveness.

Figure 3.24 shows that the pragmatic quality (1.4) and hedonic quality (1.2) categories were positively judged by the participants with scores above 1 (Van Meyel, 2021; Tolle, 2020). The hedonic quality - identity (0) and overall attractiveness (0.5) categories were judged neutral.

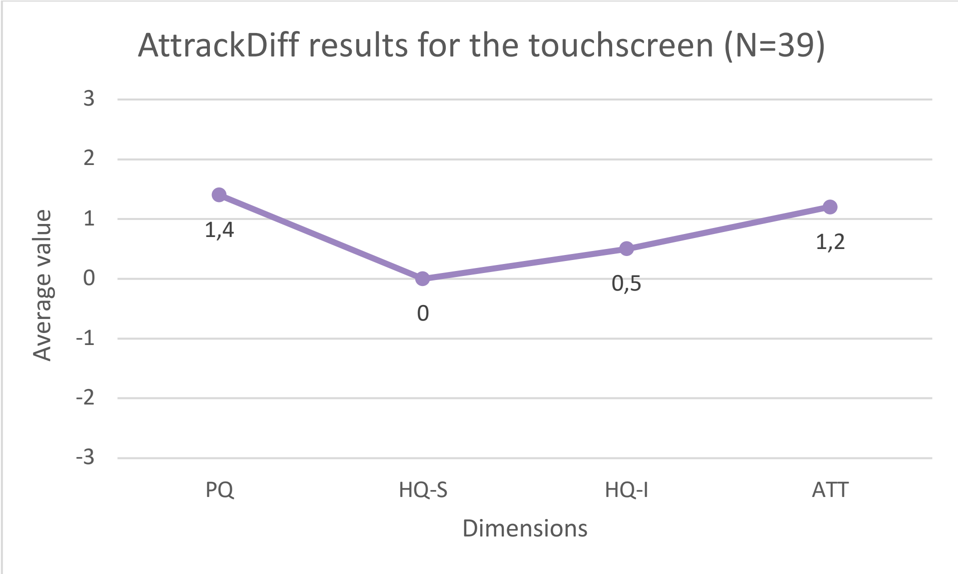


Figure 3.24: AttrakDiff results for the touchscreen.

In order to address research question 3, the scores related to the HMI will be analysed according to the participants' driving experience.

3.3.2 Effect of driving experience

3.3.2.1 On signal efficiency and satisfaction

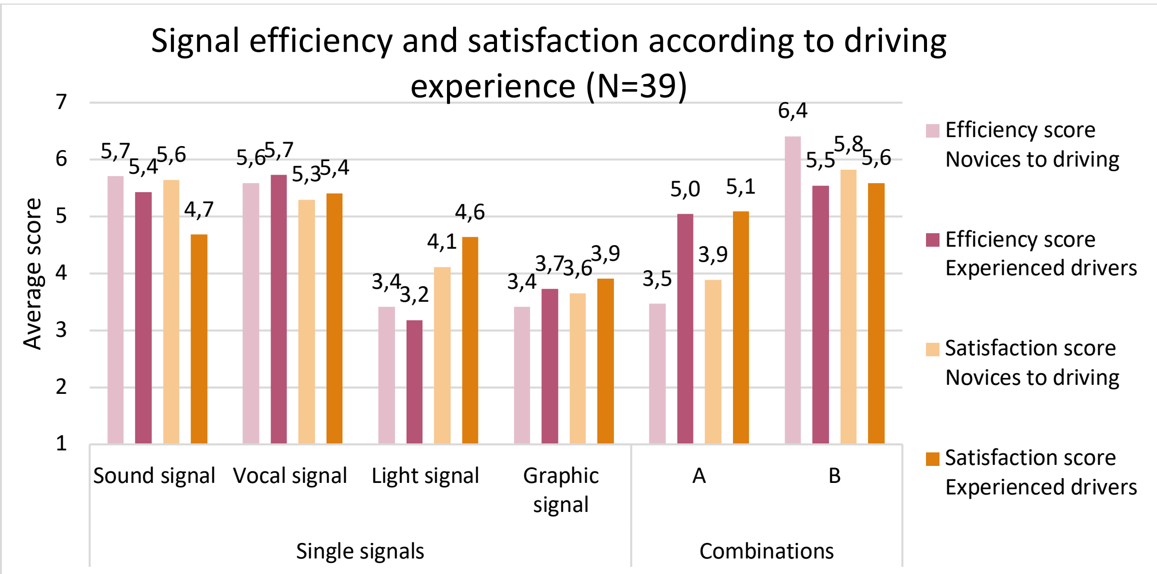


Figure 3.25: Signals efficiency and satisfaction, driving experience

A separate analysis of efficiency and satisfaction scores demonstrates that both experienced drivers and novices prefer audio signals. The most effective signals according to them are the vocal ones (5.7/7 for experienced drivers and 5.6/7 for novices). Satisfaction with sound signals is also appreciated by both types of drivers.

However, novices appreciate more the sound signal (5.3/7) perceived as "brief and pleasant" compared to experienced drivers (4.7) who express concerns about the sound signals mixing with other ones: "there are already many sound signals that sound in the car, like when you don't fasten the seat belt".

It should be noted that the experienced drivers are more sensitive to the light signal. They particularly consider the combination of sound and light signals to be more effective (5/7) than novices (3.5/7). The satisfaction score confirms this trend: 5.1/7 for experienced drivers against 3.9/7 for novices. On the other hand, novices clearly prefer the combination of sound and voice signals, particularly in terms of effectiveness (6.4 vs. 5.5/7).

3.3.2.2 HMI ease of use and attractiveness

Ease of use and attractiveness of the graphic and tactile interface (the touchscreen) were also evaluated according to the driving experience of the subjects.

UMUX questionnaire’s scores do not demonstrate a significant difference according to driving experience. However, the novices scored slightly higher (77.2/100) than the experienced drivers (75.5/100). The good perceived usability thus seems to be confirmed by the participants, specifically by the novice drivers.

This trend is confirmed by the scores in the Pragmatic Quality category of the Attrackdiff questionnaire.

As seen in Figure 3.26, novices rated the usability of the interface better (1.8) than experienced drivers (1.2), although it was rated positively in both cases. Another notable difference is the slightly better scores for novices (0.7) compared to experienced drivers (0.4) on the Hedonic Quality - Identity category. Novice drivers could identify slightly more with the interface than experienced drivers. For example, some of them referred to the interaction system as being "similar to iPhones considering the switch" for autonomous driving (dis-)engagement, so they specified "it was easy, I know it well". This trend in favour of novices is confirmed in the Overall attractiveness scores. Novices rated the overall attractiveness of the system more positively (1.5) than experienced drivers (1).

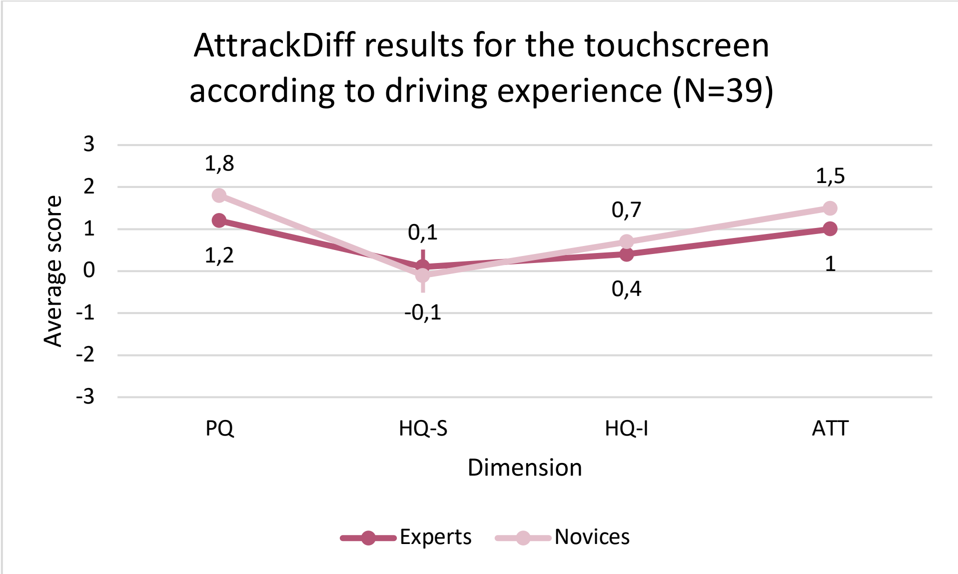


Figure 3.26: Attrackdiff results for the touchscreen, driving experience

Although the HMI is globally well accepted by all participants, we can observe some nuances. Experienced drivers are more aware to the value of the light signal than novices, and novices seem to be more comfortable with the graphical interface than experienced drivers, by grading it slightly more usable and efficient.

It is now interesting to see the impact of the information delivered by the graphic interface on the overall acceptance of the HMI.

3.3.3 Effect of a feedback about the vehicle's perception of its environment

3.3.3.1 On signal efficiency and satisfaction

Figure 3.27 confirms the trends observed in section 2.3, namely that the most effective and satisfying signals are the audio signals, remain true with or "without feedback". We can also observe that participants "with feedback" give a higher effectiveness rating to the graphic signal than participants "without feedback" condition (4 vs. 3.1/7). They find the information "complete", "well indicated", "clear", even if some of them complain "I did not have time to see everything that was displayed", "I did not pay attention to all the pictograms displayed".

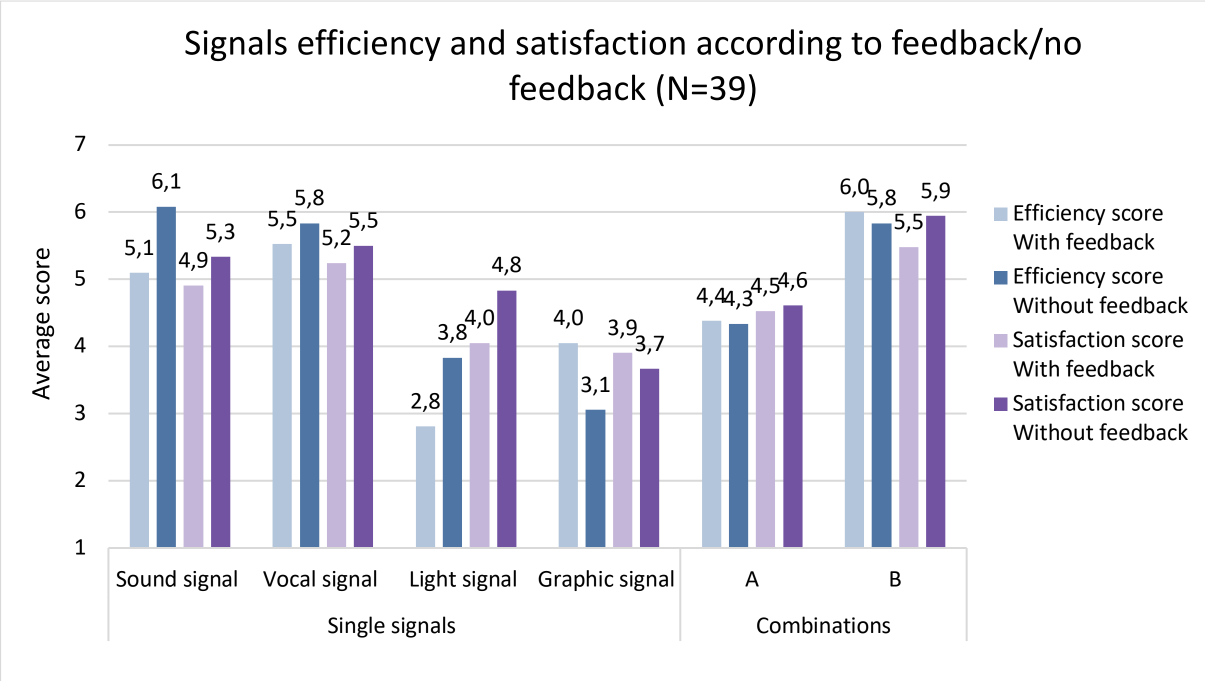


Figure 3.27: Signals efficiency and satisfaction, feedback/no feedback

3.3.3.2 On HMI ease of use and attractiveness

UMUX usability questionnaire's results indicate that participants "with feedback" grade better than ones "without feedback" (78.6 vs. 73/100). The feedback presence therefore leads to a better perception of usability.

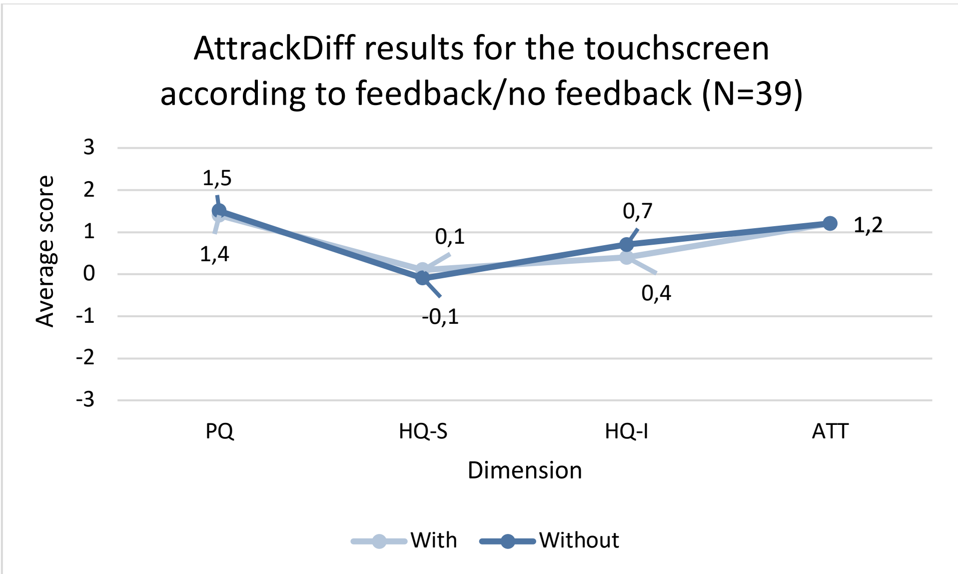


Figure 3.28: Attrackdiff results for touchscreen, feedback/no feedback

According to the SUS interpretation scale, the HMI is judged as good by both populations, the participants “with feedback” being close to the threshold of excellence (80/100).

From the AttrackDiff questionnaire point of view, as shown in Figure 3.28, no difference emerges between the 2 feedback conditions. If usability is influenced by the quantity and quality of the information displayed, it seems that attractiveness and hedonism are not directly impacted.

3.4 Ways to improve CAV design

3.4.1 General recommendations

- Novice drivers appear to be less confident into autonomous driving features, despite their young age and a supposed appetite for new technologies. They should not be neglected by information and support policies for CAVs. As initial training providers, driving schools could/should be involved.

3.4.2 Recommendations related to HMIs

- Drivers need time and habituation for trusting autonomous driving in L3 CAVs. An adaptative HMI could deliver more explanations and reinsurance to beginners then evolve to a more uncluttered display.
- Drivers ask for an easy-to-use, intuitive, and uncluttered HMI. If this is the case, experienced drivers, even older ones, are confident in dealing with it.
- Drivers ask for an easy-to-use and responsive activation/deactivation button. They express a preference for a physical control of the activation/deactivation of the autonomous mode. A control located on the steering wheel would be appreciated.
- Drivers prefer to have all information related to the car’s status and the environment displayed behind the wheel, just in front of them (0° line of sight).
- Drivers express the wish to have the following relevant information:
    - driving mode engaged (manual/autonomous driving),
    - events ahead (traffic jams, accidents, roadworks…),
    - distance (expressed in time) to these events,
    - speed cameras locations,
    - current legal speed limits.
- Drivers considered unnecessary the information about the type of road (urban, countryside, motorway…) as we displayed it.

- An indicator of the car's level of perception of its environment can generate stress for drivers. Some of them prefer not to be informed and trust the system anyway.
- Regarding the signals dedicated to HOR and TOR:
    - Drivers find audio signals more effective than visual signals,
    - Among the audio signals, the vocal signal is preferred to the sound signal (beep)
    - Some drivers ask for the vocal signal to be less robotic, more natural and more pleasant to listen to.
    - The sound signal (beep) is also appreciated, but some drivers express the risk of misunderstanding because of mixing with other sound alerts (forgotten headlights, seatbelt alert, open doors…).
    - Drivers find the light signal is not effective enough in the HOR and TOR transitions
    - Drivers prefers the combination of the vocal and graphic signals to the combination of light, sound (beep) and graphic signals
    - Some drivers wish to have the possibility of setting the signals themselves (graphic, sound, voice, light)
- Drivers consider the light signal keep relevance as a confirmation of the autonomous driving engaged status.
- Drivers' experience of autonomous driving and knowledge about CAVs should be considered. An adaptative HMI could provide suggestions and support to drivers never using autonomous driving features.
- Drivers see more value in autonomous driving for long trips. They also prefer to activate autonomous driving on motorways rather than on city roads. An adaptative HMI could improve its autonomous driving suggestions by taking these variables into account.
- Drivers would appreciate to perform additional tasks while autonomous driving is engaged (working, making phone calls, checking emails, reading...). Some infotainment features could be developed to meet these wishes.
- Drivers also wish to enjoy the environment and landscape more when they are relieved of manual driving Location-based content could be developed to deliver enriched sightseeing.

3.5 Guidelines and recommendations for pilot specifications

3.5.1 Use cases

Several use cases could eventually be explored:

	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5
• Presence of an emergency stop button within fully autonomous vehicles	x		x		
• Location-based infotainment for enriched sightseeing	x		x		
• Additional tasks (reading, mailing…) for participants	x		x		
• Autonomous driving dealing with unusual/tricky situations (traffic jams, roadworks, vehicle parked on the road…)	x	x	x	x	

3.5.2 Test variables

Different dependent and independent variables could eventually be explored:

	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5
• Feedback about the vehicle's perception of its environment vs no feedback	x		x		
• Novices vs experiences drivers		x		x	
• Participants with different levels of knowledge and experience regarding CAVs	x	x	x	x	
• Habituation (would probably need to design a longitudinal experiment)	x	x	x	x	

• Different additional tasks (reading, mailing…)	x		x		
• Attitudes to be measured	x	x	x	x	x

3.6 Conclusions

First, we were able to observe that most of the participants were confident during the experience in the driving simulator.

It also seems that there is an experiment effect, even if autonomous driving is only simulated. The participants, who had limited experience and knowledge of autonomous vehicles, were able to get a more concrete idea of how an autonomous vehicle works, what could drive to an increased acceptability, more positive attitude and feelings towards autonomous vehicles.

For addressing the first research question, the study investigated the effectiveness and acceptance of the different signals present in the CAV. The results showed that audio signals were preferred and considered the most effective by the participants. The voice signal was the most relevant signal for handover and taking over requests according to all participants in the experiment. While the experienced drivers were more sensitive to the light signal, they agreed with the novices that it was more relevant as a confirmation of autonomous driving engagement, once it has been properly activated.

The third research question concerned the effect of the driving experience on the acceptability of the CAV. As a result of the simulation experiment, we can say that experienced drivers report higher trust than novice ones. Their acceptability of CAVs seems to be higher, their attitude more positive, and their perception of the risk associated with CAVs lower. The positive effect of the simulation experiment also seems to be more effective among the experienced drivers.

Also, we have seen that providing feedback about the vehicle’s perception of its environment does not lead to higher trust from the driver. The results showed that although an information-rich HMI is better perceived in terms of usability, it does not lead to more trust for the driver. At times, the opposite is true. Some specific feedback about the car's level of perception can be perceived as a source of stress for the driver, whether experienced or novice drivers.

4 Findings from Virtual Reality Platform

4.1 Overview

4.1.1 Context

People with disabilities have considerably less mobility and transport options to access living areas and places compared to the general population (Casas, 2007). Some authors refer to this as 'transport disadvantage' (Currie and Delbosc, 2011, Currie et al., 2009, Currie and Stanley, 2007, Hine and Mitchell, 2003) to describe these mobility difficulties faced by people with disabilities.

Disadvantaged people are affected both socially and psychologically (Delbosc and Currie, 2011). In this context, one of the main claimed benefits of CAVs for disabled people is their ability to provide more extensive and convenient transport options for people who cannot currently drive (Bradshaw-Martin and Easton, 2014, Chapman, 2016, Lowe, 2017, Darcy and Burke, 2018). While much research has explored the role of general public perceptions and acceptance of CAVs (Fagnant and Kockelman, 2018; Fairley, 2018; Litman, 2018; Hohenberger et al., 2017; Bansal and Kockelman, 2017; Lavasani et al., 2016; Gold et al., 2015; Heide and Henning, 2006), the relevance to vulnerable people with disabilities remains to be established (Penmetsa et al., 2019).

Therefore, we decided to conduct an experiment with vulnerable people with disabilities to offer them a virtual encounter with future L5 autonomous transport and to measure their acceptability and attitudes towards such CAVs.

4.1.2 Purpose of the study

The overall objective of the study was to assess the behaviour, attitude, and acceptance of vulnerable people with disabilities towards fully autonomous vehicles. The specific objectives were defined as follows:

- Research Question 1
  How is the acceptability of a L5 shuttle-based transportation service for disabled people?
- Research Question 2
  How is perceived a L5 shuttle and how accessible it is in a multimodal trip context?

To answer these different research questions, a scenario was set up involving 3 different vehicles in succession for a multimodal trip:

1. a classic level 5 autonomous shuttle: no driver nor supervisor, 6 seats, cabin design similar to a conventional bus.

Figure 4.1: Classic L5 shuttle: bodywork (left) and cabin (right)

2. a conventional (level 0) bus: a driver and other passengers, conventional cabin design.

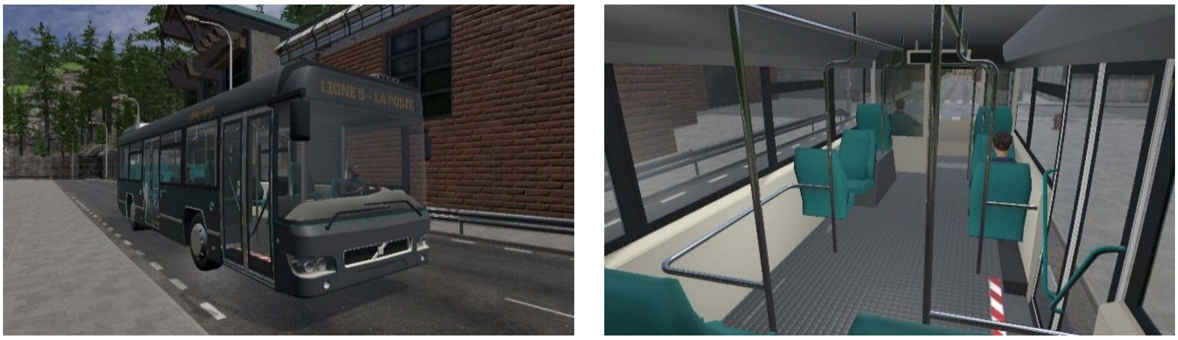


Figure 4.2: Conventional bus: bodywork (left) and cabin (right)

3. A premium level 5 autonomous shuttle: no driver nor supervisor, 4 comfy seats, a glass roof, refined materials, and individual infotainment on large foldable touchscreens.

Figure 4.3: Premium L5 shuttle: bodywork (left) and cabin (right)

Subjective and objective measurements were performed, in addition to video capture, to observe the subjects' behaviour and measure their acceptability towards fully autonomous vehicles.

4.1.3 Study population

11 vulnerable volunteers with disabilities were recruited to participate in this study. They were invited to the laboratory facilities for the experiments, which lasted around two hours (considering the completion of questionnaire and interview). All subjects gave an informed consent to participate in the study that was declared to French ethical authorities and approved by them. The population was composed of 1 female and 10 males. The average age of the participants was 36 ±11 years, (52 ±0 for women; and 34 ±11 for men). The participants were fairly diverse in terms of age and socio-professional categories. Table 4.1 presents the demographic characteristics of the study population.

Table 4.1: Demographic characteristics of the study population

Demographic Variables	Number of subjects	Percentage
Gender (n = 11)		
Male	10	(91%)
Female	1	(9%)
Age (n = 11)		
18 – 29	4	(36%)
30 - 39	4	(36%)
40 - 49	0	(0%)
50 - 59	3	(28%)
> 60	0	(0%)
Socio-professional class (n=11)		
Unemployed	3	(28%)

Senior managers and intellectual professions	2	(18%)
Employees	4	(36%)
Worker	1	(9%)
Tradesmen and business owners	1	(9%)
Educational level (n=39)		
Mid-secondary education	1	(9%)
Secondary education	3	(28%)
Post-secondary education	5	(45%)
Bachelor's or equivalent	1	(9%)
Master's Degree or higher	1	(9%)

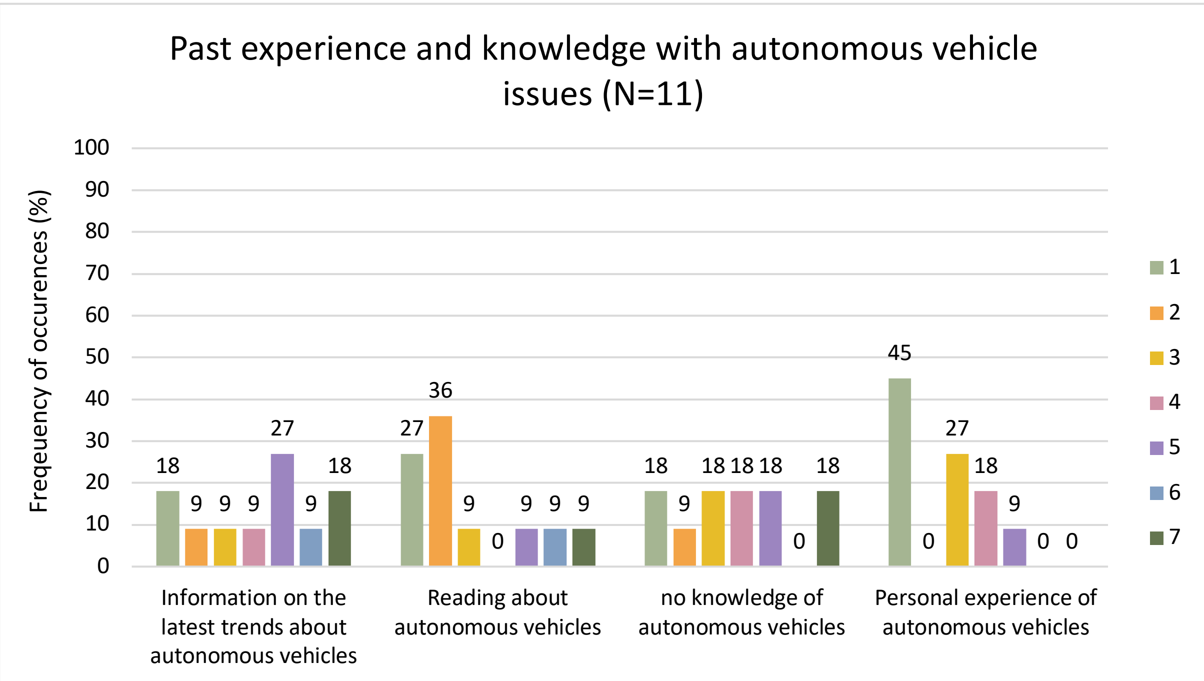


Figure 4.4: Past experience and knowledge

It is important to check the participants' experience to control for potential biases in the experiment. Figure 4.4 presents the past experience and knowledge of the 11 participants regarding autonomous vehicles.

All the participants in the experiment initially have limited experience of the autonomous vehicle (M = 2.45 / 7; SD = 1.51). In addition, they say they do not read about autonomous vehicles on a regular basis (M = 2.91 / 7; SD = 2.12). We will therefore work with a relatively naive population in terms of knowledge of autonomous vehicles. It was important to check the profile of the participants in the experiment to control for potential biases in the autonomous mobility experience.

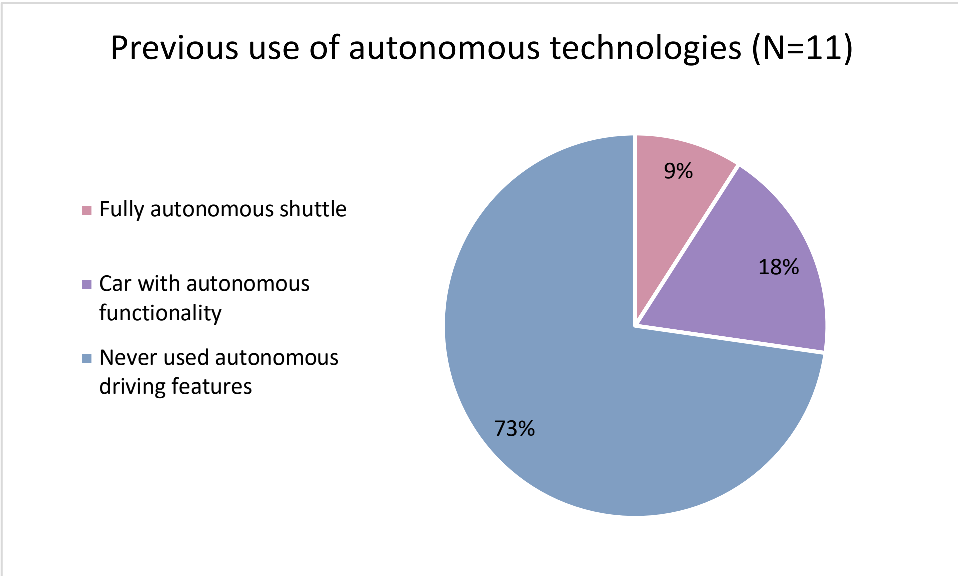


Figure 4.5: Experience of using autonomous technologies

Practically, 72.7% (n=8) of our participants have never used stand-alone features. 18.2% (n=2) said they have autonomous features in their car and 9.1% (n=1) said he/she has already used an autonomous shuttle.

4.1.4 Simulation system

This experiment was conducted on one of UTBM's virtual reality platforms. It is made of a large 50 sqm stage, a powerful PC computing the virtual reality (VR) simulation, connected to a wireless VR helmet via WiFi 5 Gbit/s and to a TV for monitoring. It delivers to the participants a 360° immersion in the VR simulation and free natural movements on the whole stage.

Figure 4.6: Virtual reality simulation room

A detailed description of the simulation system is available in the deliverable D4.1 Scenarios and experimental protocols.

4.1.5 Simulation scenario

The participants carried out a multi-modal journey by using three different vehicles to go to the city, make a first transaction, then a second, and finally return to their starting point. They only move their wheelchair to get in and out the vehicle and reach the checkpoint to validate the different stages of the scenario along a 15’ simulation.

Figure 4.7: Synchronised screenshots of the simulation and the video caption

A detailed description of the scenario is available in the deliverable D4.1 Scenarios and experimental protocols.

4.1.6 Metrics

Physiological and subjective parameters were measured in addition to video recording to characterise the subject's behaviour and to determine their acceptability regarding of the autonomous vehicle.

4.1.6.1 Subjective measures

Acceptability was measured through the questionnaires and interviews concerning attitude, trust, perceived risk, willingness to pay, change in mobility, ease of use, etc.

In addition to descriptive statistics (mean, standard deviation, percentage), a statistical analysis was carried out to specify the significant differences, but the small population (N=11) makes significant effects (p-value < 0.05) unlikely.

4.1.6.2 Objective measures

Several physiological data were also measured: Heart Rate (HR), Galvanic Skin Response (GSR), and Wrist Motility (WM).

The Absolute Cardiac Cost (ACC) directly derivates from Heart Rate: heart rate at rest. The following equation is used to determine ACC:

$$ACC_i = HR_i - HRrest$$

A detailed description of the metrics is available in the deliverable D4.1 Scenarios and experimental protocols.

4.2 Acceptability of L5 CAVs

4.2.1 Trust towards CAVs

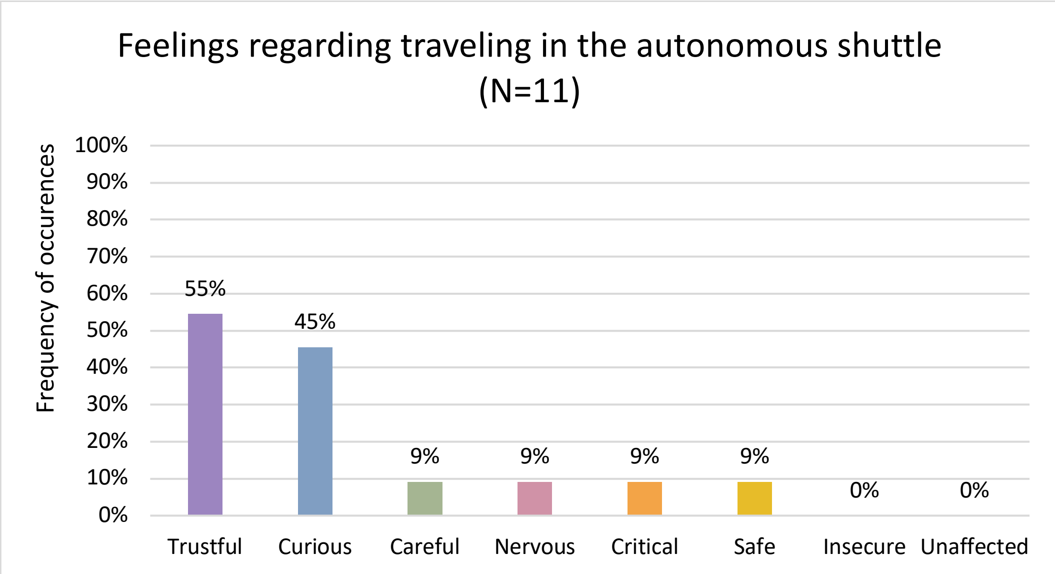


Figure 4.8: Feelings during the trip in the autonomous shuttle

It is interesting to note that half of the participants (54.5%) felt trusting in the L5 shuttles. They confirmed it by mentioning "felt comfortable", "felt good", "felt trusting", "wanted the journey to continue". Also, 45.5% were curious and said, "I was excited to see what it was like", "I came because I wanted to see how an autonomous vehicle functioned". Finally, 1 participant said he was nervous: "it's a bit strange to be in a vehicle without a driver, I'm not used to it, ones wonder about the vehicle's reactions ".

Overall, participants were confident during their shuttle trip, and the virtual reality experience allowed vulnerable participants with disabilities to experience what it feels like to ride in a shuttle.

4.2.2 Attitude towards CAVs

We observed that most subjects (63.6%) were positively surprised by the use of an L5 shuttle: "it's even better than I thought". 36.4% of the participants stated that using an autonomous shuttle corresponded to what they imagined (Figure 4.9). It appears that the L5 simulation immersed the participants in sufficiently realistic conditions to reflect their attitudes and feelings accurately.

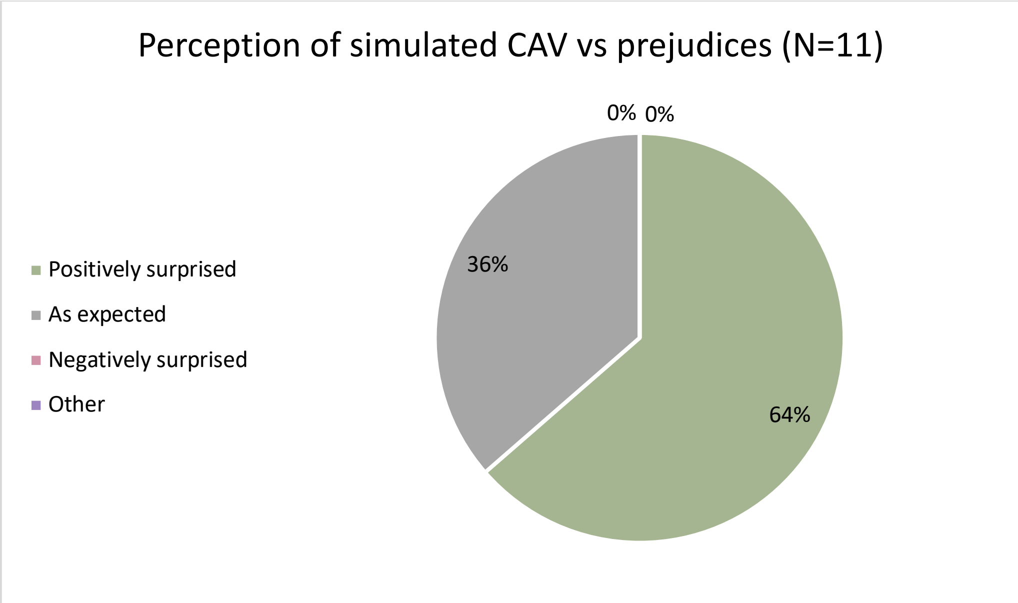


Figure 4.9: Perception of simulated CAV vs prejudices

Figure 4.10 showed the results relating to the effect of the experiment on the acceptability of the CAV by the participants. It seems that the participants' attitude changed after experiment. They more often mentioned the promising potential of CAVs after the virtual reality experience and this difference was statistically significant (6.5 vs. 4.9/7; p= 0.018). They said: "*nice interest*", "*a breakthrough for mobility*", "*great on-demand vehicle*" or "*nice, advanced technology*". Similarly, those found the CAVs very good (5.8 vs. 4.8/7) and very positive (5.7 vs. 5/7) after the simulation was higher.

The subjects also thought after experience that the shuttles would make them even more independent (5.6 vs. 4.6; p= 0.076): "*it would allow us to move around more easily*", "*it is very important for disabled people who cannot move around as they wish*". And this independence is vital for the vulnerable population (6.6 then 7/7).

In terms of transportation, participants reported an increased desire to use the shuttles after the experiment (5.5 vs. 5.1/7). It is very interesting to note that the participants feel that they would manage very well using the shuttles after the experiment, compared to their responses 15 days before experiment and this difference was statistically significant (6.2 vs. 4.9/7; p= 0.031). Some statements were: the system was "*intuitive*", "*easy to access*" or "*finally, everything is done by itself*".

The perceived usefulness of the shuttle has also increased positively. Participants felt more strongly that on-demand transport can help to customize passengers' needs after the simulation (6 vs. 5.1/7). They even believe that L5 shuttles are more a solution to specific transportation needs (6.5 vs. 5.6/7). It would be "*great for people who cannot afford to travel, for example to medical appointments*".

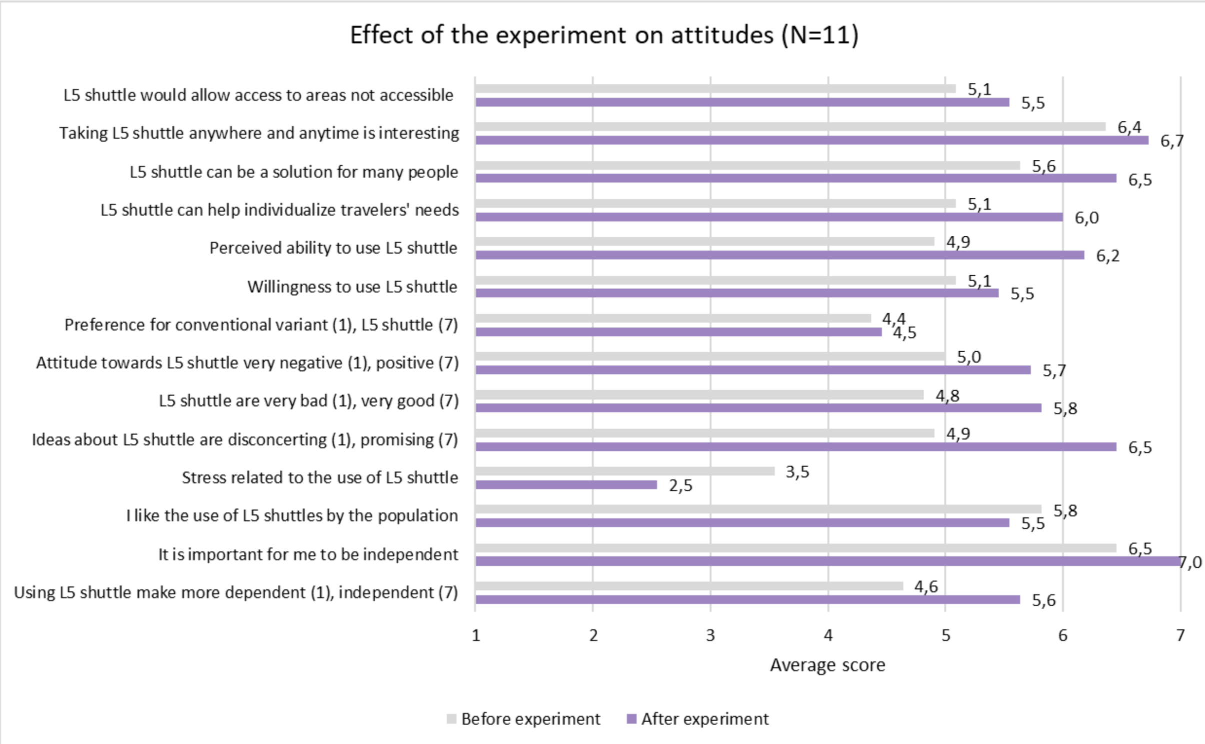


Figure 4.10: Attitudes regarding L5 shuttles

Thus, it appears that the experiment improved the acceptability of vulnerable participants regarding CAV in general and specifically for the shuttle. Participants' responses to the different dimensions of the questionnaire increased after the simulation and for some of them this increase was statistically significant. The reason we did not find a significant difference between all variables before and after the experiment could be due to the limited number of participants in the study.

The experience, even though it was conducted in virtual reality, allowed people with disabilities to see a level of operation and ease of use that they had not previously been able to fully appreciate. It is worth noting that the subjects expressed that they felt "*totally immersed*" during the simulation, that the application was "*true to life*" and that they felt "*as if they were in reality*"; for some it was "*impressive*". This statement strengthens the idea that experimentation (even simulated) enables people naive about CAVs to get a more tangible and meaningful idea of autonomous mobility systems.

The experience in the simulation changed the attitude and the feeling of the subjects about willingness to let other use autonomous solutions. We note an increase between what the participants declared 15 days before and immediately after the experiment. To the item "Would you let other members of your family or those close to you use autonomous shuttles?", 54.5% of the participants answer certainly after the experience while only 27.3% gave this response spontaneously 15 days earlier (Figure 4.11). A subject also declared "*that he saw himself getting into CAV with his parents, that he was looking forward to it*". Before experiment, 9% said they did not probably want to let their family and friends use the CAVs while no participant validated this answer after the virtual reality experience. Furthermore, participants are more ready to encourage their family and friends to use shuttles: 45.5% vs. 18.2%.

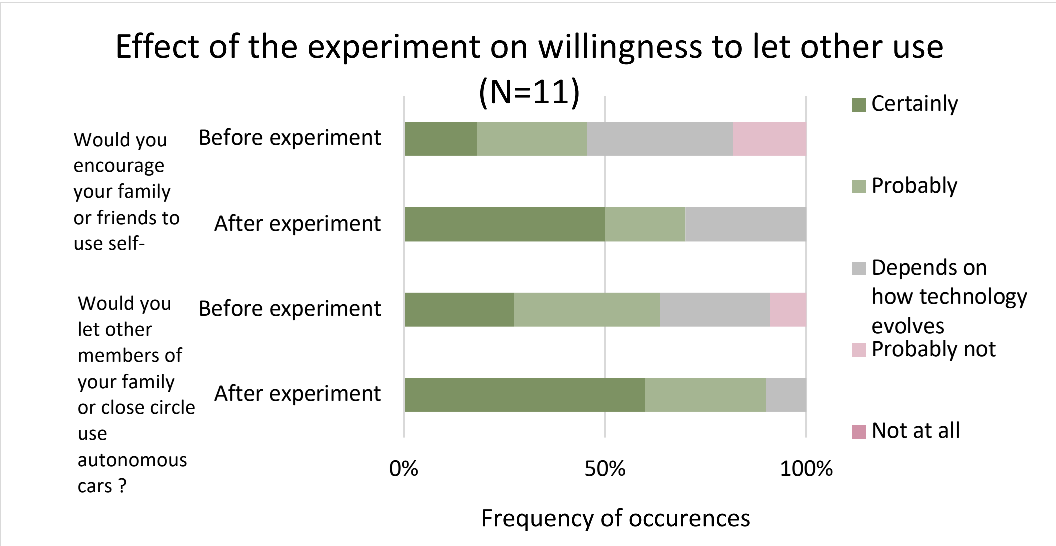


Figure 4.11: Willingness to let other use CAVs

Most of the participants express that they were confident during the travel with the L5 shuttles. 63.6% of vulnerable participants found the shuttles' reactions were very good and 36.4% noticed it safe (Figure 4.12). Subjects said that they were "safe", the shuttle seems "safe" and " it's really well done". Their feelings were again very positive, even if a subject specifies: "it is immersion, I knew that nothing could happen to me".

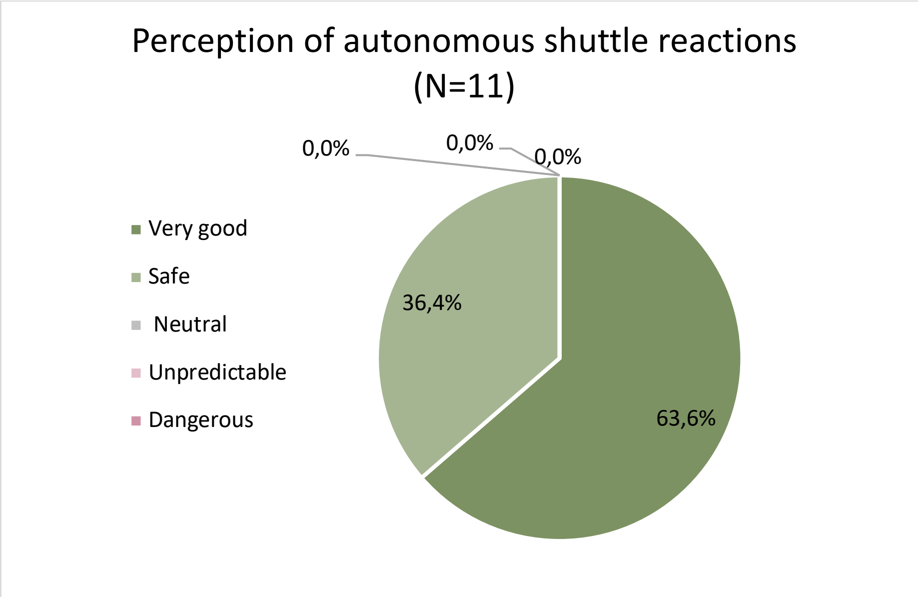


Figure 4.12: Perception of autonomous shuttle reactions

At the item "In your opinion, what would be the potential advantages of using shuttles compared to traditional public transportation?". Participants found, after experiment, higher positive outcomes for: punctuality (54.5% vs. 18.2%), safety (45.5% vs. 36.4%), traffic flow (63.6% vs. 54.5%) and less pollution (54.5% vs. 45.5%).

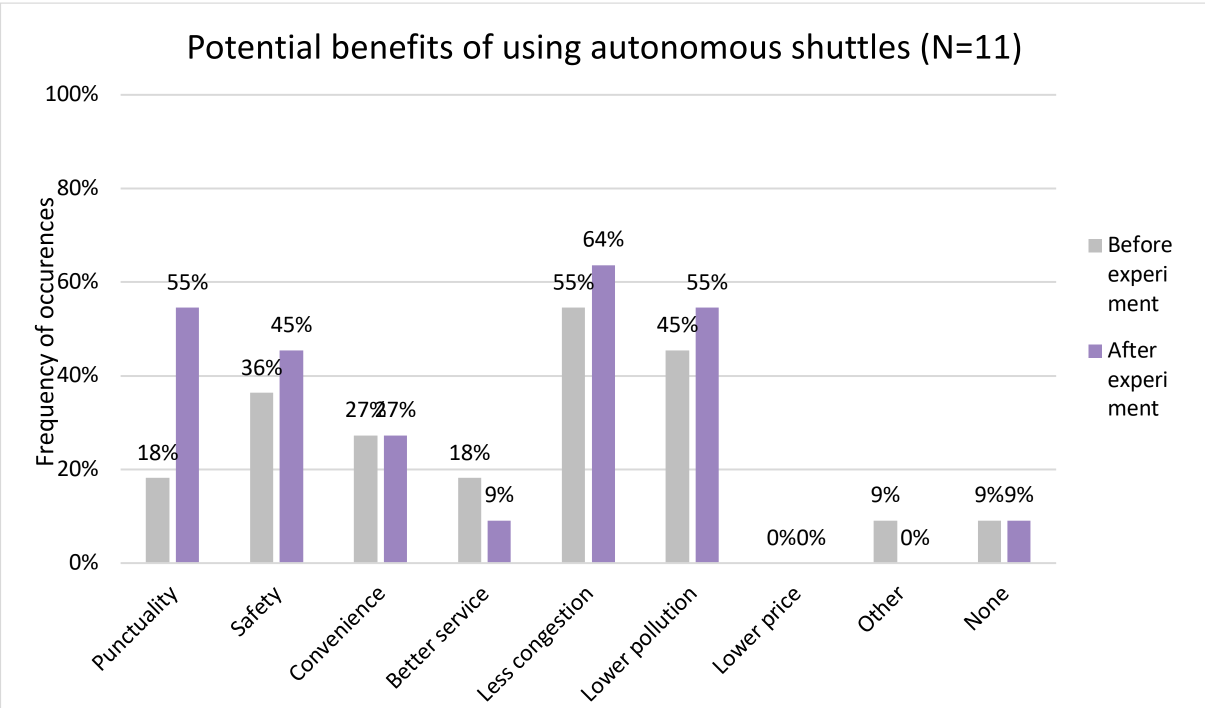


Figure 4.13: Potential benefits of using autonomous shuttles

Regarding the potential disadvantages of using the shuttles, only 18% of the participants felt after the experiment that the quality of the service offered by the CAV would be poorer, while this percentage was 27% before the experiment. The same tendency was observed for the job loss criterion, with the feeling being quite more positive (81.8% versus 72.7%). However, the participant found after the experiment that the price of such service would increase (Figure 4.14).

