

POLITECNICO DI TORINO

Laurea Magistrale in Ingegneria Informatica



Tesi Magistrale

**Comparing different approaches to notify critical events in
the operational driving domain of autonomous vehicles
using virtual reality simulation**

Supervisors

Prof. Fabrizio LAMBERTI

Dr. Filippo Gabriele PRATTICÒ

Candidate

Mattia MANFREDINI

Dicembre 2023

Abstract

The advent of autonomous vehicles (AVs) brings forth a paramount challenge for the automotive industry, encompassing the task of enhancing the overall comfort, enjoyment, and safety of travel. Specifically, in the context of conditional automated driving, the driver has the ability to engage in non-driving related tasks (NDRTs) while the vehicle operates autonomously. However, there are certain situations in which the system may encounter challenges, prompting a take-over request (TOR). In such cases, the driver must regain control of the vehicle until the situation is resolved.

The lack of trust people may have in the technology stands as one of the primary factors that may hinder a widespread adoption of autonomous vehicles. A possible solution to this issue consists of the implementation of suitable interfaces, enabling seamless communication between the driver (the passenger, in perspective) and the vehicle. These interfaces serve as a vital conduit, providing comprehensive information regarding the vehicle's status, intentions, and perceptive understanding of the surrounding environment. However, improving trust in the system goes beyond merely providing information to the occupants. Comprehending how to effectively communicate with them, including what information to provide, how to deliver it, and when to do so, is of utmost importance. This approach is essential to prevent overwhelming the occupants, as an overload of information can lead to excessive cognitive load or unwarranted stress.

This thesis work focuses, as a first step, on prototyping an on-board interface for AVs using virtual reality (VR) based on the current state of the art in the field. The interface consists of different components: an instrument cluster (IC) display that provides general information about the vehicle, as well as a navigation map and a world in miniature showing the vehicle and its surrounding; a touchscreen display that presents the NDRT; an augmented reality (AR) windshield display (WSD) which integrates screen-fixed elements and world-registered AR overlays; an auditory interface that provides, using speech and abstract sounds, explanations and alerts about incoming events. A further step of the thesis focuses on how to effectively convey those critical events that, while still manageable by the autonomous system and not prompting a TOR, may cause a sense of stress and discomfort to the driver (e.g., a jaywalker). Finally, by conducting a VR driving simulation utilizing a motion simulator, the interface is subjected to an evaluation, considering the possible approaches to represent such critical events.

| | | |
|--------|--|----|
| 1. | Introduction | 9 |
| 2. | State of the art..... | 11 |
| 2.1. | Feedback..... | 11 |
| 2.2. | Dashboards..... | 24 |
| 2.3. | HUDs..... | 34 |
| 2.4. | Auditory Interfaces | 47 |
| 2.5. | Mobile Devices..... | 59 |
| 3. | Technologies | 63 |
| 3.1. | Unity3D | 64 |
| 3.1.1. | CiThrus Traffic Simulator and Windridge City | 64 |
| 3.1.2. | NWH Vehicle Physics 2..... | 65 |
| 3.2. | Virtual Reality | 66 |
| 3.2.1. | HTC Vive | 66 |
| 3.3. | Driving Simulator | 67 |
| 3.3.1. | Forza Motorsport Wheel Bundle Fanatec | 68 |
| 3.3.2. | Atomic 3 Motion Simulator..... | 69 |
| 3.3.3. | Actuate Motion | 70 |
| 3.3.4. | Motion Compensation..... | 70 |
| 4. | Design..... | 72 |
| 4.1. | Scenario configuration | 72 |
| 4.1.1. | Graphic Restyle..... | 72 |
| 4.1.2. | Journey and Events..... | 75 |
| 4.2. | Autonomous Vehicle Interface..... | 78 |
| 4.3. | Jaywalking Notification..... | 79 |
| 5. | Implementation | 81 |
| 5.1. | Timeline event system..... | 81 |
| 5.2. | Interface..... | 82 |
| 5.2.1. | Instrument cluster | 82 |
| 5.2.2. | Touchscreen display | 83 |
| 5.2.3. | Windshield display | 84 |
| 5.2.4. | Auditory Interface | 87 |
| 5.3. | Traffic Events..... | 87 |
| 6. | Experiment | 88 |
| 6.1. | Procedure | 88 |
| 6.2. | Metrics..... | 88 |
| 6.3. | Results | 90 |

| | | |
|--------|--|-----|
| 6.3.1. | Trust, Anxiety and Mental Workload | 90 |
| 6.3.2. | Jaywalking notifications..... | 91 |
| 6.3.3. | TORs..... | 95 |
| 6.3.4. | Ranks | 96 |
| 6.3.5. | SSQ, IPQ and UEQ+ | 98 |
| 7. | Conclusions and future works | 98 |
| 7.1. | Conclusions..... | 98 |
| 7.1.1. | Trust, Anxiety, Mental Workload | 98 |
| 7.1.2. | Jaywalking Notifications..... | 99 |
| 7.2. | Future works | 100 |
| | References | 100 |

List of figures

| | |
|--|----|
| Figure 1 - SAE driving automation levels from the official website | 9 |
| Figure 2 - The simulator setup used in [6]..... | 12 |
| Figure 3 - Schematic visualization of scenarios prompting a TOR from [7]. From left to right: GPS scenario, missing-lines scenario, roadwork scenario..... | 13 |
| Figure 4 - Procedure of experiment from [7] | 14 |
| Figure 5 - Experiment setup from [8]..... | 15 |
| Figure 6 - The setup used for the experiment in [2] | 16 |
| Figure 7 - Examples of final HMI design presented to the participants in [2] | 17 |
| Figure 8 - (a) Head-up display (orange dashed outline); (b) Assistive cluster interface (blue solid outline); (c) Touchscreen display (green dotted outline) used in [3] | 19 |
| Figure 9 - Cluster's level of feedback from [3]. (a) No feedback; (b) Sensors feedback; (c) Sensors and..... | 19 |
| Figure 10 - Example of the centre of HUD with AD mode engaged from [10] | 20 |
| Figure 11 - From left to right: the world in miniature, the chauffeur avatar and the car's indicator utilized in [10] | 22 |
| Figure 12 - HMI for system maneuver: announcement, preparation, execution (from left to right) from [12]..... | 25 |
| Figure 13 - HMI for Take-Over scenario: announcement, TOR cautionary, TOR imminent (from left to right) from [12]..... | 25 |
| Figure 14 - Solution approach 1 from [13]..... | 27 |
| Figure 15 - Solution approach 2 from [13]..... | 27 |
| Figure 16 - Variants for solution approach 2 from [13]..... | 28 |
| Figure 17 - Refined HMI concept from [13]..... | 29 |
| Figure 18 - SYMBIOZ demo car from [14]..... | 30 |
| Figure 19 - SYMBIOZ demo car cockpit from [14]. When ambient light and HMI become orange, AD mode is activatedS | 30 |
| Figure 20 - L-Shape HMI of SYMBIOZ demo car examples from [14]..... | 31 |
| Figure 21 - On top, prototyping platform with wizard interface on a smartphone. On bottom, control unit inside the car prototype with lit-up buttons from [15] | 32 |
| Figure 22 - Symbols for accelerating (top left), decelerating (top right), changing lane to the left (bottom left) and changing lane to the right (bottom right) from [16] | 33 |
| Figure 23 - The implemented light pattern, embedded in the driving simulator of [16]. From left to right: accelerating, braking, changing lanes to the left..... | 34 |
| Figure 24 - Overview of HMI elements used in the simulator study from [20]..... | 35 |
| Figure 25 - AR-HUD overview from [20]..... | 36 |
| Figure 26 - Visualization in the HUD from [21] (scenario following a front vehicle) | 37 |
| Figure 27 - Differences in the visualization of the system status and maneuver between HUD (left) and ARHUD (right). From top to bottom: bottleneck caused by parked vehicles, following a front vehicle, turning left at an intersection with oncoming traffic from [21]. | 38 |
| Figure 28 - Distances preferred by participants from [23]. Dark shades indicate a high count of participants. Grey boxes represent the concepts' suggestions. | 40 |

| | |
|--|----|
| Figure 29 - Area preferred by participants from [23]. The darker the shade, the more participants chosen that area. Grey boxes represent the concepts suggestion, except for environmental data, which is distributed in the concept. | 40 |
| Figure 30 - Overview of the information provided by the WSD from [24] | 41 |
| Figure 31 - OMN design on the left, SEL design on the right, from [26]..... | 44 |
| Figure 32 - Timeline of the test scenario with simulated events from [26] | 44 |
| Figure 33 - In-vehicle setup showing the HUD, road scene and participant wearing ETG from [27] | 45 |
| Figure 34 - Experimental conditions showing HUD images superimposed on road scene (top to bottom: CV, AR, LA, LB). From [27] | 46 |
| Figure 35 - Left: a car approaching out of the fog (visual range: 200m) ahead of the driving car. Right: virtual reality study setup using Oculus Rift DK2. From [32]..... | 48 |
| Figure 36 - The avatar in the study [32] is designed as a projection to avoid confusion whether the avatar should represent a human being or a computer system. The avatar gesticulates to increase humanoid appearance..... | 49 |
| Figure 37 - The vehicle setup used in [33] | 50 |
| Figure 38 - Map of driving route and stressor for each of the 15 minutes test run from [33] . | 51 |
| Figure 39 - Lilly vs Julie passenger interactions from [33]..... | 52 |
| Figure 40 - Experimental settings considered in [34]..... | 53 |
| Figure 41 - Simulation setup used in [35][31][36] | 55 |
| Figure 42 - Experimental setup from [37][38]..... | 58 |
| Figure 43 - Setup in [40]. On top the mounted tablet and letterbox. Mobile interface underway displaying the position of the AV on the map. | 60 |
| Figure 44 - Android app featuring a split-screen design from [41]..... | 61 |
| Figure 45 - Experimental setup from [42]..... | 62 |
| Figure 46 - Concepts design employed by [42]..... | 62 |
| Figure 47 - Map centered on the car location (left), option to indicate whether the event has been detected by the user (center), confirmation of the location (right) from [43]..... | 63 |
| Figure 48 - NWH Vehicle Physics 2, Demo Scene | 66 |
| Figure 49 - HTC Vive with accessories | 67 |
| Figure 50 - Forza Motorsport Wheel Bundle Fanatec..... | 69 |
| Figure 51 - Atomic A3 Motion Simulator..... | 70 |
| Figure 52 - Vive tracker used for motion compensation | 71 |
| Figure 53 - Updated road sign assets | 73 |
| Figure 54 - Roadwork scenario | 73 |
| Figure 55 - Car crash event scenario | 74 |
| Figure 56 - Bus stop and updated pedestrian models | 74 |
| Figure 57 - Updated NPC car models | 75 |
| Figure 58 - Traffic events occurring during the simulation | 75 |
| Figure 59 Jaywalking1 event on top left, Freeway enter event on top right, Overtaking1 event on bottom | 76 |
| Figure 60 - Car crash event: TOR on the left, manual mode on the right | 76 |
| Figure 61 - Freeway exit event on the top left, Fog event on the top right, End lane event on the bottom | 77 |
| Figure 62 - Urban event on the top left, Jaywalking2 event on the top right, Crossing1 event on the bottom left, Traffic event on the bottom right | 77 |
| Figure 63 - Suburban event on the left, System failure event on the right | 78 |
| Figure 64 - Roadmap with events..... | 78 |
| Figure 65 - Timeline for the Pedestrian Crossing events..... | 81 |

| | |
|--|----|
| Figure 66 – User vehicle’s dashboard | 82 |
| Figure 67 - Touchscreen display playing a video | 83 |
| Figure 68 - TOR on the touchscreen display | 84 |
| Figure 69 - Windshield display screen fixed information and AR detection box | 84 |
| Figure 70 - Oncoming traffic indicator notifying a car approaching from the right..... | 86 |
| Figure 71 - Experimental Protocol | 89 |

List of tables

| | |
|--|----|
| Table 1 - Events for each AV Explanation condition from [6] | 12 |
| Table 2 - Driving scenarios used in [2] | 17 |
| Table 3 - Description of multimodal feedback system from [9]..... | 21 |
| Table 4 - driving events, referent pictures, and stimulus descriptions in each condition from [11] | 23 |
| Table 5 - Course and scenario from [12]..... | 26 |
| Table 6 - Requirements catalogue from [13]..... | 26 |
| Table 7 - Traffic events from [13]..... | 28 |
| Table 8 - Score for each solution from [13] | 29 |
| Table 9 - An overview of the implemented AR elements with example visualization and descriptions from [24] | 42 |
| Table 10 - Connection between time passed since the beginning of the drive in seconds and visual range at the given moment in meter | 48 |
| Table 11 - Independent and dependent variables from [33] | 51 |
| Table 12 - Driving events in the scenario (left), informative agent's script (middle), and conversational agent's script (right) from [34] | 54 |
| Table 13 - Auditory HMI speech output for system maneuvers as exploited in [35][31][36]. | 55 |
| Table 14 - Criteria and cutoff from [37][38] | 57 |
| Table 15 - Summary of literature review by automation level (SAE) from [45] | 80 |
| Table 16 - Summary of literature review by content and timing from [45] | 80 |
| Table 17 - Trust survey results | 90 |
| Table 18 - Anxiety survey results | 91 |
| Table 19 - Nasa TLX results | 91 |
| Table 20 - Ad-hoc questionnaire results for jaywalking notification..... | 92 |
| Table 21 - Intrusiveness questionnaire results for jaywalking notifications | 94 |
| Table 22 - Ad-hoc questionnaire results for TORs..... | 95 |
| Table 23 - Intrusiveness questionnaire results for TORs..... | 96 |
| Table 24 - Rank results..... | 98 |

1. Introduction

Over the last couple of decades, autonomous vehicles (AVs) have been emerging as the future of transportation, garnering significant investment from industries. Extensive research and studies are being conducted on this topic, recognizing their potential to enhance the comfort, enjoyment, and safety of travel. The Society of Automotive Engineers (SAE) has defined six levels of driving automation, ranging from 1 (basic automation) to 5 (fully autonomous) while level 0 corresponds to completely manual vehicles. Details about each level are shown in the figure below.



Figure 1 - SAE driving automation levels from the official website

The United States Department of Transportation has sanctioned this classification of AVs, making it the established reference standard in both academic and industrial sectors.

From the literature, the concept of trust emerges as one of the primary factors that may hinder the widespread adoption of this technology. One potential approach to address this concern involves integrating effective interfaces, which can relay information about the vehicle's status, intentions and surroundings, ultimately fostering trust and acceptance. Yet, enhancing trust in the system involves more than just providing information to the occupants. Understanding how to communicate with them effectively, what information to present and the appropriate delivery method and timing is crucial. Implementing this approach is essential to prevent overwhelming

the occupants, as an excessive amount of information or an inappropriate feedback method can place an unwarranted cognitive burden or unnecessary stress.