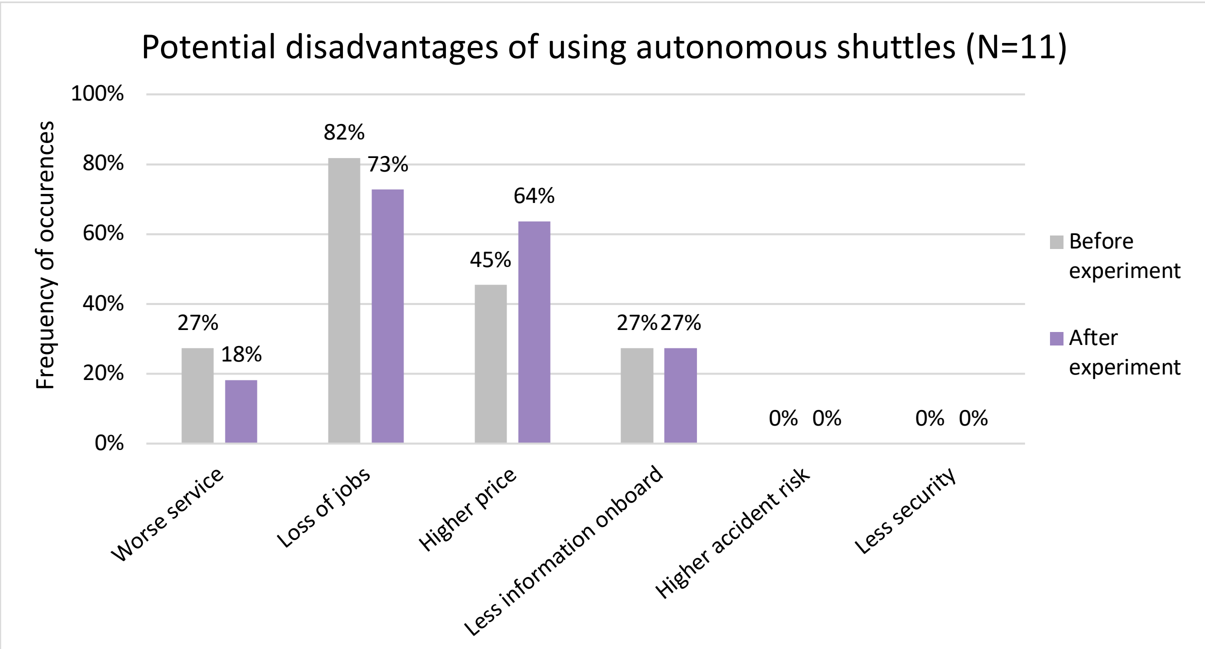


Figure 4.14: Potential disadvantages of using autonomous shuttles

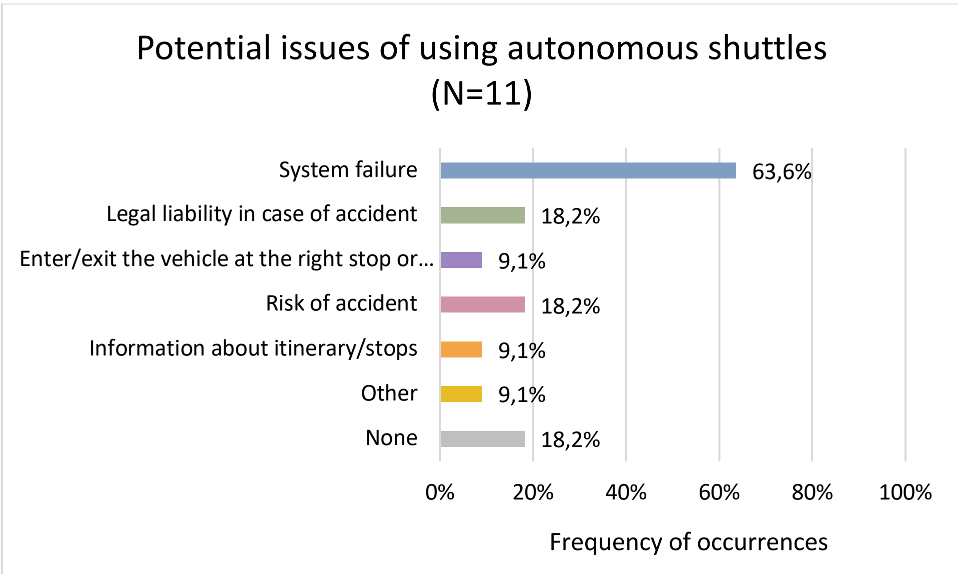
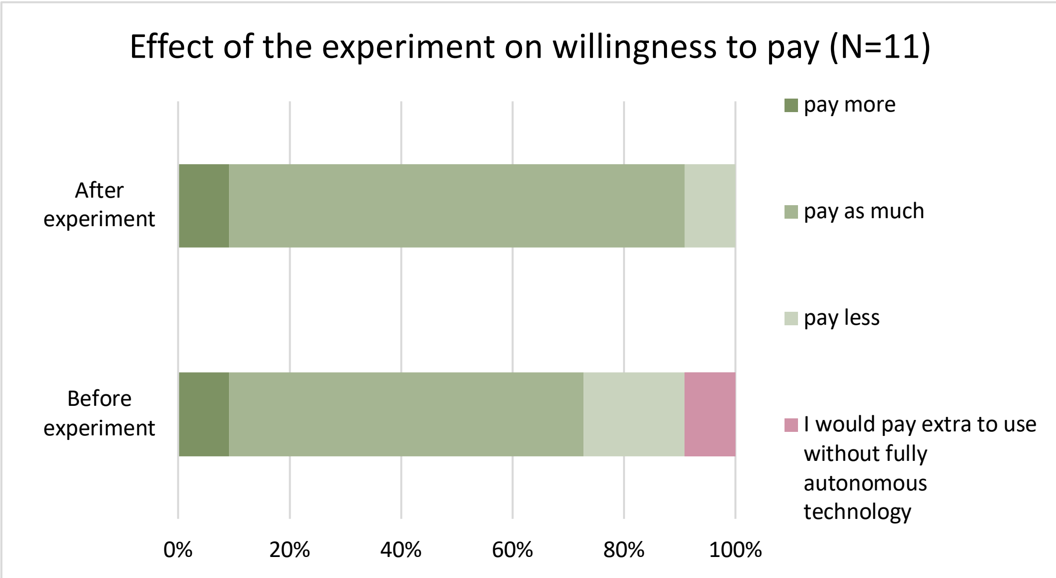


Figure 4.15: Potential issues of using autonomous shuttles

The worries of vulnerable participants regarding the use of the shuttle were the failure of the system (63.6%), the risk of an accident (18.2%) and the civil liability in the event of an accident (18.2%; Figure 4.15).

No significant effect observed on willingness to pay. The small increase after the experiment corresponds to only one more supportive participant.



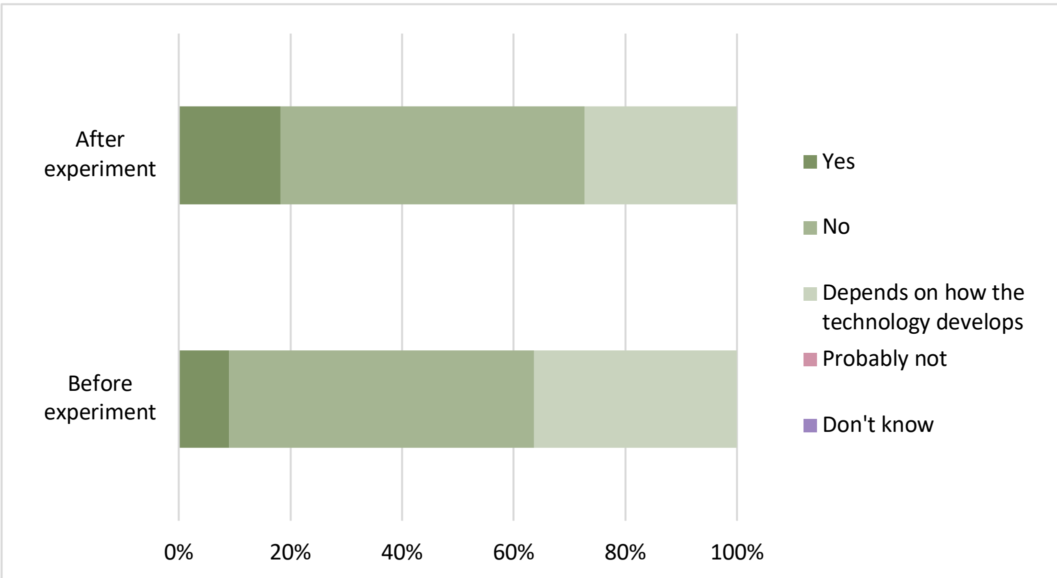


Figure 4.16: Willingness to pay for autonomous drivingfeatures

4.2.3 Comparison of L5 shuttles with a conventional bus

During the interview, the participants were asked to compare the 3 different vehicles used along the trip.

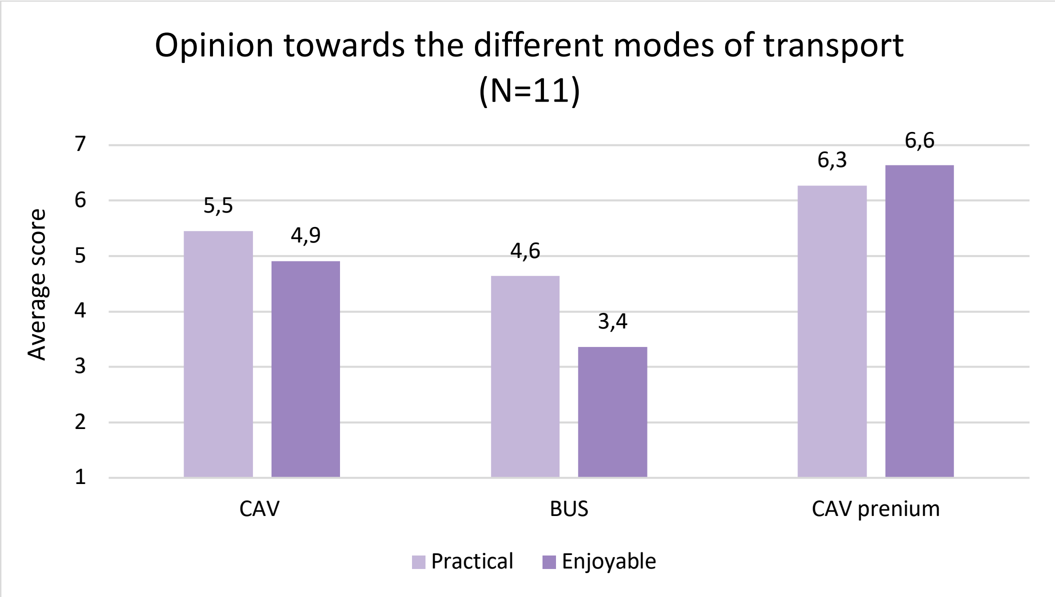


Figure 4.17: Opinion towards the different modes of transportation

Vulnerable participants much appreciated both L5 shuttles, especially the premium one (Figure 4.17). Regarding convenience, the participants had a better perception of both classic and premium shuttles (respectively 5.45/7 and 6.27/7) than the conventional bus (4.64/7). They mentioned about the premium shuttle "*A last super pleasant CAV*", "*Very comfortable vehicle, pleasant*", "*Perfect for making long journeys especially*", "*Simple interactions, easy to get on and off*". The enjoyment generates larger gaps between the for the premium shuttle (6.64/7) and the classic one (4.91/7) and also compared to the conventional bus (3.36/7). The participants believe for the premium shuttle "*maximum comfort, great parquet flooring, pleasing to the eye*", "*esthetical effort*" or "*super nice. More space, appealing design, possibility of recharging your phone*". In contrast, the conventional bus was considered "*tasteless*", "*ordinary*", a "*less pleasant experience in the bus*", "*classic bus, nothing more, not appealing*", in the end "*a less pleasant trip, more stress, and more tiring*".

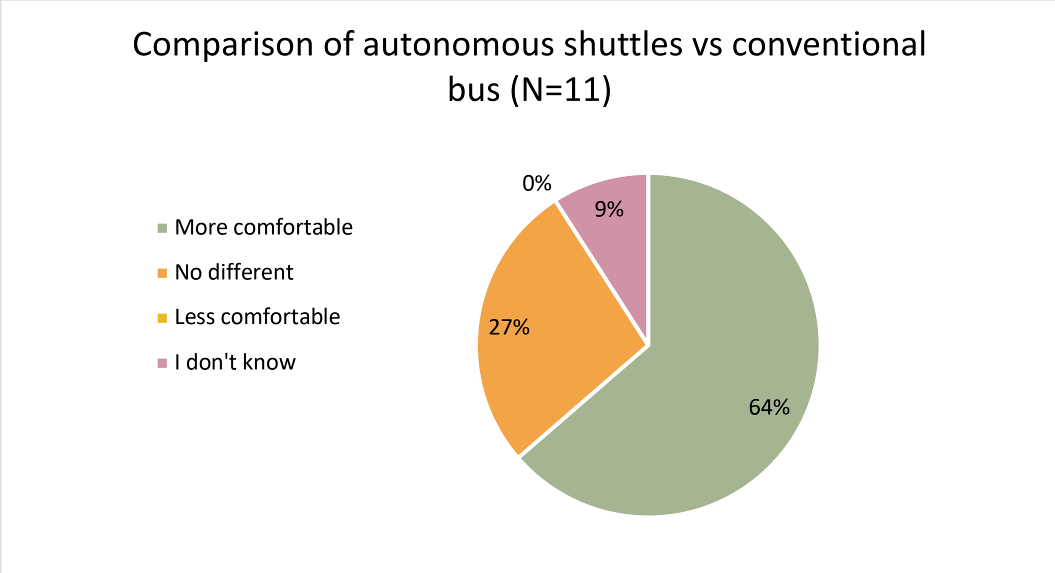


Figure 4.18: Comparison of autonomous shuttles vs conventional bus

63.6% of the vulnerable participants said they found their shuttle trip more comfortable than in a conventional bus (Figure 4.18). Subjects said they "preferred the shuttles", "the shuttles were more pleasant", "the bus isn't as fancy as the shuttle", they were " excited much more for the shuttles". 27.3% said they had travelled in equivalent comfort between the shuttles and the bus. 9.1% were unable to choose a favourite mode of transport (Figure 4.18).

Table 4.2: ACC and GSR according to type of transport

	Average	Standard Deviation	p-value
ACC (N=11)			
Conventional bus	-4.60	6.52	0.320
Classic L5 shuttle	-2.80	6.25	
Conventional bus	-4.60	6.52	0.413
Premium L5 shuttle	-5.62	4.45	
Classic L5 shuttle	-2.80	6.25	0.175
Premium L5 shuttle	-5.62	4.45	
GSR (N=9)			

Conventional bus	3.23	3.13	0.820
Classic L5 shuttle	3.12	2.87	
Conventional bus	3.23	3.13	0.652
Premium L5 shuttle	3.24	3.08	
Classic L5 shuttle	3.12	2.87	0.359
Premium L5 shuttle	3.24	3.08	

Table *4.2* displays the ACC and the galvanic skin response according to the different vehicles (classic L5 shuttle, conventional bus then premium L5 shuttle). The heart rate of the subjects was very low regardless of the type of vehicle used. As an indication, the heart rate decreases throughout the experiment and would seem to indicate an effect of the experiment, or even a possible habituation to virtual reality.

The comparison of heart rate data according to the different types of vehicles indicated that the subjects have a higher ACC in the classic L5 shuttle (-2.80; SD=6.25), a lower ACC in the conventional bus (-4.60; SD=6.52) and even lower in the L5+ premium shuttle (-5.62 SD=4.45). These results are not significant, and we cannot confirm that a small difference between the HR in the classic L5 shuttle compared to the other trials (conventional bus and premium L5+ shuttle) is induced by the mode of transport. It seems that the higher HR could be related to the subjects' discovery of the system, situations, and landscapes at the beginning of trial, so that could influence their HR. Furthermore, these results should be interpreted regarding the scenarios proposed in virtual reality: the subjects did a short trip in the city by conventional bus, whereas the classic shuttle trips were made on a more winding country road. Or the premium shuttle offers an on-board tablet on which participants can watch a video so that most of them look at the screen during the ride. This could reduce HR because the possibility of simulation illness is reduced. The GSR, which detect the difference in skin conductance when, for example, a person is stressed, are in this case very similar and not significant. The GSR data neither confirm nor contradict the other data collected. It might be interesting to replicate this experiment on a real situation to verify these results.

In general, the interviews highlighted the difference in perceived comfort and feelings concerning the 3 vehicles. The conventional bus, even if it provided "a stronger control feeling than the shuttle due to the presence of a driver. ", was perceived as "less pleasant", "more traditional" and "not very appealing" for some participants. Although the subjects felt "secure" during the conventional bus journey, some stated that "the journey was not as good as in the L5 shuttle", "less pleasant", and even "more stressful, more tiring". One participant stated quite significantly "I feel I changed a class down between 1st and 2nd vehicles".

In contrast, the shuttles seemed more "innovative", "with a futuristic feel", "more appealing", and overall, the participants "travelled well" and felt "confident". In the classic L5 shuttle, they felt "totally confident", "had total confidence in the system, even though it was an ideal mode, without any disturbance", one participant said, "all I needed was my mobile to play and it would be perfect". One participant qualified this: "I felt safe, but not totally confident because I am apprehensive about autonomous vehicles". After a reminder, he explains: "It's still impressive and unusual not to have a driver at the wheel". This first classic L5 shuttle was defined as "a vehicle that was too small, not enough room to turn around", "the design was nice but maybe a little too common", "compared to the last shuttle (the premium one), it was less classy, more common". Two subjects noted that "it lacked an emergency stop in case of a problem". Several subjects noted some positive points regarding the shuttle hygienic nature: "no need to press a button, we avoid contacts in the CAV which is good at this time"; "easy to use", "satisfied, easy to use", "interaction with the system works well. Access is easy". In conclusion, the participants therefore considered the shuttle journey to be "pleasant".

The premium shuttle is clearly reported as "the best of the 3 transports", "the best of the 3 vehicles" and "the most comfortable". One participant stated that the transport was "more pleasant than the bus but offers less visibility than the bus where there was no seat in front". The configuration of the premium shuttle was considered "super pleasant", designed to encourage "exchanges, discussion with seats facing each other". In terms of confidence, participants said they were "well confident", "confident", one subject even said he was "well confident. I trust the technology more than the human". The aesthetic aspect is also highlighted by the participants: "maximum comfort, great parquet effect on the floor, pleasant to the eye. Aesthetic effort", "panoramic vision, fancy floor, more modern, overall design that makes you want to get in, very innovative, technological". The ease of use of the premium shuttle is also highlighted through "simple interactions, convenient boarding", "no need for physical contact with someone or equipment to validate the transport" which is "important in the current pandemic context". In this sense, the subjects specify that the on-demand shuttle service is a plus because "it mixes less people, less microbes". While participants said that "too much connectivity" was not necessary, they found the proposed tablet useful: "Super comfortable. The presence of the tablet focuses attention, you don't forget that there is no driver, it's reassuring and pleasant". "The tablet is a real plus. Top service", but "for long trips, otherwise there's no point". This idea was repeated by several subjects: "Very comfortable vehicle, pleasant. Perfect for long trips especially", "it's a very interesting shuttle for a long trip, for example several hours on the motorway", "the services offered are very well adapted to long trips, I don't know if I would use all the options on a trip of a few minutes". The services offered in the premium shuttle were appreciated: "I saw the services concerning music and TV, it's a good idea, I would have liked to use them too much", "having the radio in self-service is very interesting", "all these options make you want to try everything".

At the end of the interviews, the participants were able to give more freely their general comments, their feelings, their needs, their future and potential desires in the context of autonomous transport. Concerning the potential services in the shuttles, the participants thought of "having internet, Youtube, Netflix", "the playstation 5", "being able to read multimedia" and even "casting their mobile phone". Some participants expressed the wish to be able to "check their emails" and thus "be able to work, it's in the transport that we can have time to work". It is interesting to note that several subjects would like to have a plug so that they could "recharge their electric chair". Some of the subjects who said they were "thinking big" would have liked to have "a drinks and snacks service, for example as a vending machine", "why not a mini bar, and even a spa, let's dream big".

From a social point of view, the participants mention the fact that they prefer to travel accompanied: "the idea of public transport is not to be alone", "I would have preferred not to be alone in the shuttle", "to be accompanied yes, but with a maximum of 2/3 people". Several vulnerable participants said they would prefer not to have "specific parking" or "PRM (person with reduced mobility) signs on the ground" so as not to be "stigmatised" by other passengers. The idea would be to have "a transport with space, so that several wheelchairs can take place, but without a dedicated area, it's embarrassing, we're like everyone else". In general, the vulnerable disabled person is "impatient for the shuttles to be operational, to go to the emergency room or to the doctor, or even when it's snowing. O-demand, it's the best".

Finally, some participants would like to develop an informational aspect of the proposed multimedia services. For example, "broadcasting preventive messages on the tablet such as safety reminders, cancer prevention, messages from associations, local advertisements, having books available", "having local information would be really nice and interesting", "a way of communicating about what local associations are doing". One participant even mentions the idea of "having a presentation of the city, developing tourism, making the locals discover or rediscover the environment".

4.3 Ways to improve CAV design

4.3.1 General recommendations

- Passengers need time and habituation to trust fully autonomous vehicles. An adaptative HMI could adapt to passengers' experience and knowledge to deliver more explanations and reinsurance to beginners.
- The premium L5 shuttle was more appreciated for long distance journeys. A business model based on high level of comfort and services for short hauls could be risky.
- The participants (in wheelchair) prefer to get in with the help of an automated system. Embedding a motorised ramp appears to be very popular.
- The premium L5 shuttle's aesthetics was very popular. A modern and attractive design appears to be an efficient lever for L5 shuttle's appropriation.
- Contactless interactions are positively judged by some participants, regarding to the risk of contamination. Combined with a smaller capacity, it could contribute to establish L5 shuttle as a more hygienic alternative to larger public vehicles.
- Some participants asked for an emergency stop button. It could provide reinsurance to some worried customers.

4.3.2 Social recommendations

- The participants expressed the wish not to travel alone. 3 or 4 passengers was proposed as an ideal capacity (regarding the shuttle size?) for inviting to social interactions.
- The participants asked for enough room in the cabin for 3 wheelchairs.
- The participants asked for non-specific locations for wheelchairs and no PRM markings considered to be stigmatising.

4.3.3 Service-related recommendations

- On-demand service and servicing remote areas sound both very appealing to the participants. Such arguments could contribute to fully autonomous vehicles appropriation.
- Some participants asked for an accessible socket for charging their electric wheelchair. Embedding a powerful enough socket could contribute for appealing people in electric wheelchair.
- Multimedia/infotainment embedded services appears to be popular. Some location-based contents (about touristic information and local activities) could be appreciated. Such as the ability to cast some content from passenger's phone to a larger screen.

4.4 Guidelines and recommendations for pilot specifications

4.4.1 Use cases

Several use cases could eventually be explored:

	Pilots				
	1	2	3	4	5
• Presence of an emergency stop button within fully autonomous vehicles	x		x	x	
• Location-based content for embedded infotainment (touristic information, local activities)	x		x	x	
• Experimenting a CAV as a segment of a multimodal trip	x		x	x	

4.4.2 Test variables

Different dependent and independent variables could eventually be explored:

	Pilots				
	1	2	3	4	5
• Impaired participants as part of the study population	x		x	x	
• Number of passengers on board	x		x	x	
• Habituation (would probably need to design a longitudinal experiment)	x	x	x	x	
• Different infotainment offers	x		x	x	
• Attitudes to be measured	x	x	x	x	x

4.5 Conclusions

First, we could observe that the participants, who had previously a limited experience and knowledge of CAVs, declared a high level of trust during the virtual reality experience.

The VR simulation delivered to them a more concrete idea of how works a L5 vehicle and the services it could provide. Experimenting L5 CAV shuttles was a good surprise for most of them. Their overall attitude and feelings, already positive before the experiment increased when re-measuring after, even if they still express concerns about possible technical failure, much more than about accidents or other issues.

The results also show that vulnerable disabled participants prefer shuttles to conventional buses. They felt more secure in the shuttle than in the conventional bus, their attitude and acceptability appeared to be better.

They logically expressed a clear and enthusiast preference for the premium L5 shuttle because it offers additional multimedia and infotainment services, combined with the superior design and comfort. The participants reported feeling better in the premium L5 shuttle. However, their willingness-to-pay didn't really increase while considering this option.

These results are encouraging and are consistent with research conducted on the general population. It is now necessary to confirm these effects on acceptability by testing larger panels and real-life situations.

5 Findings from Home Study Simulator

5.1 Overview

5.1.1 Purpose of Study

- *RQ1: Does driver acceptance vary after multiple exposures to an L4 vehicle?*
- *RQ2: Does driver cognitive load vary after multiple exposures to an L4 vehicle?*
- *RQ3: Does driver performance vary after multiple exposures to an L4 vehicle?*

This study explores three main research questions as outlined above via an experiment in a (simulated) level 4 vehicle over four visits (sessions or trials). It utilizes a similar approach from prior some of our prior work (Mirnig, et al., 2019). The focus being on the change during the first and last visit, with the intermediate visits existing merely to act as a way of familiarising people further with semi-autonomous vehicles. The three aspects explored are:

- Performance: based on driver performance data collected from the simulation environment;
- Cognitive load: an examination of the cognitive load experienced by the driver during the exposure to the simulator;
- Acceptance: acceptance data collected from a questionnaire and semi-structured interview technique known as repertory grid analysis.

5.1.2 Study Demographics

Subjects were recruited via social media, posters and email campaigns. The target group were those who had passed their driving test two or more years before the trial that had not taken part in other semi-automated vehicle studies as a driver. In total 8 participants took part, only four decided to complete the demographic questionnaire. Of there were two male and two female participants The age group ranged from 26 to 47 with the median age being 36.5. The years since passed the driving licence ranged from 9 to 15. While other data was collected, given the small sample size and number of responses involved it will not be reported here.

In terms of technologies which may be related to CAVs, the main aim was to understand if these were likely to shape performance and perceptions.

5.1.3 Methods and Analysis Approach

5.1.4 Simulator

The simulation environment consisted of three routes. All driving took place under EU driving conditions (e.g. right side of the road driving) with Italian road markings. Route 0 was used for training the participants to in the simulator for a period of fifteen minutes. Following on from this route 1, was used in the first and last visit to the simulator. This allows for a comparison of similar environments. In between in order to have some variation, participants visited route 2 on two occasions. In order to have some variation in conditions similar to real life the traffic density was varied, it should be noted that the traffic density applies to the left side of the road which in turn also impacts the amount of traffic at junctions.

5.1.5 Analysis Methods

The data collection methods were designed to capture subjective and objective aspects of the participants experience. The study took place over four rounds, with the first and last round being the ones where assessment using the questionnaires and repertory grids were used. Performance data was collected from all four visits, however only the first and last visits are analysed in detail but they are presented to provide some context.

5.1.5.1 Acceptance Questionnaire

The final acceptance questionnaire was modified slightly from the one presented in D4.1. This mas mainly done to improve understandability from the participants. The questions and the results reported later. The questionnaire focused on trust, general experience, usability of the autonomous vehicle and views on the experience.

Acceptance questionnaire data was compared across each question to see if the results become more positive between rounds 1 and 4. Due to the limited number of participants basic descriptive statistics are be used.

5.1.5.2 NASA Task Load indeX

The objective of the NASA TLX (NASA, 1986) (Hart & Staveland, 1988) (Hart & Staveland, 1988)questionnaire is to explore the overall cognitive load that the participant experiences for the entire driving experience. Then to compare the data from the first and last visit. We had initially discussed asking for responses for each task, however this may have resulted in them being asked to provide such data up to 40 times per session, which would have been excessive. Also asking them repeatedly even a relatively small number of times would have interrupted the driving experience. This approach means that we are assessing the level of cognitive load during manual and automated phases, and it is thus reflective of the entire level 4 experience.

NASA TLX questionnaire assesses the cognitive load of users, over 6 subjective items:

- Mental demand
- Physical demand
- Temporal demand
- Performance
- Effort
- Frustration.

Each item is rated by the user subjectively on a scale of 20. The global effort is an average of each score on a 100 scale. All 8 users filled in the NASA TLX questionnaire at the end of each of the 4 driving sessions.

We performed a "raw analysis" of the evolution of the perceived task load without weighting the different items of the tool. We compared the perceived task load for each item, as well as the global perceived effort, over the four sessions and among the different users.

5.1.5.3 Repertory Grids

Repertory grids are a form of semi-structured interview. Their objective is to collect views regarding the overall experience, through providing the participants with various elements that prompt their recollection. As noted later, the participants develop constructs based on these elements, the intention is not to assess each individual element. In order to achieve this, the participant is first given a selection of 9 pictures which represent elements of experience they encountered during that visit e.g. roundabout, intersection etc. Next participants are asked to suggest words and phrases which best describe their experience. These words and phrases are chosen completely by the participants. This is to avoid interviewer bias and as such anything about the participants experience can be uncovered. Once the words have been chosen, the participant is asked to pick three phrases, then from that pick two which are opposites. The third word is returned to the pile of words and phrases. The opposites then form constructs at opposite ends. Once all words have been selected this way or it proves no longer possible to create constructs the next phase of assessing the experience begins. For this each construct the participant is then asked to rate each of the pictures on a scale of 1 to 7 using each of the constructs they have previously developed, 1 being negative and 7 being positive.

The repertory grid data is analysed in its raw and uncoded form. This involved creating two grids, one for the first and one for the last visit. Then all constructs relating to that visit are put into one grid table. This avoid over interpreting the data by placing it into categories of constructs, however it means that gaining an overall view is a little more difficult. As noted later the primary overall analysis involved looking for related constructs. This was undertaken using the Rep Grid Plus software (University of Calgary, 2022) which can provide a list of matching constructs above a certain degree of relationship (expressed as a percentage). This approach allows for an analysis of how some constructs impact on other constructs.

5.1.5.4 Driver Behaviour and Reaction

Our primary interest relates to when there is a change in mode, e.g. from automated to manual and vice versa. A secondary interest relates to assessing the level of trust people have in the automated system, we believe this can in part be examined by looking at the number of times and duration that they turn the steering wheel and/or use the pedals while the vehicle is in automated mode. The belief being that the more they do this, the less they trust such a system. In common with the study plan only the first and last visits are assessed.

Performance measures that were assessed on a per task type basis included:

- Reaction when vehicle changes from automatic to manual mode
- Speed:
  - Average speed during a manual task
  - Number of violations above speed limits
- Pressure on the break in auto and manual mode
- Deviation from centre of the lane

The focus is on the behaviour during particular tasks during the first and last visit to the simulator. We do not take average performance measures for the routes, instead they are for specific groups of tasks, e.g. junctions.

5.1.5.5 Reaction time

This relates to the time taken for the participant to take back control of the vehicle, in essence going from automated to manual. Reaction time is calculated from the start of displaying of a warning to the time the participant responds to the event by moving the steering wheel, pressing a button or using the pedals. Within the simulator the data is stored for each participant and is time stamped. It should be noted that where the driver does not resume control the vehicle pulls into the convenient area at the side of the road.

5.1.5.6 Adherence to Speed Limit

As the vehicle changes mode frequently, taking an average across the experience is not relevant. Indeed, the vehicle speed is for large parts of the time controlled during the automated phase. As a result, speed limit adherence is calculated from the time that the driver is in manual mode for a particular task and until automated mode is resumed. We are primarily interested in if the driver exceeds the speed limit (50kmh) and their average speed for the duration of that task.

5.1.5.7 Pressure on the Break

We have calculated the average pressure on the break peddle during autonomous mode. This calculation is used to indicate an implied level of acceptance, with a higher level of pressure indicating they are not so accepting of the autonomous mode. The comparison is made for the first and last visit.

5.1.5.8 Deviation from Centre of the Lane

For this part of the study the lane deviation is again calculated per task. This is calculated from the entire duration of the manual phase and until the return of automated control.

5.2 Deviations from the Initial Plan

The Corona virus pandemic meant that were not able to recruit the initial target number of participants. Also, as the trial requires four visits to the laboratory consisting, lasting a total of up to six hours, many people were reluctant to commit. Therefore, this study is based on a smaller sample group and while the number of exposures remained the same, the duration was shortened to 2-4 weeks per participant. The intention is to continue with the trial until a reasonable number of participants are reached. Therefore, this report reflects the results of the first 8 participants. The trial is continuing and we will report on the data with a larger group at a later date. It should therefore be thought of as preliminary data. This is planned either an addendum to this deliverable and/or in future publications. Switching from home to laboratory based studies reduced health risks as it reduced the need for evaluators to enter people`s homes to set up the equipment. Also, as the study also requires visits to the laboratory in certain phases, the decision was taken to undertake all of it at LIST. This also has the advantage of being able to more easily control the study environment. A further problem arose in that LIST was organising multiple simulator studies at the same time, and broadly speaking the same group of people are interested in these studies. However, recruiting people for multiple studies proved challenging.

5.3 Results

5.3.1 Driver Behaviour and Reaction

The preliminary data points to some trend which require further explanation within a larger trials. In particular, the early findings could be used to structure new research hypothesis in the larger study. Only eight drivers were assessed using this data, therefore it is difficult to draw any meaningful and statistically valid conclusions. Further analysis in particular using pair wise data comparisons on a larger data set, would provide clearer statistical information. It is planned for more participants to take part over the remainder of the project.

In the following graphs, each comparison is that the left (blue) indicates the first trial while the right (orange) represents the fourth trial.

Adhering to speed limits (in manual mode) changed between the first and last trial (Figure 5.1). For all events, except the crossroads the number of violations fell. Following the road (following path) had the largest fall. This is not surprising as comparatively more time is spent driving on the road without a specific event e.g. a crossroads.

In terms of average speed (Figure 5.2), these remained similar for all tasks, with only small changes observed. Accurate comparative data for overtaking is not available in this and future parts of the trial.

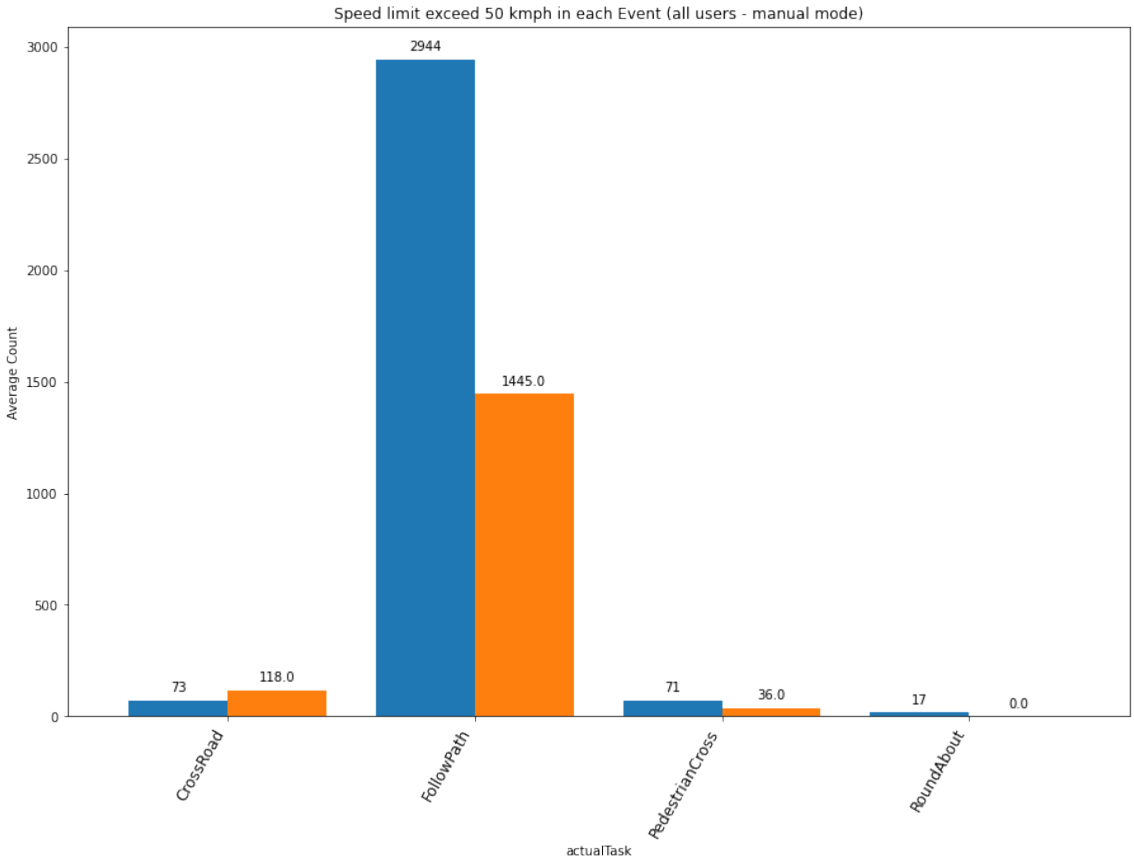


Figure 5.1: Average clock cycles per event type above speed limit

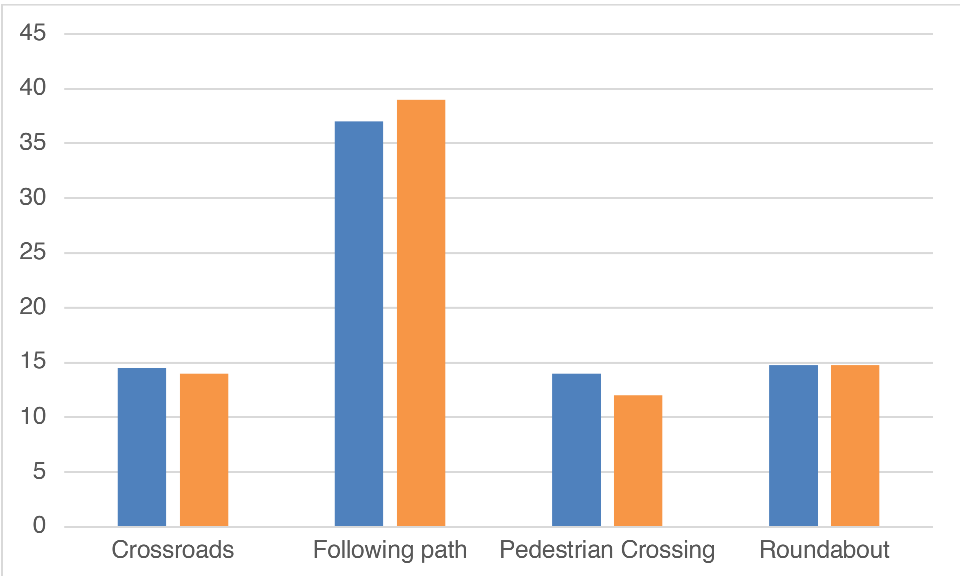


Figure 5.2: Average Speed per event Type

Use of breaks while in autonomous mode, could be taken as an indication of lack of trust and acceptance. According to the collected data (Figure 5.3), on average the participants applied less break pressure during the fourth trial, than during the first one for all events except for parking events. For overtaking, there appeared to be no difference but given the values requires further investigation.

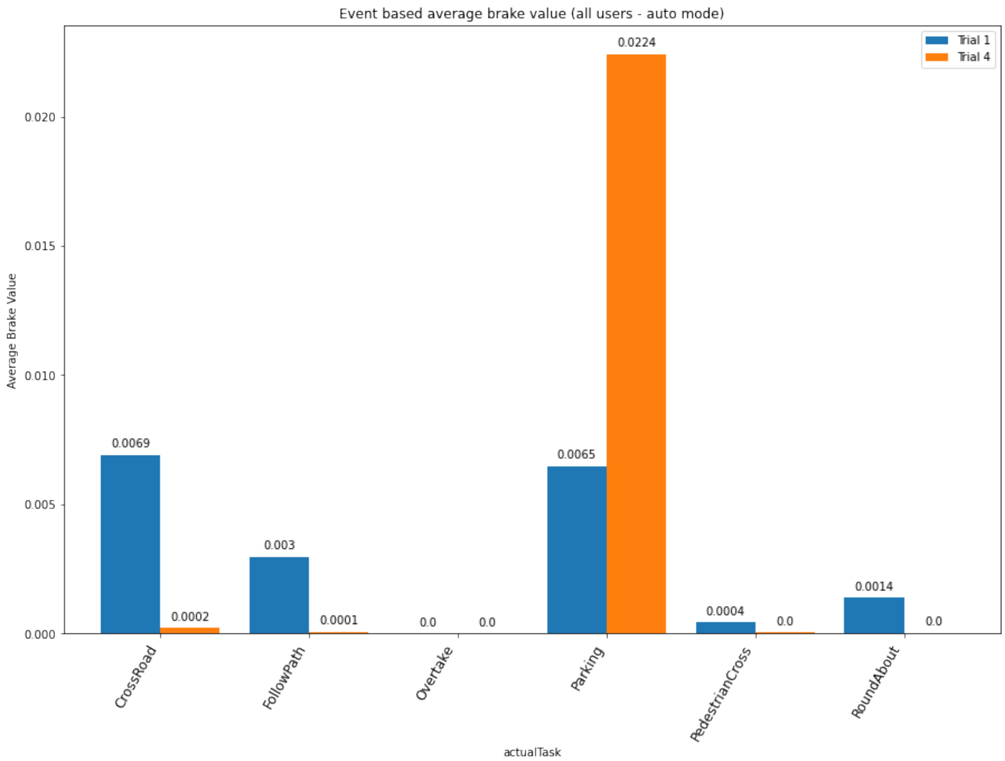


Figure 5.3: Average pressure applied on break

Responses to alerts (warning) to take back control and return to manual falls between the first and last trial, this applies to all cases except the pedestrian crossing incidents (Figure 5.4).

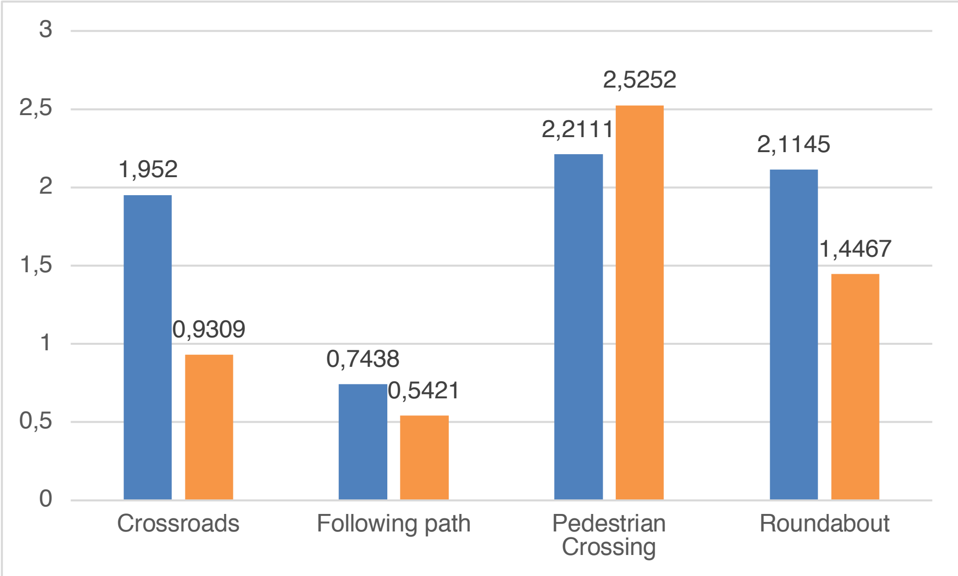


Figure 5.4: Warning Response time trial

Lane deviation (Figure 5.5) across the entire manual task, after the vehicle has handed back control to the driver also varied across tasks and between visits

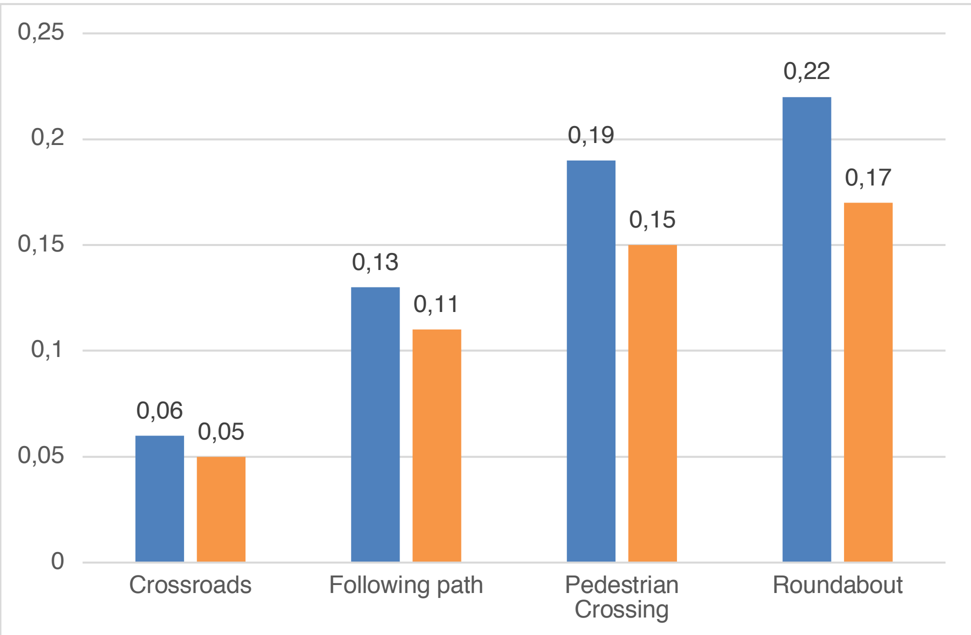


Figure 5.5: Average Deviation from lane centre trial

5.3.2 Acceptance Questionnaire Data

The following section presents the results from the acceptance questionnaire. Descriptive statistics are used as the number of participants is quite low.

Q#1 How did you feel while traveling in a CAV?