Within the context of SAE Level 3 autonomous driving, also known as conditionally automated driving, once the autonomous mode is engaged, the driver is no longer required to operate the vehicle and can engage in non-driving related tasks (NDRTs). However, there are specific scenarios in which the vehicle is unable to manage the driving task, triggering a take-over request (TOR). This necessitates the driver to regain control of the system until the autonomous mode can be reactivated. While the vehicle operates in autonomous mode without requiring supervision, research indicates that users still desire to receive updates and information from the vehicle about its actions [1] [2]. Specifically, there are certain traffic situations deemed critical due to their potential to cause user concern and anxiety, even though they fall within the operational driving domain of the autonomous vehicle, thereby not triggering a TOR (such as when an abrupt brake is necessary due to a jaywalker or a vehicle in front suddenly applying brakes). Determining the most effective notification approach is paramount. While the primary aim is to boost trust and reassure the occupants, providing too much feedback could have an adverse effect, resulting in undue stress or concern and negatively impacting trust.

Among the interfaces available for autonomous vehicles, the dashboard presents itself as a practical interface for the vehicle to convey information to the user efficiently. Due to the prevalent use of dashboards, users possess a sense of familiarity and trust in this technology, enhancing the provision of precise and user-friendly information. A head-up display (HUD) represents a display technology that superimposes visuals onto the interior of the windshield to relay information to drivers, allowing them to maintain focus on the road. It commonly presents system status, automated operations, and speed. Another category of HUDs utilizes augmented reality (AR) and is known as AR-HUD. These AR-HUDs display information while considering the driving context, integrating digital content and blending it with the real-world components visible through them. An auditory interface engages with the user via audio feedback, employing either speech or abstract sounds. The benefit of such interfaces lies in their ability to convey information without necessitating the driver to divert their visual attention. This auditory feedback can be received regardless of the driver's focal point, as it directly interfaces with the auditory system. Consequently, the driver can immerse themselves in visual tasks like reading or watching without interruption.

After a review of the state of the art of AV interfaces, this thesis work focuses on prototyping an on-board interface for AVs using virtual reality (VR). The interface developed will subsequently undergo evaluation through a VR driving simulation employing a motion simulator. The experiment's core objective is to assess how to notify those critical events that, while still manageable by the autonomous system and not requiring to take over, may have a negative impact on the driver.

2. State of the art

This chapter will dive into a key element of autonomous vehicles: user interfaces. These interfaces play a pivotal role in how the vehicle interacts with the user. By examining the state of the art of vehicle interfaces and the feedback they provide to the user, this chapter will explore how they have the potential to greatly enhance trust and acceptance. Both existing vehicle interfaces and prototyped ones will be analyzed ones, spanning from more conventional dashboard screens to advanced windshield display WSD, while also addressing auditory interfaces.

2.1. Feedback

With the advent of AVs, the concept of NDRTs has emerged, allowing passengers to engage in various activities while the vehicle handles the driving responsibilities.

Nevertheless, prior research has indicated that drivers prefer not to be entirely excluded from the driving process during autonomous driving and providing feedback is a key factor to establish trust in the system [1] [2]. However, more information available does not necessarily lead to more trust and may in fact negatively affect cognitive load and feedback without proper context, abstraction and integration may not be understandable from the driver. [3] [4] [5]. Providing feedback to users without distracting them poses a significant challenge. Within this section, we will undertake a comprehensive analysis of various feedback types, with particular emphasis on analyzing the optimal information to be presented to the user, the timing of delivery for explanations, and the appropriate modality to employ.

In their study, N. Du et al. [6] conducted a within-subject experiment aimed at examining the effects of different conditions on trust, preference, anxiety, and mental workload. In the initial condition, the AV operated without providing any explanations (NExpl) for its actions. In the subsequent conditions, the AV offered explanations either 7s before or after taking action (BExpl and AExpl respectively). Finally, in the last condition, the AV presented an explanation and allowed the driver to approve or disapprove the action after its explanation (PermReq). A total of 22 participants took part in the simulation. The study was conducted in a high-fidelity advanced driving simulator at the University of Michigan Transportation Research Institute (UMTRI) (Figure 2). The UMTRI's fixed-base simulator utilized a dedicated lab space housing a Nissan Versa sedan mockup. It was equipped with Realtime Technology's simulation engine SimCreator, version 2.63, and custom coding to incorporate automated vehicle functionalities. To provide participants with a virtual driving experience, the simulator projected forward road scenes onto three screens. Specifically for this study, the driving simulator's automation features were programmed to replicate the behavior of an SAE level 4 autonomous vehicle. In this mode, the AV assumed control of longitudinal and lateral vehicle movements, navigation, and responses to traffic control devices and other traffic elements, relieving the driver of the need for active monitoring of the environment.



Figure 2 - The simulator setup used in [6]

In each AV explanation condition, participants engaged in a 6- to 8-minute drive without the need to take over control of the vehicle. Each drive contained three unexpected events in the environments of urban, highway and rural: events by other drivers, events by police vehicles, and unexpected re-routes. The events were chosen from previous literature and corresponded to realistic unexpected situations in automated driving. Each event was distinct regarding the surrounding environment, and the order of event type was counterbalanced via a Latin square design across four AV explanation conditions. In the figure below, a table illustrating the events is provided.

| AV Explanation Conditions | Events | Explanations |
|-------------------------------|--|---|
| No explanation (NExpl) | Efficiency Route Change Swerving Vehicle Ahead Stopped Police Vehicle on Shoulder | No explanations |
| Before explanation (BExpl) | Oversized Vehicle Ahead Heavy Traffic Rerouting Police Vehicle Approaching | "Oversized vehicle blocking roadway. Slowing down." "Rerouting, traffic reported ahead." "Emergency vehicle approaching. Stopping." |
| After explanation (AExpl) | Stopped Police Vehicle on Shoulder Abrupt Stopped Truck Ahead Road Hazard Rerouting | "Emergency vehicle on shoulder, changed lanes." "Roadway obstruction, changed lanes." "Rerouted. Identified road obstruction." |
| Permission required (PermReq) | Police Vehicle Approaching Unclear Lane Markings Rerouting Vehicle with Flashing Hazard Lights Ahead | "Emergency vehicle approaching, pull over and stop?" "Unclear lane lines, reroute?" "Vehicle with hazard lights ahead, slow down?" |

Table 1 - Events for each AV Explanation condition from [6]

The dependent variables in this study are participants' subjective attitudes. The attitudinal measures include trust, mental workload, anxiety and preference. Results reveal that the AV explanation conditions had a significant main effect on trust ratings in autonomous vehicles. Post hoc comparisons indicated that participants consistently demonstrated higher levels of trust in the AV when the explanation was provided before the action. To be more specific, participants gave a higher trust ranking when they were presented with the BExpl and PermReq than NExpl and AExpl. Nonetheless, there was no significant difference of trust ranking between PermReq and BExpl. The main effect of AV explanation conditions on driver preference was significant; post hoc analysis indicated that participants preferred BExpl to NExpl and to PermReq. The average scores of mental workload survey items were not significantly different among four AV explanation conditions. However, the mental workload in BExpl had the lowest mean. Lastly, no significant differences among four AV explanation conditions were found in driver's anxiety toward the AV.

The overall mean trust of both explanation conditions was not significantly different from the NExpl condition. Given these results, the authors conclude that just providing an explanation about the vehicle's actions is not sufficient for increasing trust.

M. Körber et al. [7] investigated whether providing an explanation for a TOR in AD influences trust in automation and acceptance. Authors conducted an online pre-study to evaluate the comprehensibility and distinguishability of the chosen takeover situations based on their levels of obviousness. A total of 36 participants took part in the study, where they watched videos depicting three different takeover scenarios. These scenarios varied in their levels of obviousness: one involved missing GPS data (referred to as "GPS" with low obviousness), another depicted the absence of lane markings ("missing lines" with medium obviousness), and the third scenario portrayed road work situations (with high obviousness). No participant could name the correct reason for the system limit for the GPS scenario, 35% answered correctly for the missing-lines scenario, and 78% could name the correct reason in the roadwork scenario. In the main study, the authors employed a 2×3 mixed design. The factor "Explanation" was manipulated between subjects and comprised two groups: a control group and an experimental group ("explanation"). Participants were equally and randomly assigned to either group. The experimental group received an explanation regarding the reason for the TOR after each takeover situation, whereas the control group did not receive any explanations. The provided explanations encompassed both the external reasons for the TOR and the internal implications for the system. Each explanation followed a consistent structure and wording, with the only variation being the specific cause and effect described. The standard format used was as follows: "The take-over request was triggered because of [cause]. Due to [effect], driving in highly automated mode can temporarily not be continued". A schematic visualization of the scenarios prompting a TOR is shown in the figure below.

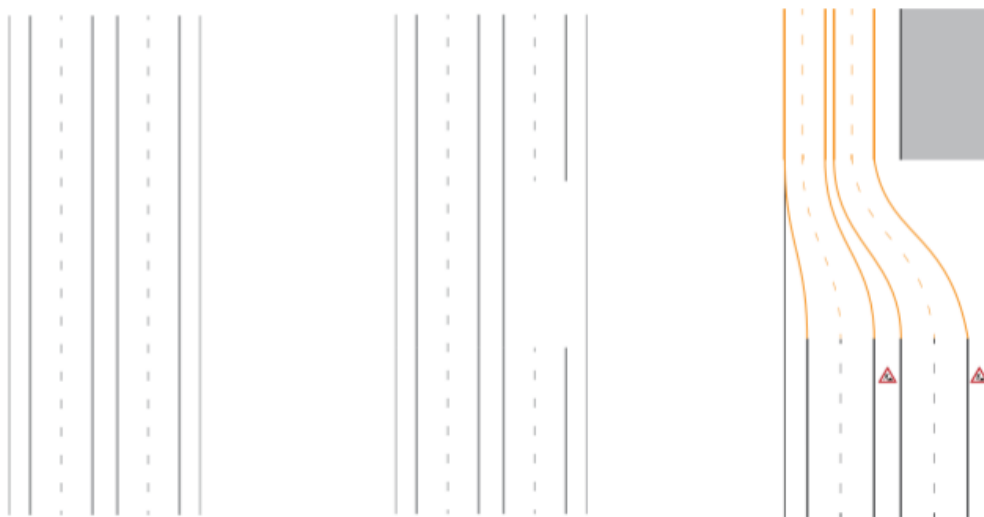


Figure 3 - Schematic visualization of scenarios prompting a TOR from [7]. From left to right: GPS scenario, missing-lines scenario, roadwork scenario.

A total of $n = 40$ participants, 20 (50%) females and 20 (50%) males, took part in the simulation. The study was conducted in a static driving simulator equipped with a BMW 6 Series mock-up. Seven projectors provided a front view of approximately 180° , side and rear

mirrors, and a mock-up of a HUD. The implemented driving automation performed on SAE Level 3, conditional automation. The participants were instructed to give their attention to the NDRT whenever it was presented. The automation could be enabled or disabled by using a button located on the steering wheel, and it would also deactivate if the participants applied steering or braking inputs. Clear instructions were provided to the participants, emphasizing the importance of activating the automation whenever it became available. The status of the automation was visibly indicated by an icon positioned in the top center of the instrument cluster. The experiment procedure is described in the figure below.

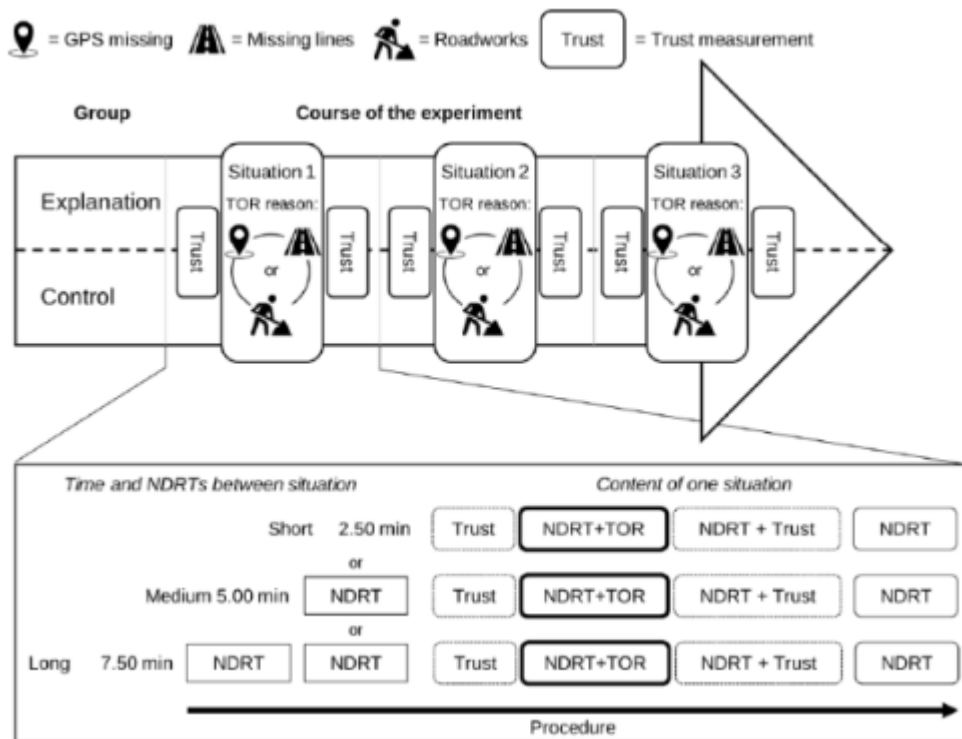


Figure 4 - Procedure of experiment from [7]

The participants were greeted by the experimenter and provided with instructions. They then completed a questionnaire regarding their demographic information. Following that, the participants engaged in a familiarization drive and practiced the NDRT. Afterward, the participants filled out the van der Laan questionnaire for the first time. The experimental drive commenced, and upon completion, the same questionnaire was filled out for the second time. The participants were then interviewed to gather their feedback on the scenario experience. Finally, a debriefing session was conducted, and the participants received their participation reward.

The results reveal that all participants displayed a high level of acceptance before and after the drive, regardless of the experimental condition. The control group's trust ratings remained unchanged in response to takeover requests across all situations. However, the experimental group exhibited decreased trust specifically after experiencing a TOR triggered by roadwork. Participants who received explanations for TOR reported a stronger sense of understanding regarding the system and the reasons behind the TORs.

In conclusion, the study findings indicate that a TOR did not result in decreased trust or acceptance among participants. Furthermore, providing explanations for TORs did not significantly impact trust or acceptance levels. However, the provision of explanations did enhance participants' perceived understanding of the system. These results suggest the importance of transparent communication and the potential benefits of explanations in fostering user understanding in automated driving systems.

Using a driving simulator, J. Koo et al. [8] tested different messages that provided advance explanation of the car's imminent autonomous action. The research explored two types of feedforward information: how the car is acting before a sudden brake (how) and why the car is acting that way (why), as well as a combination of the two pieces of information (how+why). The "how" message portrayed the action as "Car is braking," the "why" message revealed the cause as "Obstacle ahead," and the combined "how+why" message encapsulated the information as "Car is braking due to the presence of an obstacle ahead". A total of 64 students (32 female, 32 male) with a driving license took part in the driving simulation STISIM. The simulator physically consisted of a half-cut modified Ford Mustang equipped with a gas pedal and brake, a force-feedback steering wheel, and a driver's seat. During the experiment, the simulated driving scene was run on lab computers and projected onto three rear-projection screens with a total of 160° field of view (FOV) (Figure 5).

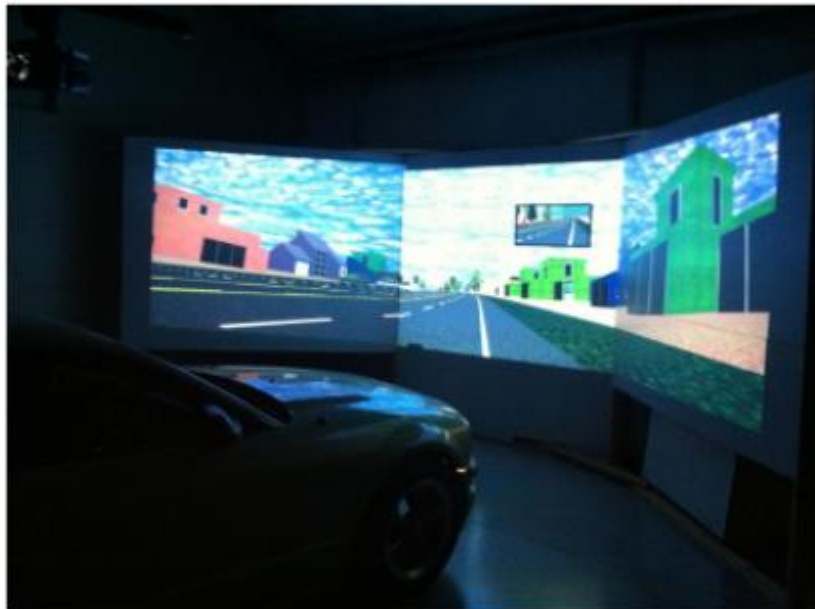


Figure 5 - Experiment setup from [8]

For the sake of this experiment, the authors constructed and programmed an automatic braking feature within the simulator, specifically designed to engage in automatic braking when faced with imminent collision scenarios, such as obstacles ahead or jaywalkers. The 12-kilometer driving route encompassed urban, suburban, and highway segments, integrating stop signs and traffic signals. The course incorporated multiple challenges, diverse traffic patterns, captivating

environmental landscapes, and dynamic driving conditions to accurately simulate a realistic and demanding driving encounter. Upon the conclusion of the driving experiment, which was preceded by a brief 5-minute training session, participants were asked to complete an online questionnaire. The questionnaire aimed to evaluate their overall driving experience during the experiment as well as their responses and handling of the warning system. The results showed that the “how” message alone leads to dangerous driving performance and the authors believe that the lack of situational awareness, which gives the driver a sense of uncertainty may be the reason behind these findings. About the “why” message alone, results show that it created the least anxiety and highest trust. The duration required to perceive and cognitively process the why-only message is approximately equivalent to the time needed to process the how-only message. However, the enhanced performance and positive attitude observed under this condition indicate that the “why” message holds greater significance and is processed more rapidly by individuals. The “how+why” message was felt more negatively by the driver, with an increase of anxiety but it led to safer driving performance. The authors conclude that both the “how” and “why” information are needed for critical safety situations, although including both information can incur a negative emotional response it improves safety performance. Alternatively, the “why” message provides a moderate amount of information without causing a negative emotional reaction and it might be optimal when the car knows it is a non-safety critical situation.

The study conducted by R. Ma et al. [1] aimed at examining the impact of visual feedback as the sole factor in influencing a driver's trust in AVs as well as to determine the optimal level of feedback required to evoke the appropriate level of trust. To explore the acceptance of AVs among users, a static driving simulator was employed to simulate various real-world driving scenarios using video presentations. Participants were instructed to imagine themselves in an autonomous vehicle while being presented with different driving scenarios. The study adopted a between-group design, assigning participants to three groups distinguished by the level of visual feedback provided. These groups consisted of no visual feedback, moderate visual feedback, and high visual feedback. After each video presentation, participants were given a trust questionnaire to gauge their trust perceptions towards the autonomous vehicle depicted in the preceding video. The figure below illustrates the setup for the simulation.



Figure 6 - The setup used for the experiment in [2]

A total of 30 participants participated in the experiment, with equal weighting in males and females. This research utilized a static driving simulator located at the Loughborough University Design School. The driving scenarios were projected onto a wall positioned

approximately 3 meters in front of the driver's seat. A total of ten videos, each lasting around 2 minutes, depicting various driving scenarios were utilized. In the table below a description of the driving scenarios is provided.

| Video No. | Driving Scenarios | Description |
|-----------|---------------------------------|---|
| 1 | Exiting Business Park | Vehicle leaving a business park; speed limit 20mph; pedestrian crossing on-road; cyclist presence; roundabouts; dry road condition; minimal traffic. |
| 2 | Dual Carriageway | Busy suburban roundabout junction; dual carriageway; speed limit 40 mph; speed cameras, dry road condition. |
| 3 | Small village | Single carriageway; small village; speed limit 30mph; narrow bridge, dry road condition. |
| 4 | Town Centre | Town centre; speed limit 20mph; busy roads; roundabout; high pedestrian level; traffic lights, dry road condition. |
| 5 | Single Carriageway | Single carriageway; speed limit 50 mph; roundabouts; low level of traffic, dry road condition. |
| 6 | Moderate Motorway Traffic | Entering motorway; speed limit 70mph; dry road condition, moderate traffic; exiting motorway at next junction. |
| 7 | Residential Area | Residential area; speed limit 30mph; dry road condition, vehicles parked in road; required vehicle to pull out into opposite lane. |
| 8 | Residential Area: Blind Corners | Residential area; speed limit 30mph; dry road condition, vehicles parked in road; on-coming vehicle wanting to pull out; two blind corners; roundabout. |
| 9 | Merging onto Motorway | Merging onto motorway; speed limit 70mph; high traffic level, dry road condition. |
| 10 | High Motorway Traffic | Motorway; speed limit 70mph; dry road condition, other vehicles merging onto motorway; high traffic level; exiting motorway at next junction. |

Table 2 - Driving scenarios used in [2]

These videos were pre-recorded using a dashboard camera and carefully selected to encompass a diverse range of scenarios. To establish an immersive autonomous experience, only road and engine sounds were audible in the videos, creating an atmosphere aligned with autonomous driving. The researchers developed the human-machine interaction (HMI) displays in accordance with the driving scenario videos. The primary objective was to create a straightforward and low-fidelity set of HMI displays, as this approach is less time-intensive and does not necessitate specialized skills and extensive resources.

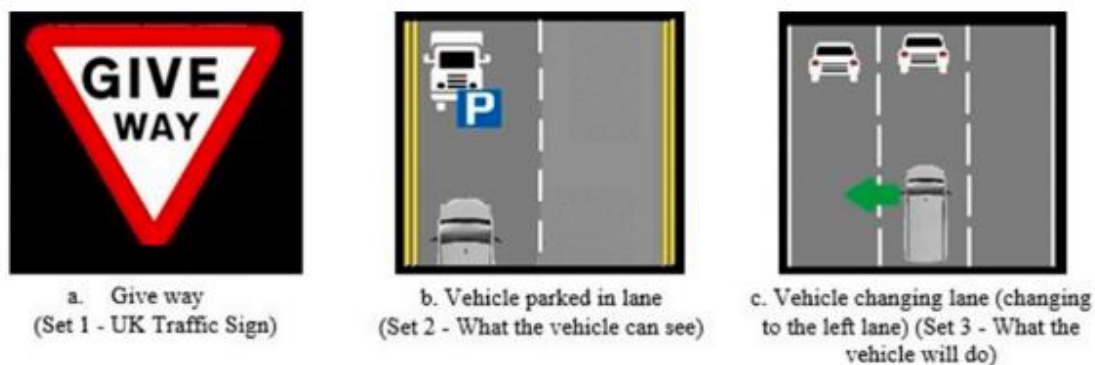


Figure 7 - Examples of final HMI design presented to the participants in [2]

In the study, three groups were established with different levels of feedback provided through HMI displays. The first group ("No feedback") did not receive any HMI display, in the second group ("Moderate feedback") participants were presented with HMI displays showing UK traffic signs and information about what the vehicle can perceive. Lastly in the third group ("High feedback") participants were provided with HMI displays showing UK traffic signs, information about what the vehicle can perceive, and details about what the vehicle is expected to do. The findings of the study suggest that the way the vehicle communicates with the driver regarding its future actions significantly impacts the level of trust drivers place in the autonomous driving (AD) system. The results highlight that participant in the high feedback group exhibited greater trust in the AD systems compared to those in the moderate feedback and no feedback groups. However, these findings appear to contradict the results of Koo et al.'s study [8], which proposed that displaying "why" messages explaining the reasons behind the vehicle's responses (similar to the moderate feedback in this study) resulted in higher levels of trust compared to displaying "how" messages detailing the vehicle's planned response (e.g., "Car is braking"), and when both "how and why" messages were presented (equivalent to high feedback in this study). One possible explanation for the contrasting outcomes could be the nature of the information provided; in Koo et al.'s study [8], verbalized messages were utilized instead of visual information. Furthermore, the results indicate that there is no notable distinction between the no feedback group and the moderate feedback group. One potential explanation for this could be the increased workload imposed on the driver, as participants in the moderate feedback group are essentially compelled to verify the driving scenarios by cross-referencing them with the displayed information.

In their experiment, A. Mackay et al. [3] aimed to evaluate drivers trust in autonomous systems depending on the feedback the vehicle provided by an assistive cluster's interface. Forty participants with a driver's license were recruited to take part in this experiment, which was conducted in a fixed-base driving simulator mockup (DSM), composed of two seats, a steering wheel, two pedals, and three monitors for rear view projection. A HUD was installed, and the assistive cluster was mounted on the right side of the steering wheel. A touchscreen display on the lower dashboard was placed within reach to prompt the visual search task (Figure 8).



Figure 8 - (a) Head-up display (orange dashed outline); (b) Assistive cluster interface (blue solid outline); (c) Touchscreen display (green dotted outline) used in [3]

Upon arrival, the participants were invited to partake in an AD experience and were provided with an informed consent form, which they read and signed. Subsequently, they were randomly allocated to one of three groups: the “No Feedback Group” (consisting of 13 participants), the “Sensors Group” (consisting of 14 participants), or the “Decision Group” (consisting of 13 participants). In the No Feedback Group, participants experienced a cluster without any information displayed. The Sensors Group received feedback regarding the proximity of surrounding vehicles, indicated by the lateral and longitudinal control lines turning yellow as other objects approached, such as during a queue or overtaking situation. The Decision Group, in addition to the sensor information provided to the Sensors Group, also received arrows indicating the vehicle's immediate future behavior concerning lane changes.



Figure 9 - Cluster's level of feedback from [3]. (a) No feedback; (b) Sensors feedback; (c) Sensors and

Each cluster in the display featured a central illustration of the vehicle and a corresponding automation bar below it, representing the different driving modes across four autonomy levels. In this study, all participants were presented with the AD mode, specifically level 4, indicated by a fully green automation bar. The HUD included a central speedometer and an AD icon (Figure 10).



Figure 10 - Example of the centre of HUD with AD mode engaged from [10]

The experimental scenario consisted of a 12-minute drive along a highway, encompassing various events such as a traffic segment with no congestion, an overtaking maneuver, and a traffic queue that necessitated the system to apply brakes. While in AD mode, participants were unexpectedly prompted with a visual search task at two specific time points: 3 minutes and 6 minutes after the start of the experiment (referred to as t1 and t2, respectively). In the search arrows task, the participant had to search for an upward facing arrow in a grid of equal but differently oriented arrows by pressing a “Yes” button (if present) or “No” (if absent). Following the completion of the drive, the car stopped at a gas station for recharging, and participants were asked to fill out a trust questionnaire administered by the experimenter. Heart rate was measured as an objective metric. The findings of the study do not clearly indicate a direct correlation between the available feedback information and the level of trust reported in the questionnaire. This could be attributed to the specific design of the interface, which only provided limited information regarding the location of surrounding vehicles and the vehicle's intended actions. It is possible that more explicit and comprehensive information is required to significantly impact trust. Additionally, the results revealed an inverse relationship between the amount of available feedback and performance on a visual search task. Hence, it appears that the assistive cluster with higher complexity induced a higher cognitive load, resulting in poorer task performance. In conclusion, it is important to note that more information does not necessarily lead to higher levels of trust and may even negatively impact cognitive load. These findings highlight the need to explore which types of feedback are more suitable and how specific design choices for the visual HMI can influence trust and cognitive load. Furthermore, alternative methods of conveying information to users should be investigated. While this experiment focused on visual feedback, it is worthwhile to explore the utilization of other sensory modalities, either in combination with visual cues or independently.

Using a driving simulator, S. Lee et al. [9] preliminary investigated the explanatory effects of multimodal feedback compared to the current feedback system in take-over transition. To address the comprehensive aspects of user experience concerning multimodal feedback, the authors collect and compare user feedback encompassing take-over response time, as well as

their subjective evaluations of annoyance and urgency levels throughout the experimental process. The control group was designed to align with the interface of Hyundai IONIQ, Tesla Model S, Volvo, and other recent vehicles from different manufacturers. On the other hand, the experimental group was developed considering several pertinent studies on in-vehicle multimodal feedback, aiming to provide a more advanced and context-specific interface compared to the current standard. The modalities are categorized into visual and auditory interfaces, where visual interfaces consist of HUD, cluster, and ambient lighting, and auditory interfaces encompass speech and non-speech audio. The ambient light feedback was devised to alert drivers about the driving situation using three distinct colors: white indicating manual driving status, green denoting auto-driving status, and red signifying urgent warning situations. The cluster interface presented textual information regarding each scenario, incorporating essential keywords to provide drivers with pertinent details. The HUD exhibited the driving status of the autonomous vehicle through context-compatible icons, conveying crucial information to the driver. The audio feedback, in the form of speech, delivered information on the vehicle's behavior and the reasons behind its execution. The non-speech audio feedback included two types of sounds: one more annoying and the other less annoying, designed to provide additional auditory cues to the driver. The table below describes the multimodal feedback system.







| Stimulus Interface | Type | Scenario | | | | |
|--------------------|------------------|--|--|---|--|---|
| | | Takeover transition suggestion | Stable autonomous driving | Short-term | Mid-term | Long-term |
| Visual | Ambient light | White (Manual driving) → Green | Green (Autonomous driving) | Red (Alert), blinking 1.125 Hz | Red (Alert), blinking 1.125 Hz | Red (Alert), blinking 1.125 Hz |
| | Cluster | "Autonomous driving possible, switch by pressing a button" | "Safe autonomous driving" | "Unexpected vehicle detected, continue driving after deceleration" | "End of autonomous driving, switch to manual by pressing the button" | "End of autonomous driving, switch to manual by pressing the button" |
| | HUD | Vehicle status  | Vehicle status   |  |  |  |
| Auditory | Speech | "Safe autonomous driving is possible. Can be switched to autonomous driving mode" | "Safe autonomous driving. Maintaining autonomous driving mode" | "Detection of an obstacle that appeared suddenly and rapid deceleration. Maintaining self-driving mode" | "Autonomous driving will be terminated due to forward construction. Switch to manual operation mode" | "Autonomous driving will be terminated due to an exit. Switch to manual operation mode" |
| | Non-speech audio | Less annoying sound | Less annoying sound | More annoying sound (1300 Hz beeps repeated 8 times per second) | More annoying sound (1300 Hz beeps repeated 8 times per second) | More annoying sound (1300 Hz beeps repeated 8 times per second) |

Table 3 - Description of multimodal feedback system from [9]

The experimental findings revealed that, in most scenarios, the reaction time decreased, while the perceived annoyance and urgency increased. The results of the experiment suggest that the inclusion of multimodal feedback in the AD system enhances the acceptance of TORs through explanations. In conclusion, although this study holds theoretical and practical implications for the design of vehicle-user interfaces, it is essential to acknowledge its limitations for future

investigations. The primary limitation arises from the small sample size of ten participants, which was constrained by time and resource limitations. Additionally, the experimental results may have been influenced by the participants' proficiency in manual driving operation. Addressing these limitations in future studies can contribute to a more comprehensive understanding of the subject matter.