Possible Answers:

- Trustful
- Careful
- Insecure
- Unsafe
- Nervous
- Curious
- Critical
- Unaffected
- Neutral

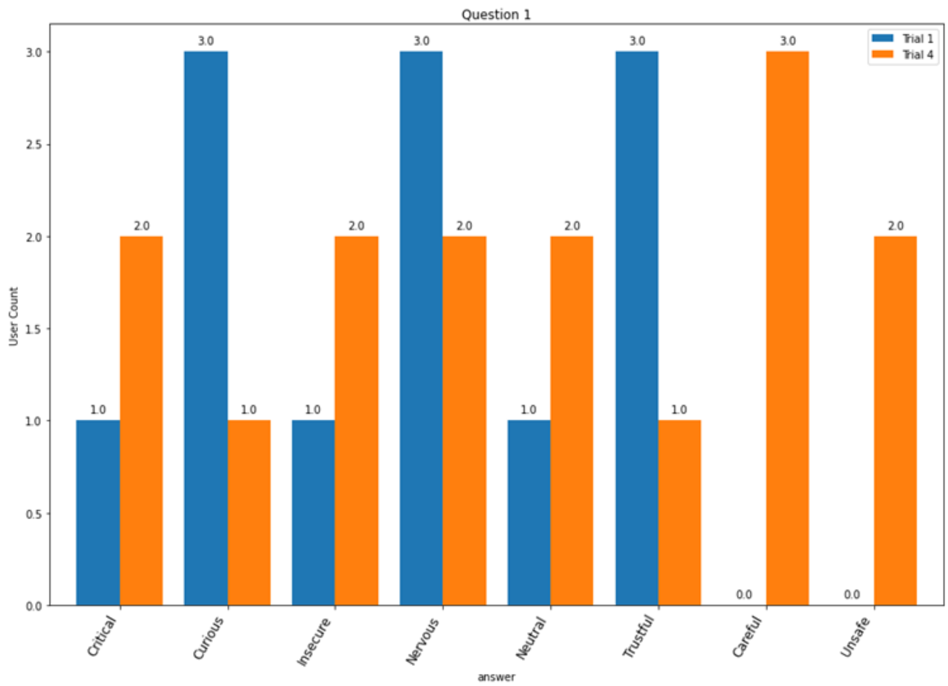


Figure 5.6: How the participants found traveling in a CAV

The overall experience (Figure 5.6) is reported by participants they felt less 'nervous' in the last trial compared to the first, which could be explained by learning effect formed throughout 4 trials and the confidence gained by experience. However, 'careful' and 'unsafe' emerged only in the last trial which means the participants felt that they must be more careful and felt unsafe regarding the CAV experience over 4 trials. One explanation could be the familiarity with the system resulted in being more immersed in the task and becoming more aware of the consequences. The experience is reported as less 'trustful', more "insecure" and more 'critical' which could be linked to the previous point as well as participants being less curious.

Q#2 How did you find the experience of using a CAV?

Possible Answers:

- Positively surprised
- Negatively surprised
- It was as I expected
- I do not know

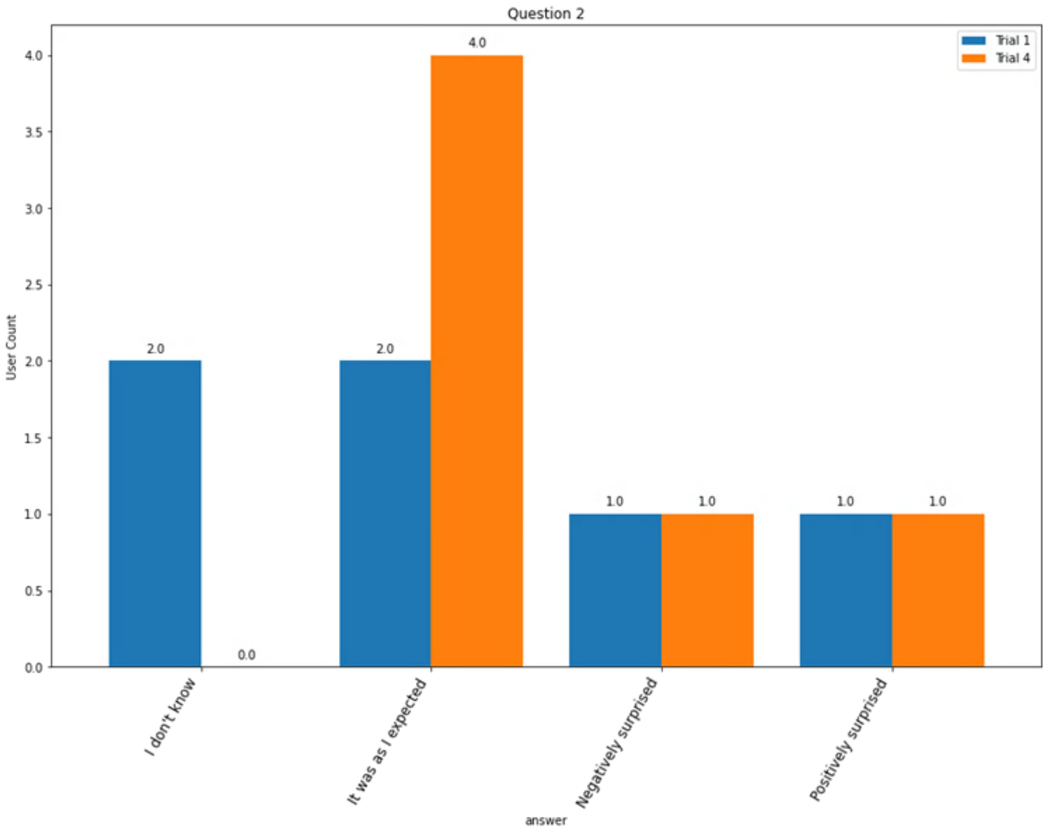


Figure 5.7: How the participants found the experience of using a CAV

Regarding the experience of using a CAV (Figure 5.7), the expectation is formed over time (they became familiar with the system and knew what to expect) and participants found it as expected. Participants responded on balance neither negatively nor positively in terms of feeling surprised. The ambiguity reported in the first visit ("I don't know") disappeared in the last one due to the familiarity with the system.

Q#3 How do you describe the reactions of the CAV?

Possible Answers:

- Very good
- Predictable
- No opinion
- Unpredictable
- Very Unpredictable

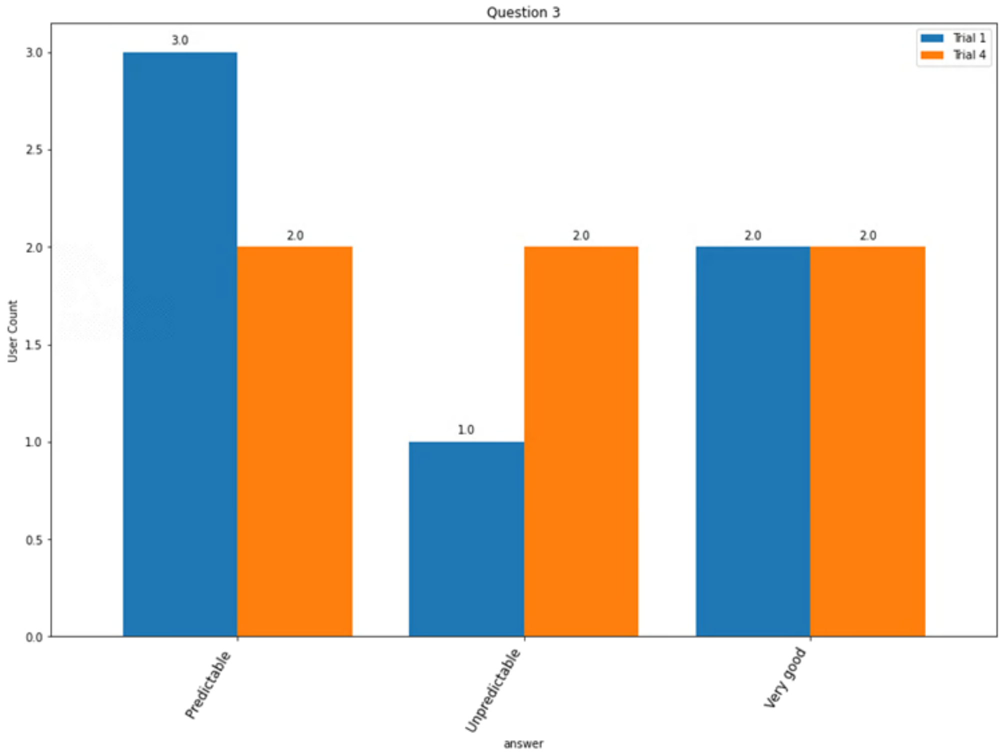


Figure 5.8: Understanding the perceived reactions of the CAV

When it comes to the reaction of the CAV (Figure 5.8), the value for 'predictable' has dropped by 1 unit and the value for 'unpredictable' increased by 1 units. The feeling of it being very good remained constant between both trials. Part of these findings could be explained reaction of the system in manual mode when they have the control and compare it to the automated experience.

Q#4 I found understanding the operating limits of the autonomous mode of the CAV?

Possible Answers:

- Very Easy
- Easy
- Neither easy nor difficult
- Difficult
- Very difficult

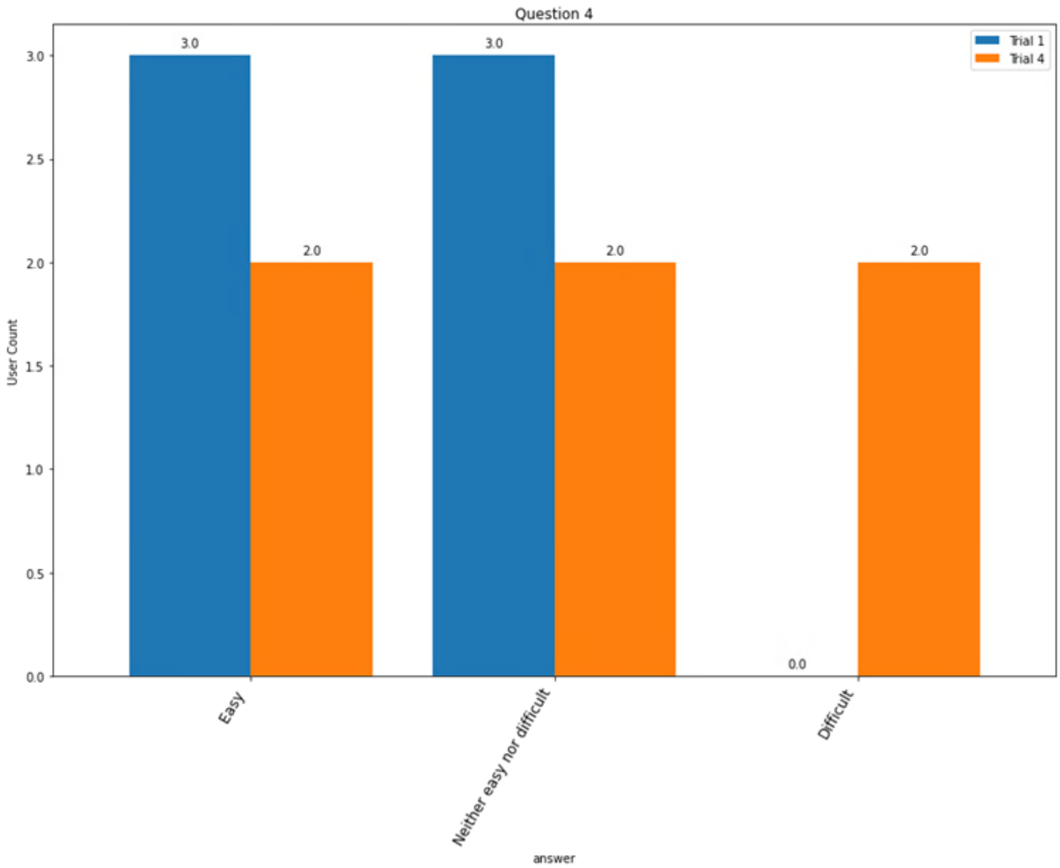


Figure 5.9: How the participant found understanding the operating limits of the CAV

Between the first and last visit it seems that the participants found understanding the operating limits of the vehicle more difficult (Figure 5.9), with ratings for easy and neither easy nor difficult falling. It is unclear why this would be the case after repeated exposures, however it points to the need for clear training on operating limits.

Q#5 Did you hear the warning signals?

Possible Answers:

- Every Time
- Some of the time
- None of the time

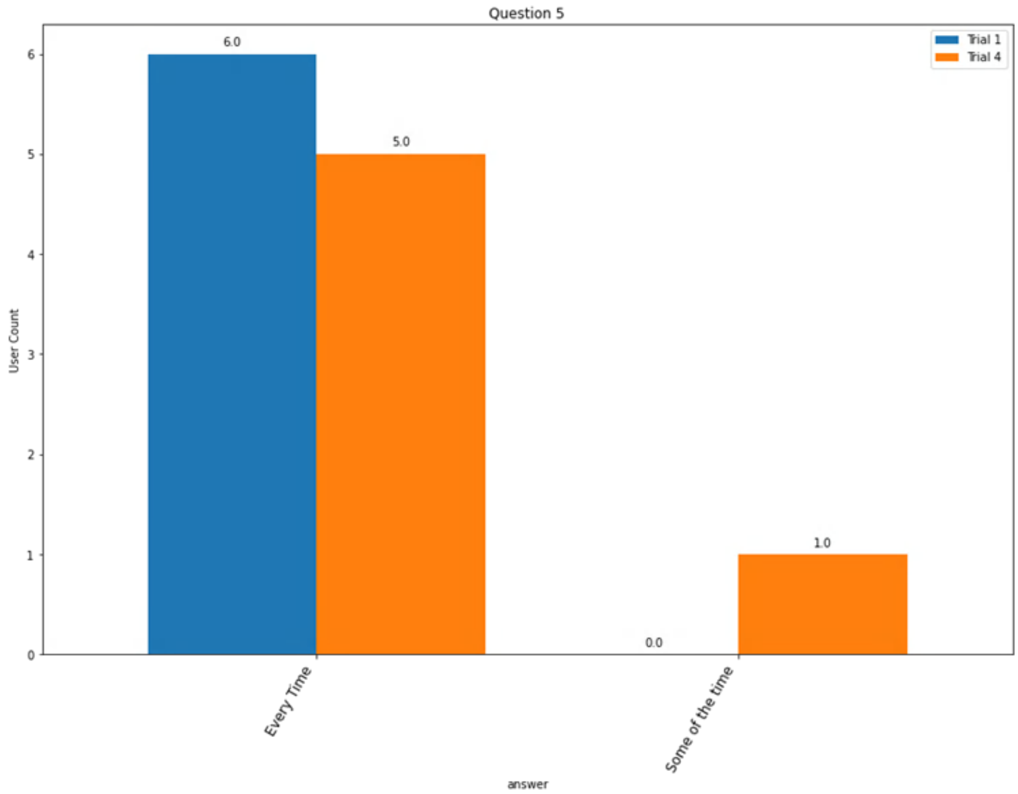


Figure 5.10: Hearing Warning Sounds

There was a small drop in the number of people hearing the sounds all of the time when the vehicle mode changed (Figure 5.10). However, more people reported hearing it some of the time, rather than providing no response.

Q#6 Was it easy to change the driving mode (autonomous/non autonomous)?

Possible Answers:

- Very easy
- Easy
- Neither easy nor difficult
- Difficult
- Very difficult

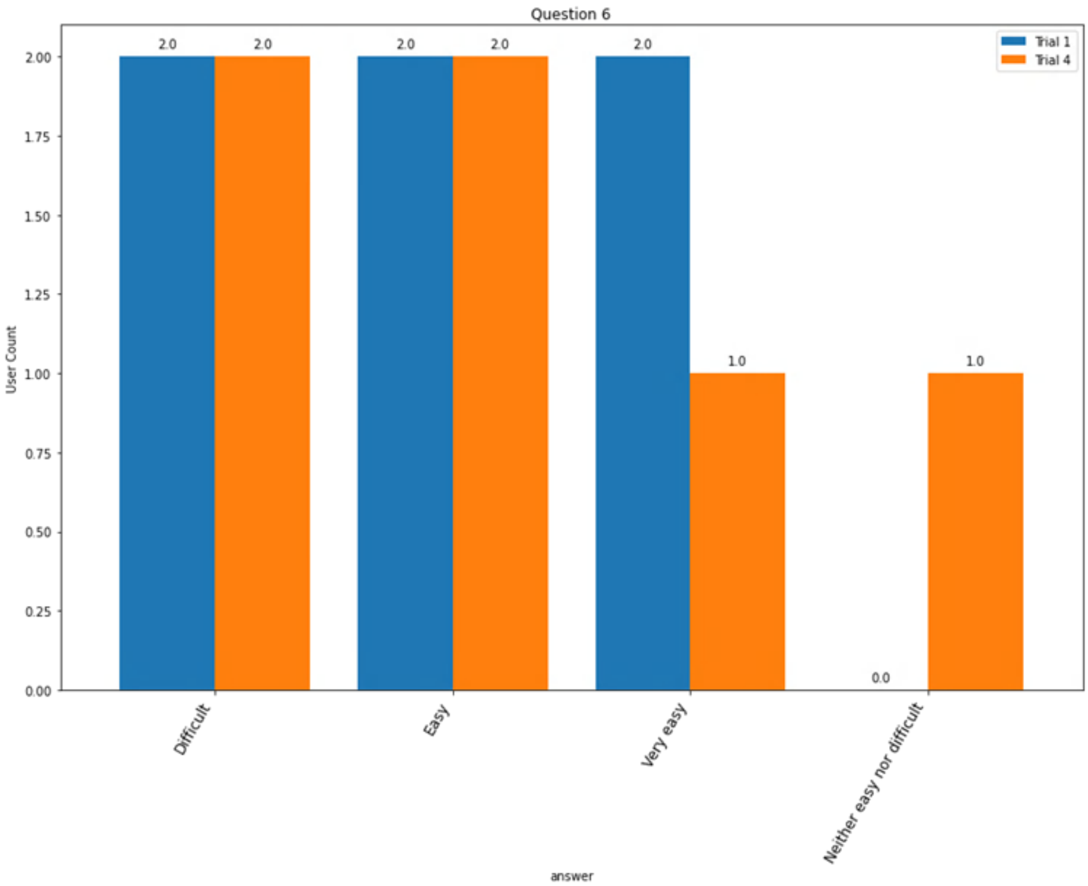


Figure 5.11: How easy was it to change the driving mode

There was a fall in the number of people reporting finding it easy to switch between driving modes (Figure 5.11), however this was a drop mainly reported in the form of neither easy nor difficult, so it requires further investigation.

Q#7 How did you find the amount of information provided by the on-board system?

Possible Answers:

- Much too little Information
- Barely too little Information
- The Right Amount of Information
- Barely too much Information
- Much too much Information

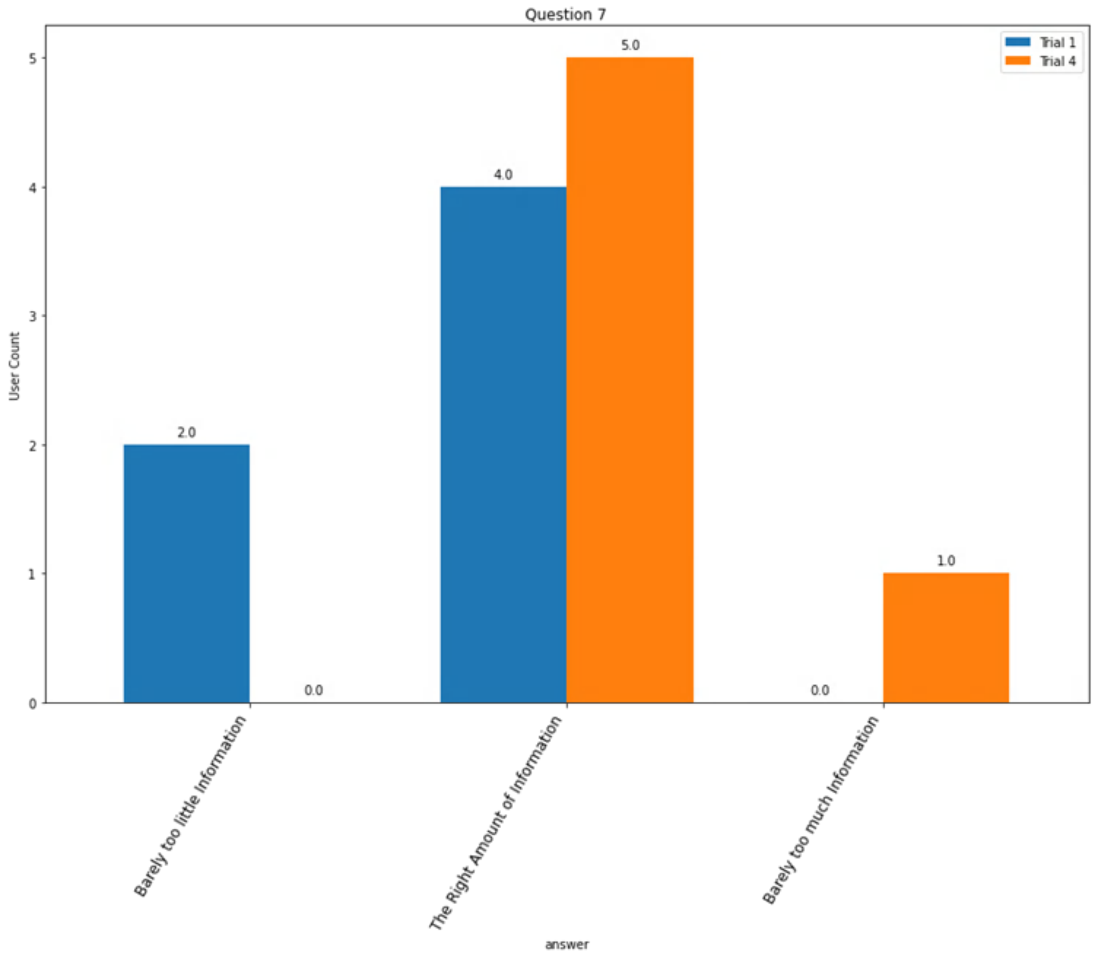


Figure 5.12: Participants views on how they found the amount of information presented by the on-board system

Regarding the amount of information provided by the on-board system (Figure 5.12), the reported value is increased in the last session which could have its root in more familiarity and confidence over the 4 visits that led them learn how to use the system.

Q#8 Did you feel confident with the onboard system?

Possible Answers:

- Very Confident
- Fairly Confident
- Neither Confident nor lacking in Confidence
- Not very Confident
- Not Confident at All

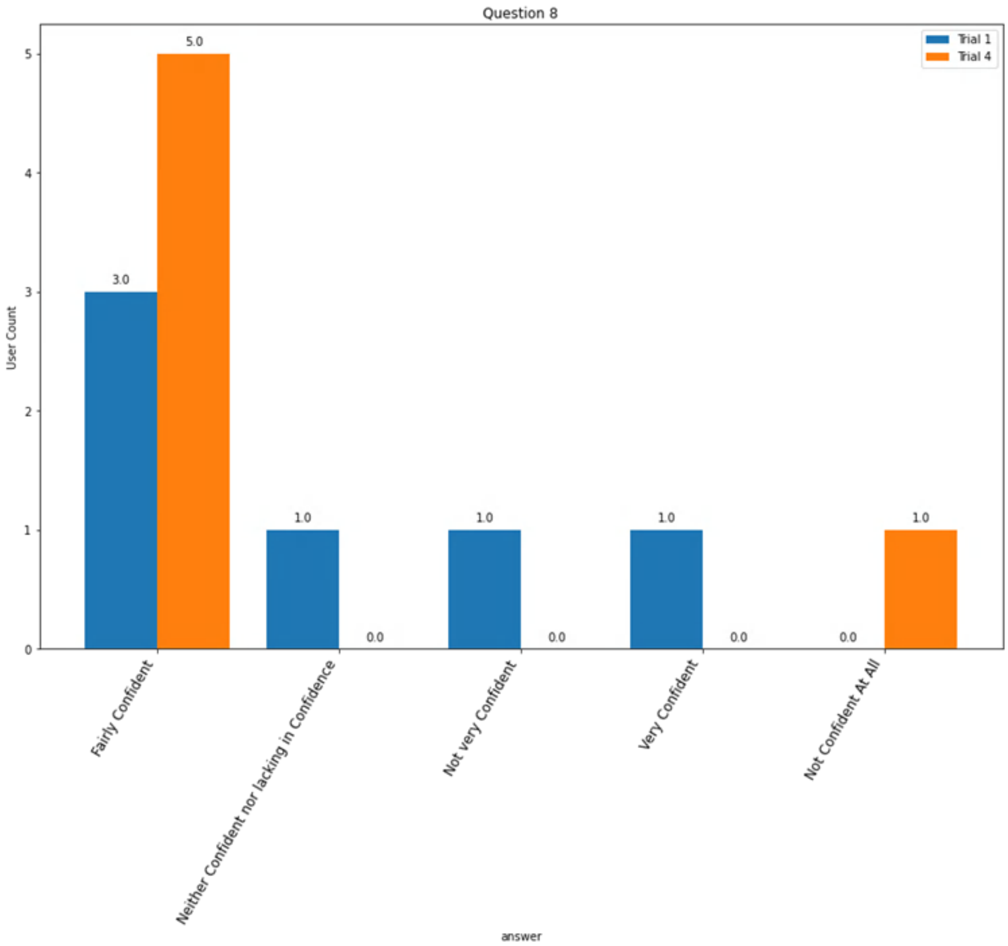


Figure 5.13: Confidence with the On-Board System

The number of responses varies between each trial (Figure 5.13), making interpretation with such a small dataset difficult. There is a rise in the number being fairly confident, but a drop in being very confident and one indicated not being confident at all. This would point to a slight drop in overall confidence.

Q#9 How safe did you feel when the vehicle was in autonomous mode?

Possible Answers:

- Very safe
- Safe
- Neither safe nor unsafe
- Unsafe
- Very unsafe

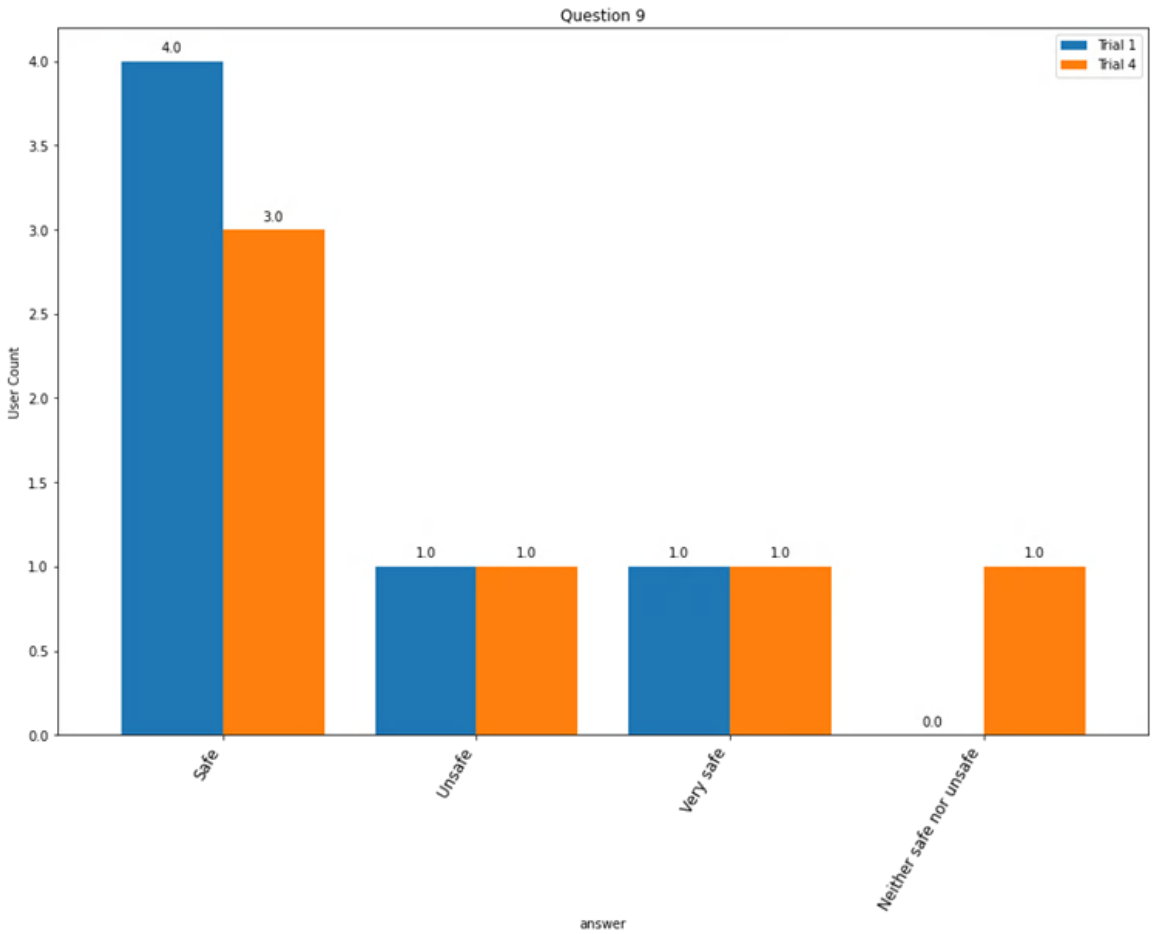


Figure 5.14: Feeling while in autonomous mode

Regarding feeling ‘safe’ when vehicle was in autonomous mode (Figure 5.14), the values has dropped in the 4th visit. The reason could be the expectation that the participants have built over the 4 visits. Comparison of their experience when they have the control of the vehicle in terms of managing the situation and when the system itself does it, could result in less confidence towards the autonomous mode.

Q#10 During the changeover to autonomous mode how safe did you feel?

Possible Answers:

- Very safe
- Safe
- Neither safe nor unsafe
- Unsafe
- Very unsafe

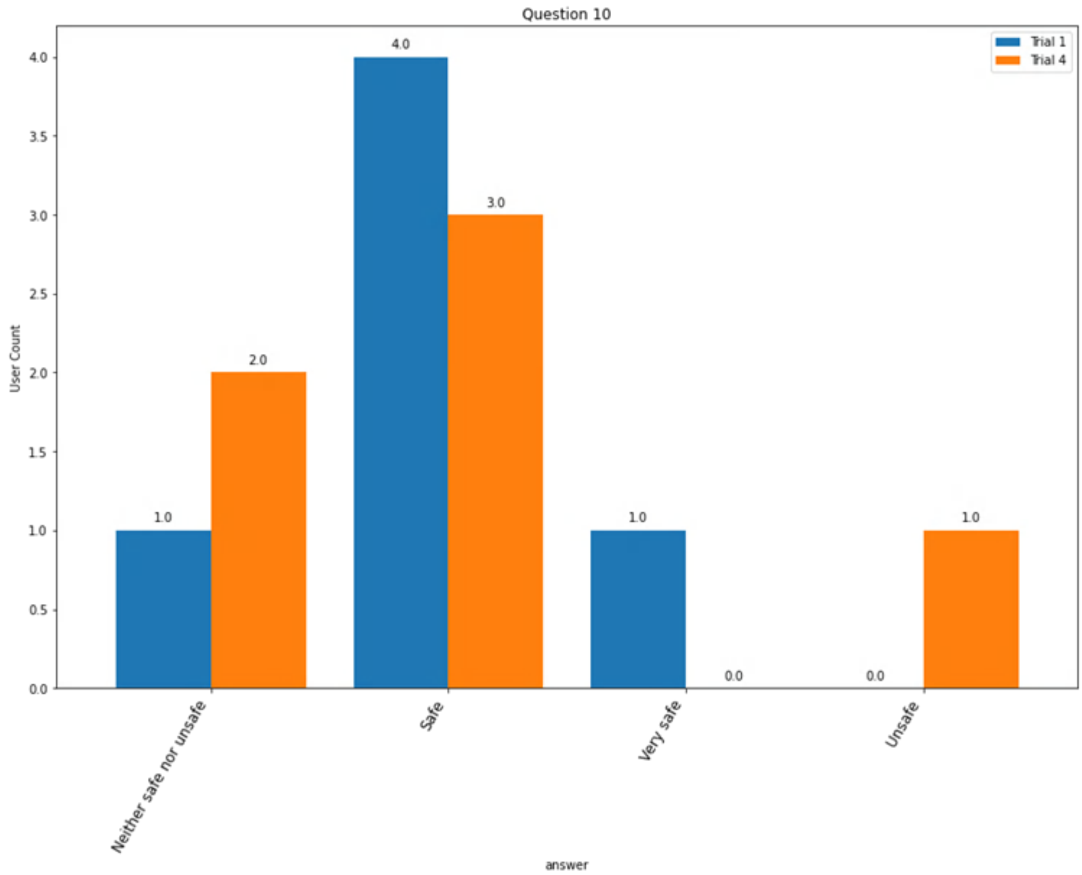


Figure 5.15: Perceptions of changeover to autonomous mode

These answers are discussed in relation to the next question.

Q#11 During the changeover to manual mode how safe did you feel?

Possible Answers:

- Very safe
- Safe
- Neither safe nor unsafe
- Unsafe
- Very unsafe

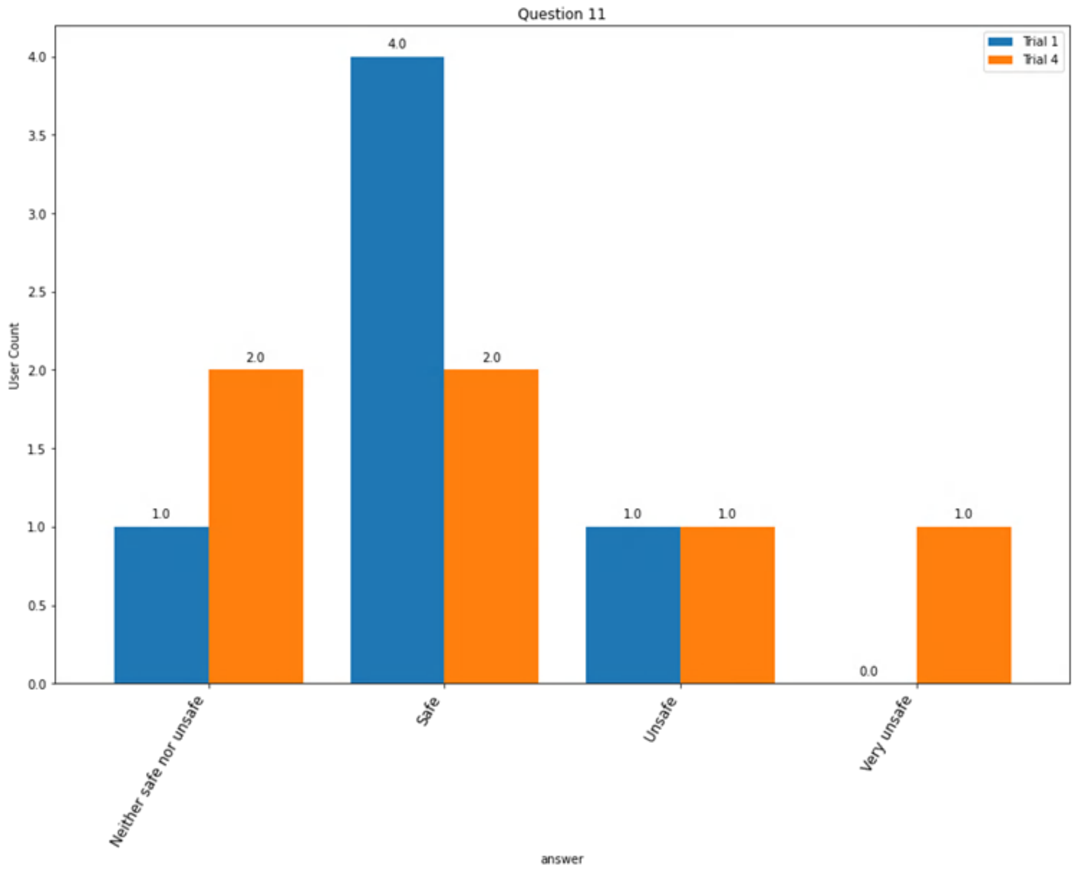


Figure 5.16: Feelings of safety when transferring to manual mode

Across both mode switch types from manual to autonomous (Figure 5.15,Figure 5.16), and autonomous to manual there is a drop in the feelings of safety. This may be due to becoming more immersed in the task, rather than focussing on the novelty of the simulator and also as perceptions change they have higher expectations.

Q#12 How safe did you feel when you were driving in manual mode?

Possible Answers:

- Very safe
- Safe
- Neither safe nor unsafe
- Unsafe
- Very unsafe

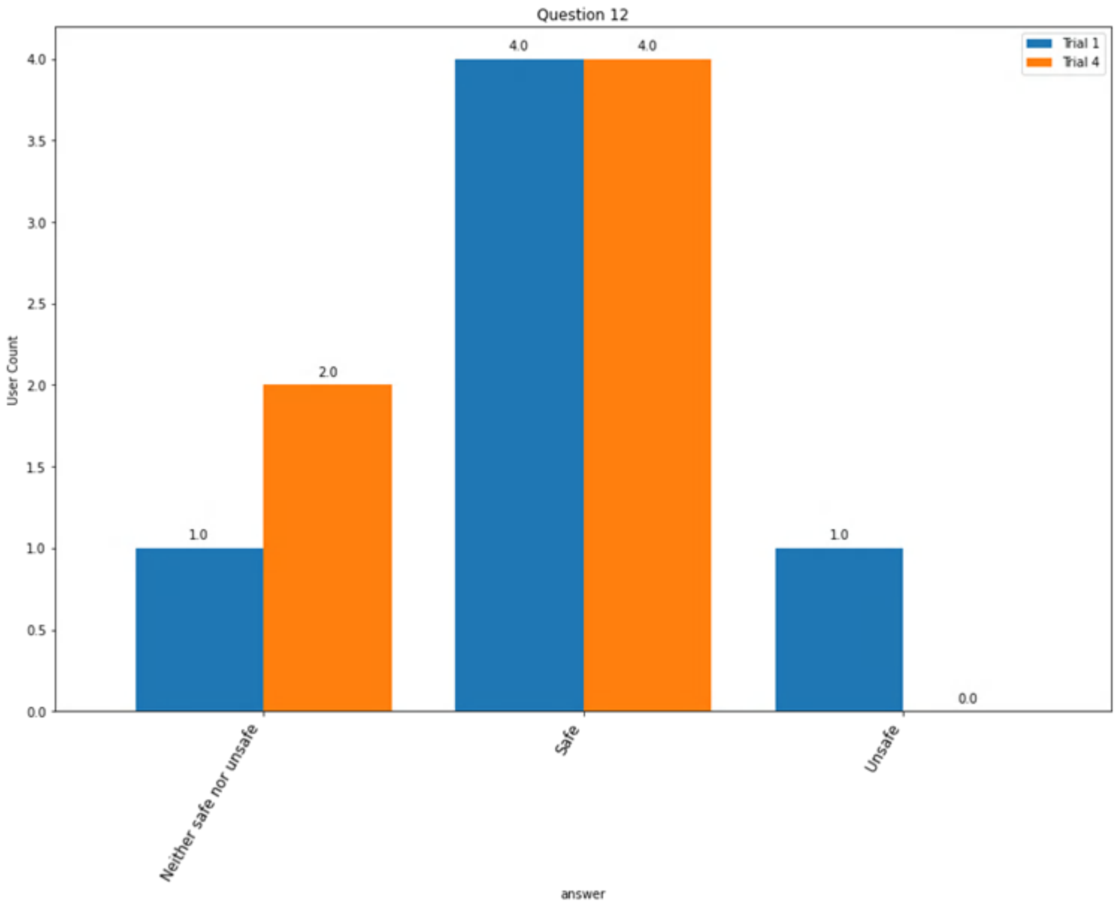


Figure 5.17: Feeling of safety when driving in manual mode

Having the control of the vehicle together with the experience coming from four trials led the participants to report the overall experience of the manual mode safer over time (Figure 5.17), this is primarily indicated in a drop in feelings of unsafety. Again, the numbers are very low so should be taken as an indication of the need for further work.

5.3.3 Cognitive Load (NASA TLX) Data

All 8 users filled in the NASA TLX questionnaire at the end of each of the 4 driving sessions. We collected 31 questionnaires (one session could not be assessed due to technical problems).

We performed a “raw analysis” of the evolution of the perceived global task load without weighting the different items of the tool. We compared the perceived task load for each item (mental demand, physical demand, temporal demand, performance, effort, frustration as well as the global perceived effort), over the four sessions and among the different users. In the following section, the question text is taken directly from the NASA TLX questionnaire (NASA, 1986).

As can be seen across the data, each scale has a maximum score of 20, with 1 being the lowest. In general, the ratings were all in the middle to lower and of the scale. The cognitive load calculated for all scales equals the sum of all scales and normalised to a score out of 100. There is no generally agreed set of thresholds (MeasuringU, 2022) for what is considered a high or low cognitive load.

5.3.3.1 Overview of the perceived task load over the sessions

The perceived cognitive load (Figure 5.18), calculated over a 100 scale, decreases along the driving sessions. This can be explained by a learning effect or progressive appropriation of autonomous driving. It should be noted that the ratings cover the entire level 4 driving experience and not any specific aspect e.g. change of mode.

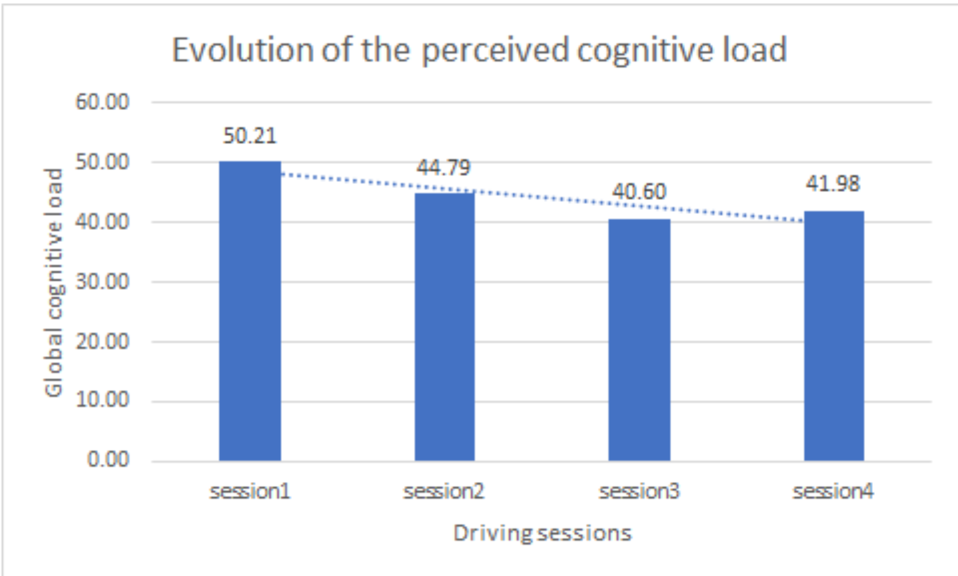


Figure 5.18: Global Rating for Cognitive Load (NASA TLX)

The largest drop in overall perceived cognitive load was between the first and second visit, although the drop remained largely in place until the last visit to the simulator.

5.3.3.2 Comparison of perceived task load compared over users

As can be expected there are discrepancies among the perceptions of users of the required task load (Figure 5.19). The largest gap in subjective perception among users is related to the perceived frustration (standard deviation of 5.27 among users), while the lowest is related to perceived effort (standard deviation of 3.55 among users).

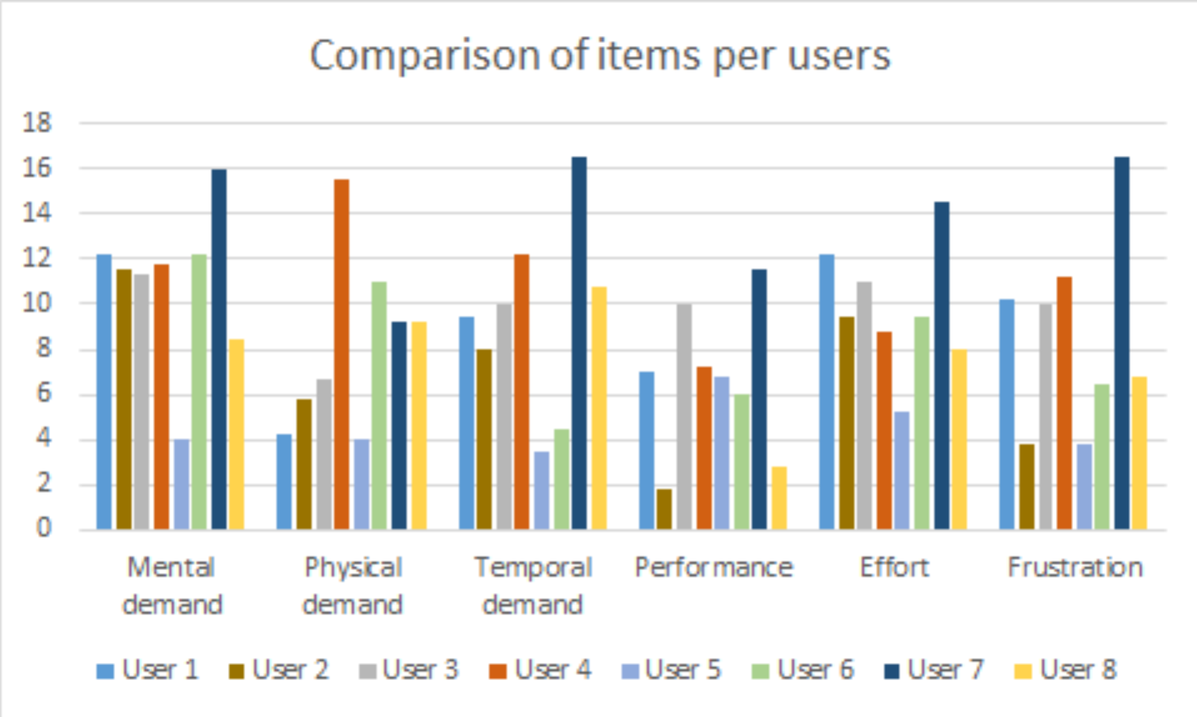


Figure 5.19: Comparisons Across Users for Different Aspects of the TLX Scale

5.3.3.3 Evolution of perceived mental demand

How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

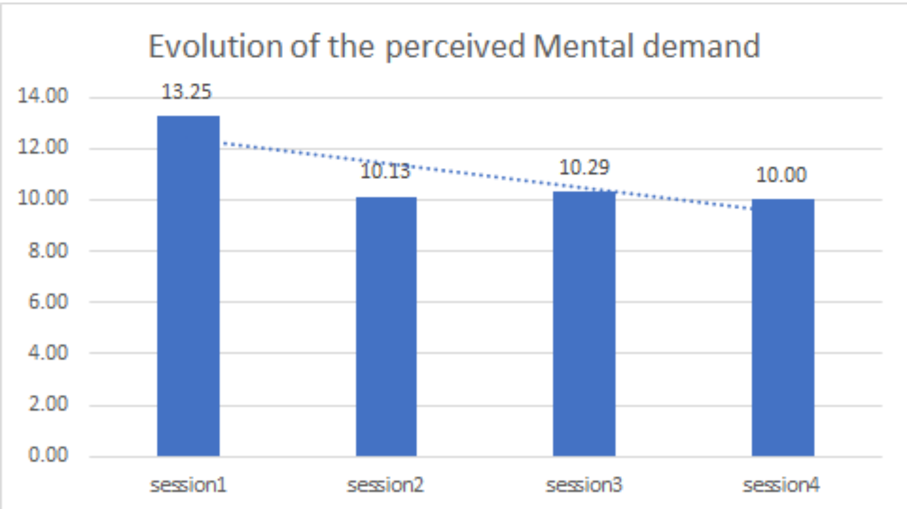


Figure 5.20: Mental Demand (NASA TLX)

Mental demand is perceived to have dropped between the first and last visit. Although it should be noted that this occurs mainly after the first visit only.

5.3.3.4 Evolution of perceived physical demand

How much physical activity was required, (e.g. pushing, pulling, turning, controlling, activating? Was the task easy or demanding? Slow or brisk, slack or strenuous, restful or laborious?

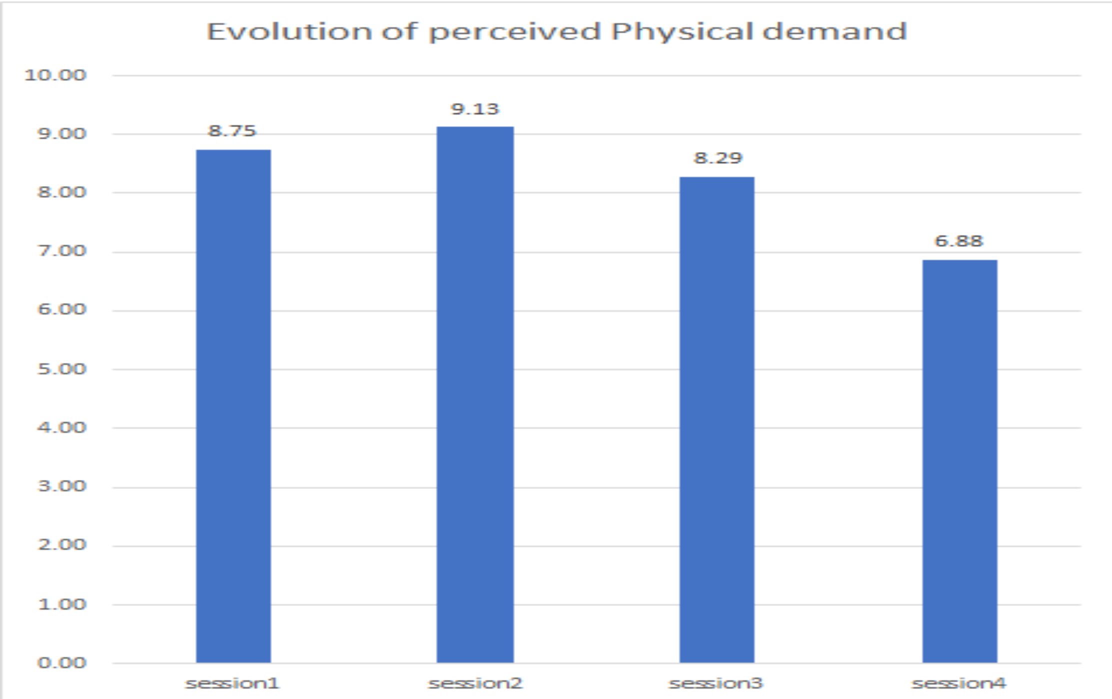


Figure 5.21: Physical Demand (NASA TLX)

Physical demand initially increased, then fell during the final visit.