R. Haeuslschmid et al. [10] argue that establishing trust in AD can be achieved through a driver interface that visually demonstrates the car's understanding of the current context and its corresponding behavioral choices. To confirm this hypothesis, they conducted a user study comparing various visualizations superimposed onto a driving scene: a chauffeur avatar, (2) a world in miniature, and (3) a display of the car's indicators as the baseline. The figure below illustrates the three visualizations.



Figure 11 - From left to right: the world in miniature, the chauffeur avatar and the car's indicator utilized in [10]

The world in miniature is a computer vision-inspired visualization that effectively portrays the car's perception of the environment, its interpretation of the surroundings, and its actions in a clean and proficient manner. The anthropomorphic visualization, represented by the chauffeur avatar, responds to the same events as the world in miniature but exhibits a greater degree of anthropomorphism, potentially evoking more emotions. As a baseline, basic visualization of the car's indicators has been used instead of no visualization.

Based on the related works, the authors assumed that the world in miniature would enhance the user's trust by effectively showcasing its competence. They also hypothesized that a human-like visualization, such as the chauffeur avatar, would increase trust more than a computer vision-style visualization.

D. McKeown and S. Isherwood [11] investigated which type of audio feedback between speech, environmental sounds, auditory icons, and abstract synthetic warnings are better candidates for within vehicle interfaces with best performance in terms of both accuracy and response time. A panel of experienced drivers initially assessed the situational urgency of nine driving events. Subsequently, the candidate auditory displays, including speech, auditory icons, nonspecific environmental sounds, and abstract sounds, were paired with the driving events, with variations in the level of priority assigned to each combination. Three distinct levels of priority were established, encompassing different degrees of urgency and action required. These levels included simple advisories, which provided general guidance; advice indicating

the need for vigilance or a specific action; and high priority warnings demanding immediate action or the utmost level of vigilance. The table below illustrates for each driving event the audio feedback provided in each condition.










| Driving Events | Pictures of Referent Events | Abstract Sounds | Auditory Icons | Nonspecific Environmental Sounds | Speech Messages |
|---|---|--|------------------------------|----------------------------------|-------------------------|
| Petrol level is low |  | Single bell ding | Water pouring | Footsteps | "Petrol is low" |
| Oil level is low |  | Low-rate tapping | Steam and water sounds | Housemartin song | "Oil is low" |
| Tire pressure is low |  | Low-rate, low-pitched warbling tone | Air release blast | Seashore lapping | "Tire pressure is low" |
| Driver door is open |  | Moderate-rate siren | Car door shutting | Cockerel | "Driver door is open" |
| Speed limit is being exceeded |  | Moderate-rate fire bell | Car speeding past | Baby sneeze | "Exceeding speed limit" |
| Hand brake is on while driving |  | Moderate-rate tone alarm | Squeaking sound | Tram passing | "Hand brake is on" |
| A car is driving in the blind spot |  | High-rate, high-pitched warbling tone | Car horn blasts | Electric warble factory sound | "Car in blind spot" |
| The car is drifting off the road |  | High-rate, high-pitched tone alarm | Driving over "rumble strips" | Small glass smashed | "Drifting off road" |
| Headway to the vehicle in front is closing fast |  | High-rate, high-pitched zapping pulses | Car crashing | Textile factory loom | "Headway closing fast" |

Table 4 - driving events, referent pictures, and stimulus descriptions in each condition from [11]

Subsequently, participants underwent testing to assess the accuracy of their learning and their response times in associating the sounds with their corresponding visual representations in a computer-based task. During the computer-based task, a central red circle with a diameter of 15 mm was displayed at the onset of every trial. To initiate the task, participants were required to move the mouse cursor to the center circle, triggering the appearance of nine pictures on the screen along with the simultaneous onset of a sound. The response time of participants was recorded from this moment onward. Each trial continued until the participant clicked on one of the nine pictures. Importantly, the order of the pictures was randomized for each trial, adding variability to the task presentation. Additionally, participants provided evaluations of the perceived urgency and pleasantness of the auditory displays both prior to ("session 1") and following their acquisition of knowledge regarding the associations between the sounds and the specific events they represented ("session 2"). The findings indicated that speech and auditory icons yielded the highest response speed and accuracy in the computer picture-matching task. In contrast, abstract sounds resulted in slower response times and lower

accuracy. Environmental sounds, despite their arbitrary associations with specific referents (although potentially more distinct than synthetic sounds), displayed an intermediate level of accuracy. However, their response times were comparable to those of abstract sounds. Moreover, a magnitude estimation measure devoid of a scale was employed to evaluate both the pleasantness and perceived urgency. The analysis revealed that speech in isolation exhibited a lack of sensitivity to both measures. The utterances consistently received moderate ratings for urgency and were consistently perceived as pleasant. Conversely, the remaining three sound categories (auditory icons, environmental sounds, and abstract sounds) exhibited a consistent correlation. Specifically, sounds associated with highly urgent events were consistently regarded as the least pleasant. The findings support the notion that urgency is primarily influenced by acoustic characteristics and tends to remain stable over time. Generally, changes in urgency ratings were not observed across sessions, except for the zapping pulses sound, which was perceived as more urgent after learning the association with "headway closing fast." In contrast, there was evidence of changes in pleasantness following learning. Specifically, two sounds from the environmental sounds group and two from the auditory icons group were rated as significantly less pleasant after learning. In the abstract sounds group, significant differences in pleasantness emerged in Session 2 between sounds associated with low and high situational urgency. Speech utterances, however, did not show changes in pleasantness. These results suggest that while urgency is predominantly governed by acoustic properties, affective properties such as pleasantness can be more flexible and responsive to associations.

2.2. Dashboards

The dashboard emerges as a viable interface through which the vehicle can effectively transmit information to the user. Given the widespread utilization of dashboards, users have developed a sense of familiarity and trust in this technology, which facilitates the delivery of accurate and comprehensible information.

This section will provide an overview of various potential or existing dashboards, delineating the information they present, and the methods employed for their presentation.

In Conditional Automated Driving (CAD), as previously shown, the user might need to take control of the vehicle. To establish a safe and seamless interaction between the driver and the system, fostering acceptance, Y. Forster et al. [12] designed and evaluated an HMI for CAD. During normal operations, the visual interface provides the "main state", which shows that lateral vehicle control and longitudinal are executed by the autonomous system by turning lane symbols from grey to blue and with a blue rectangle in front of the host vehicle respectively. Below the main state, the set speed and the current speed are displayed. When a traffic event demands speed adjustment, such as an impending speed limit change or a road curvature, the driver is alerted beforehand through a message box on the top of the main state, with a symbolic representation of the traffic event and a distance from that event. All the HMI components will be displayed in blue color if the adaption can be carried by the system. The representation of a maneuver consists of three steps. In the announcement, the type of event and the remaining distance are displayed. In the preparation, the specific type of maneuver is shown (e.g. lane

change) together with a text message before the maneuver is executed. In the execution, the maneuver is communicated with a text message (Figure 12). The arrow indicating the maneuver undergoes a color transition, shifting from turquoise (during the preparation phase) to blue.

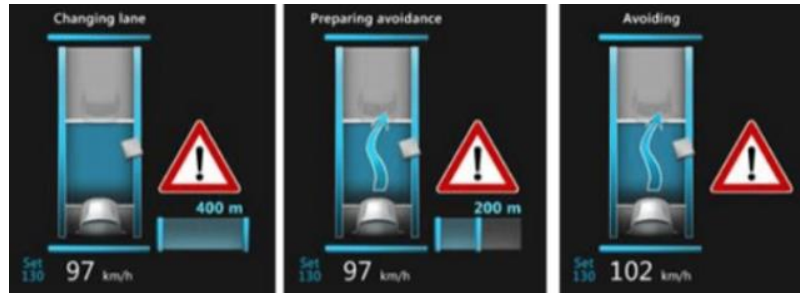


Figure 12 - HMI for system maneuver: announcement, preparation, execution (from left to right) from [12]

If the automated system is unable to handle a traffic event, a three leveled TOR is displayed, consisting of an announcement stage, a so-called soft TOR, and a so-called hard TOR. The figure below describes the HMI for take-over scenario.

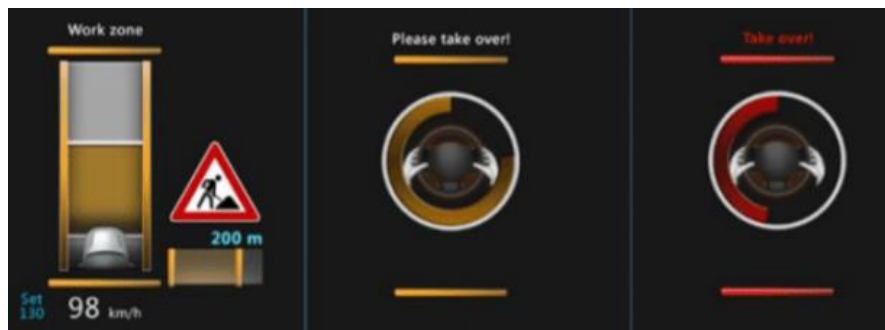


Figure 13 - HMI for Take-Over scenario: announcement, TOR cautionary, TOR imminent (from left to right) from [12]

To evaluate the HMI, a group of six experts in cognitive ergonomics took part in a static simulation that consisted of a ride on a three-lane highway, encompassing seven specific situations that required driver-system interaction.

| Situation | Description | Interaction |
|----------------------|--|--------------------|
| End of Lane Markings | Lane markings end abruptly | TOR |
| Lane Splitting | Lane Change to the right lane | System maneuver |
| Obstacle | Avoiding of lost cargo on the right roadside, | System maneuver |
| Speed Limit | Reduction of speed due to speed limit sign | System maneuver |
| Secondary Lanes | Yellow Secondary lane markings emerge | TOR |
| Curvature | High curvature. Adaptation of speed | System maneuver |
| Construction site | Construction site with lane offsets to the right | TOR |

Table 5 - Course and scenario from [12]

Afterward, participants were exposed to both quantitative and qualitative measures to evaluate the usability and pinpoint the weaknesses of the HMI's visual elements. Results showed good usability and comprehensibility of the interface.

To reduce the anxiety of passengers inside the AV, Stadler et al. [13] proposed several design concepts and variants for HMI. The HMI concepts and variants were evaluated using a simulation in VR, as well as questionnaires and interviews. The study pivots around autonomous buses in public transport and was divided into three stages: the design of HMI concepts and variants, their evaluation, and the selection and refinement of one of the HMI concepts. By conducting interviews and conducting literature research, a requirements catalog was formulated as a foundation for generating HMI concepts and it's shown in the table below.

| Must have | Should have | Could have |
|-------------------|--------------------|------------------------|
| Showing intention | Showing AV status | Information redundancy |
| Showing detection | Showing route | Notifications |

Table 6 - Requirements catalogue from [13]

Based on the requirements catalog, two solution approaches were generated:

In solution approach 1 (Figure 1), an LED stripe is installed at the bottom of the front windshield. This stripe utilizes light indications to communicate the autonomous vehicle's intentions and direction. Furthermore, it employs small light indications to visualize other road

users. In solution approach 2 (Figure 15), a semi-transparent screen-like surface is utilized to present various information during the ride, such as the intentions of the AV, the detection of obstacles, other road users, and the surrounding environment of the vehicle.

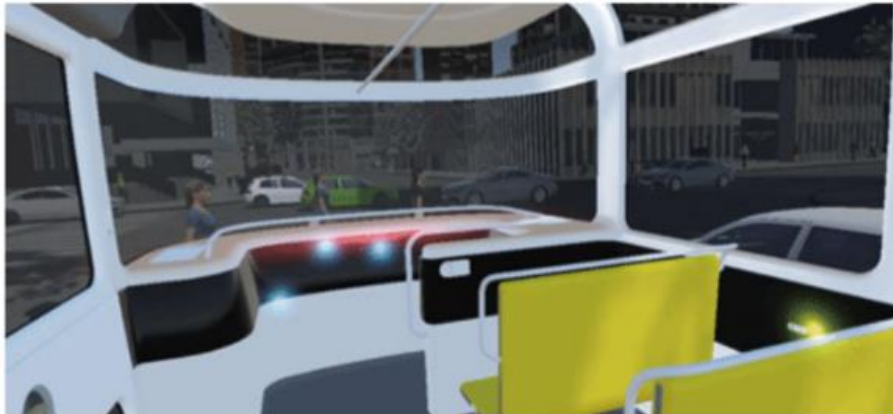


Figure 14 - Solution approach 1 from [13]

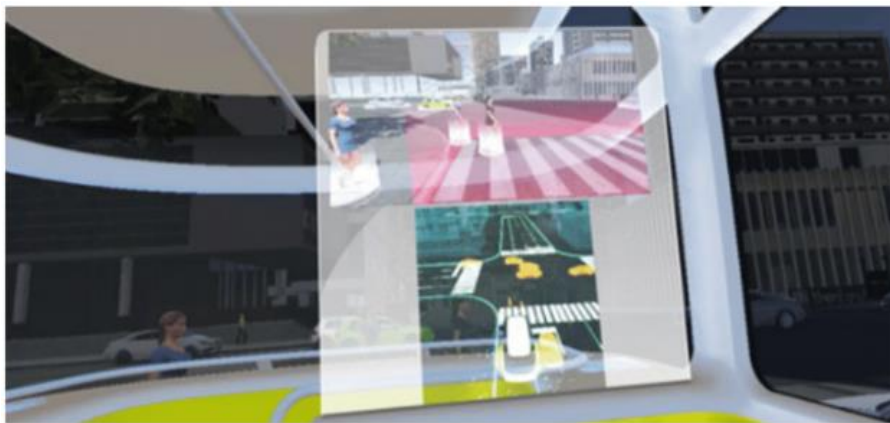


Figure 15 - Solution approach 2 from [13]

Additionally, four variations were developed for solution approach 2.

Variant A utilizes the cameras of the autonomous vehicle to portray its detected immediate surroundings, whereas variant B presents an abstracted two-dimensional map illustrating the vehicle and its nearby surroundings. In variant C, a three-dimensional bird perspective showcases the AV and an abstracted map. Lastly, variant D combines features from variants A and C (Figure 16).