5.3.3.5 Evolution of perceived temporal demand

How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

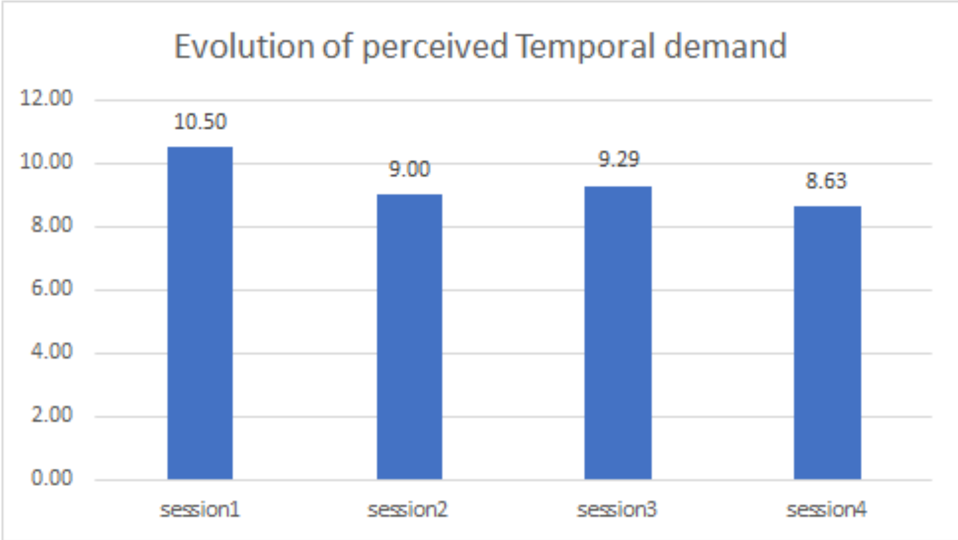


Figure 5.22: Temporal Demand (NASA TLX)

There is a drop in the feeling of being under time pressure when driving the vehicle between the first and last visit.

5.3.3.6 Evolution of perceived effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

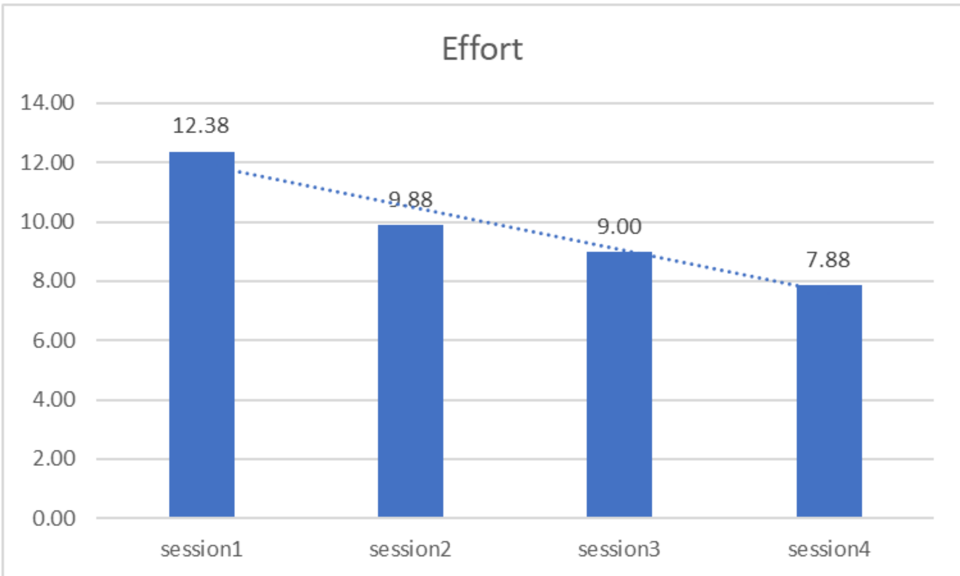


Figure 5.23: Perceived Effort (NASA TLX)

The perceived effort required to complete a task fell between the first and last visit.

5.3.3.7 Evolution of perceived performance

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? how satisfied were you with your performance in accomplishing these goals?

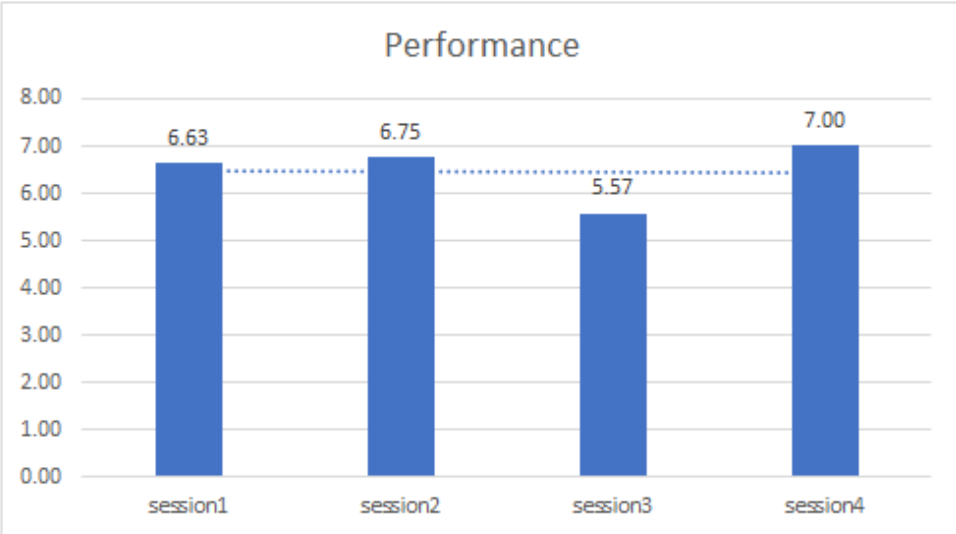


Figure 5.24: Perceived self-rating for performance (NASA TLX)

The participants indicated that they felt their performance stayed roughly the same between the first and last visit.

5.3.3.8 Evolution of perceived frustration

How insecure, discouraged irritated, stressed, and annoyed versus secure, gratified content, relaxed and complacent did you feel during the task?

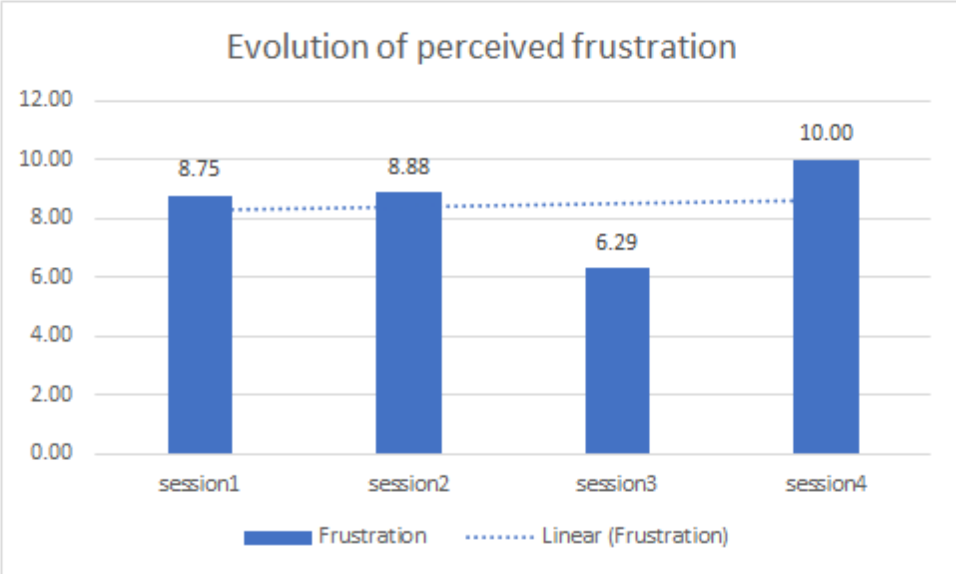


Figure 5.25: Frustration Level (NASA TLX)

Frustration between the first and last visit increase.

5.3.4 Repertory Grid Data

The raw repertory grid for participant 21 is presented in Table 5.1. In the above example there are no identical constructs, however similar issues between the first and last round are identified. For example, in both the first and last visit issues with the transition between automated and manual mode were highlighted. Also, the ability of the vehicle to obey traffic rules was noted as positive.

Table 5.1: Sample: Participant 21 Repertory Grid

First Visit		Last Visit	
Negative	**Positive**	**Negative**	**Positive**
The crash was odd, bad signal is annoying	Following traffic rules by autonomous driving was interesting	Bad signals and speed limit warning signal were annoying	Paying attention to people crossing the road without crossing lines in autonomous mode is ok

Sometimes movement, turning and breaks during autonomous driving are too harsh	Autonomous driving was fun	Transition from manual to autonomous not smooth sometimes	Pulling over when bad signal was happening was ok
I was not relaxed during autonomous driving but was ok.	Interested if autonomous driving can identify blind spot with pedestrian crossing	Some sequences like blind spot with pedestrian only happened in manual and pedestrian crossing the road without the lines only happened in auto	Obeying traffic rules in auto was interesting

While the repertory grids approach appeared useful in collecting the data, it should be noted that during the first visit participants at times found it confusing or complicated. However, during the second visit they seemed to understand its purpose better.

Table 5.2: Elements from the first (left) and last visit (right)

88.3% double lane junction (a) single lane roundabout (a)	89.2% double lane junction (a) single lane roundabout (a)
85.8% junction (a) double lane junction (a)	84.2% junction (a) double lane junction (a)
82.7% junction (a) pedestrian crossing (a)	80.2% single lane roundabout (m) double lane roundabout (m)

82.1% single lane roundabout (m) double lane roundabout (m) 81.5% junction (a) single lane roundabout (a) 80.3% pedestrian crossing (a) single lane roundabout (a)	

The matches between elements in the first visit, display a split along automated vs manual lines (Table 5.2). For example. All matching elements are either automated or manual mode, indicated as (a) or (m). This indicates, even with this limited data set that the participants separated their conceptual mapping of automated and manual modes. The same applies during the last visit, however in this case fewer elements were related. The percentages indicate the similarity in ratings provided across all constructs for that particular element. For example, 80.2% of participants in the last visit rated single lane and double lane roundabouts similarly, regardless of the underlying construct.

Data Sample

To simplify the analysis the decision was taken to only select as close as possible to the 20 strongest matching constructs from the first and last visit. This resulted in constructs with a match of 83.3% being chosen. As a result to ensure a balance in terms of match on each side, 83.3% was chosen from both visits. This results in 24 constructs from the first visit and 30 from the last visit (Table 5.3). Pure numerical comparisons are not really appropriate for repertory grid analysis, so the emphasis should focus on the qualitative aspects which are complemented by the quantitative aspects. It should be noted that the matching constructs may be from different participants. The matching process evaluates if the ratings are similar across elements. Future publications will analyse more of the repertory grid data.

During the final visit, participants seemed more able to understand how repertory grids worked, therefore the responses in general seemed to much more be along the lines of clear negative and positive constructs. During the first visit many constructs often conveyed varying degrees of negative feelings e.g. Frustration-Fear of Vehicle Behaviours. In the brackets (1,F) indicates 1st construct line and the first visit, while (3,L) indicates the 3rd construct last visit. The constructs are split across the three following tables, the term "(reversed)" is used when the constructs were rated occasionally as positive on the left rather than the right of the grid. However, the ratings of 1-7 for the elements remain the same.

The data within the table can be interpreted as follows. For example, in the first visit, first line. The construct: "Frustration - Fear of Vehicle Behaviours" is related to feeling like taking back control-Autonomous Mode", with a matching level 94.4% which indicates the ratings strongly match between the two constructs. In the following interpretation of the data, further related constructs are explored when the same one appears in another location along with another construct. Also, where a similar part of a construct appear (e.g. a positive or negative element) this may also be discussed. The left side of the construct represents a negative feeling, while the right side represents a positive feeling. It should be noted that the data presented here is verbatim, it has not been recategorized in anyway. Finally, references to a crash refer to when an error in the simulator resulted in a vehicle crash. This has been retained in the data to illustrates the impact of such events on the participants, but it should be regarded as on outlier.

Table 5.3: First 10 Constructs from the first and last visits

	First Visit	***Last Visit***
	Construct	***Construct***
1	Frustration—Fear of the vehicle behaviours Feeling like taking back control—fear of an unexpected ending of autonomous mode (reversed)	Discharge /concentration—Anticipation Discharge /concentration—Décision-making
2	Worried to create an accident—Relaxing Annoying and interruptive—Relaxing	uncertain—zen Mental Load—Control

3	Motion sickness—fun Steering wheel moving too much (playing)—Autonomous controllable	Quick Reactions Needed—Long/Slow Interrupting—Relaxing
4	Control (control)—Knowing/being aware whether the car has detected danger Leaving the roundabout was stressful—Autonomous driving the joyful in some situations/Good speed regulation in auto driving	Trust—Decision-making Anticipation—Unexpected events
5	Steering wheel moving too much (playing)—Autonomous controllable Frequent change disturbing—Fatigue after 20min	No Control—Control Quick Reactions Needed—Long/Slow
6	Stressful, not clear what to do—Easy to deal with Worried to create an accident—Relaxing	Mental Load/Hurry/stressed—Zen Mental Load—Control
7	Steering wheel moving too much (playing)—Autonomous controllable Frustration—Fear of the vehicle behaviours	Must be careful most of the time to avoid accident—Safe (traffic light) Waited for auto mode to take over—Relaxing
8	Frequent change disturbing—Fatigue after 20min Frustration—Fear of the vehicle behaviours	Mental Load/Hurry/stressed—zen Quick Reactions Needed—Long/Slow
9	Frustration—Fear of the vehicle behaviours	Trust—Decision-making Autonomy—Unexpected events

	Bad visibility/Very dangerous—In auto driving I didn't care about signals and felt little bit safe.	
10	Feeling like taking back control—fear of an unexpected ending of autonomous mode (reversed) Bad visibility/Very dangerous—In auto driving I didn't care about signals and felt little bit safe.	No Control—Control uncertain—Zen

5.3.5 First Visit

During the first visit the matching constructs contained larger amounts of more basic human instincts such a fear, frustration and aspects relating to technical matters such as the behaviour of the steering wheel and bad visibility. The latter we suspect is related to the fact the system used a single large display, which meant that participants often had to use a button on the steering wheel to emulate looking around through the window. There were only two constructs mentioning road incidents e.g. roundabout (4,F) and the zebra crossing (12,F), in both cases this was in a negative context.

In terms of vehicle behaviours, the constructs point to the participants not feeling comfortable during the first visit. This seems to be related to aspects such as unexpected behaviour or ending of autonomous modes (1,F) (24,F) which related to feelings of fear. The movements of the steering wheel (3,F)(5,F)(7,F)(20,F), which due to the nature of the system used were quite course and pronounced were mentioned as undermining the experience (5,F) and were related to feeling of frequent change. Conversely when this was not the case, the autonomous control was perceived as positive. Steering wheel movements were also perceived as being related to fear and frustration. It was also related to positive feelings of taking back control and negative feelings of an unexpected end of autonomous control (20,F). Conversely, when the steering wheel was not moving too much it gave rise to positive feelings of control.

Relaxing is noted as a positive aspect of the experience (2,F)(6,F) (14,16-17,F) and (19,F). Worrying about creating an accident, the annoying and interruptive nature of the mode change alerts gave the opposite to feeling of relaxation. Driver fatigue was rated on the positive side and may be related to feelings of relaxation (5,F), rather that its more negative safety related context. We would caution against this being seen as positive outcome.

Behaviour of the vehicle and its operating limits (18,F) (20,F) (21,F)(22,F) were indicated by a number of participants. For example (4,F) points to understanding the limits giving rise to a sense of control, which in turn is positively related to autonomous mode being joyful, relaxation and having good speed regulation. While bad visibility is related to the nature of the simulator, and is not so relevant here it is worth noting that detection of danger by the vehicle is again related to control, and in turn caring less about signalling when in automated mode (10,F)(20,F) and feelings of relaxation (16,F). Construct (9,F) relating to fear is more difficult to interpret as both poles appear negative.

Table 5.4: Second Group of Constructs, first and last visits

	First Visit	***Last Visit***
	Construct	***Construct***
11	Fear that the vehicle might not be able to detect some elements—Relaxed Bad visibility/Very dangerous—In auto driving I didn't care about signals and felt little bit safe.	No Control—Control Mental Load—Control
12	Insecure in the beginning—Safe/get used to it Traffic rules follow— No zebra/they will cross wherever they want (reversed)	Trust—Decision-making Discharge/concentration—Decision-making
13	Bad visibility/Very dangerous—In auto driving I didn't care about signals and felt little bit safe. Leaving the roundabout was stressful—Autonomous driving the joyful in some situations/Good speed regulation in auto driving	No Control—Control Mental Load/Hurry/stressed—zen

14	Frustration—Fear of the vehicle behaviours Annoying and interruptive—Relaxing	Uncertain—Zen Quick Reactions Needed—Long/Slow
15	Autonomous controllable—saccade (reversed) Felt Relaxed—Didn't feel anything special (reversed)	Discharge/concentration—Anticipation Autonomy—Unexpected events (reversed)
16	Bad visibility/Very dangerous—In auto driving I didn't care about signals and felt little bit safe. Annoying and interruptive—Relaxing	No Control—Control Interrupting—Relaxing
17	Caution (Prudence)—Visibility Worried to create an accident—Relaxing	Mental Load—Control Quick Reactions Needed—Long/Slow
18	Control (control)—Knowing/being aware whether the car has detected danger Bad visibility/Very dangerous—In auto driving I didn't care about signals and felt little bit safe.	Autonomy—Unexpected events Discharge/concentration—Decision-making
19	Feeling like taking back control—fear of an unexpected ending of autonomous mode (reversed) Annoying and interruptive—Relaxing	Mental Load/Hurry/stressed—Zen Uncertain—Zen
20	Steering wheel moving too much (playing)—Autonomous controllable	Mental Load/Hurry/stressed—zen Hurry—predictable

	Feeling like taking back control—fear of an unexpected ending of autonomous mode (reversed)	

5.3.6 Last Visit

Mental load was a dominant theme (2,L)(6,L)(8,L)(11,L)(13,L)17,L)(18,L)(19,L)(20,L),(26,L)(28,L)(29,L) although it should be noted that this construct came from a small number of participants, but in turn seems related to the responses in many constructs and elements. A high mental load, was viewed negatively by participants with aspects such as uncertainty, hurry/stress, lack/no control, the interruptive nature of the alerts and the need for quick reactions being connected to this aspect. On the positive side, where there was a perceived level of control this was related to feelings of zen, the need for longer and slower reactions, relaxation and predictability. Alongside this as the construct Mental Load-Control (2,L)(17,L)(16,FL) which also points to participants feeling that sense of control has decreased when there is a high mental load.

In addition to mental load issues, the need for quick reactions was perceived negatively (3,L)(5,L),(8,L)(14,L),(17,L) was also related negatively to aspects such as the interruptive nature of the mode changes, lack of control. Conversely, when quick reactions were not the case, this increased levels of control, Zen, perceptions of autonomy being easy (29,L) and relaxing (30,L).

Decision making undertaken by the vehicle was rated positively across many constructs (1,L)(9,L)(12,L)(18,L)(21,L). This was related to aspects such as trust, ability to anticipate the action of the vehicle, and concentration levels.

The feeling of Trust is related to decision making, appear related to aspects such as the ability to anticipate, unexpected events, decision making and autonomy (4,L)(9,L)(12,L)(21,L). Decision making was mentioned positively in the constructs, along with anticipation, with the need to concentrate being on the negative side.

Table 5.5: Remaining Rep Grids

21	Fear that the vehicle might not be able to detect some elements—Relaxed	Discharge/concentration—Anticipation Trust—Decision-making

	The crash was odd, bad signal is annoying—Following traffic rules by autonomous driving was interesting	
22	Fear that the vehicle might not be able to detect some elements—Relaxed Control (control)—Knowing/being aware whether the car has detected danger	steering wheel too reactive—slowing down Quick Reactions Needed—Long/Slow
23	Motion sickness—fun Unsecure in the beginning—Safe/get used to it	Not enough perspective—Get ready to take control No problem at the speed—Fine
24	Frequent change disturbing—Fatigue after 20min Feeling like taking back control—fear of an unexpected ending of autonomous mode (reversed)	No Control—Control Attentive—No problem at the speed
25		Bad visibility—get ready to take control Discharge/concentration—Anticipation
26		Must be careful most of the time to avoid accident—Safe(traffic light) High Mental Load—Undemanding
27		Hurry—predictable Quick Reactions Needed—Long/Slow

28		Bad visibility—Get ready to take control Mental Load/Hurry/stressed—Zen
29		Mental Load/Hurry/stressed—Zen Stressful—easy
30		Waited for auto mode to take over—Relaxing Many things to consider—Relaxing

5.3.7 Rep Grids Comparison

As vehicles will be used more than once, the primary interest is to explore similar themes/constructs remain across all present across both visits and also to explore how they could inform future design. Although there is some scope for considering the level of experience a "driver" may have with a vehicle and providing some adaptions to the experience (see later paragraph). For example, in both visits the interrupting nature of the mode change warnings remains a problem. Where this is resolved, this should reduce levels of stress and increase relaxing nature of the experience.

The perception of being in control, even in an autonomous vehicle is considered vital to the participant having a positive experience. This issue extends from how aspects such as the steering wheel behave through to the need for quick reactions, and high mental load. Feelings of control seem to improve where less quick reactions are required or uncertainty are removed.

Mental load (and the need to concentrate) is an issue even during the last visit, both from a positive and negative perspective. It is tied to issues to do with quick reactions being required, loss of control, and uncertainty. During the last visit, issues relating to the decision making capability of the vehicle become more pronounced, so it would imply that transparency relating to this is an important aspect.

In general feelings of fear and frustration were not as prevalent in the last visit. Also, there were less constructs relating to understanding the operating limits of the vehicle, and that specific incidents e.g. roundabouts were a cause for concern. This would point to a general improvement (not unexpectedly) in awareness of how the vehicle behaves. It also points perhaps to the need to consider providing clearer information during the first and early drives in such vehicles of the operating limitations.

5.4 Discussion

The overall objective of this study was to assess changes in driver perception and performance after multiple visits to a level 4 CAV. Given the sample size and number of visits were restricted due to the Covid pandemic and the multiple parallel studies, these results should be taken as indicative. Over the remainder of the project the intention is to reach a target of 30 participants.

The study used a mix of open-ended approaches (such as repertory grids), predefined questionnaires (including cognitive load via NASA TLX) and quantitative simulator data. This approach allows for an analysis of the objective changes in performance between the first and last visits and possible understandings as to what are the important changes in perception between the first and last visit.

5.4.1.1 Methodological Issues

Repertory grids provide a useful way to understand how participants conceptualise their experiences. We found them a useful way to drill down beyond the NASA TLX scores when identifying relationships between different aspects e.g. how high mental load is related to feelings of stress, or caused by the need for quick reactions. In our study we only provided participants with the elements and the participants were free to choose which constructs to use. This approach as expected makes comparisons between participants more complicated. In any future studies we may look at also providing a list of constructs or phrases to produce constructs based on the ones provided during this study. This would however be at the expense of a rich conceptual understanding. Overall all though, repertory grids have provided us with a significant dataset upon which further work can be undertaken and also an approach to develop further assessment tools based on the results.

The combination of repertory grids and NASA TLX, allowed us to address one short coming of the TLX scales, namely that they only provide a rating and not a list of possible indications as to why or how factors are related. The combination of the two allowed us to overcome part of this problem.

The length of study although shorter than expected has indicated that there are potential variations over time. This may mean that we can use shorter studies, rather than the level initially planned to evaluate such situations.

5.4.1.1.1 RQ 3.1: Does driver acceptance vary?

Repertory grids point to a change in how participants conceptualise level 4 experiences over time. There is a move from the initial set of constructs which often deal with primitives such as steering wheel movement, or emotions such a fear, through to concentrating on higher level aspects of the experience such as mental load decision making, and control. Reported feelings of safety during transitions and while in autonomous mode seemed to decline between the first and last visits. Taken together, these would point to the trajectory of the facets that make up acceptance of such vehicles changing over the number of exposures. The repertory grids and both questionnaires indicated that trust became an important issue, perhaps more so as familiarity of autonomous vehicles grows. This was indicated by a drop in trust levels in the questionnaire alongside an indication of it as a negative aspect in the repertory grids. While the CAV performed better in terms of how participants expected it to behave over time, people found it less predictable over time. Reported levels of confidence also dropped. The vehicle was seen as less predictable during the last visit.

5.4.1.1.2 RQ 3.2: Does driver cognitive load vary?

NASA TLX indicates that overall cognitive load drops between the first and last visit, although for some scales increases during the first and last visit. One interesting deviation from this is that frustration seems to rise between the first and last visit. The repertory grid analysis also identified cognitive load as an issue for the participants. Furthermore, it allows (as noted in the recommendations) for some suggestions on how to address the situation.

5.4.1.1.3 RQ 3.3: Does driver performance vary?

Due to the small data sample size it is not possible to give clear guidance here. Therefore, the results presented here should instead indicate possible future work rather than a definitive conclusion.

There was a change in driver performance between the first and last visit to the simulator. This was noted across four main aspects: adherence to speed limits (when changing to manual mode and in it), lane deviation, use of breaks while in autonomous mode and responses to warnings. All of these can be interpreted in a number of ways, for example as would be expected as people become more familiar with the system they understand how it operates and how they can manoeuvre it better. Furthermore, the drop in use of breaks can perhaps be attributed to the increasing acceptance and trust by the participant in the system.

5.5 Ways to improve CAV design

- In this study only a single alert followed by a countdown were used, which resulted in participants often feeling stressed or hurried.
  - Recommendation: timings should be increased, it may also be possible to provide phased in alerts e.g. long before control change is needed and which are based on wider contexts prior to the actual event occurring.
- Alerts are often seen as annoying and interruptive, which also had negative impacts on the participants feelings towards the CAV and increased mental load and feelings of control.
  - Recommendation: user interfaces should be used which convey the need to resume control, but which in themselves do not increase mental load and feelings of loss of control.
- Perceptions of the CAV change over time this was noted in the nature of related and matching responses which shifted often from feelings of fear (first visit) to issues to do with control and decision making (last visit). Although much more work is required, a change in driving performance may also be observed, of particular interest was the increased use of breaks in autonomous mode during the first visit, this also points to increased fear during the first visit.
  - Recommendation: CAVs need to provide more scaffolding information about its operating limits and potential behaviours during the first few visits by the driver.
- The ability to predict how and what kind of decisions the CAV will take was seen as positive. Uncertainty was perceived negatively and fear of an unexpected end to autonomous mode was present.
  - Recommendation: provide simple interfaces which aid (if people require it) to understand how decisions are made.
  - Recommendation: remove any ambiguity in expectations as to how the vehicle will behave, it should always react to a certain cue/event in the same way.

- A level 4 vehicle still requires manual intervention, which in turn places a responsibility and hence the need to be attentive at all times on the driver. This can be seen to be contributing to the level of cognitive load.
    - Recommendation: methods should be found to reduce the requirement to need to heavily concentrate on the driving experience at all times. This could be achieved both through user interface technologies but also more effective training on the operational limits of the vehicle.
- The movement of the steering wheel in autonomous mode can be quite intimidating for people during the first visit.
    - Recommendation: while the steering wheel should indicate the direction of movement, it should not be too sudden or jerky.

5.6 Guidelines and recommendations for pilot specifications

5.6.1 Use cases

In general, the recommendations above apply across all potential use cases, at least from the perspective of the driver. However, there are perhaps some specific use cases which should be explored.

	Pilots				
	1	2	3	4	5
• Emergency situations – all situations in this study were relatively common incidents with little in the way of unexpected high risk events.	X	X		X	
• Testing across “brands” – while there is a standard for each level of automation there is no uniform agreement on what specific automated features a vehicle must provide. There needs to be some work on how this potential variant impact on drivers and pedestrians.		X		X	

5.6.2 Test variables

	Pilots				
	1	2	3	4	5
• Different alert types including tones, duration and phases.		X		X	
• Provision of familiarity based interfaces e.g. changing information based on how often the driver has used the vehicle. ○ Provision of information on operating parameters during early use. ○ Provision of information on decision making processes		X		X	
• Comparison of initial expectations (e.g. aspects such as safety) and how this impact on driver performance and subjective assessment. ○ Assess what people expect from an autonomous vehicle, then assess how the impact of this being met (or otherwise) feeds into issues of trust, safety and cognitive load.		X		X	
• Steering wheel behaviour. ○ How different (strong vs more subtle) steering wheel movements impact on acceptance and trust.		X		X	
• Impact of the unknown. As it may not always be possible to know all the operating limits of a vehicle, trust is in part driven by how it may respond to unexpected events. ○ Explore perceptions and behaviours if there are a range of minor or major changes in road context which may lead to unexpected CAV behaviour.		X		X	

5.7 Conclusions

Further work with more participants is required, this will be undertaken over a period of time after the publication of this deliverable. Results in this study should therefore only be taken as indicative and areas for future work. However, the small sample group and therefore lack of quantitative data is in part compensated for with the rich and detailed qualitative data provided by the repertory grid analysis. As a result, this study can be seen to have identified further areas of work that are required, either within a continuation of this study and/or the pilots. Global level of cognitive load fell over time, however some subscales (areas) remained unchanged or increased, this requires some exploration. There is also a need to explore how expectations change over time and impact on the subjective and objective behaviours of drivers.

From a methodological perspective despite the limited number of participants, studies which take place over time like this one provide an insight into how behaviour and perception of CAVs change with time. The approach also questions the use of simulators for automated vehicle studies where there is only one visit. This is due to the fact that the perceptions people have over visits appears to change, and with it how they would rate parts of the experience. Driving behaviour also changes in some aspects.

6 Findings from Immersive Arena

6.1 Overview

6.1.1 Purpose of Study

The general purpose of this study is to measure how eHMI design impacts pedestrians' understanding, trust, receptivity and crossing behaviour.

The focus is made on the four following research questions:

- RQ 1: How the eHMI design can influence the crossing behaviour?
- RQ 2: What is the most impacting aspect of eHMIs?
- RQ 3: What is the impact of the different aspects of eHMIs?
- RQ 4: How consistent are the eHMI understanding and trustworthiness through different contexts?

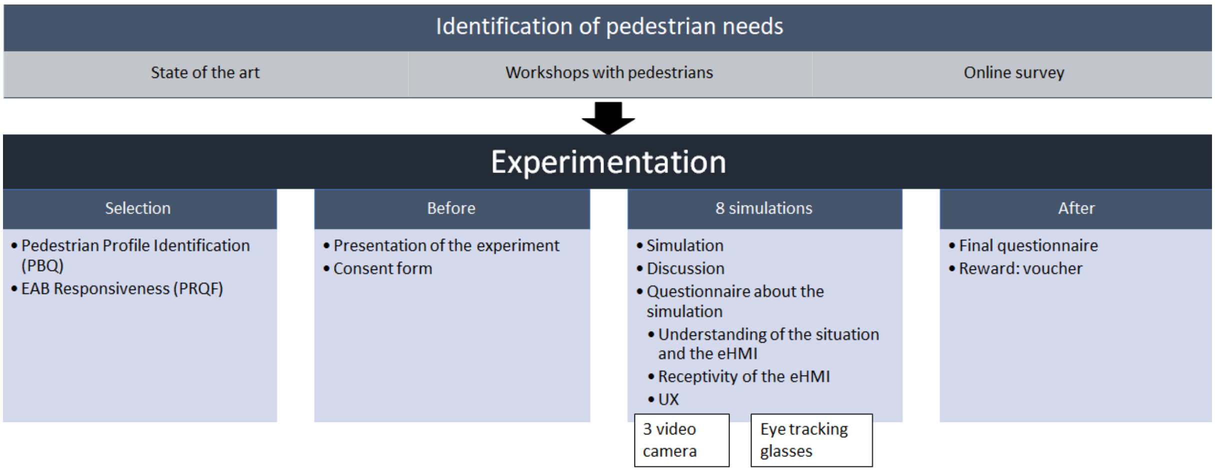


Figure 6.1: Experimentation procedure

Figure 6.2: Overview of one simulation in the Immersive arena

To answer the question a first phase of pedestrian needs identification was done thanks to a state of the art of the literature, two workshops organised with pedestrians (see D4.1) and an online survey.

Then, an experiment was done. During the experiment, to collect the participants' feelings, they are put in a pedestrian situation facing eight simulations of CAVs, in the framework of a technical simulation device (the LIST 360deg Immersive Arena) which displays a street in an urban context all around them. To mitigate learning phenomena, the scenarios are randomly distributed for the participants. After each of the 8 simulations, a questionnaire on the simulation is to be filled in with questions concerning: their understanding of the situation and the eHMI presented; their receptivity of the eHMI presented (questions from PRQF) (Deb, Strawderman, DuBien, Smith, Carruth & Garrison, 2017); the user experience of the eHMI presented (questions from meCUE) (Minge, Thüring & Wagner, 2016); their perception of risk (e.g., ease of understanding and predicting vehicle behaviour, feeling unsafe or confident, confidence in decision making, some questions are from Liu et al.).

Four situations were tested with 2 eHMIs: the CAV stops or not (because the pedestrian was hidden by a work palisade) to let the pedestrian crossing and a pedestrian crossing is paint or not on the road. The two eHMIs react differently in each situation, see Table *6.1* for more information.

The participants' behaviour was filmed during the simulations in order to observe whether they crossed, when they did so (before the CAV arrived, when it stopped, once the CAV had passed), and how long it took them to make their decision, see procedure presented in Figure 6.1.

In the following, the results of the online survey and the experimentation are presented and discussed.

Figure 6.3: Projection of a light green crosswalk on the road by the CAV

Table 6.1: Scenarios used for the simulations

	No pedestrian crossing painted on the road		**One pedestrian crossing painted on the road**	
	The CAV stops	**The CAV doesn't stop**	**The CAV stops**	**The CAV doesn't stop**
eHMI1	No interaction	The CAV horns its honk and makes a headlight call	No interaction	The CAV horns its honk and makes a headlight call
eHMI2	Projection of a green pedestrian crossing on the road	No interaction	Projection of a green pedestrian crossing on the road	No interaction

6.2 Pedestrian behaviour and reaction

6.2.1 Study demographics

39 users participated in the study, including 15 females and 24 males. The mean age was 40.8, SD = 11.6 (42.9 for female, SD = 13.3; and 39.5 for male, SD = 10.5). The profile of the participants is described below, regarding several sociodemographic characteristics:

1. Gender,
2. Country of residence,
3. Age,
4. Dependent child(ren) under 10 years,
5. Living environment,
6. Travel habit,
7. Walking time per day,
8. Walking frequency per day,
9. Difficulty walking,
10. Victim of a road accident,
11. Possession of a driver's license.

Please note that the study was conducted in French language, in the Luxembourg country, which explains the specific users' profiles (Table 6.2).

Table 6.2: users' profile for the Immersive Arena experiment

Users' profiles	**N = 39**
Gender	
Female	15
Male	14
Country of residence	
Belgium	4
France	23
Luxemburg	12
Age	
Average	40,8
Minimum	25

Maximum	72
Dependent child(ren) under 10 years	
Yes	13
No	26
Living environment	
Urban	24
Rural	15
Travel habit	
By foot	24
By bike	4
By scooter	2
By car	34
By public transport	12
Other	4
Walking time per day	
0 to 15 min.	5
15 to 30 min.	13
30 to 45 min.	9
45 to 60 min.	7
More than 60 min.	5
Walking frequency per day	
0 to 2 times	16
2 to 4 times	17
More than 4 times	6
Difficulty walking	
No	100
Yes, without assistance	0
Yes, with assistance	0

Victim of a road accident	
Never	17
Yes as a pedestrian	2
Yes as a passenger in a vehicle	20
Possession of a driver's license	
No	2
Yes for less than 3 years	1
Yes for more than 3 years	36

In order to complete the users' profiles, we asked to the participants to fill in two scales:

- The Pedestrian Behaviour Questionnaire (PBQ), in order to measure frequency of risky behaviours among pedestrians (Deb, Strawderman, DuBien, Smith, Carruth & Garrison, 2017).
- The Pedestrian Receptivity Questionnaire for FAVs (PRQF), in order to measure the willingness to cross the road in front of a FAV (Deb, Strawderman, Carruth, DuBien, Smith & Garrison, 2017).

The mean total score on the PBQ for the 39 participants is 1.87, with a standard deviation of 0.45 and a median of 1.75. The maximum total score is 3.1 (on 6), and the minimum is 1.3 (on 1). This shows that, in general, participants did not report transgressive behaviour as a pedestrian. It should also be noted that there were no significant differences according to gender, age, nationality, country of residence, whether the participants had at least one dependent child under 10 years old, or whether they had been involved in a road accident.

Regarding the PRQF, the participants obtained a mean total score of 3.61, with a standard deviation of 1.02 and a median of 3.63. The maximum total score was 6.06. (on 7), and the minimum was 1.5 (on 1). These results show a wide disparity in the responses. Several modalities of variables can explain these differences. Indeed, in a one hand, there were no significant differences in responsiveness according to gender, age and whether the subjects had been involved in a road accident. On the other hand, there was a significant difference for the following characteristics:

- subjects without dependent children under 10 years old have a higher receptivity to CAVs ($t(37) = -2.076$, $p < .05$).

- subjects of Luxembourgish and other nationalities (Bulgarian, Algerian and Portuguese) have a higher receptivity towards CAV than French and Belgian subjects ($F (3, 35) = 3.155, p < .05$);
- subjects living in Luxembourg have a higher receptivity to CAV than those living in France and Belgium ($F (2, 36) = 3.326, p < .05$).

Observation 6.1: nationality, living country and having kids under 10 seems to have an impact on CAV receptivity.

6.2.2 Observed behaviours

Several notable behaviours were observed:

- move one foot forward to mark their intention to cross and ensure better detection by the CAV;
- back up when the CAV honks its horn;
- when the CAV does not stop, follow it with their eyes until it disappears from the simulation scene and then refocus on crossing either by crossing without checking or by looking both ways before crossing or by abandoning the simulation;
- when the CAV does not stop, step back and abandon the simulation;
- when the CAV does not stop and honks, make an insulting gesture;
- but also, when the CAV does not stop, crossing after it without checking if another vehicle is coming on the other side.

6.2.3 Decision-making time

Table 6.3 presents the average decision-making time, that mean the time between the CAV stops and the pedestrian cross or the time between the CAV pass and the pedestrian cross. The decision-making time is shorter when no crosswalk is on the road and the CAV stops and it projects a green crosswalk on the road. In this situation the decision-making time is more than twice shorter with the projection of the green crosswalk on the road that if the CAV only wait without sending a message.

Observation 6.2.a: When a CAV stops to let crossing a pedestrian, it is better to send a signal that the CAV will wait the pedestrian's crossing.

The worst decision-making time is when the CAV doesn't stop, no crosswalk is on the road and the CAV doesn't emit any signal.

Observation 6.2.b: it seems better to indicate that the CAV will not stop if it detects to let a pedestrian willing cross.

Table 6.3: Average decision-making time

			Decision-making time (s)		
Crosswalk on the road	CAV behaviour	eHMI	Average	Min	Max
No	Stop	-	3.4	0.1	11.0
		CAV projects a green crosswalk on the road	1.5	0	4.4
	No stop	Headlights and horn	4.4	1.0	13.2
		-	6.1	0	16.0
Yes	Stop	-	1.9	0.3	1.9
		CAV projects a green crosswalk on the road	1.9	0.4	5.2
	No stop	Headlights and horn	4.6	1.6	9.0
		-	4.5	1.0	4.5

6.3 Acceptance of eHMIs

6.3.1 Experiments results

6.3.1.1 Participants' opinions about each scenario

After each scenario, the participants were asked to complete a questionnaire comprising 3 categories of questions, relating to:

1. CAV receptivity, i.e. the willingness to cross the road in front of the vehicle. 6 questions were taken from the pedestrian receptivity questionnaire for FAVs (PRQF);
2. Perceived user experience of the eHMI. 6 questions were taken from the meCUE questionnaire (Minge, Thüring & Wagner, 2016);
3. Perception of the road situation as a pedestrian. 6 questions related to perceived dangerousness, crossing intention, and understanding of vehicle behaviour.

For each scenario, it is then possible to draw up a score for each of these categories. The higher the score, the more positive the results were for the scenario (Figure 6.4).

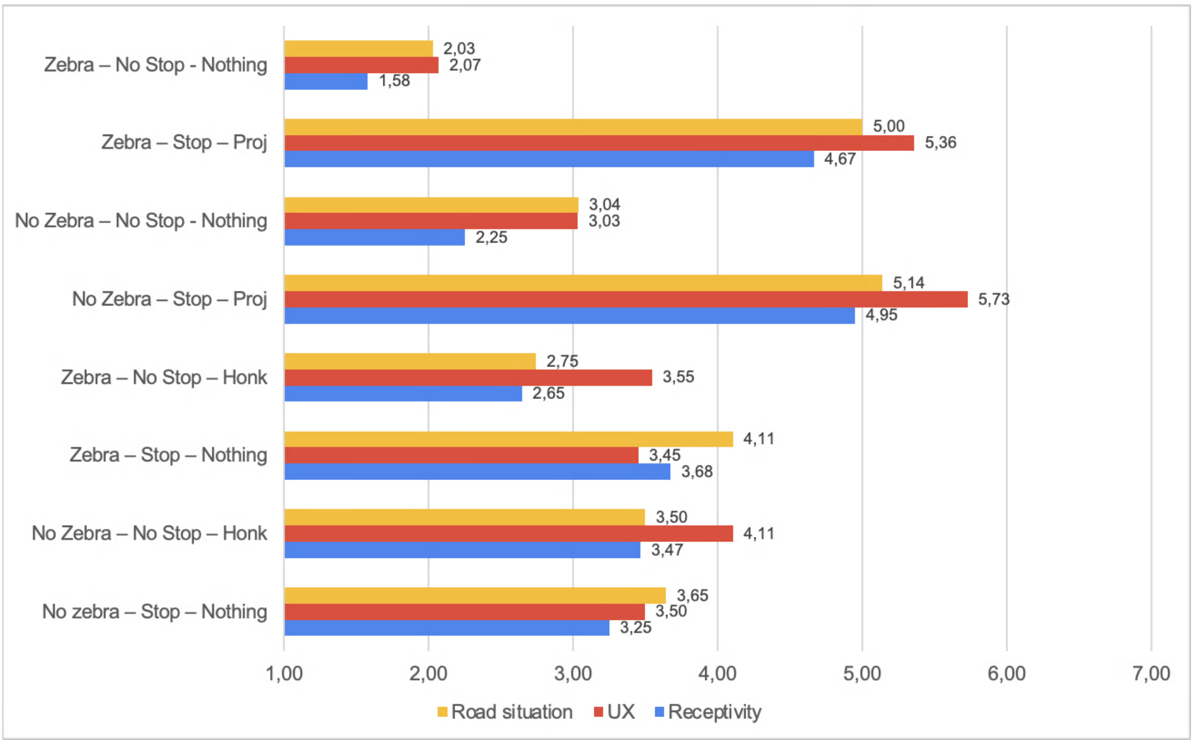


Figure 6.4: Scores for each scenario

In Figure 6.4, Zebra = a pedestrian crossing is on the road. No zebra = no pedestrian crossing is on the road. Stop = the CAV stops. No stop = the CAV doesn’t stop because it saw too late the pedestrian hidden by a work palisade. Proj = a light green pedestrian crossing is projected on the road by the CAV. Honk = the CAV makes a headlight call and honks it horn twice. Nothing = no interaction done by the eHMI.

The results show that the highest scores are obtained for the "No zebra - Stop - Proj" scenario (receptivity = 4.95; UX = 5.73; road situation = 5.14). Indeed, in this scenario, the participants expressed that the light passage projected by the vehicle was clear and very reassuring. The display makes up for the lack of road markings, and therefore allows pedestrians to understand that it is safe to cross. The other scenario which also used the projection of a pedestrian crossing by the vehicle also scores highly (receptivity = 4.67; UX = 5.36; road situation = 5.00). This shows that this eHMI is particularly appreciated by pedestrians. Nevertheless, the scores are somewhat lower than those in the situation without zebra, because the participants consider that the vehicle is obliged to stop at a pedestrian crossing. Thus, the projection of a pedestrian crossing is a reassuring and experiential element, but it is sometimes considered superfluous and unnecessary.

Observation 6.3.a: the CAV has to stop at a pedestrian crossing.

Observation 6.3.b: a signal is waited when the CAV stops to show it will wait the pedestrian's crossing. That confirms observation 6.2.a.

The scenario with the lowest scores is the one where the FAV does not stop, does not transmit any information to pedestrians, even though there is a pedestrian crossing (receptivity = 1.58; UX = 2.07; road situation = 2.03). This behaviour is considered very dangerous by the participants, which explains the very low responsiveness score. It is also noticeable that the responsiveness remains low for the scenario without zebra, and when the vehicle does not stop without transmitting any information (responsiveness = 2.25; UX = 3.03; road situation = 3.04). This indicates that pedestrians are reluctant to accept autonomous vehicles if they behave in ways that are considered dangerous, and without warning pedestrians.

Observation 6.3.c: a feedback is waited by pedestrians in all situations and particularly in dangerous ones. That confirms observation 6.2.b.

The situation with the highest scores (receptivity, user experience, situation perception and decision time) was the one where the CAV projected a crosswalk onto the road when the vehicle was stopped (see Figure 6.5 and Table 6.4). The situations with the lowest scores were those where the CAV did not stop and did not give any warning signal. In both cases (best and worst scores), the results show that the presence or absence of a crosswalk already on the road does not play a significant role.

Observation 6.3.d: presence or absence of a crosswalk already on the road does not play a significant role.