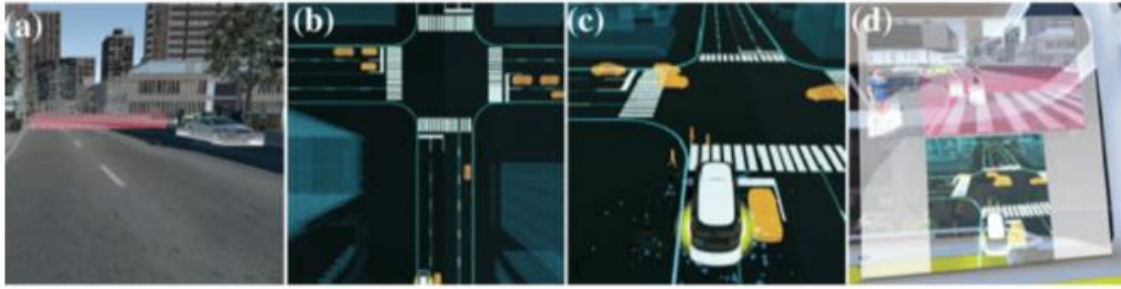


Figure 16 - Variants for solution approach 2 from [13]

The simulation consisted of a driving scenario featuring six distinct events described in the table below.

| Event | Description |
|-------------------|--|
| Start of scenario | The participants were standing in the entrance of an AV and had the task to choose a spot to sit or stand for the ride. After that, the AV started to drive. |
| Zebra crossing | The AV approached a zebra crossing with pedestrians (virtual agents) on the sidewalk. The AV stopped and the pedestrians crossed the street. |
| Change lane | The AV approached a junction at which it had to join lanes of an intersecting street. Directly afterwards, due to a temporary construction side, the AV had to change lanes again. |
| Bus stop | The AV stopped at a bus stop. |
| Emergency brake | The AV had to make an emergency brake for a jaywalker who was running across the street. |
| End of scenario | The AV returned to the starting point and the scenario ended. |

Table 7 - Traffic events from [13]

Initially, the participants were familiarized with the VR simulation through a tutorial. Following that, they went through an initial ride of the scenario without any HMI concept integrated into the AV, serving as a baseline scenario for comparison. Afterward, the scenario was conducted in a randomized order for each HMI concept that had been previously defined.

Usability scores were obtained through the questionnaires, revealing that solution approach 2 received a higher overall usability score in comparison to solution approach one. Out of the seven participants, only one rated solution approach 1 higher than solution approach 2. The table below illustrates the score that each participant assigned to each solution.

| Participant | Solution approach 1 | Solution approach 2 |
|---------------|---------------------|---------------------|
| 1 | 37.5 | 82.5 |
| 2 | 37.5 | 77.5 |
| 3 | 75 | 75 |
| 4 | 65 | 65 |
| 5 | 75 | 42.5 |
| 6 | 72.5 | 80 |
| 7 | 55 | 72.5 |
| Average Score | 59.6 | 70.7 |

Table 8 - Score for each solution from [13]

Another question involved participants selecting their preferred variant from solution approach 2. Five out of the seven participants chose variant C as their favorite concept. With the help of participants' feedback, the preferred HMI concept was refined. The final HMI concept, positioned above the windshield, incorporates additional information such as current speed, overall route, and real-time notifications (Figure 17).



Figure 17 - Refined HMI concept from [13].

Through the HMI presentation of the SYMBIOZ demo car, H. Drezet, et al. [14] present Renault's vision of HMI for AVs. Renault style and engineering teams, in collaboration with their partners (LG, Ubisoft, Devialet, Sanef, TomTom, and IAV), studied and developed the running autonomous car named SYMBIOZ demo car (Figure 18).



Figure 18 - SYMBIOZ demo car from [14]

The interior layout of the cabin has been thoughtfully designed to adapt to various driving modes and Level 4 mind-off automation, enabling the driver to disengage from road attention and engage in other activities. The car presents no center console, the air vent system has been redesigned to achieve a flat floor all the way through to under the dashboard. Furthermore, to create more space on the sides, the storage areas in the dashboard have been relocated and the door panels have been optimized with built-in lighting. A view of the car cockpit is shown in the figure below.



Figure 19 - SYMBIOZ demo car cockpit from [14]. When ambient light and HMI become orange, AD mode is activated

The interior configuration of the SYMBIOZ demo car adapts automatically based on the selected driving mode; When in AD mode, the steering wheel and dashboard slide back by 12 centimeters, creating extra space in front of the driver. In the dynamic mode, the driver's seat provides enhanced lateral support, delivering a bucket seat experience.

The car features an L-shape display concept, consisting of three fully reconfigurable displays controlled by a high-performance embedded computer. In non-autonomous mode, the L-shape displays conventional feedback such as speed, car status, navigation, and music information. However, during autonomous tasks, the L-shape concept transforms into a robust entertainment

system, showcasing movies and providing information about points of interest. Example views of the L-Shape display are shown below.

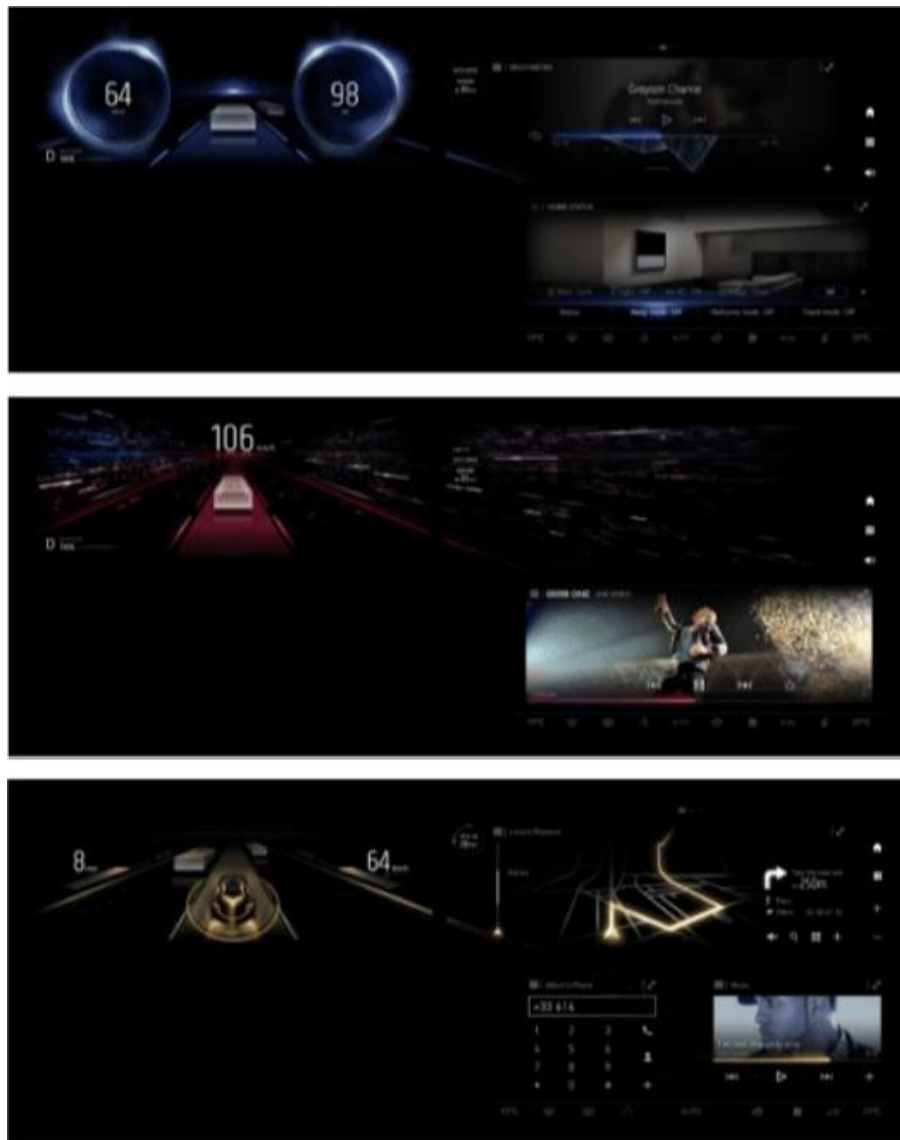


Figure 20 - L-Shape HMI of SYMBIOZ demo car examples from [14]

The SYMBIOZ demo car is also equipped with an AR-HUD designed to bolster the driver's confidence in the autonomous system.

Through the SYMBIOZ demo car, Renault has been able to offer drivers an immersive HMI experience during real road tests. The valuable feedback collected from these drivers will play a pivotal role in enhancing Renault's core HMI concepts.

In a Wizard of Oz (WOZ) based on-the-road-driving emulation of an autonomous vehicle, H. Sandhaus and E. Hornecker [15] developed an ambient light display that communicates driving decisions in advance via light signals in passengers' peripheral vision. The display visualizes both planned events such as taking turns, speeding up and slowing down, as well as unplanned

event such as swerving because of unexpected obstacles. The events are portrayed with animations that dynamically conform to the shape of the light display. Urgent events are depicted using intense light, vibrant colors, and rapid animations, accentuating their significance. Conversely, less critical events are represented through subtle and subdued animations, emphasizing their lower importance. Whenever feasible, self-explanatory colors and animations were employed to enhance clarity and intuitiveness. For instance, yellow was utilized to indicate turn signals, red was chosen to signify braking actions, and a pulsating green light was employed to represent traffic signals. The light display serves the additional purpose of functioning as a reading light. Consequently, it must be positioned on the ceiling in a manner that allows convenient access from all seats. To achieve optimal functionality, an oval shape was deemed ideal. To diffuse the LED light smoothly, white pipe insulation was found to be an effective solution. The prototype utilized an RGBW LED strip, with 60 LEDs per meter. For the final prototype, which required more complex software integration across multiple components, a Raspberry Pi was employed. The Raspberry Pi controlled two parallel 2-meter chains and was housed within the control unit. The control unit also featured a user-friendly interface for passengers, comprising a segment display and buttons for starting, emergency stop, and toggling the reading light and light signals (Figure 21).



Figure 21 - On top, prototyping platform with wizard interface on a smartphone. On bottom, control unit inside the car prototype with lit-up buttons from [15]

Using a counterbalanced, 2x2 experimental design, a group of 12 participants engaged in a series of test rides. Initially, each participant experienced either route A or route B, with the feedforward information either enabled or disabled. Subsequently, they rode the inverse

configuration. Each test ride had a duration of 45 minutes. The participants filled out one pre-study questionnaire, two identical trust in automation questionnaires and a post-study questionnaire.

The findings of the research show promising results. Participants reported a slight increase in trust and a decrease in mistrust towards the vehicle, accompanied by positive emotional experiences (as self-reported). The light display effectively captured participants' attention, directing their focus towards the screen and then to the external environment. One limitation of the study was the inherent variability among test drives, stemming from the fluctuating traffic conditions and differing positions of the sun throughout the rides. For future research endeavors, it is recommended to incorporate a light display that adjusts its luminance based on the prevailing light intensity. Furthermore, the ambient display should be seamlessly integrated into a comprehensive multimodal system that encompasses both feedforward and feedback mechanisms. To obtain more comprehensive insights, conducting a full-scale study with a diverse range of participants is crucial. Additionally, it is advisable to incorporate physiological measures to assess factors such as stress levels and potential car sickness.

To communicate the short terms plans or intentions of the AV, A. Löcken et al. [16] designed and implemented an ambient light display inside the vehicle. As a first step towards the interface, the authors conducted semi-structured interviews with seven people. The interview investigated what task the participants would engage in an AV, some questions regarding possible design as well as which maneuver, they would like to be notified about. After conducting interviews, the researchers utilized the gathered insights to develop two prototypes. The first prototype involved an iconic display employing an RGB-LED Matrix positioned at the central console. Symbols for accelerating, decelerating and changing lane are shown in the figure below.

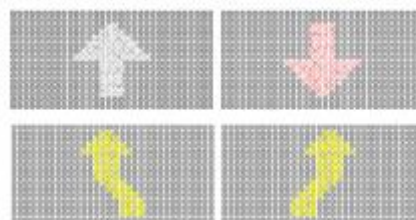


Figure 22 - Symbols for accelerating (top left), decelerating (top right), changing lane to the left (bottom left) and changing lane to the right (bottom right) from [16]

The second prototype incorporated LED-strips located between the A- and B-Pillars on both sides of the display. Interestingly, only one participant suggested the inclusion of the LED-strips (prototype b), whereas four out of the seven participants recommended incorporating a spatial dimension in the display. Different arrows using various colors were designed to indicate different maneuvers. The proposed location would be the central console. Although the light emitted from the display might be visible in the peripheral vision of the driver, decoding the information would require focused attention on the display itself. Notably, this

design does not incorporate a spatial dimension, and therefore, we did not pursue further exploration of this particular design.

The participants put forward three suggestions for utilizing the spatial dimension: installing a LED-strip below the windscreen, equipping the side mirrors with LEDs, or placing LED-strips below the left and right windows. The authors opted to utilize LED-strips located at the doors. This choice was made due to the possibility of creating light patterns that dynamically move along the longitudinal axis. To assess the usability of their interface design, the researchers conducted a static VR-based driving simulator study involving six participants.



Figure 23 - The implemented light pattern, embedded in the driving simulator of [16]. From left to right: accelerating, braking, changing lanes to the left

The participants perceived the patterns as both intuitive and inconsistent. However, it is important to note that the inconsistencies stemmed from the triggers of the patterns and the lack of information regarding aborted maneuvers, rather than being inherent flaws of the patterns themselves. Preliminary results suggest that the light display was easy to perceive while remaining unobtrusive. Furthermore, the second experiment demonstrated that the light patterns were correctly interpreted and aided in understanding upcoming maneuvers. For future iterations of this design, it is crucial to consider making the patterns configurable for drivers without sacrificing their intuitiveness. Additionally, revisiting the triggers for each pattern is necessary to reduce the perceived inconsistencies. Furthermore, the initial ideas regarding the location of the display and the patterns should be considered as a starting point, and further exploration of the design space is recommended to enhance the overall design.

2.3. HUDs

With their ability to seamlessly integrate crucial information into the driver's line of sight, HUDs have revolutionized driving, enabling drivers to effortlessly access essential data without diverting their attention from the road. Many studies reveal that the use of HUDs significantly enhances driving experiences and performance in comparison to head-down displays (HDDs), resulting in a decrease in driving errors [17], improved reaction times [18], and reduced cognitive load [19]. In recent years, AR-HUDs have emerged as a supplementary tool, seeking to augment the physical world with virtual information, potentially offering a range of benefits. This section will delve into the examination of traditional HUDs and AR-HUDs designs, conducting comprehensive comparisons among them as well as with other types of interfaces.

N. Schömig et al. [20] investigated whether an AR-HUD supports usability and reduces visual demands during CAD. 24 drivers took part in a driving simulator study, which consisted of several driving scenarios. The drivers completed a ride with a fully developed HMI designed for AD, featuring an instrument cluster (IC), ambient LEDs, and providing acoustic and haptic feedback. In another drive they were also supported with an AR-HUD concept, providing static and dynamic visual feedback. The following figure gives an overview of the HMI concept used in the simulator study.



Figure 24 - Overview of HMI elements used in the simulator study from [20]

One trial involved a 30-minute highway journey that featured various traffic events, including overtaking, construction sites, and lane changes, managed by the autonomous system. The drivers encountered a total of 4 TORs, requiring them to assume control of the vehicle to manage the following situations: two instances of blocked ego lanes at construction sites and two broken-down vehicles positioned behind a curvy hill. When the TORs occurred, the drivers were initially unaware of the reason behind them. To distract the drivers during these instances, a NDRT was introduced.