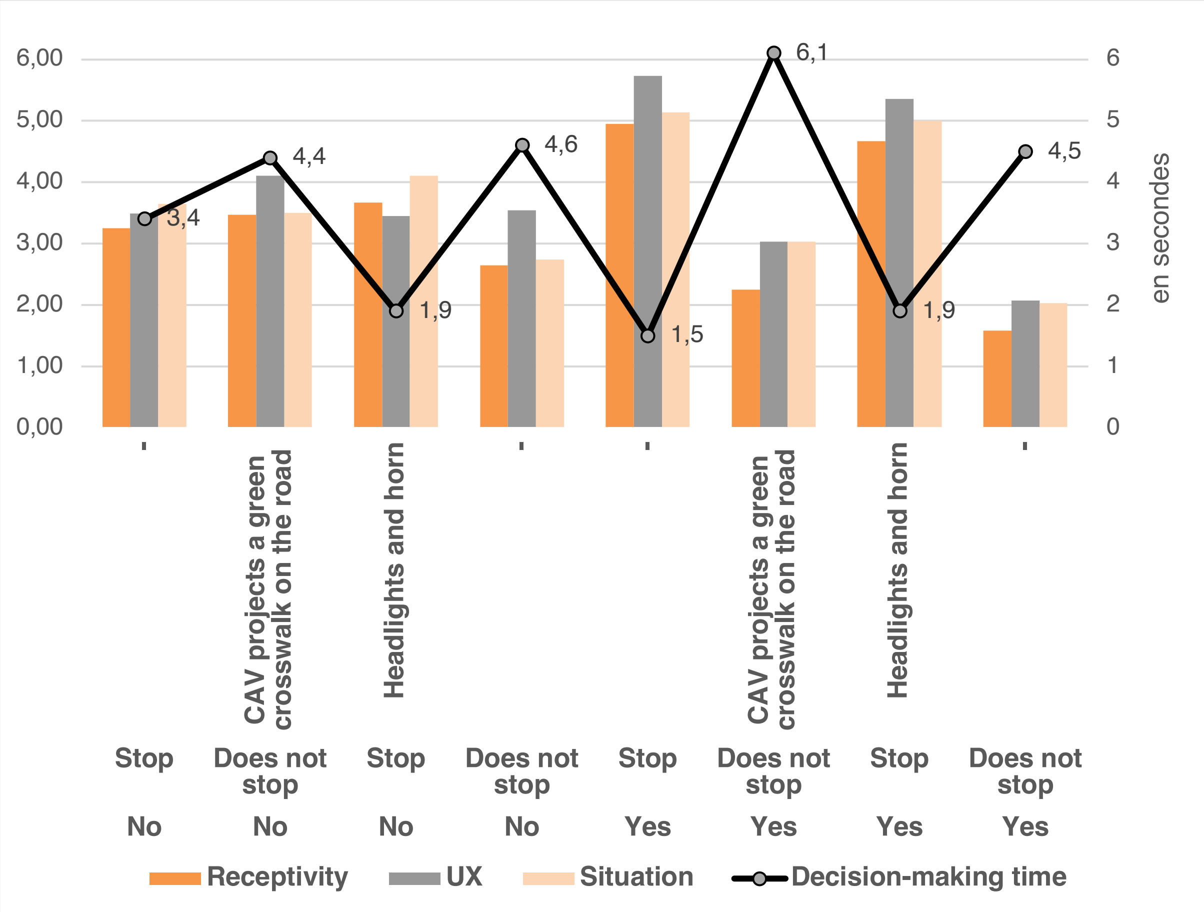


Figure 6.5: *Synthetic view of the participants' behaviour and feelings for each simulation*

Table 6.4: Participants' perception of the different simulations and observed decision-making time

Crosswalk painted on the road	CAV Behaviour	eHMI	Receptivity	User Experience	Perception of the situation	Decision-making time
No	Stop	-	3.25	3.5	3.65	3.4
		CAV projects a green crosswalk on the road	4.95	5.73	5.14	1.5

Crosswalk painted on the road	CAV Behaviour	eHMI	Receptivity	User Experience	Perception of the situation	Decision-making time
	No stop	Headlights and horn	3.47	4.11	3.5	4.4
		-	2.25	3.03	3.04	6.1
Yes	stop	-	3.68	3.45	4.11	1.9
		CAV projects a green crosswalk on the road	4.67	5.36	5	1.9
	No stop	Headlights and horn	2.65	3.55	2.75	4.6
		-	1.58	2.07	2.03	4.5

6.3.1.2 Participants' preferences about eHMIs

If there is NO pedestrian crossing on the road, when the CAV stops, participants prefer that the CAV projects at least a zebra on the road to indicate that it will let them cross (37 of 39 respondents, see Figure 6.6). But 3 respondents would like that the projected pedestrian crossing is augmented by a visual (e.g., written message, LED or headlight call) or audio feedback. One respondent stated that the CAV should only project if the pedestrian is very close. Two respondents would prefer that the CAV doesn't stop to respect the road regulation and one precise that he/she prefer a sound signal to prevent from crossing. 3 respondents prefer when the CAV does not emit a signal.

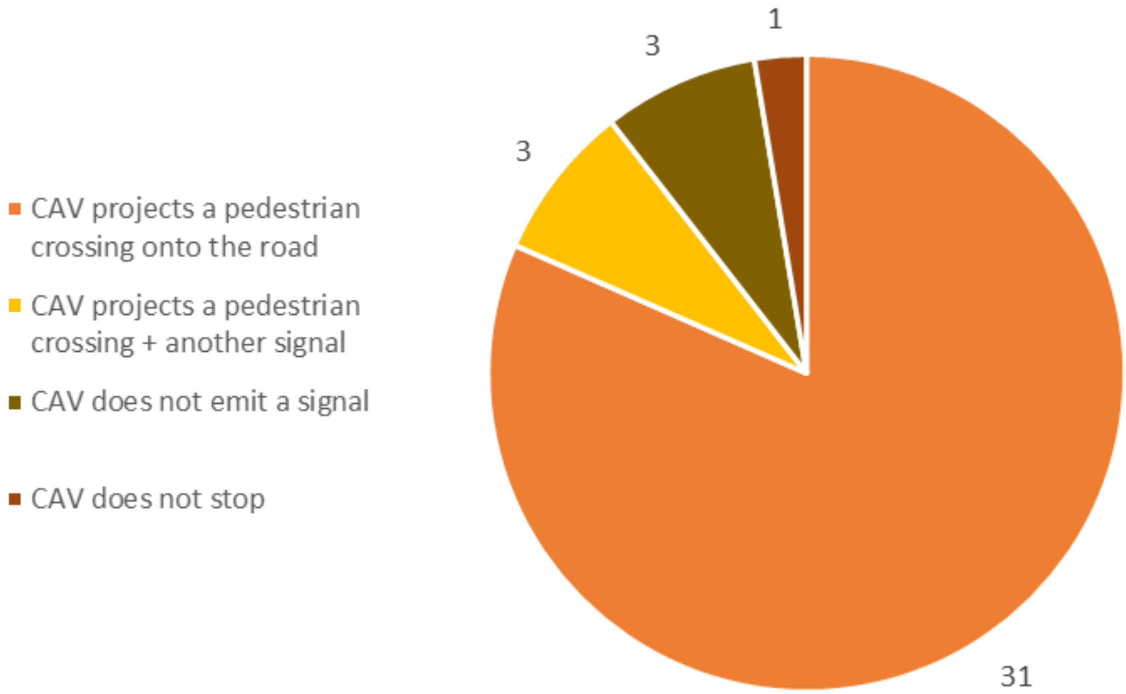


Figure 6.6: When the CAV stops and NO pedestrian crossing painted on the road, pedestrians would prefer that a pedestrian crossing is projected.

If ONE pedestrian crossing on the road, when the CAV stops, 32 (of 39) participants would prefer that the CAV projected at least a signal on the road, see Figure 6.7. Two of them would prefer another signal that a pedestrian crossing (e.g., a green pedestrian). 4 respondents would like that the projection is accompanied by another signal like sound or something visual on the CAV, particularly to indicate before the stop that the CAV detected them. 5 respondents would prefer that the CAV does not emit a signal. And 2 respondents would prefer that the CAV gives another signal that the projection on the road, like sound or visual signal on the CAV.

Observation 6.4.a: When the CAV stops the use of a signal to show that the CAV is waiting that the pedestrian cross is needed. The projection on the road is well accepted in the cases of a pedestrian crossing is painted or not on the road. That confirms observations 6.2.a and 6.3.b.

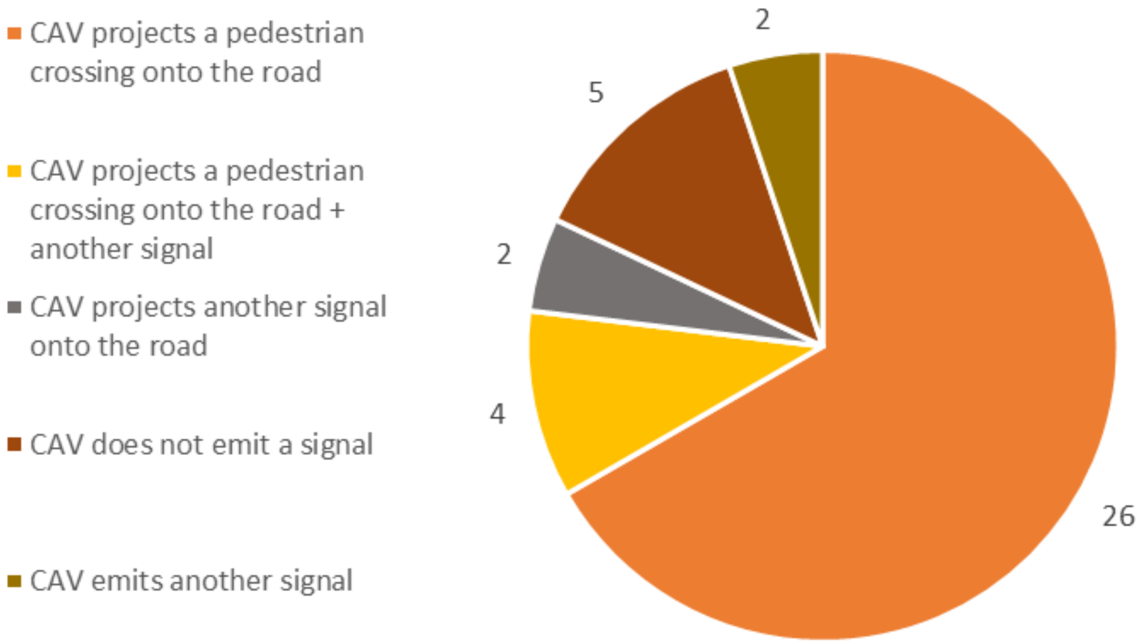


Figure 6.7: When the CAV stops and ONE pedestrian crossing painted on the road, pedestrians would prefer that a pedestrian crossing is projected.

If there is NO pedestrian crossing on the road, when the CAV doesn't stop (e.g., because the CAV detects the pedestrian too late to stop, or because a stop could generate more damages to more people), participants prefer that the CAV doesn't emit a signal (26, N=39, see Figure 6.8). This can be explained by the fact that, if no pedestrian crossing is painted on the road, pedestrians expect that the CAV doesn't stop in coherence with the road regulation. As it is coherent with the road regulation pedestrians don't expect any signal from the CAV. 14 respondents would prefer that the CAV flashes its headlights and honks it horn. But the horn is judged aggressive (1) and can be replaced by another sound (1) or just deleted (1). Some respondents insist on the fact that a signal like headlights and particularly horn should be launch only in real critical situations (2) like when the pedestrian doesn't see the CAV or is too close of the CAV to stop it. Some respondents would prefer projection on the road of a red pedestrian crossing (1). 3 other respondents would like to have a red LED signal on the CAV (3), eventually in addition of the headlights and horn (1).

Observation 6.4.b: when no pedestrian crossing is painted on the road, the pedestrians mostly expect that the CAV doesn't stop. Thus, no signal seems needed in this case. Or a discrete signal without honk can be used like a red light on the VAE or projected on the road. (to be confirmed)

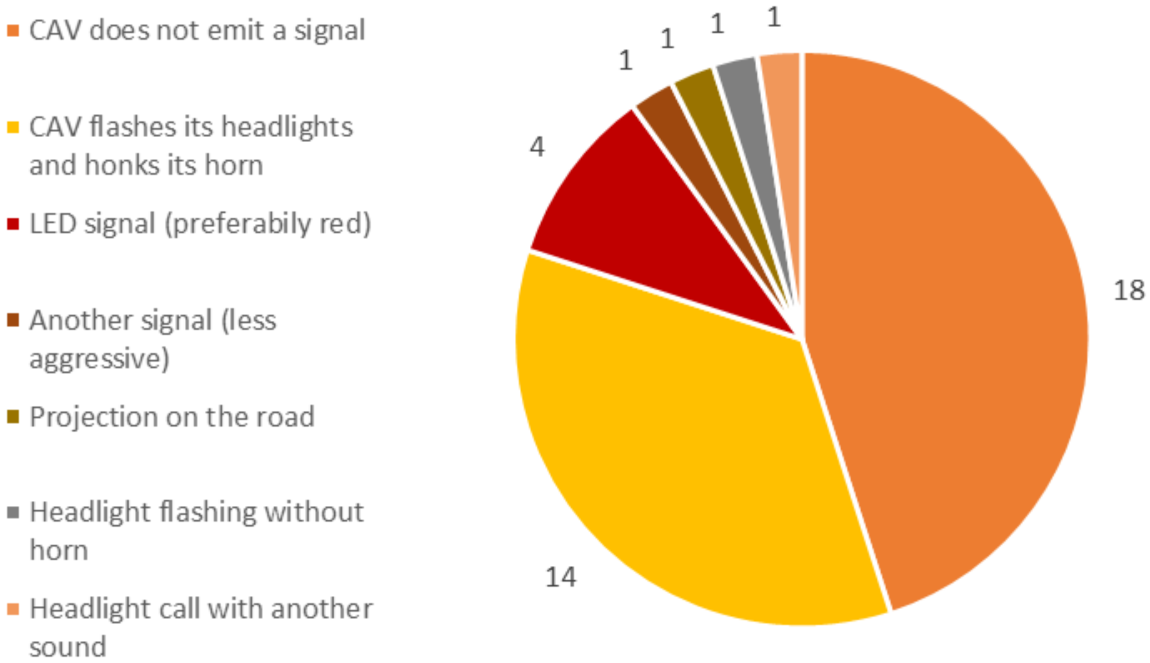


Figure 6.8: When the CAV doesn't stop and NO pedestrian crossing is painted on the road, pedestrians would prefer that CAV does not emit a signal.

If ONE pedestrian crossing is on the road, when the CAV doesn't stop, 23 (of 39) participants would prefer that the CAV flashes its headlights and honks its horn, see Figure 6.9Figure 6.9: . But 11 respondents prefer that the CAV doesn't emit a signal, this answer is often a reaction to the fact that the horn is judge too aggressive. 1 respondent indicates that only the horn should be used, as the headlights can be interpreted as the CAV gives way. For another participant, the decrease of the speed and the headlights should be used. 2 respondents mentioned that a red LED signal should be used in addition or not of the headlights and/or horn. Finally, one respondent would prefer the projection of a red pedestrian crossing on the road by the CAV.

Observation 6.4.c: If a pedestrian crossing is painted on the road, pedestrians expect that the CAV stops. So, if for a reason the CAV is not able to stop, a signal is expected. The horn is not well accepted because it is judge too aggressive and during discussions, majority of participants is against the use of a sound signal.

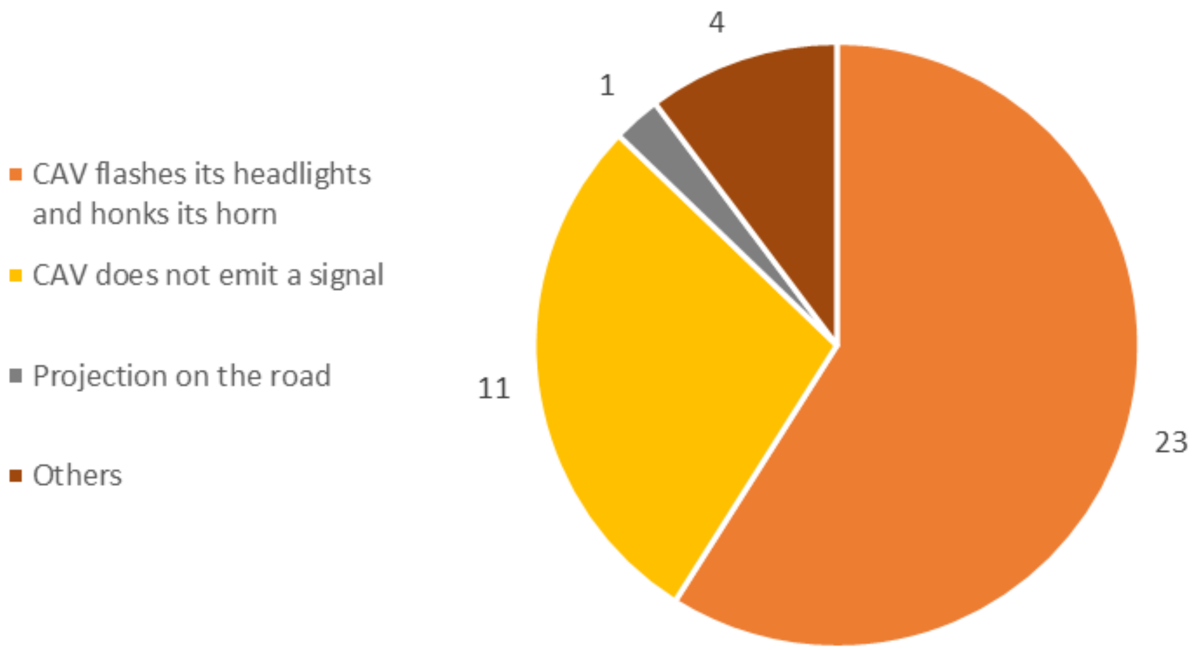


Figure 6.9: When the CAV doesn't stop and ONE pedestrian crossing is painted on the road, pedestrians would prefer that CAV flashes its headlights and honks its horn.

A blue LEDs stripe was displayed on the car to indicate that it's a CAV. We test dark blue and turquoise colours. But majority of participants didn't see it or associate it to the fact that the car is electric or to car tuning. But some participants asked to have an indication that a car is a CAV, moreover when normal cars and CAV will circulate at the same time on roads. Some participants ask that the CAV has a specific colour like for taxis, or a message written on indicating that's it a CAV or that the automatic mode is on, like on some taxi to indicate that it is free or note.

Observation 6.4.d: the CAV has to be easily identified in the traffic.

6.3.2 Participants' needs in information and trainings

After the experiments, several questions have been asked to participants concerning their need for information and training. Three questions have been asked: what kind of information they need, how they prefer to receive information, and who should communicate what kind of information. Results are discussed below.

6.3.2.1 Information needed by pedestrians in cohabitation with CAVs

The most important information for the 39 participants to this study are (see Figure 6.10): how the **infrastructure and equipment** have been adapted to match the introduction of the CAV (average 4.26 min=1, max=5), this concerns among others the pedestrian crossings; how the different **CAV interfaces** work (average 4.10); and **problems detected** in CAV release areas (average 3.92).

Other important information needed are: **CAV release schedule, areas and stages** (average 3.62); **Insurance conditions** communicated by the insurers (conditions and amount of coverage in case of accidents, amount of insurance premiums, etc.) (average 3.41); The **law** and how to interpret it regarding liability in the event of an accident with a CAV (average 3.36); Publication of the **algorithms** and programmed behaviours of the CAVs in different situations (what decisions the CAVs can make according to a situation) (average 3.18); and **statistics** on the risk of accidents with CAVs (average 3.10).

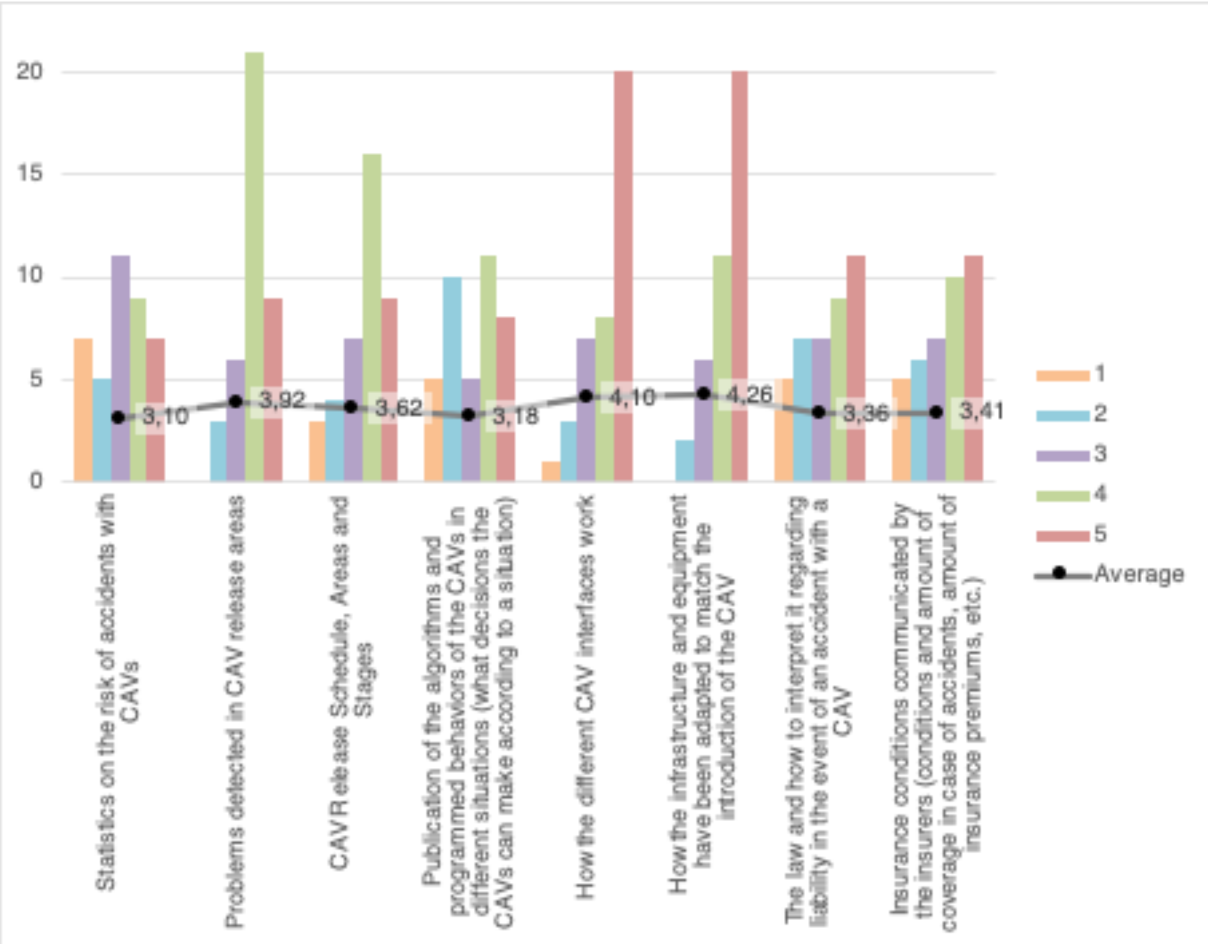


Figure 6.10: What information do you think you need to better understand the behaviour of CAVs? Bars show the number of answers for each agreement value (from 1 - Strongly disagree to 5 - Strongly agree) and the line shows the average.

Pedestrians seem to need full information about CAVs as no kind of information has an average less than 3. It is surprising that statistics, which are the most actually used to inform about CAV nowadays, is the last needed information.

Observation 6.5.a: Pedestrians seem to need full information about CAVs including: adaptation of infrastructure, eHMIs explanations, problems detected in areas where CAVs have been introduced, planning and areas of CAVs introduction, insurance conditions in case of accidents, law texts about liability in accidents and how to interpret them, algorithms and behaviours of CAV and statistics about accidents with CAVs.

Participants also asked for a **regulation or standardization of CAV interfaces** to ensure uniformity regardless of the manufacturer at least in Europe but better in the World and to be informed about it (4 respondents mentioned it as other information needed). One participant suggests that this standard should be available to everyone.

One participant also asks to have **tests results done by independent organizations** as EURO NCAP[7].

One participant insists on the need to know the **responsibility scheme in case of an accident** between a CAV and a pedestrian.

Observation 6.5.b: regulation and standardization of eHMIs are needed to ensure uniformity regardless of the manufacturer and improve predictivity, understanding and so acceptation of CAVs.

6.3.2.2 Ways to inform pedestrians

The 39 participants to the study prefer that information about CAV is distributed through: a **school-based training** for minors (18 years old in France) (average 4.62 min=1, max=5); an **official communication campaign** in the media (as for road safety, for example) (average 4.54); and a specific training during the **preparation of the driver's license** (Average 4.41), see *Figure 6.11*.

Other proposed ways to be informed have all average bigger than 3: the **possibility to test** the "driving" of a CAV (average 4.23); a communication campaign in the media (television, radio, newspapers) in general (average 4.18); **free training for adults** (e.g., training in a driving school) (average 3.08).

[7] https://www.euroncap.com/en

The mean of information “official communication campaign **by mail or e-mail**” received only an average of 2.95.

Observation 6.5.c: Participants would prefer that **young people are informed in priority**. Then, **for the adults** already holding a driving license, a **massive communication campaign**, preferably **official**, in all **standards media** is preferred. The **possibility to test CAV** and training sessions for this public is nice to have but do not have to be mandatory.

Regarding the other answers that participants have added, about training and informing minors, one participant suggests that the delivery of the training should be done **at/by school**. The **driving licence training is the second preferred moment** to be informed, trained and to test CAV as driver and pedestrian.

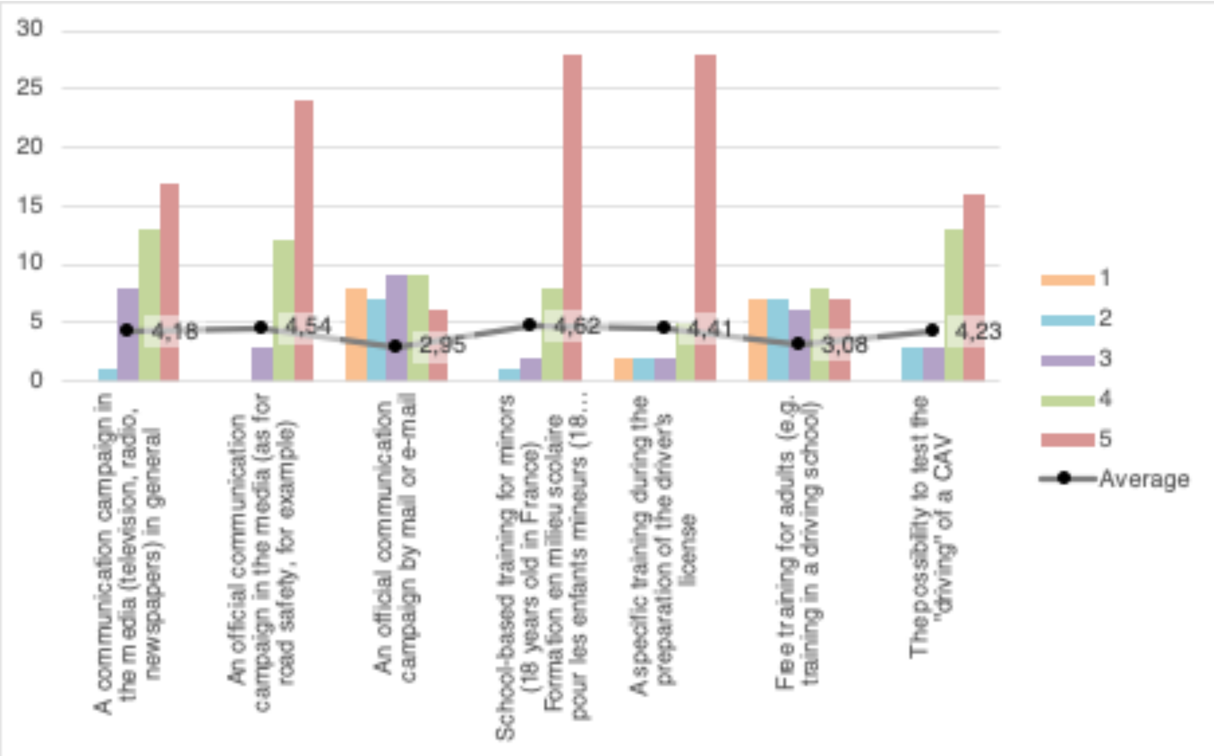


Figure 6.11: In your opinion, would each of the following proposals allow pedestrians (or any other road user) to better understand the behaviour of CAVs? Bars show the number of answers for each agreement value (from 1 - Strongly disagree to 5 - Strongly agree) and the line shows the average.

Two participants specify that CAVs should be **tested in real condition** on a circuit or in a dedicated zone in a city **as pedestrian**. The test of CAV can help to know the system, understand its limits and to combat preconceived ideas, as one participant pointed out.

Again, the need of **standardization** and update of the **traffic regulations** is underlined by two participants.

One participant would like that everybody should have to follow mandatory training, not only minors.

One participant insists on the fact that the **information campaign should be repeated**.

6.3.2.3 Authorities in charge of information for pedestrians

Participants to the study would prefer that insurances (37 out of 39) communicate on insurance conditions when CAVs will be involved in an accident.

Concerning the law, the main information provider should be the State (35 out of 39).

The adaptation of infrastructure should be explained by local authorities (30 out of 39) and State (24 out of 39).

CAVs' interfaces should be explained by manufacturers (24 out of 39), the media (19 out of 39) and the scientists and experts (18 out of 39).

The algorithms and programmed behaviours must be described by scientists and experts (30 out of 39) and manufacturers (20 out of 39).

CAV areas and stages of deployment as well as the release agenda should be given by the State (31 out of 39), the local authorities (30 out of 39) and the media (24). But emerging problems during the release phase should be communicated mainly by local authorities (29 out of 39), then by the State and media (20 out of 39) and by scientists and experts (18 out of 39).

Statistics about accident risks with CAVs should be spread by scientists and experts (30 out of 39), the State (22 out of 39), the media (20 out of 39) and local authorities (18 out of 39).

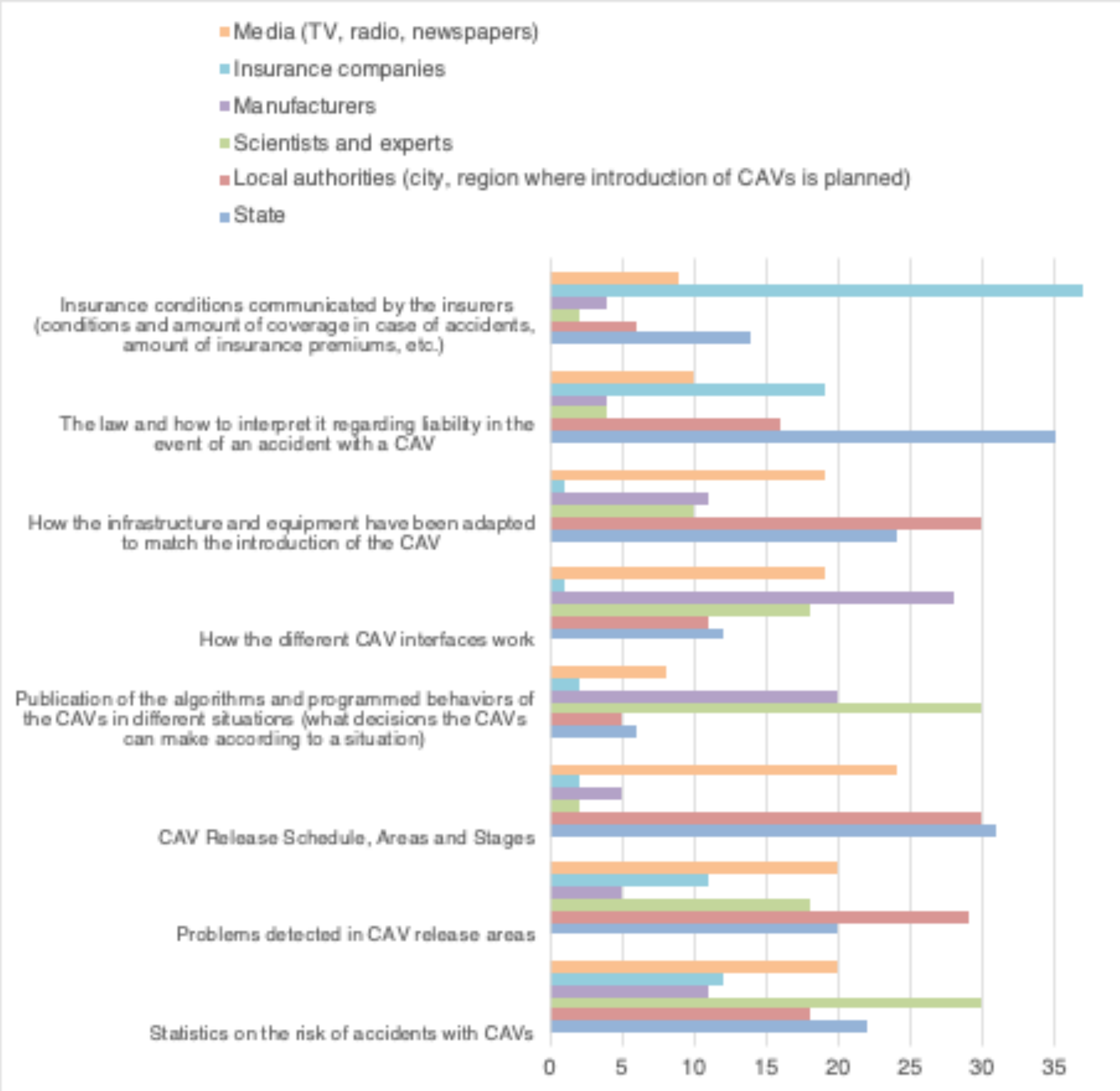


Figure 6.12: Who do you think should provide this information about the behaviour of CAVs? Bars show the number of answers for each authority (from 1 - Strongly disagree to 5 - Strongly agree).

6.3.3 Results from the online survey on eHMIs

6.3.3.1 Structure of the survey

In parallel to the experiments, we conducted a survey to gather the perceptions of pedestrians on a selection of 21 eHMIs (Figure 6.13 and Annex 1 for more details). The aim of this survey was to gather the opinions of pedestrians about the perceived understanding of the eHMIs, and their user experience.

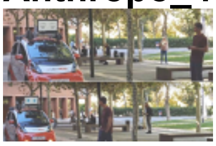
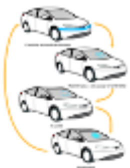
Anthropo-morphic	Iconic	Luminous	Textual	Others
Anthropo_1 	Iconic_1  	Luminous_1 	Textual_1 	Other_1 
Anthropo_2 	Iconic_2 	Luminous_2 	Textual_2 	Other_2 
Anthropo_3  	Iconic_3 	Luminous_3 	Textual_3  	Other_3 
Anthropo_4 	Iconic_4 	Luminous_4 	Textual_4 	
		Luminous_5 		
		Luminous_6 		

Figure 6.13: The eHMIs presented in the survey

The survey was structured in two parts. The first part was dedicated to collect the profile of the pedestrians, with the following questions:

- questions on pedestrians walking habits,
- pedestrians behaviour using the "Pedestrian Behaviour Questionnaire" (PBQ) questionnaire (Deb, Strawderman, DuBien, Smith, Carruth & Garrison, 2017),
- receptivity toward the autonomous vehicles as pedestrians, using the "Pedestrian Receptivity Questionnaire for FAVs" (PRQF) questionnaire (Deb, Strawderman, Carruth, DuBien, Smith & Garrison, 2017).

The second part presented the eHMIs extracted from one of the five following categories. Note that each respondent was randomly assigned to one of these categories:

- 4 anthropomorphic eHMIs,
- 4 iconic eHMIs,
- 6 luminous eHMIs,
- 4 textual eHMIs,
- 3 other eHMIs.

For each eHMI, the participants had to indicate their perception on it, using 8 items extracted from the meCUE questionnaire on User Experience (Minge, Thüring & Wagner, 2016). 4 items concerned the ease of use and ease of understanding of the interface, 4 other items concerned the emotional value and the confidence given to this interface. 8 other items, extracted from the PRQF questionnaire, collected the pedestrians' receptivity toward the FAV.

6.3.3.2 Respondents' profile

The survey received 224 answers. As respondents were randomly assigned to one of the 5 eHMI categories, the number of responses differed from one category to another:

- N = 42 for the anthropomorphic eHMIs,
- N = 38 for the iconic eHMIs,
- N = 51 for the luminous eHMIs,
- N = 47 for the textual eHMIs,
- N = 50 for the other eHMIs.

Table 6.5 describes the age of the participants for each of the eHMI categories.

Table 6.5: Age of participants for the 5 eHMIs categories

	Anthropomorphic	Iconic	Luminous	Textual	Other
N	42	38	51	47	50
Mean	33.64	33.10	30.92	32.93	35.26
SD	14.97	11.61	9.958	11.49	13.18
Min	18.00	19.00	18.00	18.00	16.00
Max	69.00	65.00	56.00	65.00	70.00

Table 6.6 describes the gender of the participants for each of the eHMI categories.

Table 6.6: Gender of participants for the 5 eHMIs categories

eHMI	Gender	N	Percent
Anthropomorphic	Female	34	80.95%
	Male	8	19.04%
	Total	**42**	**100%**
Iconic	Female	24	63.15%
	Male	14	36.84%
	Total	**38**	**100%**
Luminous	Female	30	58.82%
	Male	21	41.17%
	Total	**51**	**100%**
Textual	Female	36	76.59%
	Male	11	23.40%
	Total	**47**	**100%**

Others	Female	33	66.00%
	Male	17	34.00%
	Total	**50**	**100%**

The "Pedestrian Behaviour Questionnaire" (PBQ) aims at measuring frequency of risky behaviours among pedestrians. The higher is the score of PBQ, the riskier is the pedestrian's behaviour. The PBQ has 20 items in its short version, measuring five dimensions:

- Violations,
- Errors,
- Lapses,
- Aggressive Behaviours,
- Positive Behaviours.

The results are presented in Table 6.7, by differentiating scores according to gender.

Table 6.7: Scores of the PBQ by gender

	PBQ_Violation		PBQ_Errors		PBQ_Lapses		PBQ_Aggressive		PBQ_Positive		PBQ_Global	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Valid	157	71	157	71	157	71	157	71	157	71	157	71
Missing	0	0	0	0	0	0	0	0	0	0	0	0
Mean	2.831	3.366	1.755	2.310	1.414	1.345	1.462	1.606	2.589	2.732	2.010	2.272
Std. Deviation	1.269	1.159	0.824	0.954	0.673	0.497	0.629	0.712	1.057	1.098	0.503	0.522

Note. Excluded 1 rows from the analysis that correspond to the missing values of the split-by variable Genre

There are gender differences in scores. However, none of these differences are statistically significant.

Regarding the pedestrian receptivity questionnaire for FAVs (PRQF), Table 6.8 presents the results according to the gender. The PRQF aims at measuring the acceptance of FAV by the pedestrians, considering five dimensions:

- ➢ Compatibility,
- ➢ System Effectiveness,
- ➢ Social Norm,
- ➢ Attitude,
- ➢ Trust.

Table 6.8: Scores of the PRQF by gender

	PRQF_Attitude		PRQF_SocialNorms		PRQF_Effectiveness		PRQF_Trust		PRQF_Compatibility		PRQF_Global	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Valid	157	71	157	71	157	71	157	71	157	71	157	71
Missing	0	0	0	0	0	0	0	0	0	0	0	0
Mean	3.124	3.766	3.637	3.704	3.573	3.979	2.698	3.432	2.688	2.953	3.116	3.568
Std. Deviation	1.393	1.386	1.051	0.928	1.413	1.369	1.346	1.484	1.252	1.423	1.104	1.064
Minimum	1.000	1.200	1.000	1.330	1.000	1.500	1.000	1.000	1.000	1.000	1.000	1.690
Maximum	7.000	7.000	6.000	6.330	7.000	7.000	7.000	7.000	7.000	6.000	6.190	5.690

Note. Excluded 1 rows from the analysis that correspond to the missing values of the split-by variable Genre

As the PBQ, we observe gender differences in scores. However, none of these differences are statistically significant.

6.3.3.3 Survey analysis

The analysis of the eHMIs was done by distinguishing between the scores obtained for the user experience measure, and the scores obtained for the items on the receptivity of pedestrians to FAVs. The results show (*Figure 6.14*) that the eHMI with the highest UX score is Anthropomorphic_3 (4.7/7). This interface offers an accompaniment of the pedestrian crossing by two eyes integrated into the vehicle's headlights. This interface is considered the most pleasant and useful. Conversely, the eHMI Other_3, which describes an interface on the pedestrian's mobile device capable of informing him or her about the behaviour of an autonomous vehicle, is judged the least experiential. (2,3/7). Indeed, this interface raises many questions for the participants, particularly in terms of reliability in the absence of a network.

The eHMI with the highest receptivity score is Other_1, with a score of 4.1/7. This original interface proposes to modify the infrastructure to enhance pedestrian safety. The zebra is then a plate that lifts up when the pedestrian wants to cross the street, thus preventing any vehicle from passing at that moment. This interface, which can be considered as a deported interface, offers a better acceptance of FAVs. In contrast, the interface with the lowest receptivity score is Anthropo_2 (2.4/7), which displays two eyes instead of headlights. Users consider this interface to be unattractive, which does not favour its acceptance.

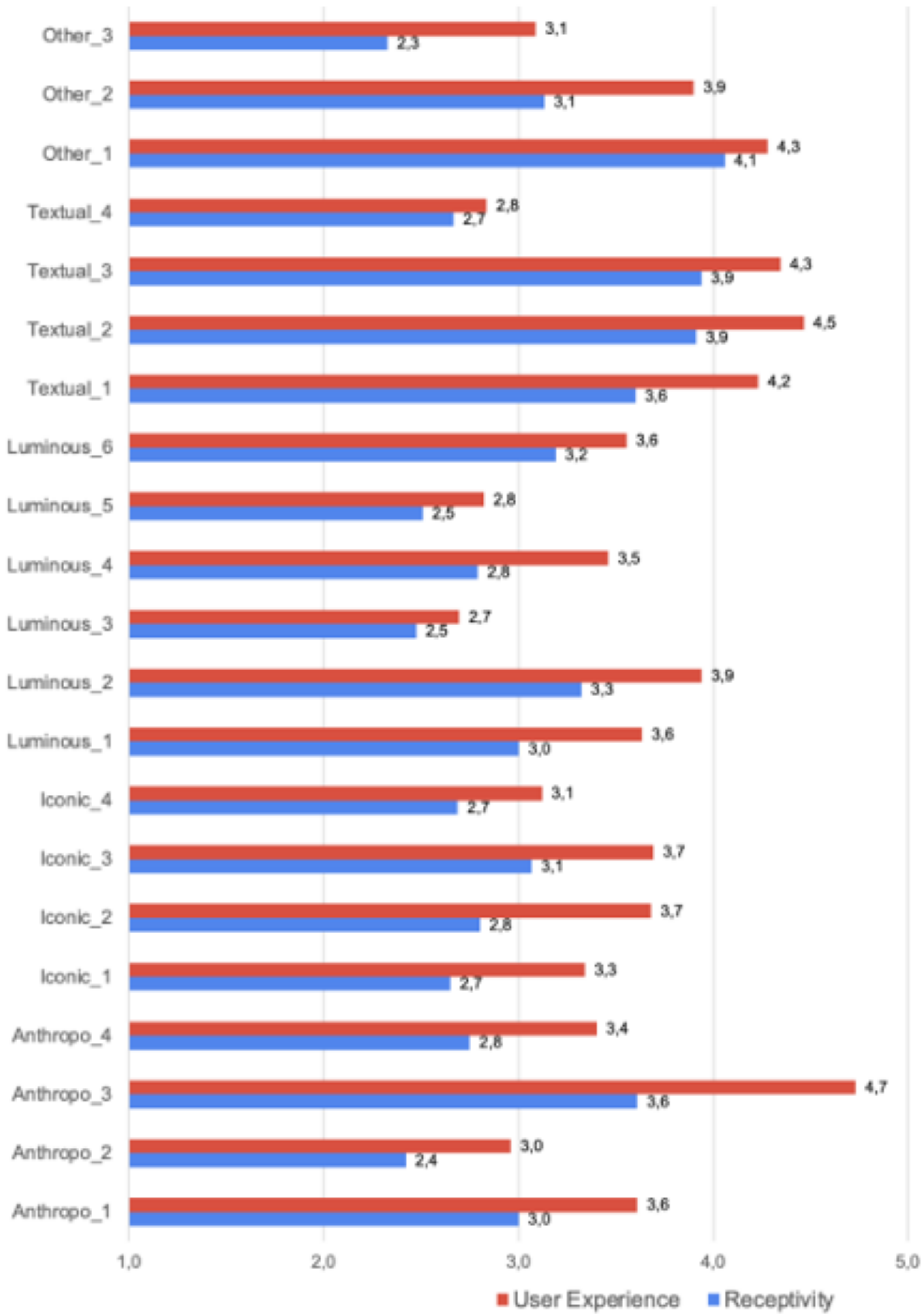


Figure 6.14: Receptivity toward FAV and perceived user experience of each eHMI

The results for responsiveness and user experience scores by interface category (*Figure* 6.15) show that text-based interfaces score best on average for both criteria. In general, users report that text-based information is clear and unambiguous. Some issues were raised, such as the language used for the text, or the legibility of the information depending on its location on the vehicle. A display on the windscreen is considered difficult to see for wheelchair users, children, or people of small stature. Conversely, information on the vehicle's grille is considered too low for most pedestrians. Despite these counter-indications, text remains the safest way to inform pedestrians of the autonomous vehicle's attention and behaviour.

The category with the lowest receptivity score is the iconic interfaces. CAVs equipped with these eHMIs are indeed considered less reliable than CAVs equipped with other eHMIs, as the icons are perceived as too open to interpretation. The lighted interfaces have the lowest scores in terms of user experience. Most of the interfaces in this category are considered unintuitive, and need to be explained according to the users. This then goes against what is expected from eHMIs, which should be immediately understood and unambiguous for pedestrians.

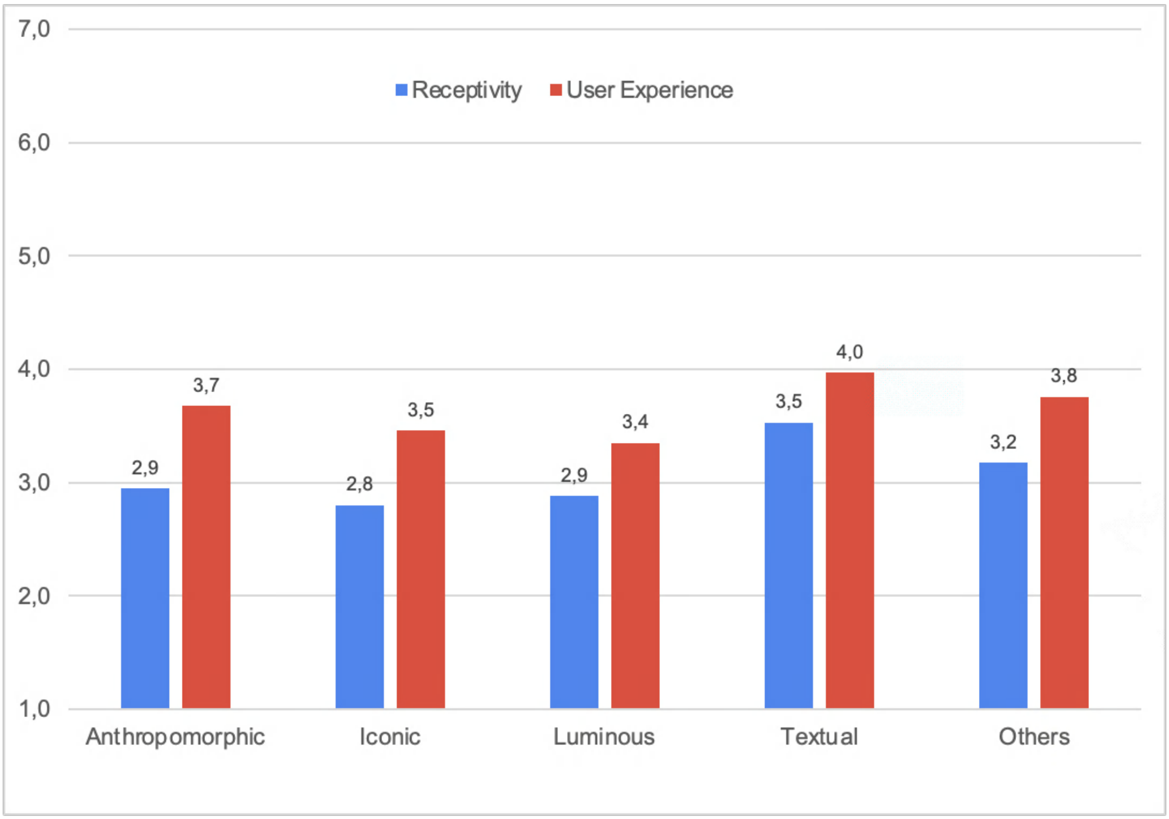


Figure 6.15: Receptivity toward FAV and perceived user experience of each category of eHMIs.

It is interesting to note that there is a very strong correlation between the responsiveness and user experience scores. The calculated correlation coefficient is equal to 0.84. This shows that the easier to understand and more elegant the eHMI is perceived to be (our user experience criteria), the better the acceptance of the CAVs equipped with it. These results are in line with the main models of technology acceptance, such as the TAM (Davis, 1989), which include ease of use and usefulness of the technology in their explanatory factors.

Observation 6.6.a: the more promising eHMI in terms of UX and receptivity are text-based interfaces, but it raised some issues to be understood by everybody including visually impaired, illiterate, kids, persons not able to read the language used.

Observation 6.6.b: the easier to understand and more elegant the eHMI is perceived to be, the better the acceptance of the CAVs equipped with it.