The AR-HUD HMI variant incorporated additional information displayed on the windshield, alongside the elements of the AD-HMI. The HUD display presented static information regarding speed, current speed limitations, and the reason behind TORs or road-based regulations of the AD. Dynamic information regarding the automated drive trajectory, along with real-time markings of objects critical for collision avoidance and successful take-over, was presented through augmented reality in the driving scene. The driving scene incorporated dynamic AR elements placed in relation to relevant objects. These elements included a continuous path element on the road, referred to as "fish-bones," which visually indicated the planned trajectory of the system while it remained active. This feature aimed to provide assistance, particularly in situations where the automation retained control unexpectedly. In the event of a broken-down vehicle, a red triangle continued to be displayed on the road, even when the driver took over control of the vehicle and approached the situation manually. The path followed by the automated vehicles during lane changes, along with the time remaining until the end of the operating driving domain (ODD), were also displayed. An overview of the AR-HUD is provided in the figure below.

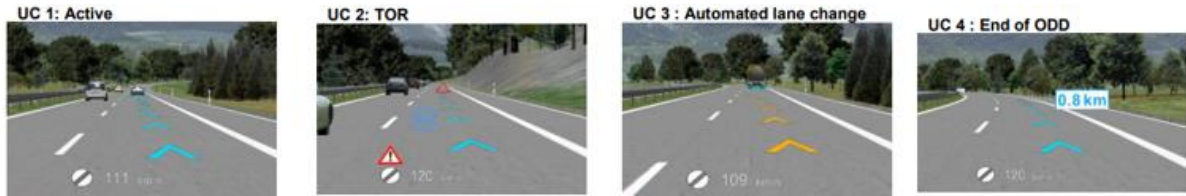


Figure 25 - AR-HUD overview from [20]

Following the two drives, the drivers evaluated the perceived understandability, usefulness, and visual workload of the HMI feedback using a 16-point scale and provided a preference rating. As quantitative indicators of visual demand, the percentage of gazes directed toward the IC was analyzed using a four-camera system from SmartEye. Additionally, in the AR-HUD drive, the drivers were requested to self-assess the level of attention they devoted to both the IC and the windshield.

The findings indicated that the self-reported visual workload was consistently low for both HMI variations, reaching a comparable level. However, the inclusion of AR was perceived as more understandable. Among the drivers who completed drives with both HMI variants, the AR-HUD was favored by 83% (20 out of 24 participants).

Participants mentioned two primary reasons for their preference: the AR-HUD makes it easier to perceive relevant information as it remains within their field of view, allowing them to maintain focus on driving. Additionally, they found the additional information provided by the AR-HUD particularly useful in takeover situations.

The results suggest that the inclusion of the AR-HUD as an addition to the AD-HMI is widely embraced and highly regarded within automated driving systems. It is preferred over the existing AD-HMI, which is already perceived as highly comprehensible and beneficial. However, a more extensive analysis is required, considering the driver's objective takeover performance.

A. Feierle et al. [21] conducted a driving simulation study with 30 participants to explore the potential of an AR-HUD in comparison to a conventional HUD. The HUD concept comprised a traditional HUD with an extra IC (instrument cluster) component, whereas the AR-HUD concept encompassed the AR-HUD along with selected HUD elements providing a preview of navigation and speed display, along with the IC. The IC's display remained consistent across both concepts. The visualization of system status and maneuvers in the IC closely resembled the visualization presented in the HUD. The HUD consisted of three separate sections, each serving a specific purpose. In the left section, the current speed and speed limit was prominently displayed, while the middle section relayed essential details about the status and maneuvers of the automated driving system. The right section was dedicated to presenting pertinent navigation-related information.

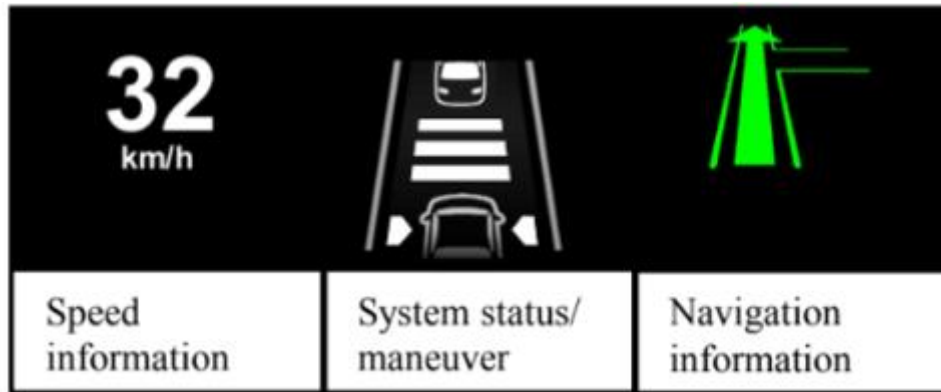


Figure 26 - Visualization in the HUD from [21] (scenario following a front vehicle)

In the AR-HUD concept, the visualization of the left and right sections of the HUD was replicated to ensure an equivalent amount of information was displayed in both concepts. To visualize the planned trajectory of the AV, the authors used a boomerang chain, which has a lower risk of concealing real objects than a drawn-through arrow [22]. When the automated vehicle had to slow down, the arrow edges of the boomerang chain transitioned from green to yellow. The presence of oncoming or cross traffic is indicated by the display of a rectangular bracket around the object, while a round bracket is used to highlight static and dynamic obstacles within the vehicle's lane, providing additional information about the lateral distance during passing. The following figure provides an overview of the differences in visualization between the two configurations in different scenarios.

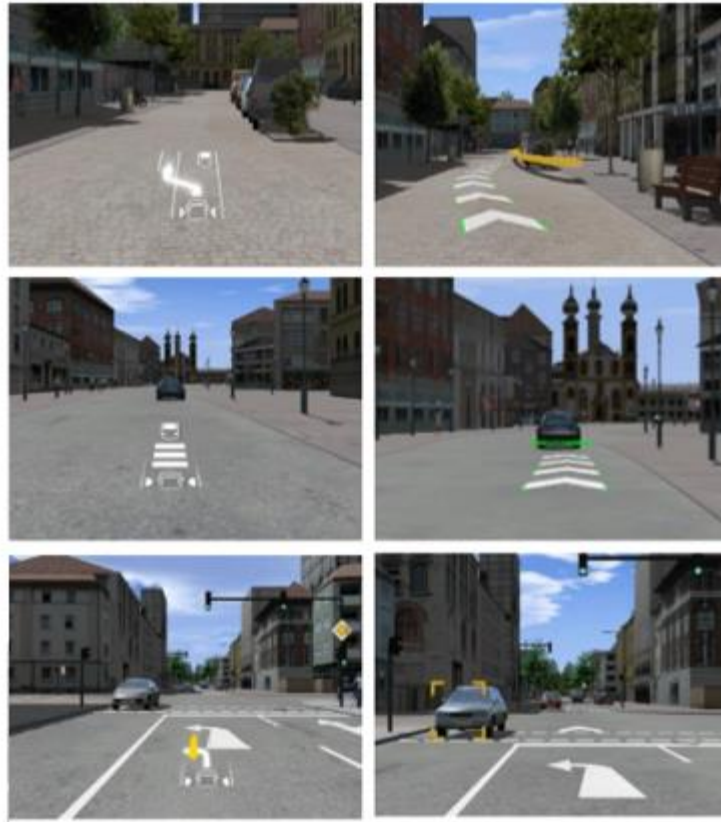


Figure 27 - Differences in the visualization of the system status and maneuver between HUD (left) and ARHUD (right). From top to bottom: bottleneck caused by parked vehicles, following a front vehicle, turning left at an intersection with oncoming traffic from [21].

A total of 30 participants completed the experiment, 13 of those already had experience with HUDs. In contrast, only 3 participants already experienced AR-HUDs. Each drive took about 14 minutes. After the first minute of manual control, the user could switch to autonomous mode and no TORs occurred. During AD, various events occurred such as bottlenecks, overtaking of cyclists, following vehicles in front, and intersections with traffic lights. To evaluate the HMI configurations, gaze behavior was measured as an objective variable, as well as subjective variables using questionnaires. The results highlighted a significant difference in gaze behavior when using AR-HUD compared to HUD. The disparity can be attributed to the substantial amount of information presented in the HUD concept case, whereas the AR-HUD primarily displays relevant information within the area of interest (AOI), which corresponds to the road ahead. The statistical analysis unveiled a notable decrease in the subjective workload for the AR-HUD concept, indicating that participants experienced less mental strain. This suggests that the information presented in the AR-HUD required less cognitive effort to interpret, potentially leading to improved attention retention. The result of trust in AV supports the finding of [10] according to which an AR visualization of sensor data in the driver FOV can increase trust in AD. Participants reported that both HMI concepts provided effective, efficient, and satisfactory support in monitoring the AV, but also that they needed to remember more details in the AR-HUD, partly perceiving it as overloaded in individual traffic situations., which could lead to potential cognitive stress and danger of concealing real objects.

Considering this aspect, the authors advocate for future research to carefully manage and reduce the simultaneous display of information in the AR-HUD.

In order to examine the optimal information selection and placement on a WSD, R. Haeuslschmid et al. [23] developed a view management concept that introduced zones and designated areas to ensure the effective presentation of specific information on the WSD.

Areas are defined as 2D spaces facing the driver and zones add a third dimension which together defines 3D spaces for the display of information according to the task. The following 3D zones have been defined: the personal zone, ranging from 0 to 50 cm, containing personal information such as private messages; the social zone, ranging from 50 cm to 290 cm, designed for business-type interactions; the vehicle zone, partially matching the dimensions of a car, primarily displaying information about the car and navigation. Lastly, the environment zone ranges from 290 cm to infinity and displays information related to the environment in a world-fixed manner.

Moving forward, the authors proceeded to divide the windshield into several areas based on the context and priority of the content:

The notification area overlays urgent warnings related to the primary task at hand and potentially the secondary task but does not have any relevance to the external environment. The vehicular area presents information about the own car exclusively; The personal area presents information related to the driver; The ambient area encompasses everything that is not of specific relevance or has low priority. The reading area supports the reading of continuous texts on the WSD. The environmental area is related to the surroundings in a world-fixed manner.

The authors undertook an initial exploratory VR-based study involving 21 participants. The simulation presented to the participants a driving scene and example WSD information on a Google Cardboard with an attached Huawei Mate 7.

The study conductor overlaid supplementary information across the majority of the windshield at various distances. This additional information was intentionally made visible for a limited duration of 5 seconds. Participants were instructed to memorize the information and later provide their recollections through a questionnaire. The procedure has been repeated a total of six times, with each repetition featuring a different set of five information pieces and distinct distributions. Following a break of 5 to 10 minutes, participants were then asked to indicate their preferred placement of information through a questionnaire.

From the results, it was discovered that out of the 21 individuals observed, only 13 utilized the complete range. Among these, 7 participants confined their usage of WSD information to a maximum distance of 10 meters, while 3 participants within this subset limited it to 5 meters. Astonishingly, these individuals also restricted their perception of environmental information to the same distance. Furthermore, 9 participants suggested employing only two or three depth levels, with no participant opting for a single layer exclusively. This demonstrates that there is a diversity of perspectives regarding the utilization of multiple depth layers or even continuous

depth. The figure below illustrates the preferred distances for each type of information from participants suggestions.

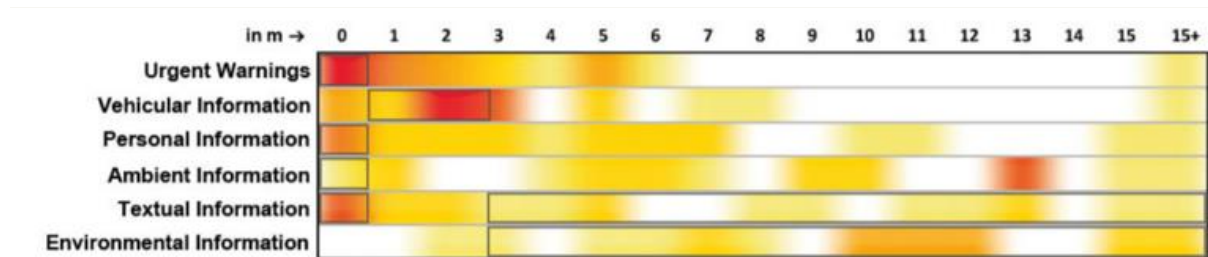


Figure 28 - Distances preferred by participants from [23]. Dark shades indicate a high count of participants. Grey boxes represent the concepts' suggestions.

About the areas, participants displayed a noticeable reluctance to distribute information evenly throughout the entire designated space. It was observed that 8 participants confined the information solely to the left and middle areas of the windshield, with 7 of them not utilizing the complete range of distances. Several participants chose to position content along the edges and in the corners of the windshield, which are significantly distant from the driver's line of sight. However, the area directly within the driver's focus point was consciously avoided, although all but 2 participants utilized the immediate vicinity surrounding it.

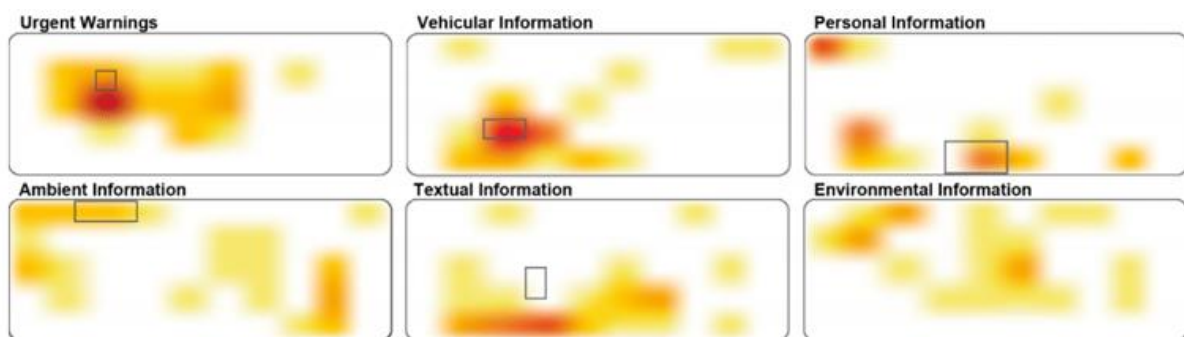


Figure 29 - Area preferred by participants from [23]. The darker the shade, the more participants chosen that area. Grey boxes represent the concepts' suggestion, except for environmental data, which is distributed in the concept.

The researcher introduced a view management concept for windshield displays that proposes zones derived from the theory of Proxemics. However, a comprehensive study involving a substantial sample of participants is necessary to examine the recognition, response times, and potential distractions associated with this view management concept.

Urban traffic environments can cause unpredictable hazardous situations that can cause discomfort and possibly loss of trust in the driver of an AV. To increase driver situational awareness in such scenarios, P. Lindemann et al. [24] simulated an explanatory windshield display (WSD) interface for AC in a mixed-reality driving simulation. When designing the interface, authors followed recommendations for the placement of information in large-sized from [23] and sickness prevention in self-driving car user interfaces [25]. The user interface (UI) consists of two elements: screen-fixed information visualization and world-registered augmentation. The authors based the selection of elements and content in part on preliminary

feedback. An overview of the information displayed by the WSD can be seen in the figure below.