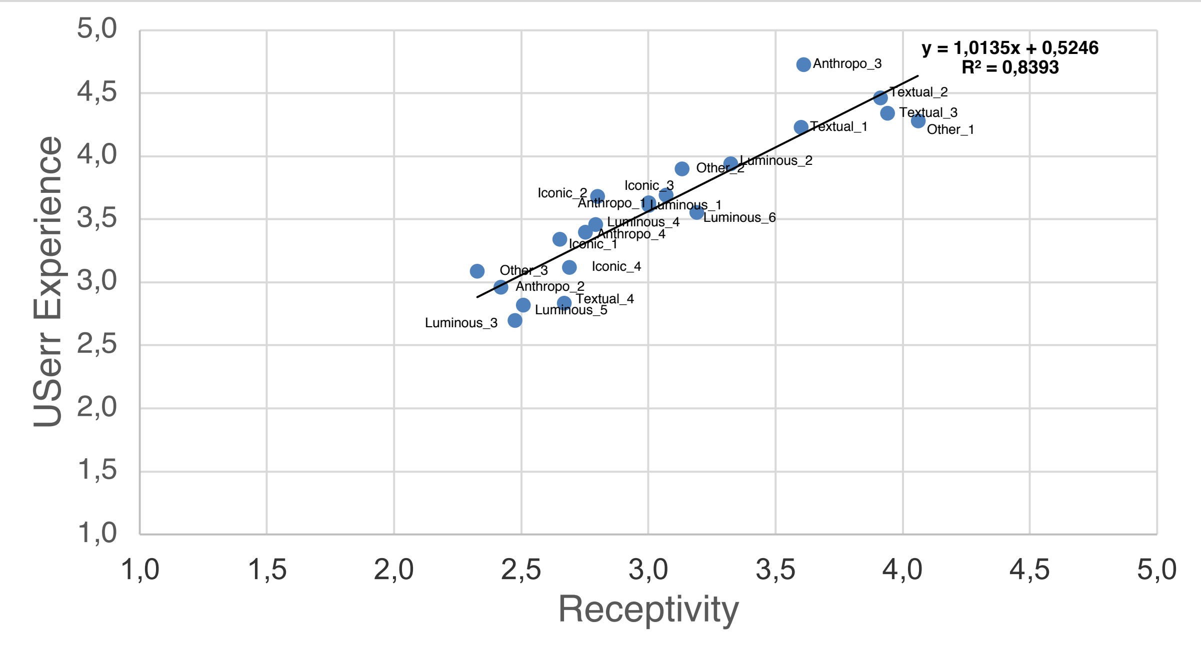


Figure 6.16: Link between the receptivity toward FAV and the perceived user experience

6.4 Conclusions

The main observations done during this study are:

Observation 6.1: Luxembourgish citizen having no kid under 10 seems have a better receptivity about CAV than the others.

Observation 6.2.a: When a CAV stops to let crossing a pedestrian, it is better to send a signal that the CAV will wait the pedestrian's crossing.

Observation 6.2.b: it seems better to indicate that the CAV will not stop if it detects to let a pedestrian willing cross.

Observation 6.3.c: a feedback is waited by pedestrians in all situations and particularly in dangerous ones. That confirms observation 6.2.b.

Observation 6.3.d: presence or absence of a crosswalk already on the road does not play a significant role.

Observation 6.4.a: When the CAV stops the use of a signal to show that the CAV is waiting that the pedestrian cross is needed. The projection on the road is well accepted in the cases of a pedestrian crossing is painted or not on the road. That confirms observations 6.2.a and 6.3.b.

Observation 6.4.b: when no pedestrian crossing is painted on the road, the pedestrians mostly expect that the CAV doesn't stop. Thus, no signal seems needed in this case. Or a discrete signal without honk can be used like a red light on the VAE or projected on the road. (to be confirmed)

Observation 6.4.c: If a pedestrian crossing is painted on the road, pedestrians expect that the CAV stops. So, if for a reason the VAE is not able to stop, a signal is expected. The horn is not well accepted because it is judge too aggressive and during discussions, majority of participants is against the use of a sound signal.

Observation 6.4.d: the CAV has to be easily identified in the traffic.

Observation 6.5.a: Pedestrians seem to need full information about CAVs including: adaptation of infrastructure, eHMIs explanations, problems detected in areas where CAVs have been introduced, planning and areas of CAVs introduction, insurance conditions in case of accidents, law texts about liability in accidents and how to interpret them, algorithms and behaviours of CAV and statistics about accidents with CAVs.

Observation 6.5.b: regulation and standardization of eHMIs are needed to to ensure uniformity regardless of the manufacturer and improve predictivity, understanding and so acceptation of CAVs.

Observation 6.5.c: Participants would prefer that **young people are informed in priority**. Then, **for the adults** already holding a driving license, a **massive communication campaign**, preferably **official**, in all **standards media** is preferred. The **possibility to test CAV** and training sessions for this public is nice to have but do not have to be mandatory.

Observation 6.6.a: the more promising eHMI in terms of UX and receptivity are text-based interfaces, but it raised some issues to be understood by everybody including visually impaired, illiterate, kids, persons not able to read the language used.

Observation 6.6.b: the easier to understand and more elegant the eHMI is perceived to be, the better the acceptance of the CAVs equipped with it

Furthermore, the authorities who should communicate information about CAVs are, concerning:

- The law, the main information provider should be the State.
- The adaptation of infrastructure should be explained by local authorities.
- The eHMIs should be explained by manufacturers, media and scientists and experts (18 out of 39).
- The algorithms and programmed behaviours must be described by scientists and experts and manufacturers.
- The areas and stages of deployment of CAVs, as well as the release agenda should be given by the State, the local authorities, and the media.
- The emerging problems during the release phase should be communicated mainly by local authorities, then by the State and media and by scientists and experts.
- Statistics about accident risks with CAVs should be spread by scientists and experts, the State, media, and local authorities.

To be able to trust a machine such as a CAV, it is necessary to be able to anticipate its reactions, to understand how it works and to be able to check that it has detected a pedestrian about to cross the road. For this, the simplest way is to share a common frame of reference. In the case of pedestrian/CAV interaction, the common frame of reference is the traffic regulations. Therefore, pedestrians expect the traffic rules to be applied. Thus, the presence of a crosswalk gives priority to the pedestrian and the absence of a crosswalk gives priority to the vehicle as it is specified in traffic regulations of Luxembourg and France. Pedestrians seem to value predictive behaviour over the practicality of being able to cross at any time.

Finally, what seems to worry the participants of this study is the fact that the design of eHMI and the choice of the behaviour of AEVs is left to the manufacturers without any regulation and homogenization, with the risk that aesthetics will prevail over utility and comprehensibility.

6.5 Ways to improve CAV design

To summarise, there are several steps that break down the interaction between the CAV and the pedestrian. These steps can be summarised as follows:

- The car is driving in its lane. It has not detected any obstacles.

- The car detects a pedestrian who intends to cross the road.

- The car decelerates to the height of the pedestrian.

- The car stops.

- The car lets the pedestrian pass.

- The car starts again after the pedestrian has passed (Figure 6.17).

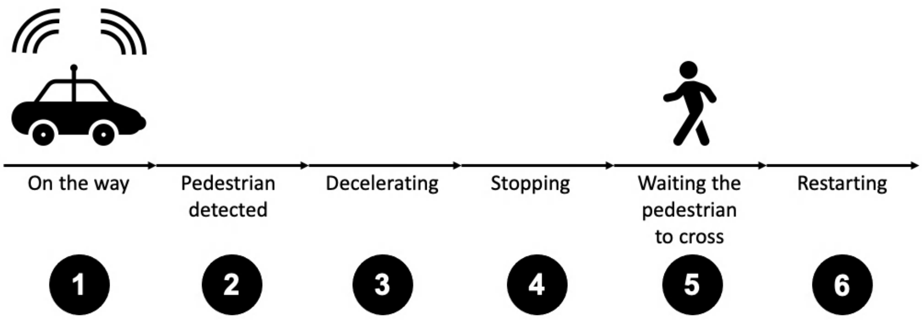


Figure 6.17: Main steps that break down the interaction between the CAV and the pedestrian

Some of these 6 main steps require the CAV to inform the pedestrian of the behaviour:

- The CAV should give a signal when a pedestrian is detected (step 2).

- The CAV should make it clear to the pedestrian that it is letting them pass, and that it will not restart until they have crossed (step 5).

Furthermore, the CAV should inform the pedestrian if it is not able to stop, either because it detected the pedestrian too late or for technical reasons.

Finally, the pedestrian should be able to clearly identify that he/she is interacting with a level 5 (driverless) CAV, in order to enable him/her to adopt his/her behaviour.

6.6 Guidelines and recommendations for pilot specifications

6.6.1 Use cases

Several use cases could eventually be explored:

	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5
• The CAV stops or does not stop (for different reasons).	x	x	x	x	
• The CAV respects the traffic rules and does not let pedestrians pass outside the crossings.	x	x	x	x	
• Textual eHMIs could be tested, with the will to homogenise the interfaces to define a standard eHMI.	x		x	x	x

6.6.2 Test variables

Different dependent and independent variables could eventually be explored:

	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5
• Nationality.	x	x	x	x	x
• Country of residence.	x	x	x	x	x
• Dependence on one or more children under 10 years of age.	x			x	x

7 Findings from Heliflight-R

7.1 Simulation Trials Setup

7.1.1 Simulation system

The trials to be run at the University of Liverpool will make use of the HELIFLIGHT-R full motion flight simulator (White et al., 2013) (Figure 7.1).

Figure 7.1: Heliflight-R Full Motion Simulator

The outside world scene is rendered using one of several image generators, projected onto a 12 ft diameter dome by three High Definition projectors. The commercially available simulation software package X-Plane (https://www.x-plane.com/) has been selected to generate an outside world image for this project due to the availability of models of the

City of Liverpool and Liverpool John Lennon Airport, purchased from OrbX. The output from each display channel is warped and blended to create a seamless image on the surface of the dome covering a field of view (FoV) of approximately 210° (horizontal) by 70° (vertical, Figure 7.2). This FoV is extended in the region ahead of the pilot. The instrument panel uses LCD displays, featuring two touch screens, which are also user-programmable. The motion platform features 6 Degrees of Freedom, using 24-inch electric actuators. The motion range is documented in Table 7.1. The base accommodates an 1800 Kg payload.

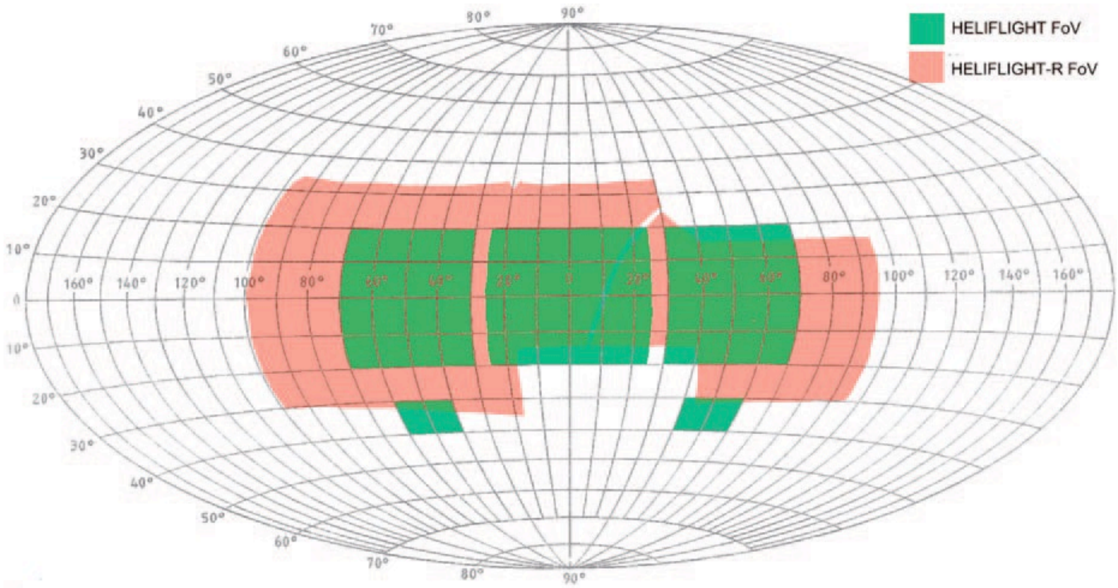


Figure 7.2: *Heliflight-R field of View*

Table 7.1: Heliflight-R Performance Envelope

	Displacement	Velocity	Acceleration
Pitch	-23.3°/25.6°	±34 °/s	300 °/sec^2
Roll	-23.2°	±35 °/s	300 °/sec^2
Yaw	±24.3°	±36 °/s	500 °/sec^2
Heave	±0.39 m	±0. 7 m/s	+/- 1.02 g
Surge	-0.46 /+0.57 m	±0. 7 m/s	+/- 0.71 g
Sway	±0.47 m	±0. 5 m/s	+/- 0.71 g

HeliFlight-R is reconfigurable for flight dynamics engineering and training applications. Aircraft specific cabs can be implemented to allow for authentic fixed wing or rotorcraft simulation environment. This included the use of two pilots, with extra capacity for in-flight analysis. Authentic controls allow for re-configurable force-feedback and re-configurable instruments. All physical switches and levers are user-programmable as required.

7.1.2 Research Questions (RQ)

The simulation experiments will aim to evaluate attitudes of passengers in a Level 3 or Level 4 autonomous urban air vehicle, their willingness to make control decisions, should this be required or convenient and whether the level of information about the system provided to the passenger affects their attitudes and behaviour towards this.

- RQ 1: How much the level of comfort (subjective and objective), ease-of-use, and perception of risk varies between Levels 3 and 4?
- RQ 2: What behaviour and level of acceptance will passengers have in a Level 3 Personal Aerial Vehicle (PAV) flying in an urban environment?
- RQ 3: What behaviour and level of acceptance will passengers have in a Level 4 PAV flying in an urban environment?
- RQ 4: Does the level of information about how to operate the system received by the passengers make a difference to their behaviour and level of acceptance in a PAV flying in an urban environment?

7.1.3 Scenario Development

In order to be able to answer the research questions posed, a number of scenarios have been developed in order to expose autonomous aerial vehicle participants to a variety of situations.

For each of the scenarios, it is possible that participants will need to be able to control the vehicle in the event of the system relinquishing control. For road-based autonomous systems, passengers are much more likely to be experienced in driving and can intervene by taking control from the autonomous system when required or desired. For the airborne system equivalent to a road vehicle, passengers would require an equivalent of a pilot license, or have experience of flight. For this work, this would have meant training the occupants to fly the vehicle should they have to intervene. Even experienced pilots would need to be trained to a level of competence with the myCopter configuration. All of this would be time-consuming.

The solution to this was to provide the air system with a high level of automatic functionality i.e. the user need only select an option such as (change of) destination airfield if prompted or desired. This level of functionality exists on current manned and unmanned aircraft with automatic waypoint navigation, flightpath, landing systems etc. Therefore, the user need only make high level decisions about the journey e.g. select the destination rather than intervene in the flying task. Passengers will have to receive a short training brief or leaflet detailing how to use the display in order to reach their destinations. They can be provided with varying levels of detail regarding how the system works and its limitations, how they can take control and when can they do it in order to test if this impacts their level of acceptance of PAV and their behaviour during the flight.

The scenario is that on arrival at Liverpool John Lennon airport, a passenger (subject or participant) will take an Unmanned Air Taxi from the airport to the Pier Head on the edge of Liverpool City Centre. This route would take approximately 40 minutes by bus or 30 minutes using a taxi or private car. Performing the same trip on the simulated flying taxi will only take around 5 minutes.

The route from the airport towards the City Centre is depicted in Figure 7.3. Departure from the airport is illustrated in Figure 7.4. Three fictitious vertiports serve the City Centre region (Figure 7.5). The first is the Pier Head, close to the waterfront and central tourist attractions and further transport links. The second is to the east of the City Centre, in the knowledge quarter close to the cathedrals and University and easy to reach by foot or bus from Pier Head. The third, Seacombe Terminal, is on The Wirral across the river Mersey from Liverpool City Centre. This is more difficult to reach from Pier Head, requiring the use of public transport, either bus or ferry.

The passenger selects the Pier Head vertiport as the destination. Four possible scenarios might then unfold.

7.1.3.1 Scenario 0: Flight to destination without any adverse events.

The first (baseline) run of the simulation will utilize either CAV Level 3 or Level 4 in a scenario where the flight takes place as planned.

Figure 7.3: Unmanned air Taxi route from Liverpool airport to the City Centre

Figure 7.4: Liverpool John Lennon airport

Figure 7.5: Liverpool City Region vertiports

7.1.3.2 Scenario 1 – CAV Level 4, "correct" decision

For this scenario, City Centre airspace, which includes the Pier Head Vertiport, will be closed, as illustrated in Figure 7.6, as the vehicle approaches the City Centre. The autonomous system onboard reroutes to the Knowledge Quarter and continues on the journey with no need for the subject to intervene as s/he can easily walk or transit towards his final destination.

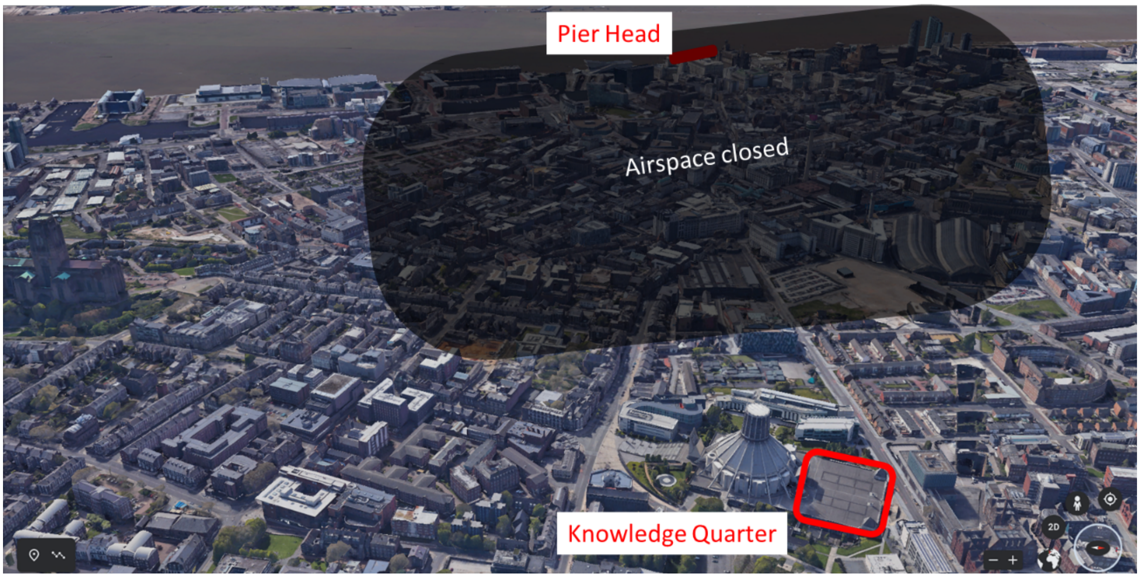


Figure 7.6: City Centre airspace which includes the Pier Head Vertiport closed – Divert to the Knowledge quarter

7.1.3.3 Scenario 2 – CAV Level 4, "inconvenient" decision

In this scenario, the airspace is closed as per Scenario 1. However, in this case, the system chooses to divert to the Seacombe Terminal (Figure 7.7). This choice is less practical for the subject as an additional journey is required to cross the river to reach the City Centre. Although the CAV Level 4 system can continue on the journey, the occupant can intervene to change the destination to the Knowledge Quarter vertiport, as long as this action is performed before reaching the point-of-no-return waypoint.

7.1.3.4 Scenario 3 – CAV Level 3, user needs to make a decision

In this final scenario, for a Level 3 CAV, the airspace is once again closed as the vehicle approaches the City Centre. This time, the autonomous system does not have the capability to select a new route. Therefore, the default option for the vehicle is to continue to the initial destination unless the subject selects a new destination prior to any of the point-of-no-return decision waypoints. Voice notifications will prompt the user to do so.

Figure 7.7: City Centre airspace which includes the Pier Head Vertiport closed – Divert to Seacombe Terminal

7.1.4 Progress in Experimental Set-up

For this study the Matlab-based flight dynamics model from the EU project myCopter is to be used (Perfect et al., 2015). This has been upgraded with autopilot functions as well as pre-planned routes from Liverpool John Lennon Airport to three fictitious vertiports around Liverpool City Centre as part of the PAsCAL project.

The routes are defined as a series of waypoints leading to the destination vertiport at either Pier Head, Knowledge Quarter or Seacombe Terminal (see Figure 7.8). Each waypoint is defined by five coordinates indicating target position in three-dimensional space, target speed and heading, which the vehicle's autopilot will follow as closely as possible (Figure 7.9).

Depending on the destination, there are two decision waypoints, after which the vehicle is committed to its final destination and the user will not be able to change the destination. These are waypoint number 4 if the new destination is Knowledge Quarter and waypoint 6 if the destination is either Seacombe Terminal or Pier Head.

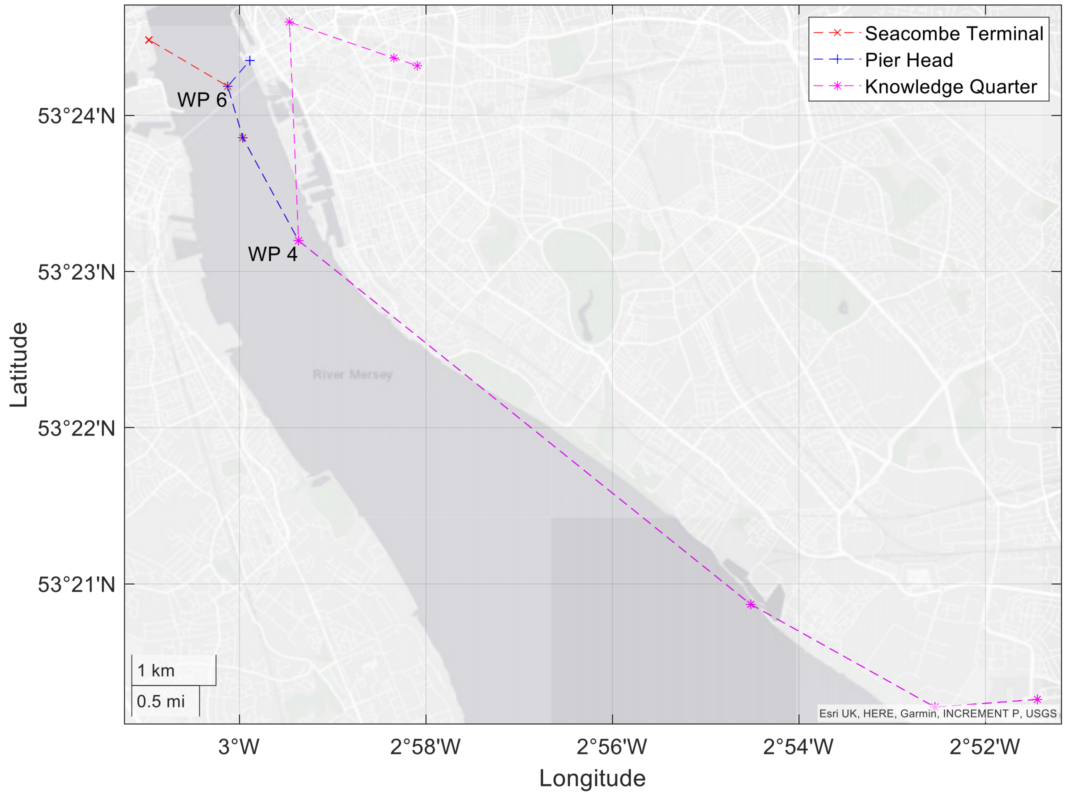


Figure 7.8: Routes and waypoints from Liverpool Airport (bottom right) to the three possible heliport destinations. Decision waypoints 4 and 6 are labelled.

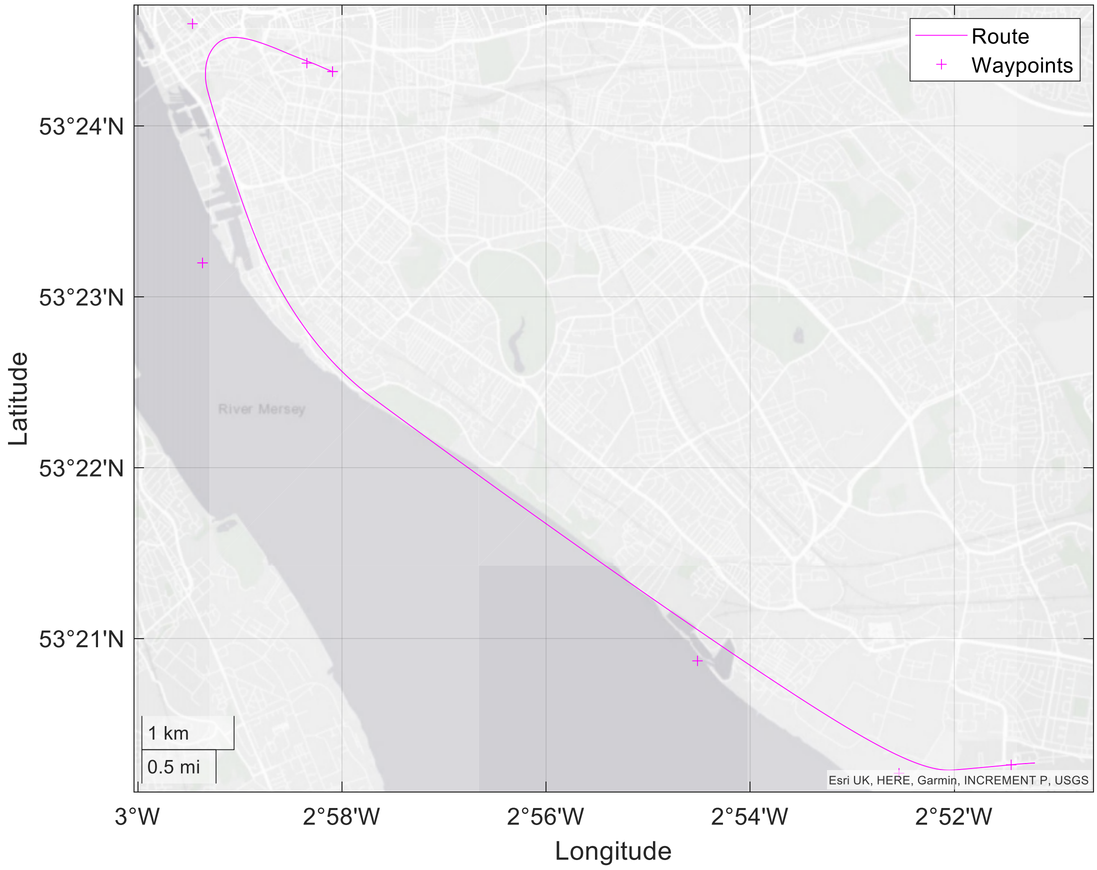


Figure 7.9: Actual route to Knowledge Quarter flown by vehicle.

The passenger is presented with a tactile display from which the preferred destination vertiport can be selected. By clicking on each option, the passenger interface will present a brief description of the destination. Once the passenger has chosen the destination, the journey can be started by pushing the 'Depart' button (see Figure 7.10). A recorded voice notification will inform the passenger that the vehicle is departing and confirm its destination.

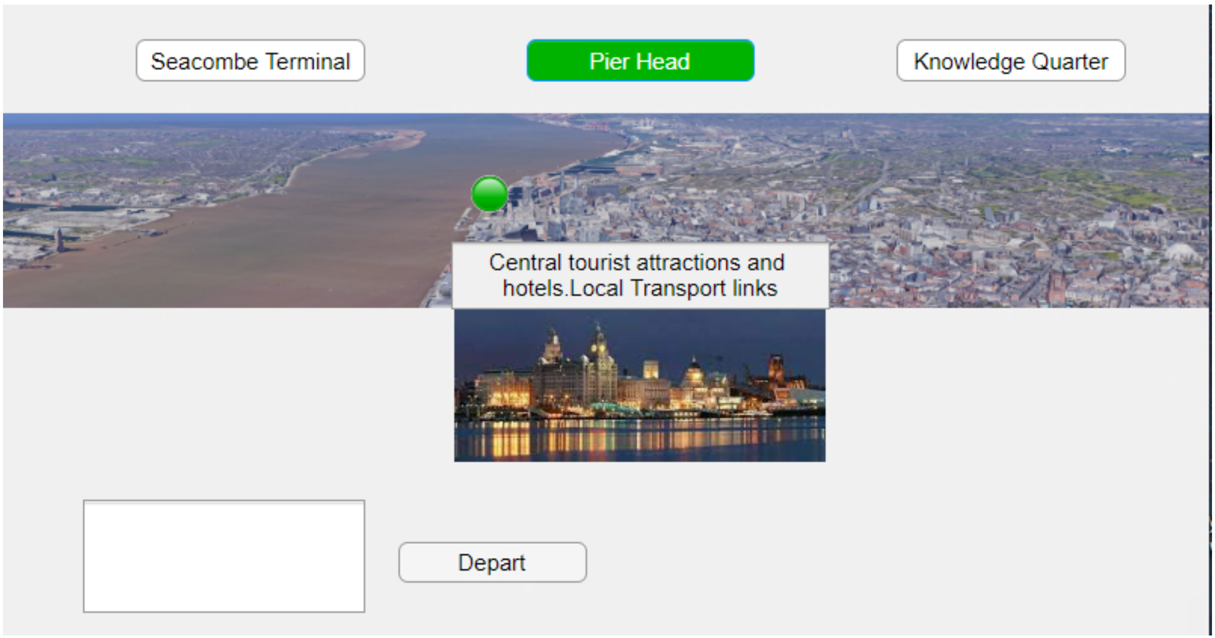


Figure 7.10: Passenger terminal

Once the flight has started, the passenger has the option to select an alternative destination and update the flight plan as long as the vehicle has not yet reached the point-of-no-return decision waypoint.

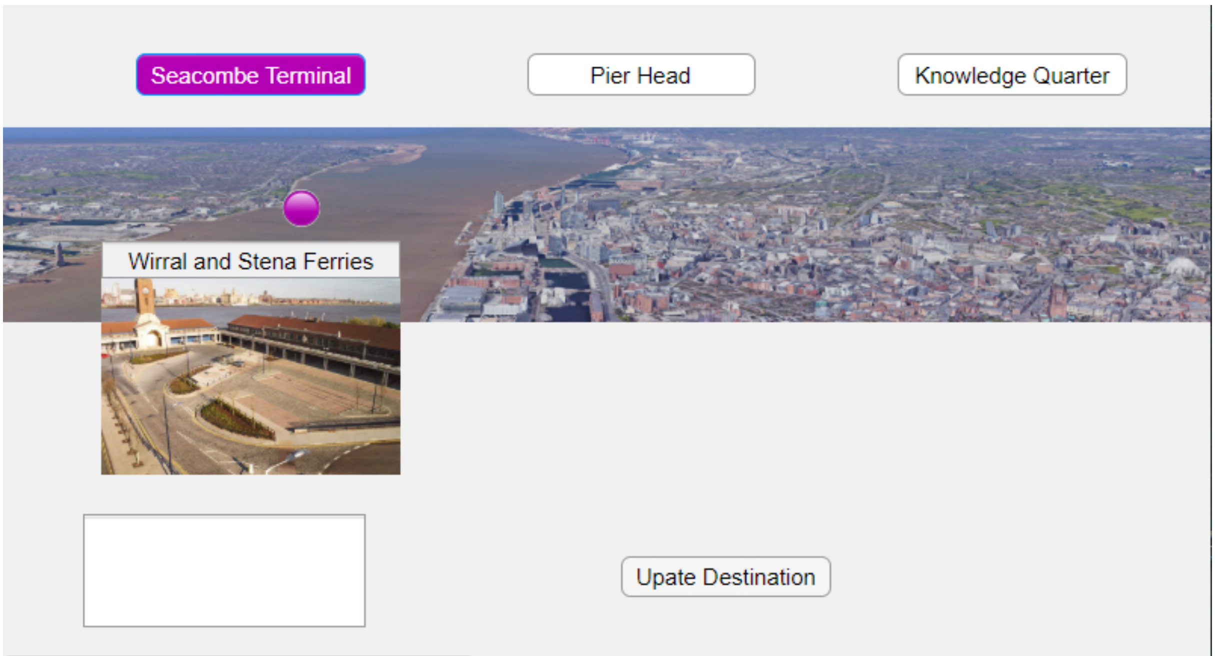


Figure 7.11: Option to update destination

The display will also provide information about any incident that will impede the aircraft from reaching its destination and, in the case of a simulated Level 4 autonomy vehicle of the alternative destination selected. The passenger will also be informed by voice, and in the case of Level 3 autonomy, the user will be prompted to select an alternative destination.

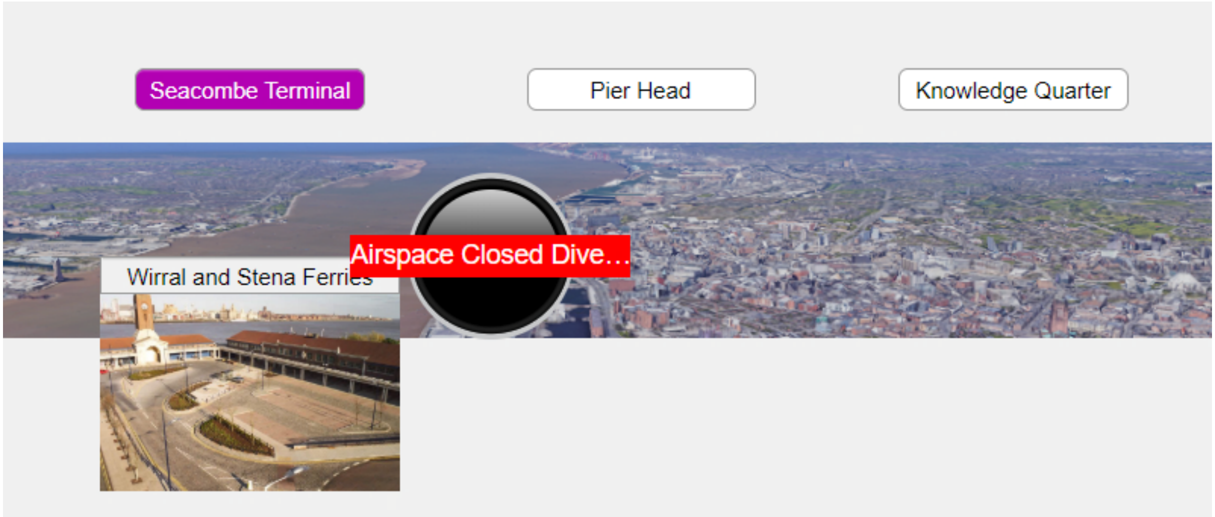


Figure 7.12: Display informing the user of airspace closure at the destination and of the alternative destination chosen by the system (CAV level 4)

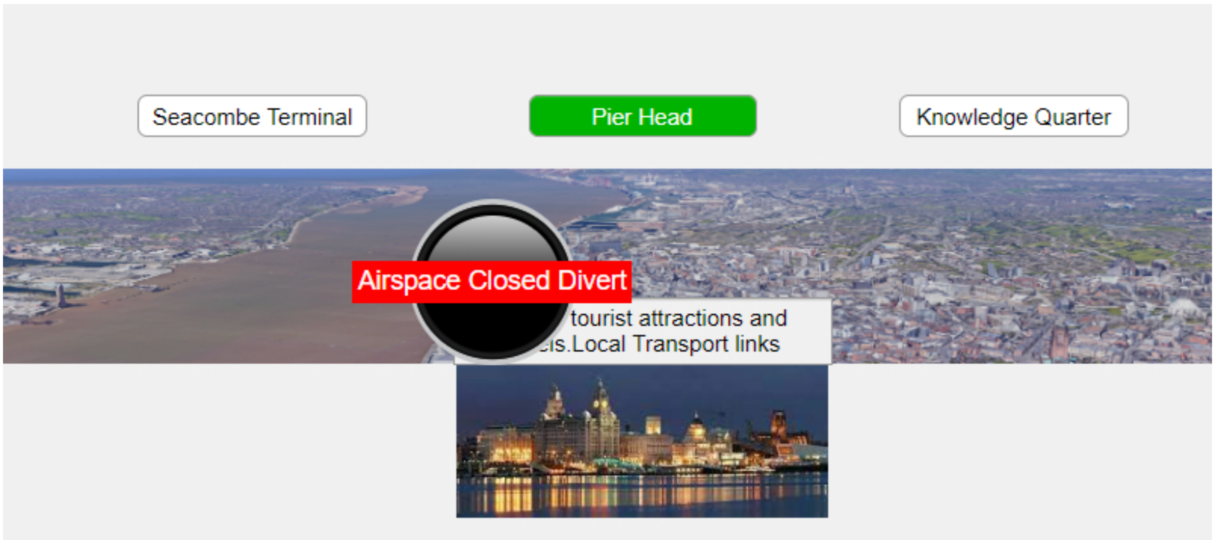


Figure 7.13: Display informing the user of airspace closure at the destination and prompting the user to divert (CAV Level 3)

Unexpected events (unexpected to the user) and vehicle decisions are controlled by the test operator from the simulator control room. Who will close the airspace over the Pier Head and select the alternative destination or prompt the user to do so. The test operator has a specific interface with three options, to select one of the three possible scenarios for this experiment.

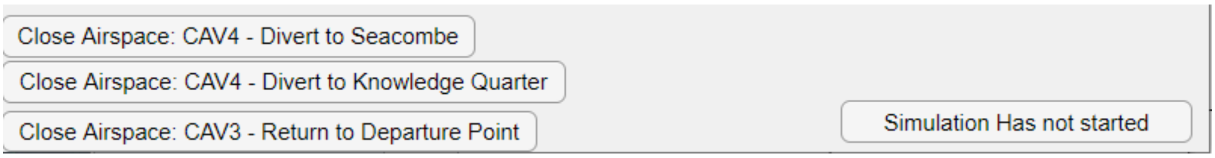


Figure 7.14: Control interface for the test operator.

The control interface also allows the test operator to start or interrupt the simulation before its conclusion if this is required for safety or any other reasons.

The system architecture is illustrated in Figure 7.15.

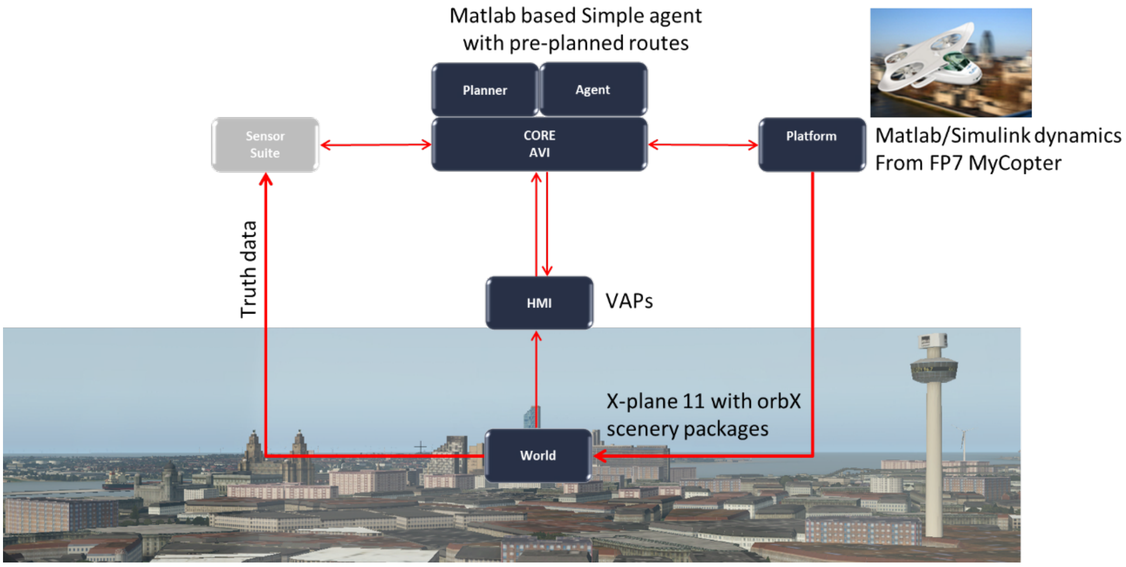


Figure 7.15: System Architecture

7.2 Experiment design

The experiment will evaluate passenger acceptance of autonomous air vehicles and their capability or willingness to take control should this be necessary or convenient.

7.2.1 Measurement package

The independent variables in this experiment are:

- Gender
- Age
- Piloting skill level
- General attitude towards autonomous vehicles

- Training received towards operating the vehicle
- Time histories of vehicle parameters (trajectory, control inputs (autonomous or human operator)
- Cockpit video/audio recording

7.2.2 Questionnaires:

A standard background questionnaire will be used to collect demographic data such as age, gender, driving and flying experience and experience/perception of eVTOL / PAV.

After each simulation run the subject will be asked to provide feedback on their comfort level with the run and decision options.

After the simulation runs, elements of the NASA TLX rating scale (Hart and Staveland, 1988) will be completed by the occupant detailing:

- Mental Demand - How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?
- Physical Demand - How much physical activity was required? Was the task easy or demanding, slack or strenuous?
- Temporal Demand - How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?
- Overall Performance - How successful were you in performing the task? How satisfied were you with your performance?
- Effort - How hard did you have to work (mentally and physically) to accomplish your level of performance?
- Frustration Level - How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?

7.2.3 Experimental protocol

Each participant session will last approximately 2 hours and consist of a briefing session (approximately 30 mins), the simulation experiments (approximately 1 hour) and a debriefing session (approximately 30 mins).

1) Briefing session: The subject will be briefed on the project and goals of the simulations. The subject will be asked about their level of experience and comfort with flight and of autonomous systems. The experiments will be described and the options available in each test case and required subject inputs and outputs. The safety procedures and enhanced covid protocols will be described (including participants wearing gloves in the simulator and enhanced cleaning

processes between participants). The subject will be told that some subjects are prone to experiencing some motion sickness and must let the simulation controller know if this occurs.

2) Simulation Experiments: Each subject will first fly the nominal flight (scenario 0), and then at least one of the other possible scenarios which contemplate an airspace closure. The simulation experiments will be run in the same order for all subjects. The order of participant's age/sex/piloting skill is not important to the experiment. Each task is briefed again as the task is about to be undertaken. Subject's perception and comfort will be ascertained after each run.
3) Debrief session: the subject will be asked to complete the NASA TLX scale and provide any additional feedback.

7.3 Future work

Immediate work is focusing on finalising setup of the testing environment, develop a series of briefing and de-brief questionnaires and obtaining approval from the University ethics committee. Once this approval has been granted, recruitment of volunteers will begin. It is anticipated that initial testing will begin in early December 2021 and that this will conclude by the end of January 2022. The target population of test participants will be 30 in total, balanced, as far as is possible, in terms of age and gender. The results will be reported in D7.3 which brings all integrated data analysis together to address the objectives of PAsCAL.

8 Enrichment of public acceptance maps

This chapter showcases comparisons between the participants to the simulations and those panel participants to the large-scale acceptance assessment of WP3 (see. D.3.1). We take advantage of the fact that for some simulations, specifically XP1 (driving simulator) and XP2 (the VR system), identical measures for acceptance were employed.

8.1 Insights from XP1

In XP1, participants (regular drivers) experienced a L3 car in a driving simulator. Their responses before (pre) and after (post) using the driving simulator are examined. Furthermore, the results of XP1 are compared with findings of a panel study.

The first dependent variable considered in more detail was "attitude" (see Figure 8.1). Overall, it can be stated that the panellists had the least positive average attitude (*M* = 3.81, *SD* = 1.65). The fact that the panellists had a less positive average attitude than the XP1 participants before the driving simulation (Pre: *M* = 4.76, *SD* = 1.14) can be partly attributed to a selection bias in XP1. Interestingly, the average attitude differed between XP1 participants before and after using the driving simulator: Their attitude was more positive after using the driving simulator (Post: *M* = 5.21, *SD* = 0.96). The same pattern emerged when looking at the results separated by gender (see Figure 8.2) and past experience (see Figure 8.3). From this, it can be inferred that direct experience with an autonomous vehicle improves people's attitude.

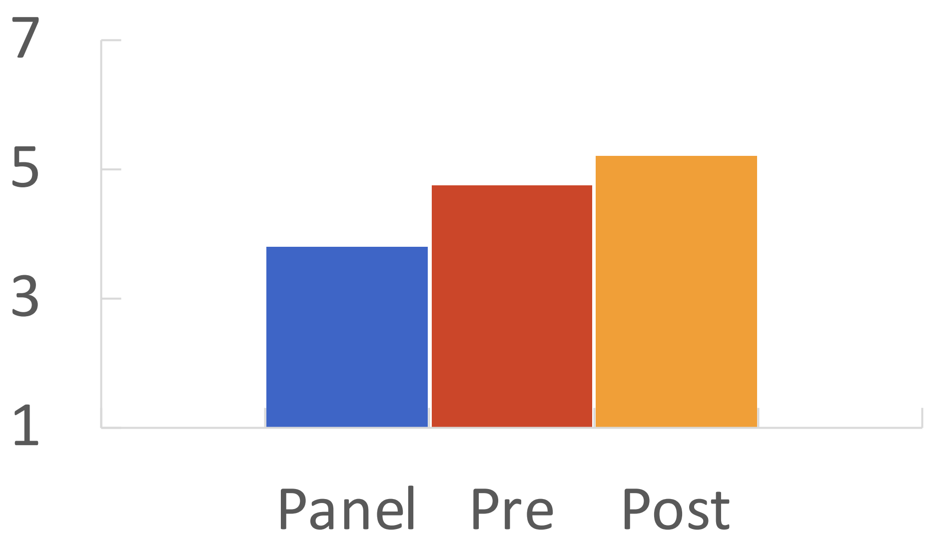


Figure 8.1: Average “attitude”

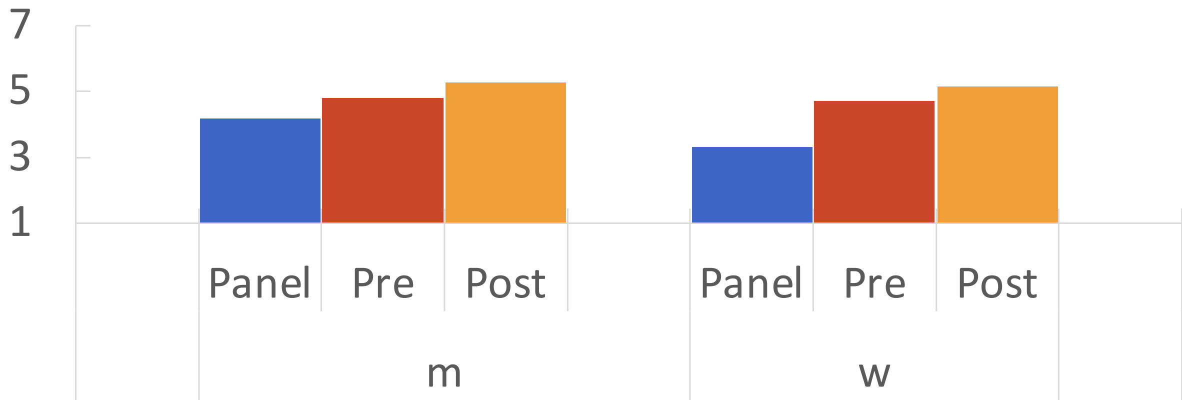


Figure 8.2: Average “attitude” separated by gender (m = male, w = female)

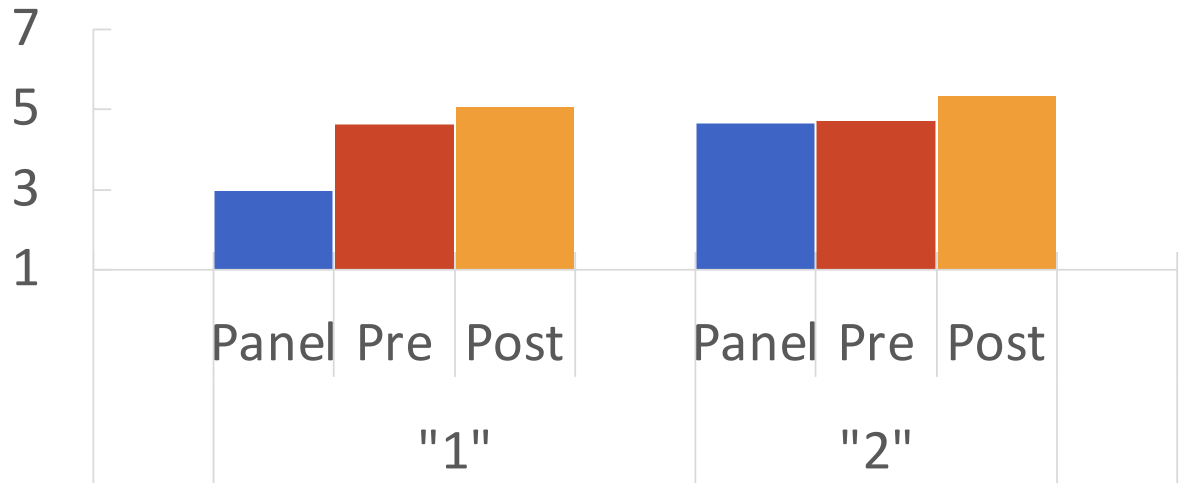


Figure 8.3: Average “attitude” separated by past experience (1 = little experience, 2 = much experience)

The second dependent variable considered in more detail was "intention to use" (see Figure 8.4). Overall, it can be stated that the panellists had the lowest average intention to use (*M* = 3.06, *SD* = 1.28). The fact that the panellists had a lower average intention to use than the XP1 participants before the driving simulation (Pre: *M* = 3.79, *SD* = 0.77) can be partly attributed to a selection bias in XP1. Interestingly, the average intention to use differed between XP1 participants before and after using the driving simulator: Their intention to use was stronger after using the driving simulator (Post: *M* = 4.38, *SD* = 0.77). The same pattern emerged when looking at the results separated by gender (see Figure 8.5) and past experience (see Figure 8.6). From this, it can be inferred that direct experience with an autonomous vehicle reinforces people's intention to use.