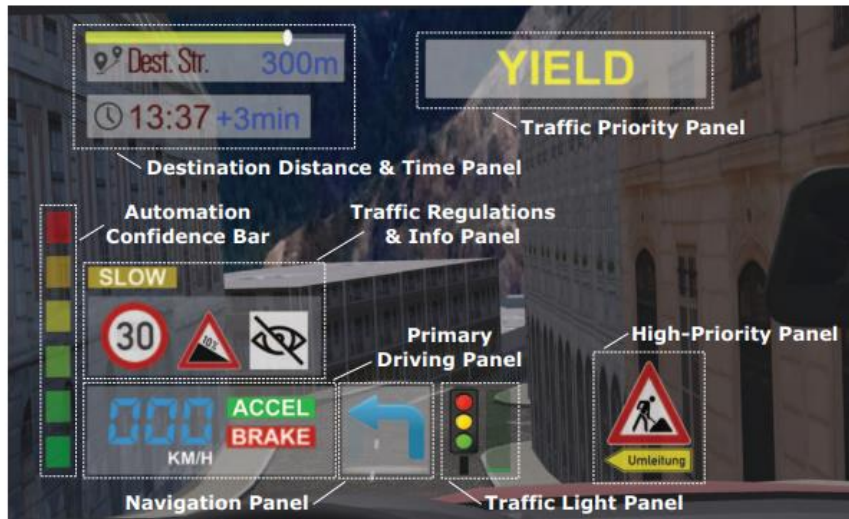


Figure 30 - Overview of the information provided by the WSD from [24]

The destination and time information as well as the automation confidence bar have been introduced because of preliminary user feedback. The primary driving panel shows the current car speed and binary indicators showing an imminent acceleration or brake action by the car. The navigation panel provides information about the upcoming turn direction and assists in assessing circumstances where the automated vehicle veers off the driver's expected path. The traffic light panel shows the remaining green or red-light time for an upcoming intersection, assuming the availability of a vehicle-to-infrastructure communication system (V2X). The traffic regulations panel presents the active rules and environmental conditions that are in effect and being taken into consideration by the AV. Furthermore, under specific circumstances, such as when the preceding vehicle exhibits abnormal behavior, a warning message is displayed, indicating the unusually slow driving of the automated vehicle. The panel for traffic priority presents the right-of-way status for the upcoming intersection, informing whether the autonomous vehicle plans to yield or continue driving upon reaching it. Lastly, the high-priority panel contains specific indicators for rare occurrences, e.g. unplanned deviations due to roadworks or an approaching emergency vehicle. Initial feedback indicated that users expressed a preference for maintaining a partially unrestricted field FOV. As a result, aside from the main driving and destination panels, all other components only display content dynamically when relevant. These elements are deliberately positioned away from the central FOV of the driver to minimize any interference with the world-registered augmentations described in more detail in the figure below.






| Visual Example | AR Element Description |
|--|---|
| <p>Threat & warning markers</p>  | <p>Colored object outlines warning drivers of potential hazards. Possible threats are marked in yellow. Direct threats in the vehicle's path are marked in red. In case of an imminent threat, other WSD elements are temporarily hidden to guide attention to the position of the hazard.</p> |
| <p>Oncoming traffic indicators</p>  | <p>Superimposed on the road when the AC approaches an intersection. Red arrows indicate from where oncoming traffic is detected. Their current distance is roughly indicated by the amount of arrowheads (up to 3). It allows drivers to notice approaching cars without leaning or head-turning.</p> |
| <p>Brake and stopping bar</p>  | <p>A stopping bar is superimposed on the road whenever it comes to a situation in which the vehicle must brake hard due to unforeseen circumstances. It indicates where the car is expected to stop completely. It may reassure drivers by giving a point of reference for safe stopping but could also facilitate to grasp the cause of the situation.</p> |
| <p>Moving object markers</p>  | <p>Rotating cubic markers are overlaid above moving objects if they are relevant to driving decisions. E.g., drivers ignoring the right of way and becoming dangerous are highlighted by a red cube. On the other hand, a car classified as safe (e.g. if it yields) might be marked in green to reassure the driver.</p> |
| <p>Road sign overlay</p>  | <p>Yield, priority or stop road signs in front of intersections are superimposed in large scale onto the road surface. For the yield/stop signs, the position of the overlay also marks where the car may need to stop. Priority signs might reassure drivers in cases where the AC is not waiting to yield.</p> |

Table 9 - An overview of the implemented AR elements with example visualization and descriptions from [24]

The implemented world-registered traffic object augmentations are directly displayed on or near objects they are highlighting in the world space. The system uses colored outlines on objects to caution drivers about possible hazards. Yellow outlines indicate potential threats, while red outlines signify direct threats directly in the vehicle's trajectory. In critical situations, where a threat is imminent, other elements of the WSD are briefly concealed to guide the driver's attention specifically to the position of the hazard. When the AV approaches an intersection, indicators of oncoming traffic are overlaid on the road. Red arrows are used to show the direction from which oncoming traffic is detected, while the number of arrowheads roughly represents their current distance. In situations where the vehicle must abruptly brake due to unforeseen circumstances, a stopping bar is overlaid on the road. This bar indicates the expected location for the car to come to a complete stop. To aid in driving decisions, relevant moving objects are enhanced with rotating cubic markers positioned above them. For instance, a red cube is used to highlight drivers who ignore the right of way and become hazardous. Conversely, a green cube is employed to mark vehicles classified as safe, such as those that yield appropriately, aiming to instill confidence in the driver. Yield, priority, or stop road signs located in front of intersections are projected onto the road surface in a large-scale overlay. The

position of the overlay not only displays the sign itself but also indicates where the car may need to come to a stop, particularly in the case of yield or stop signs.

To test the proposed interface, 32 participants took part in a driving simulation. The independent variables encompassed the WSD Interface, which could be toggled on or off, and the visibility factor, which varied between high and low due to the presence of fog. The authors hypothesized that when the WSD interface is active, the situational awareness would show a significant increase, and the decline in driver situational awareness from high to low visibility would be significantly less pronounced compared to when the WSD interface is inactive. Results highlighted a significant difference between drives with the WSD interface and those without the WSD interface under high visibility with a medium effect size and under low visibility with a large effect size. Additionally, there was a significant difference between high and low visibility conditions when not using the WSD interface (small effect size). On the other hand, there was no significant difference between visibility levels under enabled WSD. The authors believe that added situational awareness in critical situations with sudden maneuvers may prevent loss of trust or even increase trust through positive situational experiences.

L. Morra et al. [26] introduced a methodology that validates the user experience in AVs through the analysis of objective data obtained from physiological parameters. This methodology is implemented while the user actively engages in a dynamic VR-based driving simulation, which incorporates additional audio cues. This methodology has been applied to an AR-HUD that provides information about vehicle planning and sensory system. The AR-HUD is derived from the combination of various HMI designs sourced from prototype solutions, commercial products, and research endeavors. The virtual reality simulation was developed using the HTC Vive ecosystem. The AR-HUD has been developed with two distinct designs: omnicomprehensive (OMN) and selective (SEL). In the OMN version, the system presents information by superimposing visual indicators onto all dynamic elements (vehicles and pedestrians) within a 150-meter detection radius of the vehicle. Static objects are excluded from the display, unless they become dangerous, according to the principle of cognitive load, which emphasizes the importance of creating an interface that is understandable and user-friendly while minimizing excessive visual complexity. In the SEL version, only information relevant to the user's specific interests is displayed. The underlying principle was to choose information related to environmental elements that directly impact the behavior of the AD system at any given moment. Considering road signs, if a road sign is in the radius of detection of the vehicle but that information is not helpful at that given moment, it would not be displayed. In the SEL variation, more specifically, only cars that are ahead of the current vehicle or intersect its current trajectory are visually outlined with a bounding box. Pedestrians and other stationary or moving objects are identified only in the event of potential danger, specifically when a collision becomes possible. Navigation lines for other vehicles are selectively shown based on the vehicle's assessment, such as at intersections for priority determination. Nevertheless, traffic light information, tracking of the vehicle's navigation line, and the road center line remain unchanged in this variant.



Figure 31 - OMN design on the left, SEL design on the right, from [26]

In the simulation, participants experience a 12-minute urban ride that includes the sudden crossing of a dog (Dog), a child on the sidewalk throwing a ball onto the street (Ball), scooters and cars adeptly splitting lanes while driving (Scooter, Car1, and Car2), as well as pedestrians crossing the road (Man1 and Man2). A timeline with the events simulated within the test scenario is provided in the figure below

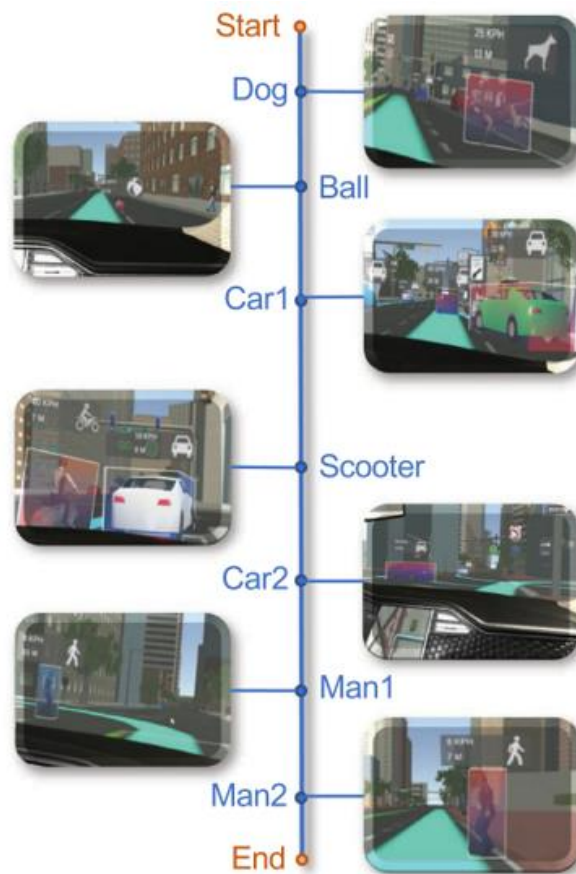


Figure 32 - Timeline of the test scenario with simulated events from [26]

User feedback is gathered through subjective analysis, which involves questionnaire-based judgments, as well as objective measurements using physiological signals such as Galvanic

Skin Response and Heart Rate Variability. According to the authors' findings, the delivery of explanatory "why" information is pivotal in assuring users of the system's capability, fostering trust, and enhancing situational awareness. The research also shows that an OMN display may place an excessive amount of cognitive overload on the user. An effective strategy to mitigate information overload involves filtering and displaying solely the key visual cues that hold significance within the current context which researchers call selective (SEL).

When faced with potentially dangerous situations, users reported feeling less stressed while using OMN AR-HUD as they trusted its capability to effectively manage unexpected changes in the environment, even if the information it provides is felt slightly excessive. This result was highlighted by both subjective and objective metrics. The proposed methodology allows for the evaluation of various autonomous vehicle types by modifying their interior features and behavioral aspects. Moreover, it provides a means to investigate the effects of possible faults on the user's experience and the level of trust placed in the system.

In the field of landmark-based navigation research, most previous studies centered around incorporating text or icon-based landmark information on dashboard displays. In a departure from this convention, A. Bolton et al. [27] proposed a different methodology. They utilized a HUD to provide landmark-based navigation information, leveraging AR to highlight the landmarks and improve their overall usefulness. In the experiment, 20 participants took part in a medium-fidelity, fixed-based driving simulator. The driving scenario was created using STISIM Version 3 software and reproduces a suburban road scene presenting a challenging navigation environment. Navigation cues were created using Synfig Studio animation software independently of the road scene and displayed using a HUD unit affixed in place of the driver's sun visor.



Figure 33 - In-vehicle setup showing the HUD, road scene and participant wearing ETG from [27]

Participants completed the driving scenario using each of the presentations of navigational instructions: conventional (CV), AR arrows (AR), landmark arrows (LA), and landmark boxes (LB). The CV displayed the distance-to-turn in meters along with the turn direction, using a fixed arrow either on the right or on the left of the HUD. During the AR condition, an augmented arrow dynamically displayed in the middle of the road pointed towards the

designated turn, continuously changing its size to reflect the proximity of the vehicle to the junction. Landmarks were highlighted either using an arrow pointing at the landmark (LA) or by using boxes that enclosed the landmark (LB). Distance-to-turn information was displayed in each condition. An overview of each of the presentations of navigational instructions can be seen in the figure below.



Figure 34 - Experimental conditions showing HUD images superimposed on road scene (top to bottom: CV, AR, LA, LB). From [27]

Drivers encountered eight decision points where they were asked to select the correct turning by using the information presented on the HUD. No voice information was provided as navigational cues, to ensure participants were fully reliant on visual representation. The following measures were captured: response time, success rate, subjective workload, driver preference, eye glance behavior, and general comments. The results revealed a strong preference and navigational performance benefits (in terms of speed of response and accuracy) when using AR landmark presentation. Particular benefits were associated with using a box to highlight a landmark located at the required intersection, improving response time, and success rate but keeping the lowest rating of workload and being ranked as preferred. The study also revealed that the usability and utility of landmarks, that were previously assumed to be

unsuitable navigational candidates, could be enhanced by highlighting them using AR. In future research, it is important to take precautions in order to minimize the distractions caused by presenting additional information to drivers and to avoid overwhelming them with excessive information.

Furthermore, it is crucial to investigate the effects on driving performance and assess the implications of enhanced depth perception in a real-world setting, rather than relying solely on simulated environments.

2.4. Auditory Interfaces

Auditory interfaces present a promising solution to communicate with passengers without distracting them, offering voice warnings and alerts that not only mitigate anxiety but also enhance alertness and restore a sense of control for passengers [28]. Additionally, audio interfaces that utilize speech-based interactions have the potential to enhance passengers' trust in autonomous driving systems through a combination of anthropomorphism, clarity, and system transparency [29] [30] [31]. By providing clear and concise verbal communication, these interfaces offer real-time updates, warnings, and explanations regarding the vehicle's actions, intentions, and surrounding conditions. In the upcoming section, different auditory interface designs will be discussed, centering on the differentiation between informational voice agents and conversational voice agents. Additionally, we will explore the effectiveness of employing either virtual or physical avatars in these interfaces.

In their study, P. Hock et al. [32] conducted a VR static simulation with 38 participants. The objective of the study was to examine the impact of HMI design and system feedback on the driver's trust in the AV and their willingness to maintain automation in safety-critical situations. A total of 38 participants took part in the simulation. At the beginning, a demographic questionnaire was filled out and an introduction to the study procedure was provided in written text. The study setting was conducted in a VR driving simulator by using an Oculus Rift DK2 and a Logitech G27 steering wheel, as shown in the figure below. To introduce participants to the VR device and to make them familiar with the study environment, three practice scenarios were conducted beforehand. These practice scenarios were designed specifically to familiarize participants with the process of overtaking cars within the VR setup, ensuring their comfort and readiness for the actual experiment.