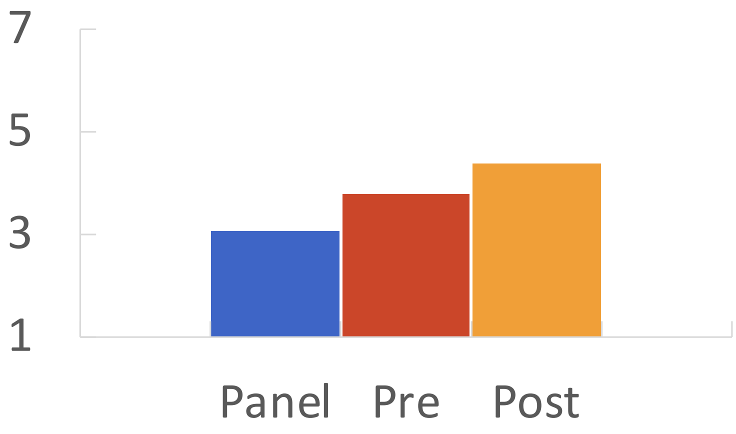


Figure 8.4: Average "intention to use"

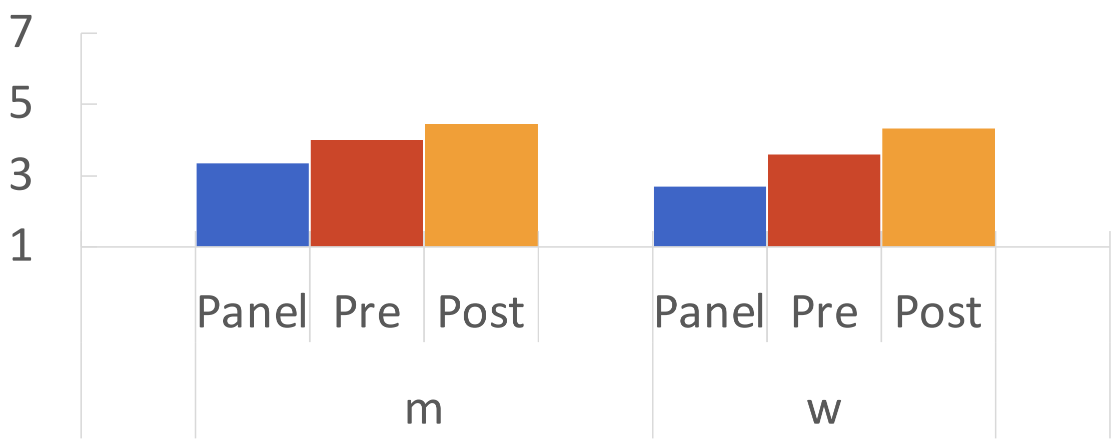


Figure 8.5: Average "intention to use" separated by gender (m = male, w = female)

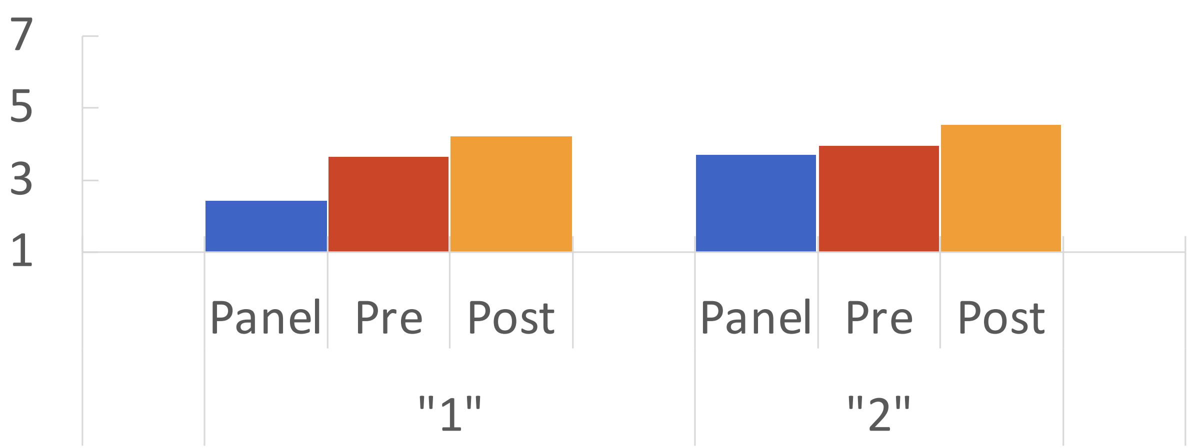


Figure 8.6: Average "intention to use" separated by past experience (1 = little experience, 2 = much experience)

The third dependent variable considered in more detail was "perceived risk" (see Figure 8.7). Higher values indicate less perceived risk. Overall, it can be stated that the panellists had the highest average perceived risk (*M* = 4.43, *SD* = 1.80). The fact that the panellists had a higher average perceived risk than the XP1 participants before the driving simulation (Pre: *M* = 5.26, *SD* = 1.14) can be partly attributed to a selection bias in XP1. Interestingly, the average perceived risk differed between XP1 participants before and after using the driving simulator: Their perceived risk was lower after using the driving simulator (Post: *M* = 5.64, *SD* = 0.96). Nearly the same pattern emerged when looking at the results separated by gender (see Figure 8.8) and past experience (see Figure 8.9). From this, it can be inferred that direct experience with an autonomous vehicle decreased people's perceived risk.

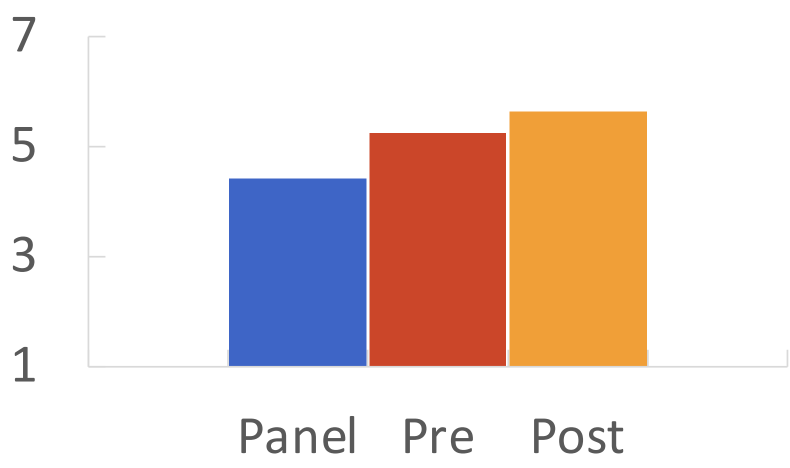


Figure 8.7: Average "perceived risk" (higher values = less risk)

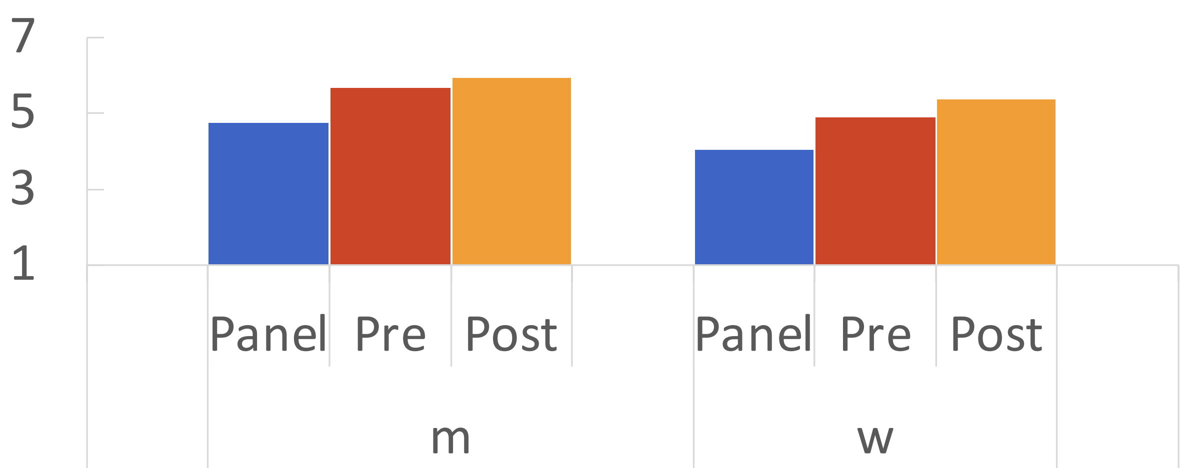


Figure 8.8: Average "perceived risk" separated by gender (m = male, w = female; higher values = less risk)

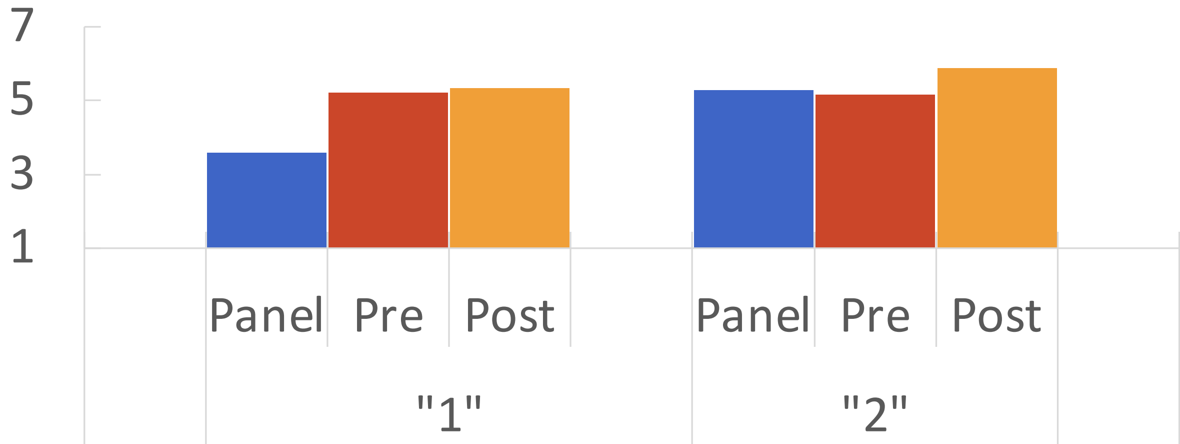


Figure 8.9: Average "perceived risk" separated by past experience (1 = little experience, 2 = much experience; higher values = less risk)

The fourth dependent variable considered in more detail was "perceived risk of accident" (see Figure 8.10). Again, higher values indicate less perceived risk. Overall, it can be stated that the panellists had the highest average perceived risk of accident (*M* = 4.67, *SD* = 1.93). The fact that the panellists had a lower average perceived risk of accident than the XP1 participants before the driving simulation (Pre: *M* = 5.69, *SD* = 0.89) can be partly attributed to a selection bias in XP1. Interestingly, the average perceived risk of accident differed between XP1 participants before and after using the driving simulator: Their perceived risk of accident was lower after using the driving simulator (Post: *M* = 5.97, *SD* = 0.87).

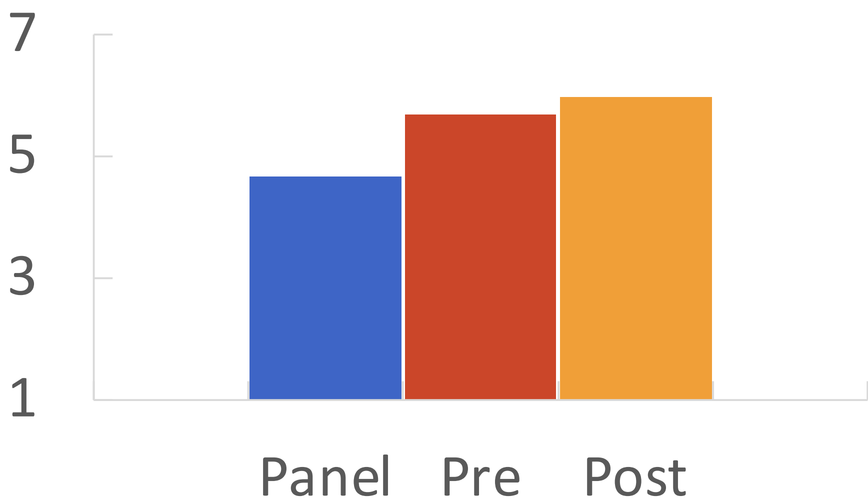


Figure 8.10: Average "perceived risk of accident" (higher values = more positive)

The same pattern emerged when looking at the results separated by gender (see Figure 8.11) and past experience (see Figure 8.12). From this, it can be inferred that direct experience with an autonomous vehicle decreases people's perceived risk of accident (coding: higher values are more positive).

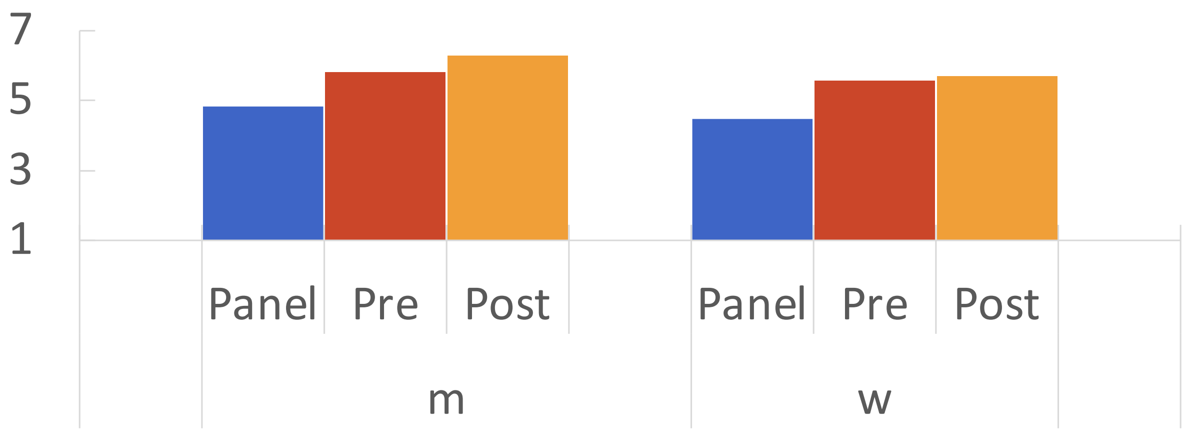


Figure 8.11: Average "perceived risk of accident" separated by gender (m = male, w = female; higher values = less risk)

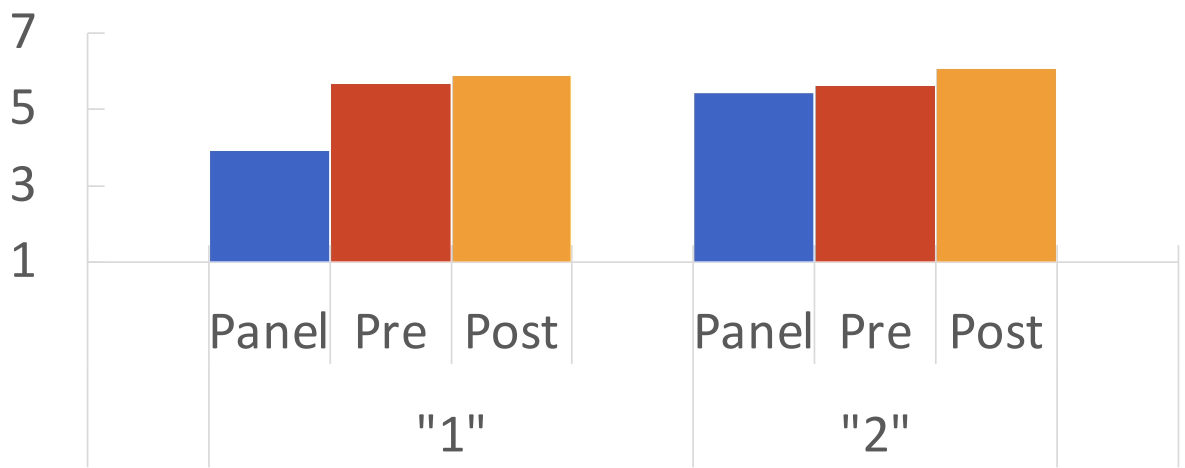


Figure 8.12: Average "perceived risk of accident" separated by past experience (1 = little experience, 2 = much experience; higher values = less risk)

The fifth dependent variable considered in more detail was "expected pleasantness" (see Figure 8.13). Overall, it can be stated that the panellists had the lowest average expected pleasantness (M = 4.30, SD = 1.83). The fact that the panellists had a lower average expected pleasantness than the XP1 participants before the driving simulation (Pre: M = 5.49, SD = 1.30) can be partly attributed to a selection bias in XP1.

Interestingly, the average expected pleasantness differed between XP1 participants before and after using the driving simulator: Their expected pleasantness was (minimally) greater after using the driving simulator (Post: M = 5.56, SD = 1.21).

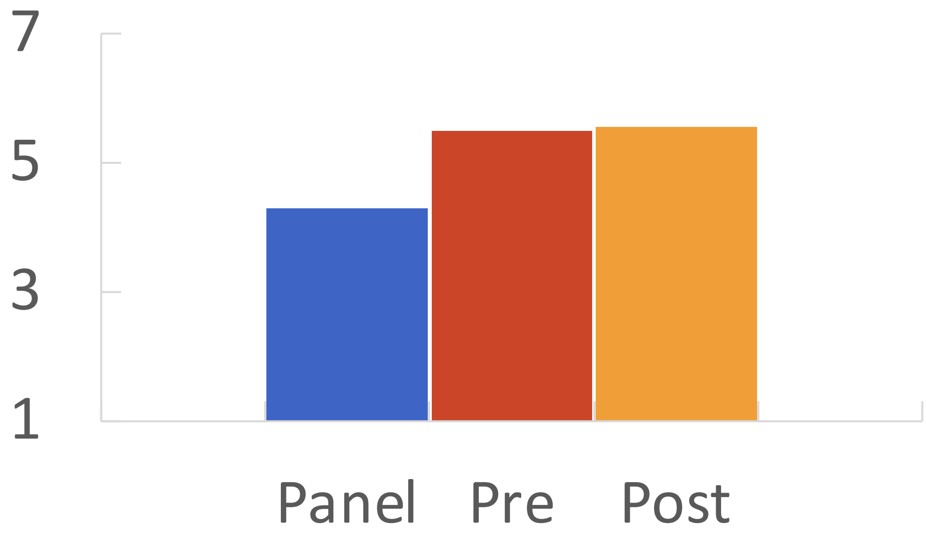


Figure 8.13: Average "expected pleasantness"

Nearly the same pattern emerged when looking at the results separated by gender (see Figure 8.14) and past experience (with the exception of people with much experience; see Figure 8.15). From this, it can be inferred that direct experience with an autonomous vehicle magnifies people's expected pleasantness (albeit minimally).

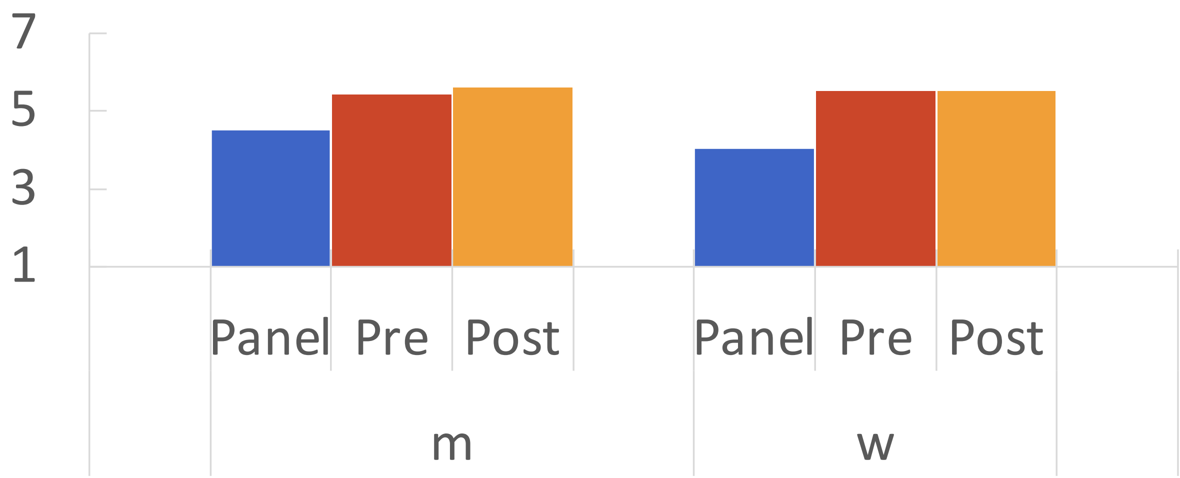


Figure 8.14: Average "expected pleasantness" separated by gender (m = male, w = female)

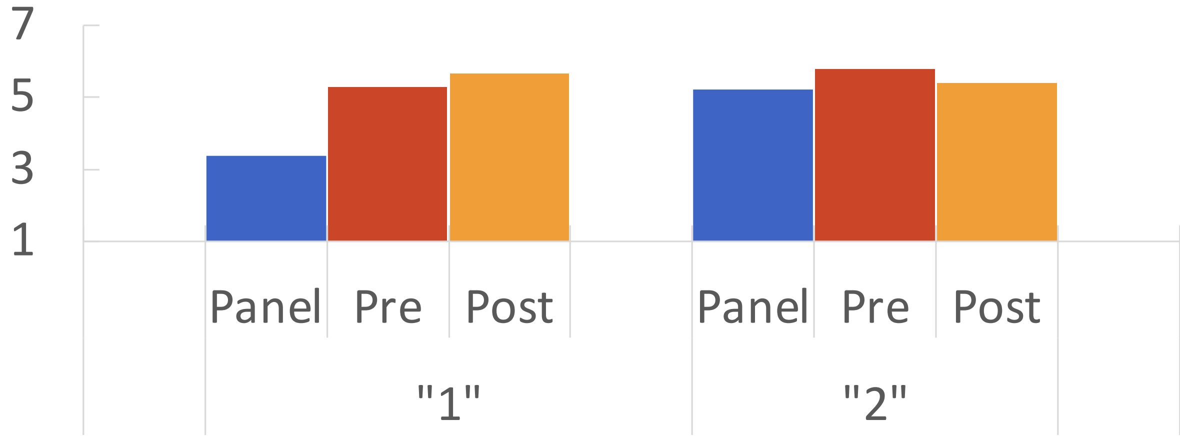


Figure 8.15: Average “expected pleasantness” separated by past experience (1 = little experience, 2 = much experience)

The sixth dependent variable considered in more detail was "affective reaction" (see Figure 8.16). Overall, it can be stated that the panellists had the least positive average affective reaction (M = 4.49, SD = 1.98). The fact that the panellists had a less positive average affective reaction than the XP1 participants before the driving simulation (Pre: M = 4.79, SD = 1.66) can be partly attributed to a selection bias in XP1. Interestingly, the average affective reaction differed between XP1 participants before and after using the driving simulator: Their affective reaction was more positive after using the driving simulator (Post: M = 5.13, SD = 1.76).

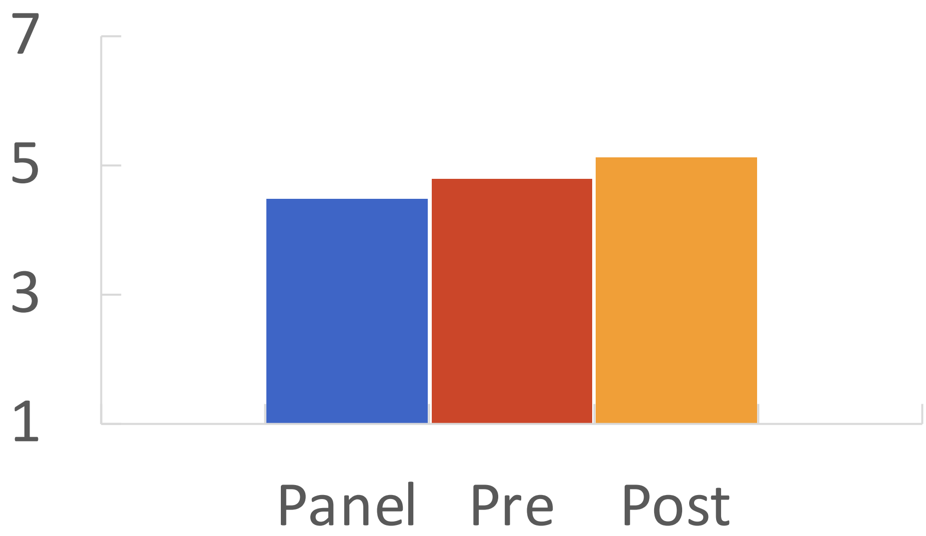


Figure 8.16: Average “affective reaction”

A different pattern emerged when looking at the results separated by gender (see Figure 8.17) and by past experience (see Figure 8.18). Only for men and those with prior experience the XP1 experience made the affective reaction more positive.

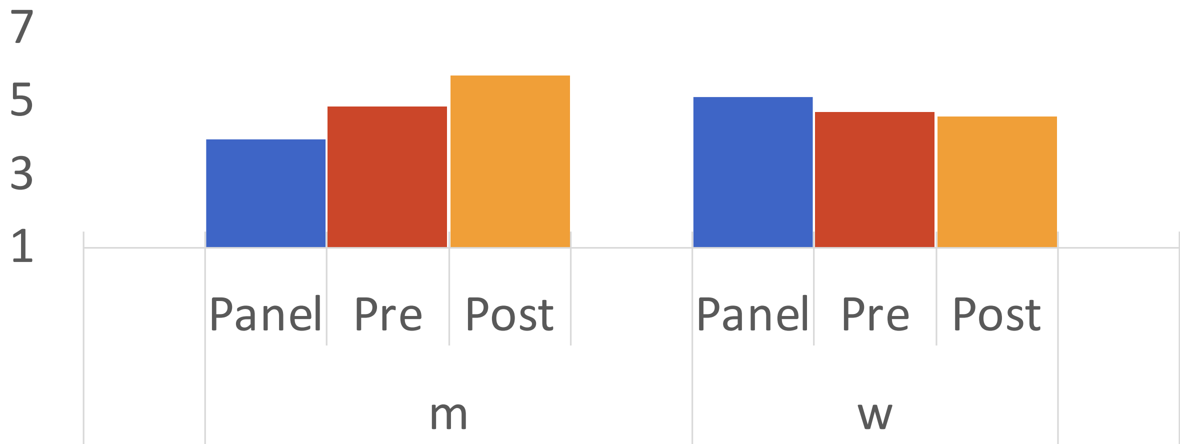


Figure 8.17: Average "affective reaction" separated by gender (m = male, w = female)

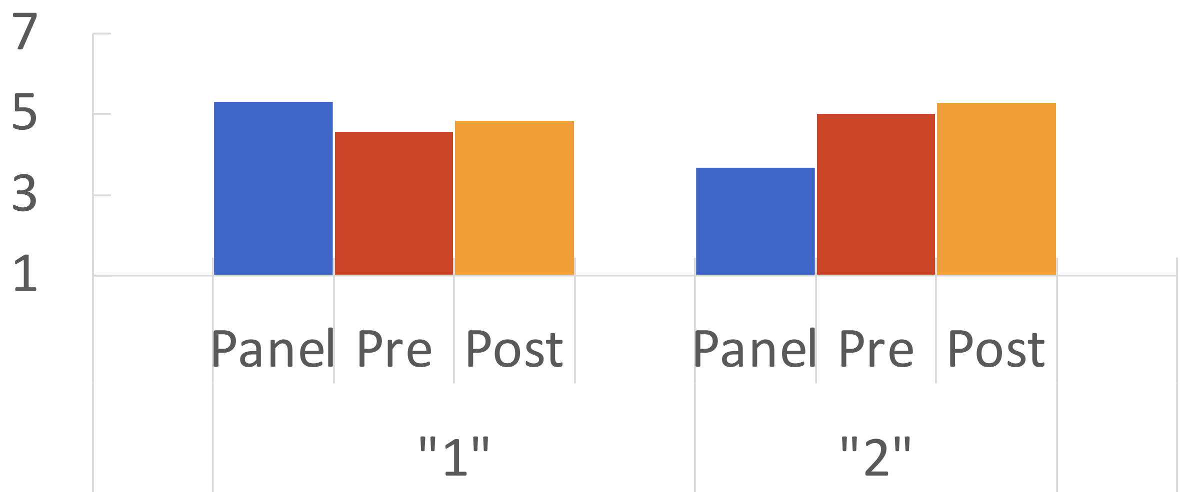


Figure 8.18: Average "affective reaction" separated by past experience (1 = little experience, 2 = much experience)

The last dependent variable considered in more detail was "willingness to pay" (see Figure 8.19). Overall, it can be stated that the panellists had the lowest (even negative) average willingness to pay (M = -0.16, SD = 0.52). The fact that the panellists had a lower average willingness to pay than the XP1 participants before the driving simulation (Pre: M = 0.28, SD =

0.69) can be partly attributed to a selection bias in XP1. Interestingly, the average willingness to pay differed between XP1 participants before and after using the driving simulator: Their willingness to pay was higher after using the driving simulator (Post: M = 0.33, SD = 0.62).

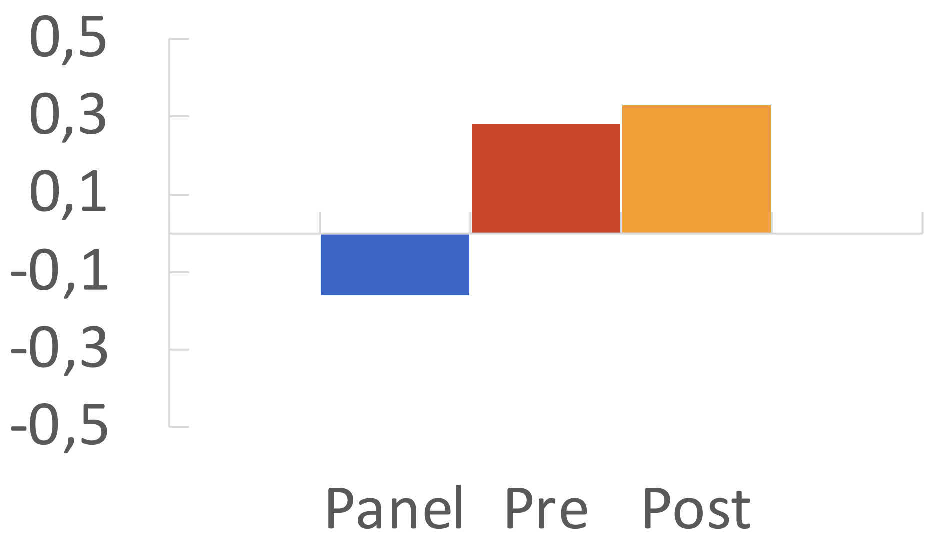


Figure 8.19: Average "willingness to pay"

Nearly the same pattern emerged when looking at the results separated by gender (see Figure 8.20) and past experience (with the exception of people with little experience; see Figure 8.21). From this, it can be inferred that direct experience with an autonomous vehicle increases people's willingness to pay, except for those with little prior experience.

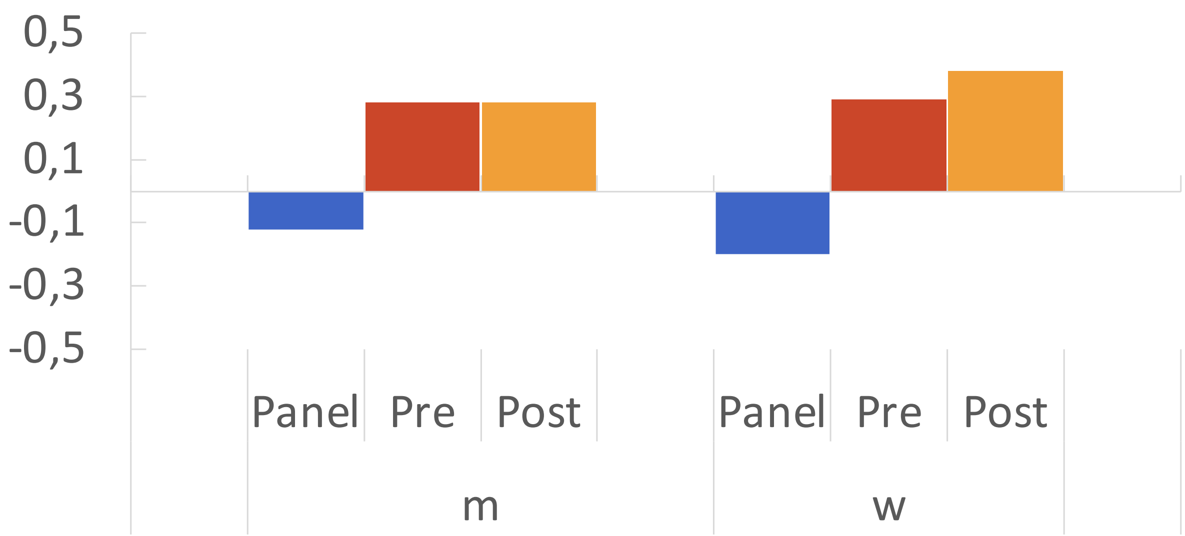


Figure 8.20: Average "willingness to pay" separated by gender (m = male, w = female)

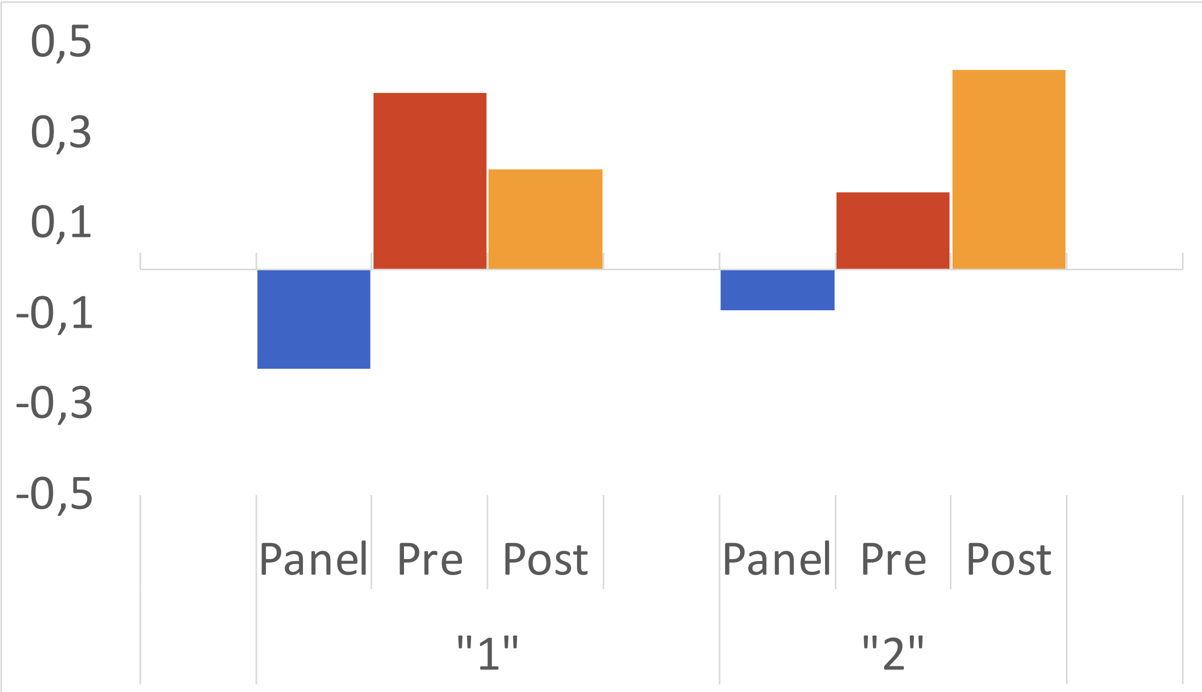


Figure 8.21: Average "willingness to pay" separated by past experience (1 = little experience, 2 = much experience)

8.2 Insights from XP2

In XP2, participants (with motor disability) experienced L5 shuttles on a Virtual Reality (VR) platform. Their responses before (pre) and after (post) using the VR system are examined. Furthermore, the results of XP2 are compared with findings of a panel study. In contrast to section 8.1, this section does not consider the results separated by gender since the sample of XP2 includes only one woman.

The first dependent variable considered in more detail was "attitude" (see *Figure 8.22*). Overall, it can be stated that the panellists had the least positive average attitude (*M* = 4.47, *SD* = 1.82). The fact that the panellists had a less positive average attitude than the XP2 participants before using the VR system (Pre: *M* = 4.77, *SD* = 1.57) can be partly attributed to a selection bias in XP2. Interestingly, the average attitude differed between XP2 participants before and after using the VR system: Their attitude was more positive after using the VR system (Post: *M* = 5.61, *SD* = 1.16).

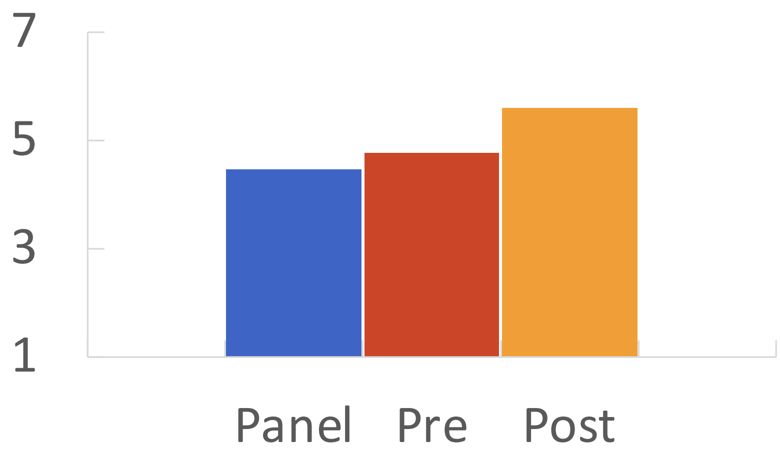


Figure 8.22: Average "attitude"

The same pattern emerged when looking at the results separated by past experience (see *Figure 8.23*). From this, it can be inferred that direct experience with an autonomous vehicle improves people's attitude.

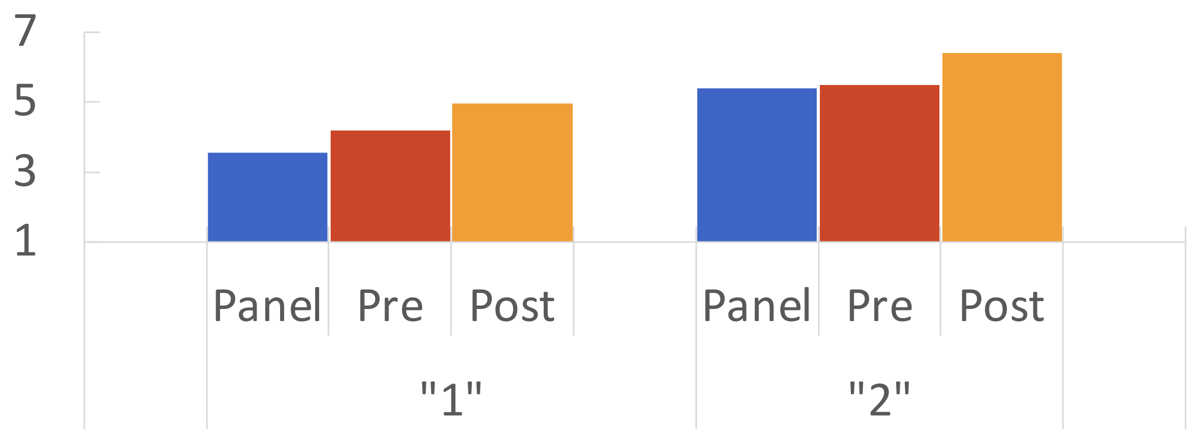


Figure 8.23: Average "attitude" separated by past experience (1 = little experience, 2 = much experience)

The second dependent variable considered in more detail was "intention to use" (see Figure 8.24). Overall, it can be stated that the panellists had the lowest average intention to use (M = 3.72, SD = 1.28). The fact that the panellists had a lower average intention to use than the XP2 participants before using the VR system (Pre: M = 5.00, SD = 1.95) can be partly attributed to a selection bias in XP2. Interestingly, the average intention to use differed between XP2 participants before and after using the VR system: Their intention to use was stronger after using the VR system (Post: M = 5.82, SD = 1.40).

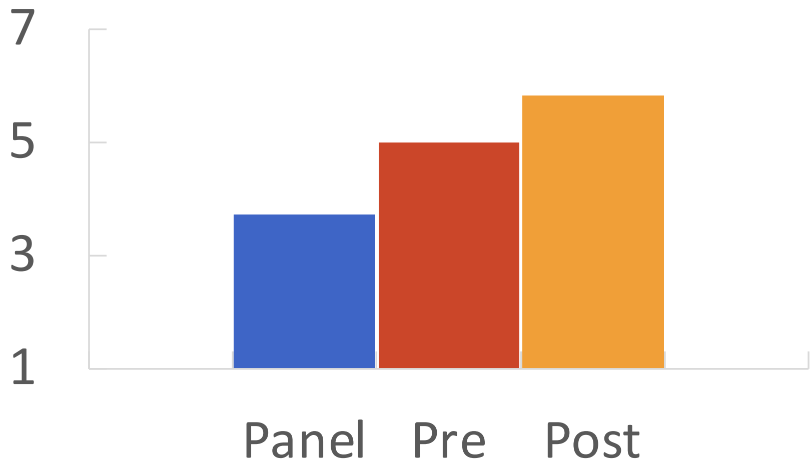


Figure 8.24: Average "intention to use"

The same pattern emerged when looking at the results separated by past experience (see Figure 8.25). From this, it can be inferred that direct experience with an autonomous vehicle reinforces people's intention to use.

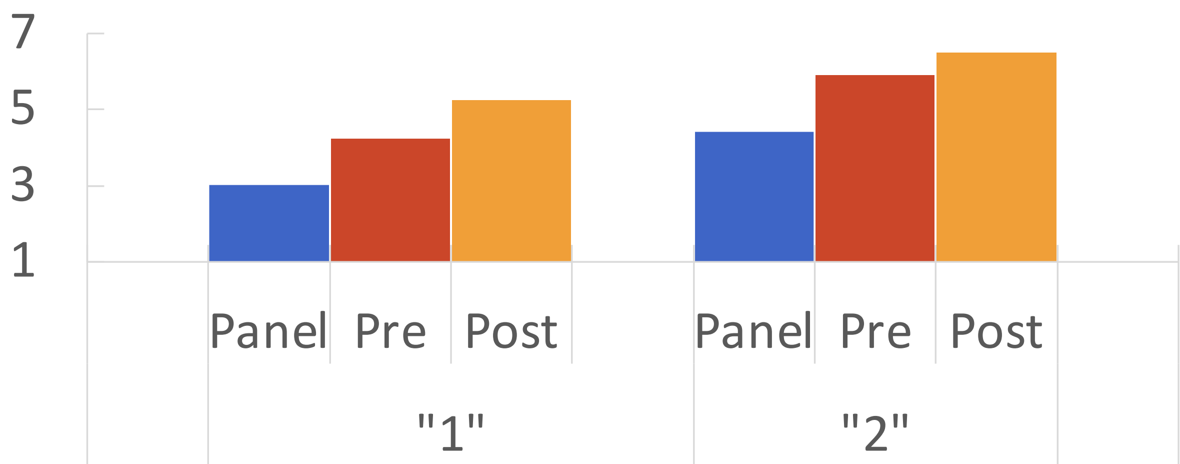


Figure 8.25: Average "intention to use" separated by past experience (1 = little experience, 2 = much experience)

The third dependent variable considered in more detail was "affective reaction" (see Figure 8.26). Overall, it can be stated that the panellists had the least positive average affective reaction (M = 4.17, SD = 1.14). The fact that the panellists had a less positive average affective reaction than the XP2 participants before using the VR system (Pre: M = 5.14, SD = 1.27) can be partly attributed to a selection bias in XP2. Interestingly, the average affective reaction differed between XP2 participants before and after using the VR system: Their affective reaction was more positive after using the VR system (Post: M = 5.50, SD = 1.60).

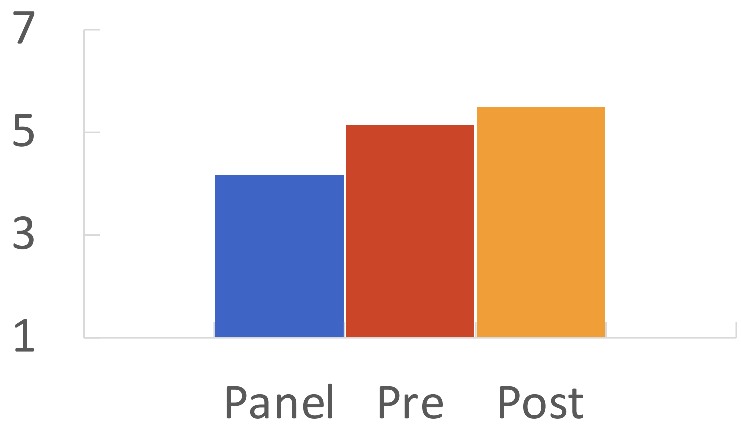


Figure 8.26: Average "affective reaction"

The same pattern emerged when looking at the results separated by past experience (see Figure 8.27). From this, it can be inferred that direct experience with an autonomous vehicle improves people's affective reaction.

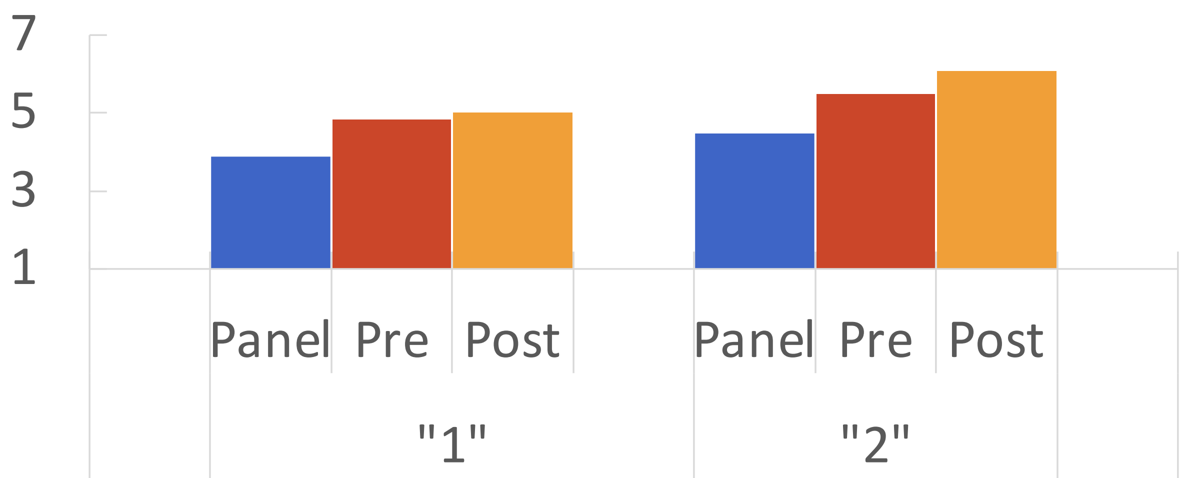


Figure 8.27: Average "affective reaction" separated by past experience (1 = little experience, 2 = much experience)

The last dependent variable considered in more detail was "willingness to pay" (see Figure 8.28). Overall, it can be stated that the panellists had the lowest average willingness to pay (M = -0.31, SD = 0.53). The fact that the panellists had a lower average willingness to pay than the XP2 participants before using the VR system (Pre: M = -0.18, SD = 0.60) can be partly attributed to a selection bias in XP2. Interestingly, the average willingness to pay differed between XP2 participants before and after using the VR system: Their willingness to pay was higher (or not negative) after using the VR system (Post: M = 0.00, SD = 0.45).

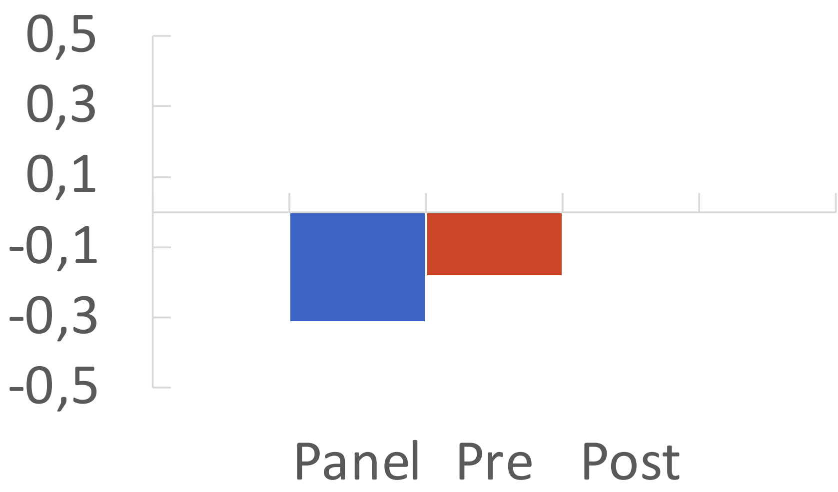


Figure 8.28: Average "willingness to pay"

Nearly the same pattern emerged when looking at the results separated by past experience (see Figure 8.29). From this, it can be inferred that direct experience with an autonomous vehicle increases (or at least neutralizes) people's willingness to pay.

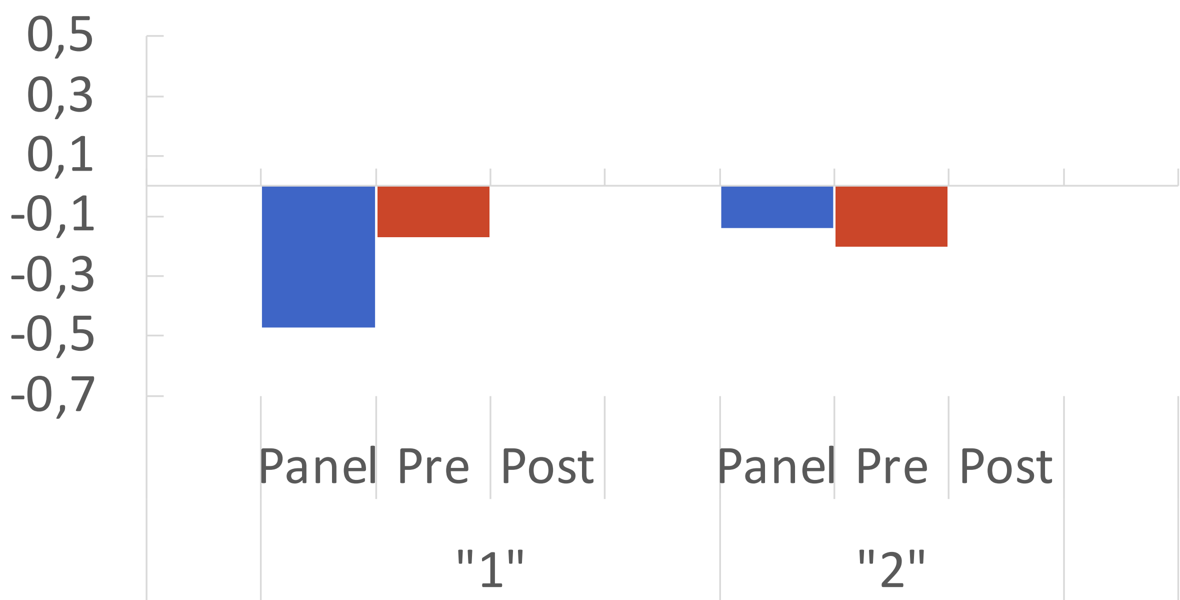


Figure 8.29: Average "willingness to pay" separated by past experience (1 = little experience, 2 = much experience)

Overall, it can be concluded that the participants' responses were more positive after they had used the driving simulator (XP1) or the VR system (XP2). This finding has important implications for practice.

8.3 Conclusions

Overall, we find that direct experience with CAV simulations, increases acceptance, including attitudes, affective reactions, intention to use and willingness to pay. This extends previous research illustrating that direct experience with real-life prototypes increases acceptance (e.g. Liu & Xu, 2020). Further we find that some degree of previous experience is necessary for furthermore immersive experience to yield positive effects.

These findings have strategic implications. First, to increase acceptance simulators might offer a cost efficient and safe alternative to on-the-road prototypes. Second, “phasing-in” autonomous features stepwise, for example by exposure to partially automated vehicle features, seems more advisable than direct confrontation with L5 systems.

Though in line with previous research, these findings are limited due to small sample sizes, brought about by COVID-19 restrictions, and to only two simulations, due to a lack of comparable items. Future research should address the effects of experience with CAV simulations in more detail.

9 Conclusions

This deliverable reports the analysis of five simulation experiments of the interactions of the autonomous vehicle with different road users including non-drivers. While focusing on different users, levels of automation and driving situations, these experiments have carried out several common tasks, namely (1) correlation and analysis of driver behaviour/reaction under different scenarios, (2) assessment of the acceptance of new interfaces integrated in the simulators, including information feedback and entertainment systems, (3) recommendations describing ways to improve the CAVs design for future drivers' trainings, and (4) guidelines for WP6 pilot specifications and demonstrations. The main findings of these simulation experiments are summarised as follows.

Findings of "DRIVING SIMULATOR":

- It was observed that those participants who had some experience and knowledge of autonomous vehicles were able to get a more concrete idea of how an autonomous vehicle works, what could drive to an increased acceptability, more positive attitude and feelings towards autonomous vehicles.
- The results from the study of the effectiveness and acceptance of the different signals present in the CAV showed that audio signals were preferred and considered the most effective by the participants. The voice signal was the most relevant signal for handover and taking over requests according to all participants in the experiment. While the experienced drivers were more responsive to the light signal, they agreed with the novices that it was more relevant as a confirmation of autonomous driving engagement, once it has been properly activated.
- The results from the analysis of the effect of the driving experience on the acceptability of the CAV showed that experienced drivers report higher trust than novice ones, with higher acceptability, more positive attitude, and lower perception of the risk associated with CAVs, which emphases the importance of knowledge transfer, training/education, and awareness of CAVs.
- This experiment also showed that although an information-rich HMI is better perceived in terms of usability, it does not lead to more trust for the driver. At times, the opposite is true. Some specific feedback about the car's level of perception can be perceived as a source of stress for the driver, for both experienced and novice drivers.

Findings of "VIRTUAL REALITY PLATFORM":

- Like the previous driving simulator, the VR experiment also observed that the participants who had some experience and knowledge of CAVs declared a high level of trust during the VR experience.
- The VR simulation delivered to them a more concrete idea of how works a L5 vehicle and the services it could provide. Experimenting L5 CAV shuttles was a good surprise for most of them.
- The overall attitude and feelings of most of the participants, who were already positive before the experiment, increased when re-measuring after.
- The results also showed that vulnerable disabled participants preferred shuttles to conventional buses.
- Participants were in favour of premium L5 shuttles for the multimedia and infotainment services, combined with their superior design and comfort. Their willingness-to-pay, however, didn't increase while considering this option.
- Further research is needed to confirm these findings of acceptability by testing larger panels and real-life situations.

Findings of "HOME STUDY SIMULATOR":

- In this study only a single alert followed by a countdown were used, which resulted in participants often feeling stressed or hurried.
- Alerts often seen as annoying and interruptive had negative impacts on the participants feelings towards the CAV and increased mental load and feelings of control.
- Perceptions of the CAV change over time from feelings of fear (first visit) to issues to do with control and decision making (last visit).
- The ability to predict how and what kind of decisions the CAV will take was seen as positive. Uncertainty was perceived negatively and fear of an unexpected end to autonomous mode was present.
- As a level 4 vehicle still requires manual intervention, it places a responsibility and hence the need to be attentive at all times on the driver.
- Further work with more participants is required to obtain sufficient quantitative data, in combination with the rich and detailed qualitative

data provided by the repertory grid analysis, to provide scientific evidence for assessment of real "driver" behaviours towards CAVs.

Findings of "IMMERSIVE ARENA":

This experiment has produced a number of observations including:

- Nationality, living country and having young kids seems to have an impact on CAV receptivity.
- When a CAV stops to let crossing a pedestrian, it is better to send a signal that the CAV will wait the pedestrian's crossing.
- A feedback is waited by pedestrians in all situations and particularly in dangerous ones.
- Presence or absence of a crosswalk already on the road does not play a significant role.
- When the CAV stops the use of a signal to show that the CAV is waiting that the pedestrian cross is needed. The projection on the road is well accepted in the cases of a pedestrian crossing is painted or not on the road.
- When no pedestrian crossing is painted on the road, the pedestrians mostly expect that the CAV doesn't stop. Thus, no signal seems needed in this case. Or a discrete signal without honk can be used like a red light on the VAE or projected on the road.
- If a pedestrian crossing is painted on the road, pedestrians expect that the CAV stops.
- The CAV has to be easily identified in the traffic.
- Regulation and standardization of eHMIs are needed to ensure uniformity regardless of the manufacturer and improve predictivity, understanding and so acceptation of CAVs.
- The more promising eHMI in terms of UX and receptivity are text-based interfaces, but it raised some issues to be understood by everybody including visually impaired, illiterate, kids, persons not able to read the language used.
- The easier to understand and more elegant the eHMI is perceived to be, the better the acceptance of the CAVs equipped with it.

Findings of "HELIFLIGHT-R":

Immediate work was focusing on finalising setup of the testing environment, developing a series of briefing and de-brief questionnaires and obtaining approval from the University ethics committee. No experimental data has been collected. Once this approval has been granted, recruitment of volunteers will begin.

Other work carried out in WP4:

WP4 has carried out a State-of-the-Art (SoA) review of **lessons learned** and **results found in other projects**, complementing the findings obtained from the aforementioned experiments. It focuses on several human-vehicle interactions such as:

- Interaction of the human driver with the autonomous vehicle, focusing on HMI designs for TOR, their impact on behaviour and acceptance.
- Driver training: Given the novelty of the systems, drivers need to have accurate expectations and mental models. Studies in the existing literature have investigated the impact of training on driver behaviour and acceptance of autonomous vehicles.
- Interactions with pedestrians, focusing on the use, efficiency and acceptance of eHMI that aim in facilitating interactions of pedestrians with autonomous vehicles.
- Autonomous public transport, investigating the acceptance of autonomous shuttles after experiencing the system, and the needs of peoples with disabilities.
- Issues related to the acceptance of autonomous urban air mobility.

The results from these simulation experiments were also used to enrich the **multidimensional map of public acceptance** developed in WP3. It was found that direct experience with CAV simulations, increases acceptance, including attitudes, affective reactions, intention to use and willingness to pay. It was also found that some degree of previous experience is necessary for furthermore immersive experience to yield positive effects. These findings have strategic implications. First, to increase acceptance simulators might offer a cost efficient and safe alternative to on-the-road prototypes. Second, "phasing-in" autonomous features stepwise, for example by exposure to partially automated vehicle features, seems more advisable than direct confrontation with L5 systems.

10 References

Ahmed, S. S., Fountas, G., Eker, U., Still, S. E., & Anastasopoulos, P. C. (2021). An exploratory empirical analysis of willingness to hire and pay for flying taxis and shared flying car services. *Journal of Air Transport Management, 90*. doi:10.1016/j.jairtraman.2020.101963

Al Haddad, C., Chaniotakis, E., Straubinger, A., Plötner, K., & Antoniou, C. (2020). Factors affecting the adoption and use of urban air mobility. *Transportation Research Part A: Policy and Practice, 132*, 696-712. doi:10.1016/j.tra.2019.12.020

Almallah, M., Hussain, Q., Reinolsmann, N., & Alhajyaseen, W. K. M. (2021). Driving simulation sickness and the sense of presence: Correlation and contributing factors. *Transportation Research Part F: Traffic Psychology and Behaviour, 78*, 180-193. doi:10.1016/j.trf.2021.02.005

Anderson, J., Kalra, N., Stanley, K., Sorenson, P., Samaras, C., & Oluwatola, O., (2014). Autonomous Vehicle Technology: A Guide for Policymakers (No. RR443-RC). RAND Corporation: Santa Monica, CA. https://www.rand.org/pubs/research_reports/RR443-2.html

Astfalk, S., Silberer, J., Planing, P., & Müller, P. (2021). The effect of a functional prototype on user acceptance in transportation: Assessing the level of acceptance before and after the first demonstration flight of an air taxi. *Transportation research interdisciplinary perspectives, 11*. doi:10.1016/j.trip.2021.100444

Aydin, B. (2019). Public acceptance of drones: Knowledge, attitudes, and practice. *Technology in Society, 59*. doi:10.1016/j.techsoc.2019.101180

Bansal, P., Kockelman, K.M., Singh, A., 2016. Assessing public opinions of and interest in new vehicle technologies: an Austin perspective. Transp. Res. Part C: Emerg. Technol. 67, 1–14. http://dx.doi.org/10.1016/j.trc.2016.01.019

Bansal P., Kockelman K.M. (2017). Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. Transp. Res. A Policy Pract., 2017 (95) (2017), pp. 49-63.

Bansal, P., & Kockelman, K. (2018). Are we ready to embrace connected and self-driving vehicles? A case study of Texans. Transportation, 45, 641-675. https://doi.org/10.1007/s11116-016-9745-z

Bazilinskyy, P., Kooijman, L., Dodou, D., & de Winter, J. C. F. (2021). How should external human-machine interfaces behave? Examining the effects of colour, position, message, activation distance, vehicle yielding, and visual distraction among 1,434 participants. *Appl Ergon, 95*, 103450. doi:10.1016/j.apergo.2021.103450

Bazilinskyy, P., Petermeijer, S. M., Petrovych, V., Dodou, D., & de Winter, J. C. F. (2018). Take-over requests in highly automated driving: A crowdsourcing survey on auditory, vibrotactile, and visual displays. *Transportation Research Part F: Traffic Psychology and Behaviour, 56*, 82-98. doi:10.1016/j.trf.2018.04.001

Becker, F., & Axhausen, K. (2017). Literature review on surveys investigating the acceptance of automated vehicles. Transportation, 44, 1293-1306. https://doi.org/10.1007/s11116-017-9808-9

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

Bellem, H., Kluver, M., Schrauf, M., Schoner, H. P., Hecht, H., & Krems, J. F. (2017). Can We Study Autonomous Driving Comfort in Moving-Base Driving Simulators? A Validation Study. *Hum Factors, 59*(3), 442-456. doi:10.1177/0018720816682647

Bennett, R., Vijaygopal, R., & Kottasz, R. (2020). Willingness of people who are blind to accept autonomous vehicles: An empirical investigation. *Transportation Research Part F: Traffic Psychology and Behaviour, 69*, 13-27. doi:10.1016/j.trf.2019.12.012

Bernhard, C., Oberfeld, D., Hoffmann, C., Weismüller, D., & Hecht, H. (2020). User acceptance of automated public transport. *Transportation Research Part F: Traffic Psychology and Behaviour, 70*, 109-123. doi:10.1016/j.trf.2020.02.008

Blaauw, G. J. (1982). Driving experience and task demands in simulator and instrumented car: a validation study. *Human factors, 24*(4), 473-486.

Borojeni, S. S., Chuang, L., Heuten, W., & Boll, S. (2016). *Assisting drivers with ambient take-over requests in highly automated driving.* Paper presented at the Proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications.

Bradshaw-Martin H., & Easton C. (2014). Autonomous or "driverless" cars and disability: A legal and ethical analysis. European Journal of Current Legal Issues, 20(3). Retrieved from http://webjcli.org/article/view/344.

Brandenburg, S., & Chuang, L. (2019). Take-over requests during highly automated driving: How should they be presented and under what conditions? *Transportation Research Part F: Traffic Psychology and Behaviour, 66*, 214-225. doi:10.1016/j.trf.2019.08.023

Brinkley, J., Huff, E. W., Posadas, B., Woodward, J., Daily, S. B., & Gilbert, J. E. (2020). Exploring the Needs, Preferences, and Concerns of Persons with Visual Impairments Regarding Autonomous Vehicles. *ACM Transactions on Accessible Computing, 13*(1), 1-34. doi:10.1145/3372280

Casas (2007). Social exclusion and the disabled: an accessibility approach Profess. Geogra., 59 (4) (2007), pp. 463-477.

Chang, C.-M., Toda, K., Sakamoto, D., & Igarashi, T. (2017). *Eyes on a Car*. Paper presented at the Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Chapman L. (2016). What do self-driving vehicles mean for disabled travellers? Disabled World, 19 December 2016, www.disabled-world.com.

Chen, C.-F. (2019). Factors affecting the decision to use autonomous shuttle services: Evidence from a scooter-dominant urban context. *Transportation Research Part F: Traffic Psychology and Behaviour, 67*, 195-204. doi:10.1016/j.trf.2019.10.016

Clothier, R. A., Greer, D. A., Greer, D. G., & Mehta, A. M. (2015). Risk Perception and the Public Acceptance of Drones. *Risk Anal, 35*(6), 1167-1183. doi:10.1111/risa.12330

Colley, M., Walch, M., Gugenheimer, J., Askari, A., & Rukzio, E. (2020). *Towards Inclusive External Communication of Autonomous Vehicles for Pedestrians with Vision Impairments*. Paper presented at the Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.

Collingwood, L. (2017). Privacy implications and liability issues of autonomous vehicles. Information & Communications Technology Law, 26(1), 32–45. https://doi.org/10.1080/13600834.2017.1269871

Cordts, P., Cotten, S. R., Qu, T., & Bush, T. R. (2021). Mobility challenges and perceptions of autonomous vehicles for individuals with physical disabilities. *Disabil Health J, 14*(4), 101131. doi:10.1016/j.dhjo.2021.101131

Currie G., Stanley J. (2007). No Way To Go: Transport and social disadvantage in Australian Communities, Monash University ePress, Melbourne, Australia.

Currie G., Richardson T., Smyth P., Vella-Brodrick D., Hine J., Lucas K., Stanley J. (2009). Investigating links between transport disadvantage, social exclusion and well-being in Melbourne—Preliminary results Transp. Policy, 16 (3), pp. 97-105

Currie G., Delbosc A. (2011). Transport Disadvantage: A Review New Perspectives and Methods in Transport and Social Exclusion Research Emerald Group Publishing Limited, Bingley, BD UK, pp. 15-26.

Darcy S., Burke P. (2018). On the road again: The barriers and benefits of automobility for people with disability. Transp. Res. (A), 107 (2018), pp. 229-245.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319. https://doi.org/10.2307/249008.

Deb, S., Strawderman, L., DuBien, J., Smith, B., Carruth, D. W., & Garrison, T. M. (2017). Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the U.S. population. Accident Analysis and Prevention, 106(January), 191–201. https://doi.org/10.1016/j.aap.2017.05.020.

Deb, S., Strawderman, L., Carruth, D. W., DuBien, J., Smith, B., & Garrison, T. M. (2017). Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles. Transportation Research Part C: Emerging Technologies, 84, 178–195. https://doi.org/10.1016/j.trc.2017.08.029.

Deb, S., Strawderman, L. J., & Carruth, D. W. (2018). Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. *Transportation Research Part F: Traffic Psychology and Behaviour, 59*, 135-149. doi:10.1016/j.trf.2018.08.016

Delbosc A. and Currie G. (2011). The spatial context of transport disadvantage, social exclusion and well-being. Journal of Transport Geography. Volume 19, Issue 6, November 2011, Pages 1130-1137. doi.org/10.1016/j.jtrangeo.2011.04.005.

Dogan, E., Rahal, M.-C., Deborne, R., Delhomme, P., Kemeny, A., & Perrin, J. (2017). Transition of control in a partially automated vehicle: Effects of anticipation and non-driving-related task involvement. *Transportation Research Part F: Traffic Psychology and Behaviour, 46*, 205-215. doi:10.1016/j.trf.2017.01.012

Ebnali, M., Hulme, K., Ebnali-Heidari, A., & Mazloumi, A. (2019). How does training effect users' attitudes and skills needed for highly automated driving? *Transportation Research Part F: Traffic Psychology and Behaviour, 66*, 184-195. doi:10.1016/j.trf.2019.09.001

Edwards, T., & Price, G. (2020). eVTOL passenger acceptance.

Eriksson, A., Banks, V. A., & Stanton, N. A. (2017). Transition to manual: Comparing simulator with on-road control transitions. *Accid Anal Prev, 102*, 227-234. doi:10.1016/j.aap.2017.03.011

Fagnant, D. and Kockelman, K. (2018). Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas Transportation, 45 (1) (2018), pp. 143-158.

Fairley P. (2018). The self-driving car's bicycle problem IEEE Spectrum, https://spectrum.ieee.org/cars-that-think/transportation/self-driving/the-selfdriving-cars-bicycle-problem, Accessed 17th Jul 2018.

Ferenchak, N., & Shafique, S. (2021). Pedestrians' Perceptions of Autonomous Vehicle External Human-Machine Interfaces. *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*. doi:10.1115/1.4051778

Fisher, D. L., Caird, J. K., Rizzo, M., & Lee, J. D. (2011). Handbook of driving simulation for engineering, medicine and psychology: an overview. *Handbook of driving simulation for engineering, medicine, and psychology*.

Forster, Y., Hergeth, S., Naujoks, F., Krems, J., & Keinath, A. (2019). User Education in Automated Driving: Owner's Manual and Interactive Tutorial Support Mental Model Formation and Human-Automation Interaction. *Information, 10*(4). doi:10.3390/info10040143

Forster, Y., Naujoks, F., Neukum, A., & Huestegge, L. (2017). Driver compliance to take-over requests with different auditory outputs in conditional automation. *Accid Anal Prev, 109*, 18-28. doi:10.1016/j.aap.2017.09.019

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

Gold, C., Korber, M., Hohenberger C., Bengler K. (2015). Trust in automation – before and after the experience of take-over scenarios in a highly automated vehicle. Procedia Manufacturing, 3, pp. 3025-3032.

Gold, C., Korber, M., Lechner, D., & Bengler, K. (2016). Taking Over Control From Highly Automated Vehicles in Complex Traffic Situations: The Role of Traffic Density. *Hum Factors, 58*(4), 642-652. doi:10.1177/0018720816634226

Goyal, R. (2018). Urban air mobility (uam) market study.

Hart, S. & Staveland, L., 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Advances in Psychology, pp. 139-183.

Heide A., Henning K. (2006). The "cognitive car": a roadmap for research issues in the automotive sector. Annu. Rev. Control., 30 (2) (2006), pp. 197-203.

Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior Familiarization With Takeover Requests Affects Drivers' Takeover Performance and Automation Trust. *Hum Factors, 59*(3), 457-470. doi:10.1177/0018720816678714

Hilgarter, K., & Granig, P. (2020). Public perception of autonomous vehicles: A qualitative study based on interviews after riding an autonomous shuttle. *Transportation Research Part F: Traffic Psychology and Behaviour, 72*, 226-243. doi:10.1016/j.trf.2020.05.012

Hine, J. and Mitchell, F. (2003), Transport Disadvantage and Social Exclusion: Exclusionary Mechanisms in Transport in Urban Scotland, Aldershot: Ashgate, 162 pp., £42.50 hbk, ISBN: 0-7546-1847-1. Journal of Social Policy, 33(3), 525-526. doi:10.1017/S004727940434794X

Hochman, M., Parmet, Y., & Oron-Gilad, T. (2020). Pedestrians' Understanding of a Fully Autonomous Vehicle's Intent to Stop: A Learning Effect Over Time. *Front Psychol, 11*, 585280. doi:10.3389/fpsyg.2020.585280

Hohenberger C., Spörrle M., Welpe I. (2017). Not fearless, but self-enhanced: the effects of anxiety on the willingness to use autonomous cars depend on individual levels of self-enhancement. Technol. Forecast. Soc. Chang., 116 (2017), pp. 40-52.

Howard, D., & Dai, D. (2015). Public perceptions of self-driving cars: the case of Berkeley, California. Presented at the 93rd Annual Meeting of Transportation Research Board, Washington, DC. https://www.ocf.berkeley.edu/~djhoward/reports/Report%20-%20Public%20Perceptions%20of%20Self%20Driving%20Cars.pdf

Hwang, J., Li, W., Stough, L., Lee, C., & Turnbull, K. (2020a). A focus group study on the potential of autonomous vehicles as a viable transportation option: Perspectives from people with disabilities and public transit agencies. *Transportation Research Part F: Traffic Psychology and Behaviour, 70*, 260-274. doi:10.1016/j.trf.2020.03.007

Hwang, J., Li, W., Stough, L. M., Lee, C., & Turnbull, K. (2020b). People with disabilities' perceptions of autonomous vehicles as a viable transportation option to improve mobility: An exploratory study using mixed methods. *International journal of sustainable transportation*, 1-19. doi:10.1080/15568318.2020.1833115

Jamson, H. (1999). Curve negotiation in the Leeds Driving Simulator: the role of driver experience. *Engineering in psychology and cognitive ergonomics, 3*, 351-358.

Janotta, F., & Hogreve, J. (2021). *Understanding user acceptance of air taxis-Empirical insights following a flight in virtual reality*. Retrieved from

Jeon, M. (2019). *Multimodal Displays for Take-over in Level 3 Automated Vehicles while Playing a Game*. Paper presented at the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems.

Kaleefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & Winter, J. d. (2020). External Human–Machine Interfaces Can Be Misleading: An Examination of Trust Development and Misuse in a CAVE-Based Pedestrian Simulation Environment. *Human factors*, 0018720820970751.

Kassens-Noor, E., Cai, M., Kotval-Karamchandani, Z., & Decaminada, T. (2021). Autonomous vehicles and mobility for people with special needs. *Transportation Research Part A: Policy and Practice, 150*, 385-397. doi:10.1016/j.tra.2021.06.014

Kelly G. (1955). The repertory test. The psychology of personal constructs. 1. A theory of personality. New York: WW Norton and Company.

Kennedy, C. (2016). New threats to vehicle safety: how cybersecurity policy will shape the future of autonomous vehicles. Mich. Telecomm. & Tech. L. Rev., 23, 343.

Korber, M., Prasch, L., & Bengler, K. (2018). Why Do I Have to Drive Now? Post Hoc Explanations of Takeover Requests. *Hum Factors, 60*(3), 305-323. doi:10.1177/0018720817747730

Krampell, M., Solis-Marcos, I., & Hjalmdahl, M. (2020). Driving automation state-of-mind: Using training to instigate rapid mental model development. *Appl Ergon, 83*, 102986. doi:10.1016/j.apergo.2019.102986

Kyriakidis, M., Happee, R., & de Winter, J. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. Transportation research part F: traffic psychology and Behaviour, 32, 127-140. https://doi.org/10.2139/ssrn.2506579

Large, D. R., Burnett, G., Salanitri, D., Lawson, A., & Box, E. (2019). *A Longitudinal Simulator Study to Explore Drivers' Behaviour in Level 3 Automated Vehicles*. Paper presented at the Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Lavasani M., Jin X., Du Y. (2016). Market Penetration Model for Autonomous Vehicles Based on Previous Technology Adoption Experiences. Presented at 95th Annual Meeting of the Transportation Research. D.C., Washington.

Lehtonen E. Lappi O. Koirikivi I. Summala H. (2014). Effect of driving experience on look-ahead fixations in curve driving. Accident Analysis and Prevention, 70, 195–208.

Li, Y., Dikmen, M., Hussein, T. G., Wang, Y., & Burns, C. (2018). *To Cross or Not to Cross*. Paper presented at the Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Litman, T. 20218. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. Victoria Transport Policy Institute, 28.

Liu, P., Yang, R., & Xu, Z. (2018). How Safe Is Safe Enough for Self-Driving Vehicles? Risk Analysis, 39(2), 315–325. https://doi.org/10.1111/risa.13116

Liu, P., & Xu, Z. (2020). Public attitude toward self-driving vehicles on public roads: Direct experience changed ambivalent people to be more positive. *Technological Forecasting and Social Change, 151*. doi:10.1016/j.techfore.2019.119827

Löcken, A., Golling, C., & Riener, A. (2019). *How Should Automated Vehicles Interact with Pedestrians?* Paper presented at the Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Low, W.-Y., Cao, M., De Vos, J., & Hickman, R. (2020). The journey experience of visually impaired people on public transport in London. *Transport Policy, 97*, 137-148. doi:10.1016/j.tranpol.2020.07.018

Lowe, E., 2017. Driverless cars could transform the lives of disabled people, Huffington Post, 8 January 2017, www.huffingtonpost.co.uk.

Lu, Z., Zhang, B., Feldhütter, A., Happee, R., Martens, M., & De Winter, J. C. F. (2019). Beyond mere take-over requests: The effects of monitoring requests on driver attention, take-over performance, and acceptance.

Transportation Research Part F: Traffic Psychology and Behaviour, 63, 22-37. doi:10.1016/j.trf.2019.03.018

Lundgren, V. M., Andersson, J., Enerbäck, O., & Dolins, S. (2020). *User acceptance of mixed-traffic autonomous shuttles in Gothenburg, Sweden.* Paper presented at the IOP Conference Series: Earth and Environmental Science.

Madigan, R., Louw, T., Wilbrink, M., Schieben, A., & Merat, N. (2017). What influences the decision to use automated public transport? Using UTAUT to understand public acceptance of automated road transport systems. *Transportation Research Part F: Traffic Psychology and Behaviour, 50*, 55-64. doi:10.1016/j.trf.2017.07.007

Marayong, P., Shankar, P., Wei, J., Nguyen, H., Strybel, T. Z., & Battiste, V. (2020). *Urban Air Mobility System Testbed using CAVE Virtual Reality Environment.* Paper presented at the 2020 IEEE Aerospace Conference.

Matthews, M., Chowdhary, G., & Kieson, E. (2017). Intent communication between autonomous vehicles and pedestrians. *arXiv preprint arXiv:1708.07123*.

McGehee, D. V., Mazzae, E. N., & Baldwin, G. S. (2000). *Driver reaction time in crash avoidance research: Validation of a driving simulator study on a test track.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

MeasuringU, 2022. 10 Things to know about the NASA TLX. https://measuringu.com/nasa-tlx/.

Merriman, S. E., Plant, K. L., Revell, K. M. A., & Stanton, N. A. (2021). Challenges for automated vehicle driver training: A thematic analysis from manual and automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 76*, 238-268. doi:10.1016/j.trf.2020.10.011

Miller, E. E., & Boyle, L. N. (2019). Adaptations in attention allocation: Implications for takeover in an automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour, 66*, 101-110. doi:10.1016/j.trf.2019.08.016

Minge, M., Thüring, M., & Wagner, I. (2016). Developing and validating an English version of the meCUE questionnaire for measuring user experience. Proceedings of the Human Factors and Ergonomics Society, 60(1), 2063–2067. https://doi.org/10.1177/1541931213601468.

Mirnig, A. et al., 2019. Driving Simulator Studies at Home: Promises, Potholes, and Pitfalls. s.l., ACM, pp. 221-226.

Mohebbi, R., Gray, R., & Tan, H. Z. (2009). Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone. *Hum Factors, 51*(1), 102-110. doi:10.1177/0018720809333517

Molina, N., Navas, F., González, D., Mahtout, I., & Milanés, V. (2021). Impressions after an automated mobility experience: An acceptance study. *Transportation Research Part F: Traffic Psychology and Behaviour, 81*, 27-40. doi:10.1016/j.trf.2021.05.005

Morales-Alvarez, W., Sipele, O., Léberon, R., Tadjine, H. H., & Olaverri-Monreal, C. (2020). Automated Driving: A Literature Review of the Take over Request in Conditional Automation. *Electronics, 9*(12). doi:10.3390/electronics9122087

Mueller, A. S., Cicchino, J. B., Singer, J., & Jenness, J. W. (2020). Effects of training and display content on Level 2 driving automation interface usability. *Transportation Research Part F: Traffic Psychology and Behaviour, 69*, 61-71. doi:10.1016/j.trf.2019.12.010

NASA, 1986. NASA TLX Manual Version 1.0. https://humansystems.arc.nasa.gov/groups/TLX/downloads/TLX.pdf.

NHTSA, (2008). National Motor Vehicle Crash Causation Survey. National Highway Traffic Safety Administration, U.S. Department of Transportation. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811059

Nordhoff, S., de Winter, J., Payre, W., van Arem, B., & Happee, R. (2019). What impressions do users have after a ride in an automated shuttle? An interview study. *Transportation Research Part F: Traffic Psychology and Behaviour, 63*, 252-269. doi:10.1016/j.trf.2019.04.009

Nordhoff, S., Malmsten, V., van Arem, B., Liu, P., & Happee, R. (2021). A structural equation modeling approach for the acceptance of driverless automated shuttles based on constructs from the Unified Theory of Acceptance and Use of Technology and the Diffusion of Innovation Theory. *Transportation Research Part F: Traffic Psychology and Behaviour, 78*, 58-73. doi:10.1016/j.trf.2021.01.001

Paddeu, D., Parkhurst, G., & Shergold, I. (2020). Passenger comfort and trust on first-time use of a shared autonomous shuttle vehicle. *Transportation Research Part C: Emerging Technologies, 115*. doi:10.1016/j.trc.2020.02.026

Paddeu, D., Tsouros, I., Parkhurst, G., Polydoropoulou, A., & Shergold, I. (2021). A study of users' preferences after a brief exposure in a Shared Autonomous Vehicle (SAV). *Transportation research procedia, 52*, 533-540.

Park, J., & Chowdhury, S. (2018). Investigating the barriers in a typical journey by public transport users with disabilities. *Journal of Transport & Health, 10*, 361-368. doi:10.1016/j.jth.2018.05.008

Paschalidis, E., Hajiseyedjavadi, F., Wei, C., Solernou, A., Jamson, A. H., Merat, N., . . . Boer, E. R. (2020). Deriving metrics of driving comfort for autonomous vehicles: A dynamic latent variable model of speed choice. *Analytic methods in accident research, 28*, 100133.

Patel, R. K., Etminani-Ghasrodashti, R., Kermanshachi, S., Rosenberger, J. M., & Weinreich, D. (2021). *Exploring Preferences towards Integrating the Autonomous Vehicles with the Current Microtransit Services: A Disability Focus Group Study.* Paper presented at the International Conference on Transportation and Development 2021.

Payre, W., Cestac, J., Dang, N.-T., Vienne, F., & Delhomme, P. (2017). Impact of training and in-vehicle task performance on manual control recovery in an automated car. *Transportation Research Part F: Traffic Psychology and Behaviour, 46*, 216-227. doi:10.1016/j.trf.2017.02.001

Penmetsa, P., E. K. Adanu, D. Wood, T. Wang, and S. L. Jones. 2019. "Perceptions and Expectations of Autonomous Vehicles: A Snapshot of Vulnerable Road User Opinion." Technological Forecasting & Social Change 143: 9–13.

Perfect, P., Jump, M., & White, M. (2015). Handling Qualities Requirements for Future Personal Aerial Vehicles. Journal of Guidance, Control, and Dynamics, 38(12), 2386-2398. doi:10.2514/1.G001073

Perfect, P., Jump, M., & White, M. D. (2017). Towards the development of a flight training programme for future personal aerial vehicle users. *CEAS Aeronautical Journal, 8*(3), 541-560. doi:10.1007/s13272-017-0255-2

Persson, D. (2019). Passenger Flight Experience of Urban Air Mobility. In.

Petermeijer, S., Bazilinskyy, P., Bengler, K., & de Winter, J. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Appl Ergon, 62*, 204-215. doi:10.1016/j.apergo.2017.02.023

Petermeijer, S., Doubek, F., & de Winter, J. (2017). *Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants.* Paper presented at the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC).

Piatkowski, D. P. (2021). Autonomous Shuttles: What Do Users Expect and How Will They Use Them? *Journal of Urban Technology*, 1-19. doi:10.1080/10630732.2021.1896345

PytlikZillig, L. M., Duncan, B., Elbaum, S., & Detweiler, C. (2018). A Drone by Any Other Name: Purposes, End-User Trustworthiness, and Framing, but Not Terminology, Affect Public Support for Drones. *IEEE Technology and Society Magazine, 37*(1), 80-91. doi:10.1109/mts.2018.2795121

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How Traffic Situations and Non-Driving Related Tasks Affect the Take-Over Quality in Highly Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58*(1), 2063-2067. doi:10.1177/1541931214581434

Ragbir, N. K., Rice, S., Winter, S. R., Choy, E. C., & Milner, M. N. (2020). How weather, distance, flight time, and geography affect consumer willingness to fly in autonomous air taxis. *The Collegiate Aviation Review International, 38*(1).

Rahimian, P., O'Neal, E. E., Yon, J. P., Franzen, L., Jiang, Y., Plumert, J. M., & Kearney, J. K. (2016). *Using a virtual environment to study the impact of sending traffic alerts to texting pedestrians.* Paper presented at the 2016 IEEE Virtual Reality (VR).

Rahimian, P., O'Neal, E. E., Zhou, S., Plumert, J. M., & Kearney, J. K. (2018). Harnessing Vehicle-to-Pedestrian (V2P) Communication Technology: Sending Traffic Warnings to Texting Pedestrians. *Hum Factors, 60*(6), 833-843. doi:10.1177/0018720818781365

Ramseyer, R., Cimmino, F., Emery, L., Grèzes, S., Grèzes, V., Nanchen, B., Fragnière, E. (2018). *Using phenomenology to assess risk perception of a new technology in public transportation the case of the autonomous vehicles as mobility as a service (MaaS) in Switzerland.* Paper presented at the 2018 3rd International Conference on System Reliability and Safety (ICSRS).

Reddy, L. B., & DeLaurentis, D. (2016). Opinion Survey to Reduce Uncertainty in Public and Stakeholder Perception of Unmanned Aircraft. *Transportation Research Record: Journal of the Transportation Research Board, 2600*(1), 80-93. doi:10.3141/2600-09

Rice, S., Winter, S. R., Mehta, R., & Ragbir, N. K. (2019). What factors predict the type of person who is willing to fly in an autonomous commercial airplane? *Journal of Air Transport Management, 75*, 131-138. doi:10.1016/j.jairtraman.2018.12.008

Riggs, W., & Pande, A. (2021). Gaps and Opportunities in Accessibility Policy for Autonomous Vehicles.

Risto, M., & Martens, M. H. (2014). Driver headway choice: A comparison between driving simulator and real-road driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 25*, 1-9.

Rombaut, E., Feys, M., Vanobberghen, W., De Cauwer, C., & Vanhaverbeke, L. (2020). *Experience and Acceptance of an Autonomous Shuttle in the Brussels Capital Region.* Paper presented at the 2020 Forum on Integrated and Sustainable Transportation Systems (FISTS).

Rudin-Brown, C. M., Williamson, A., & Lenne, M. (2009). Can driving simulation be used to predict changes in real-world crash risk?[Monash University Accident Research Centre-Report 299].

Sadeghian Borojeni, S., Boll, S. C., Heuten, W., Bülthoff, H. H., & Chuang, L. (2018). *Feel the movement: Real motion influences responses to take-over requests in highly automated vehicles.* Paper presented at the Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.

Sahaï, A., Barré, J., & Bueno, M. (2021). Urgent and non-urgent takeovers during conditional automated driving on public roads: The impact of different training programmes. *Transportation Research Part F: Traffic Psychology and Behaviour, 81*, 130-143. doi:10.1016/j.trf.2021.06.001

Salonen, A., & Haavisto, N. (2019). Towards Autonomous Transportation. Passengers' Experiences, Perceptions and Feelings in a Driverless Shuttle Bus in Finland. *Sustainability, 11*(3). doi:10.3390/su11030588

Salonen, A. O. (2018). Passenger's subjective traffic safety, in-vehicle security and emergency management in the driverless shuttle bus in Finland. *Transport Policy, 61*, 106-110. doi:10.1016/j.tranpol.2017.10.011

Scott, J. J., & Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Hum Factors, 50*(2), 264-275. doi:10.1518/001872008X250674

Shaw, E., Large, D., & Burnett, G. (2020). Driver Training for Future Automated Vehicles: Introducing CHAT (CHeck, Assess, Takeover).

She, J., Neuhoff, J., & Yuan, Q. (2021). Shaping Pedestrians' Trust in Autonomous Vehicles: An Effect of Communication Style, Speed Information, and Adaptive Strategy. *Journal of Mechanical Design, 143*(9). doi:10.1115/1.4049866

So, J. J., Park, S., Kim, J., Park, J., Yun, I., & Keyvan-Ekbatani, M. (2021). Investigating the Impacts of Road Traffic Conditions and Driver's Characteristics on Automated Vehicle Takeover Time and Quality Using a Driving Simulator. *Journal of Advanced Transportation, 2021*, 1-13. doi:10.1155/2021/8859553

Sportillo, D., Paljic, A., & Ojeda, L. (2018). Get ready for automated driving using Virtual Reality. *Accid Anal Prev, 118*, 102-113. doi:10.1016/j.aap.2018.06.003

Stadler, S., Cornet, H., & Frenkler, F. (2019). Towards user acceptance of autonomous vehicles: a virtual reality study on human-machine interfaces. *International Journal of Technology Marketing, 13*(3-4), 325-353.

Stojmenova, K., Sanghavi, H., Nadri, C., Ko, S., Tomažič, S., Sodnik, J., & Jeon, M. (2020). Use of spatial sound notifications for takeover requests in semi-autonomous vehicles-a cross-cultural study. In: ICIST.

Summala H., Nieminen T., Punto M. (1996). Maintaining lane position with peripheral vision during in-vehicle tasks Hum. Factors, 38, pp. 442 451, 10.1518/001872096778701944

Tennant, C., Howard, S., Franks, B., & Bauer, M. W. and Stares, S. (2016). Autonomous Vehicles - Negotiating a Place on the Road: A study on how drivers feel about Interacting with Autonomous Vehicles on the road (executive summary). http://www.lse.ac.uk/websitearchive/newsAndMedia/PDF/AVsnegociating-aplace-on-the-road-1110.pdf

Triggs, T. (1996). Some critical human factors issues and challenges in simulation and training. *Simtect 96 Proceedings*, 21-26.

Wang, P., Motamedi, S., Qi, S., Zhou, X., Zhang, T., & Chan, C.-Y. (2021). Pedestrian interaction with automated vehicles at uncontrolled intersections. *Transportation Research Part F: Traffic Psychology and Behaviour, 77*, 10-25. doi:10.1016/j.trf.2020.12.005

Wang, Y., Mehler, B., Reimer, B., Lammers, V., D'Ambrosio, L. A., & Coughlin, J. F. (2010). The validity of driving simulation for assessing differences between in-vehicle informational interfaces: A comparison with field testing. *Ergonomics, 53*(3), 404-420.

White, M.D., Perfect, P., Padfield, G.D., Gubbels, A.W. and Berryman, A.C. 2013. Acceptance testing and commissioning of a flight simulator for rotorcraft simulation fidelity research. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering. 227(4), pp.663–686.

Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. *Accident Analysis & Prevention, 40*(1), 274-288.

Yang, Y., Karakaya, B., Dominioni, G. C., Kawabe, K., & Bengler, K. (2018). *An hmi concept to improve driver's visual behavior and situation awareness in automated vehicle.* Paper presented at the 2018 21st International Conference on Intelligent Transportation Systems (ITSC).

Yeo, D., Kim, G., & Kim, S. (2020). *Toward Immersive Self-Driving Simulations: Reports from a User Study across Six Platforms*. Paper presented at the Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.

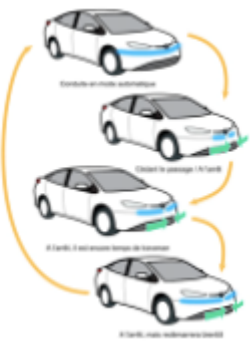
Yoon, S. H., Kim, Y. W., & Ji, Y. G. (2019). The effects of takeover request modalities on highly automated car control transitions. *Accid Anal Prev, 123*, 150-158. doi:10.1016/j.aap.2018.11.018

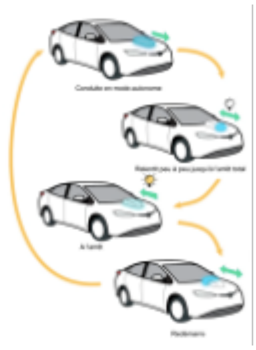
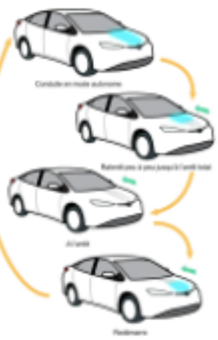
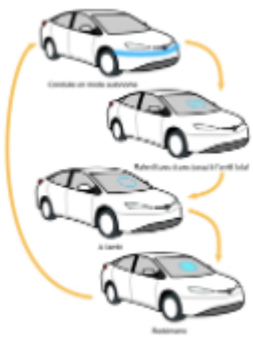
Yun, H., & Yang, J. H. (2020). Multimodal warning design for take-over request in conditionally automated driving. *European Transport Research Review, 12*(1). doi:10.1186/s12544-020-00427-5.

Annexes

Annex 1

eHMI	Illustration	Reference
Anthropo_1		https://www.jaguarlandrover.com/2018/virtual-eyes-have-it
Anthropo_2		Chang, C.-M., Toda, K., Sakamoto, D., & Igarashi, T. (2017). Eyes on a Car: an Interface Design for Communication between an Autonomous Car and a Pedestrian. In Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '17. https://doi.org/10.1145/3122986.3122989
Anthropo_3		https://media.daimler.fr/smart-vision-eq-fortwo-un-concept-car-autonome/
Anthropo_4		Alvarez, W. M., Moreno, F. M., Sipele, O., Smirnov, N., & Olaverri-Monreal, C. (2020). Autonomous driving: Framework for pedestrian intention estimation in a real world scenario. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 39-44). IEEE.
Iconic_1		Clamann, M., Aubert, M., & Cummings, M. L. (2017). Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles (No. 17-02119).

Iconic_2		https://semcon.com/smilingcar/
Iconic_3		https://www.mercedes-benz.com/en/innovation/autonomous/research-vehicle-f-015-luxury-in-motion/
Iconic_4		https://www.mitsubishielectric.com/news/2015/1023.html
Luminous_1		Dey, D., Martens, M., Wang, C., Ros, F., & Terken, J. (2018). Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018 (pp. 82–86). https://doi.org/10.1145/3239092.3265946
Luminous_2		Dey, D., Martens, M., Wang, C., Ros, F., & Terken, J. (2018). Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018 (pp. 82–86). https://doi.org/10.1145/3239092.3265946

Luminous_3		Dey, D., Martens, M., Wang, C., Ros, F., & Terken, J. (2018). Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018 (pp. 82–86). https://doi.org/10.1145/3239092.3265946
Luminous_4		Dey, D., Martens, M., Wang, C., Ros, F., & Terken, J. (2018). Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018 (pp. 82–86). https://doi.org/10.1145/3239092.3265946
Luminous_5		Dey, D., Martens, M., Wang, C., Ros, F., & Terken, J. (2018). Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018 (pp. 82–86). https://doi.org/10.1145/3239092.3265946
Luminous_6		Li, Y., Dikmen, M., Hussein, T. G., Wang, Y., & Burns, C. (2018). To cross or not to cross: urgency-based external warning displays on autonomous vehicles to improve pedestrian crossing safety. In Proceedings of the 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive UI'18), September 23-25 (pp. 188–197). Toronto, Canada.
Textual_1		https://media.daimler.fr/smart-vision-eq-fortwo-un-concept-car-autonome/

Textual_2		Nissan IDS Concept (2015)
Textual_3		https://www.dailymail.co.uk/sciencetech/article-6008793/The-self-driving-car-screens-warn-pedestrians-Drive-ai-launches-standout-cars-Texas.html
Textual_4		https://www.mercedes-benz.com/en/innovation/autonomous/research-vehicle-f-015-luxury-in-motion/
Other_1		https://www.facebook.com/watch/?v=774925076753556 Crédits: SAAQ et Driver's School
Other_2		Dey, D., Martens, M., Wang, C., Ros, F., & Terken, J. (2018). Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In Adjunct Proceedings - 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2018 (pp. 82–86). https://doi.org/10.1145/3239092.3265946
Other_3		

PAsCAL

Enhance driver behaviour & Public Acceptance
of Connected & Autonomous vehicles