Figure 35 - Left: a car approaching out of the fog (visual range: 200m) ahead of the driving car. Right: virtual reality study setup using Oculus Rift DK2. From [32]

Participants were seated in an automated car that traveled autonomously at a speed of 100 km/h on a two-lane rural road. The study utilized fog to limit their visual range to a specific distance, starting at 200m. The fog conditions were adjusted periodically throughout the experiment to serve as a controlled variable for visual range. Participants were informed that the vehicle's sensor range matched the human visual range. During the scenario, a car approached in the opposite lane after 15 seconds to demonstrate the presence of oncoming traffic. Then, after 25 seconds, a car traveling at 60 km/h emerged from the fog as a lead vehicle. The automated car caught up to the car ahead after 37 seconds and reduced its speed to follow it. At the 45-second mark, the system provided feedback on why overtaking the car ahead was not possible. This feedback was repeated 40 seconds later, with no changes in the fog conditions. Subsequently, the fog conditions changed every 40 seconds, altering the visual range to 500m, 800m, and 1100m, respectively, as shown in the following table.

| time passed (s) | visual range (m) |
|-----------------|------------------|
| 0s | 200m |
| 120s | 500m |
| 160s | 800m |
| 200s | 1100m |
| 240s | End |

Table 10 - Connection between time passed since the beginning of the drive in seconds and visual range at the given moment in meter

For the evaluation, three conditions were utilized. In the "none" group, participants did not receive any feedback. In the "audio" condition, the system offered auditory feedback concerning the current condition, elucidating the reasons why the automation did not perform an overtaking maneuver. In the "co-driver" condition, participants received the same audio feedback as in the audio condition, along with the presence of a virtual co-driver seated in the front passenger seat (Figure 36). To eliminate any ambiguity, the authors chose to present the co-driver as a projection, underscoring its artificial humanoid nature and emphasizing that it serves as a representation of the computer system and the automation, rather than being perceived as a human individual. The audio voice was the same in both "audio" and "co-driver" conditions.

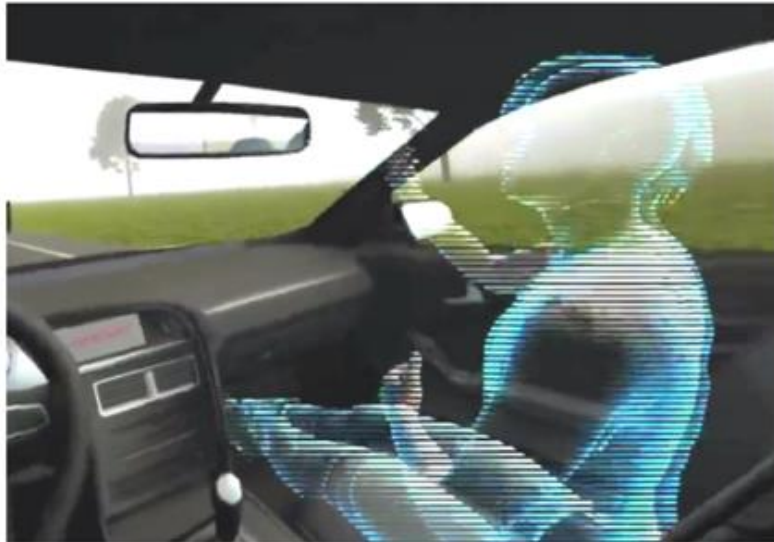


Figure 36 - The avatar in the study [32] is designed as a projection to avoid confusion whether the avatar should represent a human being or a computer system. The avatar gesticulates to increase humanoid appearance.

The study revealed that passengers in the audio or co-driver conditions surpassed the duration of automation override by 86 seconds compared to passengers in the none condition. This indicates that audio information effectively enhanced the user's inclination to sustain automation for a longer period. However, qualitative feedback indicated that the avatar in the co-driver condition was often perceived as displeasing and distracting, suggesting that an avatar may not always possess a more compelling character than a mere voice.

S. Alpers et al. [33] investigated the user experience implications of information delivery and delivery styles by an in-vehicle digital assistant in a mock AV through an on-road-driving study. The research study was conducted over a two-year period from April 2017 to April 2019. The authors initially conducted a vehicle set-up and believability study, followed by another study consisting of testing and analysis. The vehicle exterior has been modified to include a loud branding, non-functioning light detection and ranging (LIDAR) system to provide the illusion of active remote sensing, side-tinted windows to prevent participants from seeing a human driver, and a GoPro video camera mounted in front of the vehicle. The interior of the van was equipped with various components, including an electronic stop button positioned on the left side, an interior GoPro camera to record participants' verbal responses and facial expressions, a live-stream video monitor displaying data from an external GoPro camera, a sound dampening partition to prevent participants from hearing the human driver, and a human driver dressed in a car seat-like suit.



Figure 37 - The vehicle setup used in [33]

Eleven adults were recruited to participate in the first phase. Initially, participants were introduced to the vehicle and stop button and completed an interview about their perception of technology and AVs. The AV then drove to the curb in front of the building's front entrance. Following this pre-ride, the mock AV drove in a loop on surface streets for about 15 minutes at a speed not exceeding 35mph and ending at the initial starting point. Video and biometrics data were captured. Following each ride, each test subject was given a brief written survey and exit interview to verify if participants believed the vehicle to be an actual AV. The second phase focused on examining the effects of a digital assistant on the user experience within a mock AV, with a secondary focus on capturing passengers' physiological stress and anxiety responses. The independent and dependent variables used in the experiment can be seen in the figure below.

| X (independent variables) | Y (dependent variables) |
|--|--|
| X₁: Lilly Digital Assistant | Y₁: Overall User Experience |
| X _{1,1} : Provides 5 points of information | Y _{1,1} : Comfort & relaxation |
| X _{1,2} : Lacks human-like interaction qualities | Y _{1,2} : Personal safety |
| X₂: Julie Digital Assistant | Y _{1,3} : Safety for others |
| X _{2,1} : Provides 13 points of information | Y _{1,4} : Vehicle trustworthiness |
| X _{2,2} : Exhibits human-like interaction qualities | Y _{1,5} : Willingness to ride again |
| | Y₂: Confidence in Vehicle's Capability |
| | Y _{2,1} : AV control |
| | Y _{2,2} : AV driving capability |
| | Y _{2,3} : AV decision-making ability |
| | Y₃: AV Anthropomorphic Characteristics |
| | Y₄: AV Riding Experience |
| | Y _{4,1} : Digital assistant interaction |
| | Y _{4,2} : Interior vehicle accommodations |
| | Y _{4,3} : Drive and driving route |
| | Y _{4,4} : Awareness of distractions and stressors |
| | Y₅: Value/Usefulness of Information Provided |
| | Y _{5,1} : Future preference toward digital assistant |

Table 11 - Independent and dependent variables from [33]

During each test run, a participant was subjected to four stressors, intended to elicit an anxiety response while riding in an AV. These included the van circling a roundabout multiple times, the vehicle reaching a dead-end, the van detecting an object (gravel) in a roundabout, and an aggressive stop at a stop sign (Figure 38). Speed never exceeded 35mph.

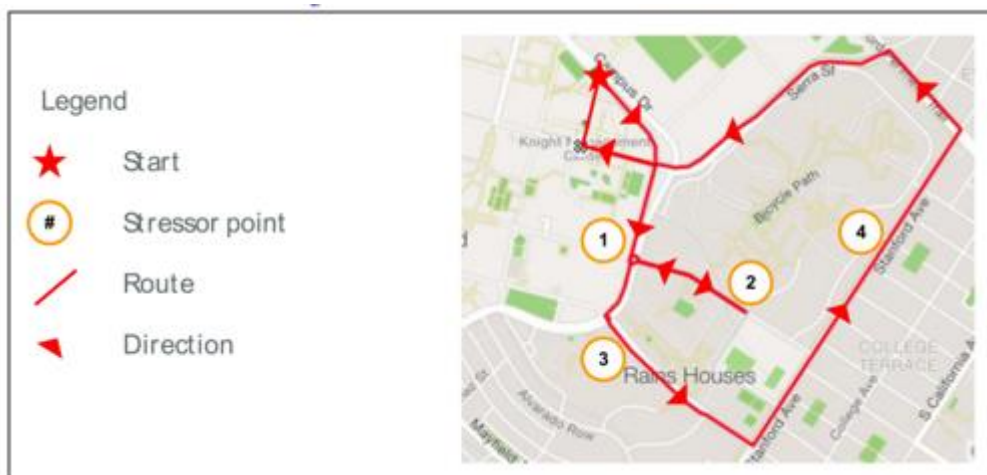


Figure 38 - Map of driving route and stressor for each of the 15 minutes test run from [33]

Each participant was assigned to either the Lilly or Julie digital assistant. Both digital assistants had the same female-speaking voice. Each digital assistant had a distinct persona, with human vs. non-human interaction qualities. The interaction of the digital assistants with the passenger is shown in the figure below.



Figure 39 - Lilly vs Julie passenger interactions from [33]

A total of 32 participants took part in the second phase. After the test, the authors analyzed the data collected from the participants, including survey responses, interviews, video recordings, and biometric data. The results revealed that passengers assigned to Julie, the digital assistant providing more human-like information, exhibited lower levels of anxiety and higher engagement compared to those assigned to Lilly. Based on these findings, it was concluded that an in-vehicle digital assistant that presents abundant information in a human-like manner

enhances passengers' trust in the perceived driving capability of a fully autonomous vehicle (FAV) and their willingness to rely on it.

To investigate the impact of agents on passengers' perceptions of a FAV's intelligence, competence, and likability, S. C. Lee et al. [34] conducted a comparative analysis involving three distinct agents: the informative voice agent (IVA), informative robot agent (IRA), and conversational robot-agent (CRA). A group of twelve experienced drivers participated in a screen-based driving simulation with a motion platform. In the two robot agent configurations, a humanoid robot was positioned adjacent to the driver's seat (Figure 40).



Figure 40 - Experimental settings considered in [34]

The primary objective of these agents was to furnish information to enable the driver to comprehend the intentions and status of the FAV. Participants completed one driving trip with each of the configurations. The driving scenario encompassed eight distinct events aimed at showcasing the FAV's proficiency in effectively managing various driving situations. During the simulation, the participants were allowed to engage in NDRTs. Before and after each event, each agent provided verbal feedback to the user. The IVA had no physical form, communicating simply with speech while both the IRA and CRA presented the humanoid robot form. The script used by each agent for each event can be seen in the table below.

| Event list | Informative script list | Conversational script list |
|-------------------------------------|---|--|
| 1. Exit a road and enter a new road | 1. Exit ahead | 1. We are entering a new road |
| 2. Road construction | 2. Road construction ahead | 2. We are slowing down because of road construction |
| 3. Swerving a car | 3. A car is swerving | 3. I am sorry for the sudden slow down. A car swerved into traffic |
| 4. Tunnel | 4. Tunnel ahead | 4. We are entering a tunnel |
| 5. Jaywalking | 5. Jaywalker ahead | 5. Are you okay? A man suddenly popped out onto the road |
| 6. Waiting for traffic signal | 6. Red traffic light | 6. We are waiting for the signal to turn green |
| 7. Turning left/right | 7. Turning left/right ahead | 7. We are turning left/right |
| 8. All way stop intersection | 8. This is a four-way stop intersection | 8. We've reached a four-way stop intersection. We are waiting for other cars to go first |
| 9. Fog | 9. Fog ahead | 9. Wow, it's really foggy outside |

Table 12 - Driving events in the scenario (left), informative agent's script (middle), and conversational agent's script (right) from [34]

The authors employed questionnaires to gather subjective evaluations, while a preference ranking method was utilized to assess riders' perceptions of the three agents. Upon analyzing the questionnaires, it was revealed that most participants regarded the conversational robot agent (CRA) either as the best or worst option, describing it as friendly and natural or verbose and distracting. Interestingly, the results indicated no notable distinctions among the three scenarios, suggesting that equipping the voice agent with a physical robot presence might not contribute further to enhancing the user's perception of the FAV.

In their research, Forster et al. [35] [31] [36] examined the potential utilization of voice output to strike a balance between the driver's inclination to engage in NDRTs and the necessity to stay informed regarding the AV's status and intentions. A group of 17 participants engaged in driving a Level 3 AV on a highway with moderate traffic using a motion-based driving simulator and the SILAB simulation program. Throughout the experiment, the participants encountered three distinct driving scenarios that necessitated communication between the automation system and the driver. The AV's HMI included a visual-auditory interface, whereby the auditory component was either generic feedback (referred to as the "generic" condition) or additional voice output (referred to as the "speech + generic" condition). In the generic condition, a generic auditory output comprising two high-frequency tones (duration: 150ms, frequency: 1000 Hz sinus, interval: 150ms) was presented along with the visual notification on

the display whenever a traffic event occurred. Conversely, in the speech + generic condition, the generic output was followed by a speech output using a Dictaphone recording of a female voice. The voice output was utilized to articulate information concerning future system maneuvers. The simulation setup is shown in the figure below.



Figure 41 - Simulation setup used in [35][31][36]

All the participants engaged in the simulation twice. In the first simulation, speech output was provided, while in the second, no speech output was given. The driving events were presented in a random order to ensure variability. As part of the study, participants were instructed to read a magazine during the ride, allowing the researchers to observe the frequency of interruptions in performing non-driving tasks. The researchers closely monitored the participants' gaze behavior as an indication of how often they diverted their attention from the NDRTs to check on the vehicle's automation. The table below shows the wording used by the speech output.

| Situation | Wording |
|---|---|
| <i>Avoidance</i> : lost cargo on the right lane. The CAD system adapts its lateral position and avoids the obstacle on the road. | "Obstacle ahead. Adjusting lane position to avoid." |
| <i>Speed limit</i> : adaptation of the host vehicle's set speed (from 120 km/h to 80 km/h) due to a speed limit change. | "Speed limit ahead. Adjusting speed." |
| <i>Lane change</i> : CAD system recognizing highway intersection ahead and changing lane to the right in order to follow the route. | "Intersection ahead. Changing lane and following navigation." |

Table 13 - Auditory HMI speech output for system maneuvers as exploited in [35][31][36]

To assess the impact on NDRTs, a chi-square test was conducted on the frequency of behavioral observations under both HMI conditions. The results of the analysis indicated no significant

differences in the frequency of interruptions between the two HMI situations. However, it was observed that participants allocated more gaze time to the HMI in the generic condition, suggesting that when speech output was provided, they required less time to gather relevant information from the visual interface. These findings suggest that increased speech output can reduce distractions caused by the visual interface, allowing drivers to focus more on NDRTs. Nevertheless, incorporating NDRT presentation within the visual component of the HMI may render speech output unnecessary, as it eliminates the need to divert attention away from the display.

With a series of three studies, B. A. Lewis et al [37] [38] analyzed the acoustic characteristics that contribute to a sound being unambiguously perceived as an urgent alarm within a vehicle context. Urgency mapping is a traditional technique employed to ensure a suitable alignment between the urgency associated with a situation and the innate urgency conveyed by a provided alert E. Hellier and J.Edworthy [39]. A total of 21 individuals participated in the initial experiment, in which 29 sounds were produced using Adobe Audition CS5.5, and Audacity software was integrated for the perceptual sorting task. These stimuli were selected from a diverse range of categories such as forward collision warning, park-assist, door open, and more. The stimuli were standardized to have a consistent loudness level of 84 dB sound pressure level. Participants were required to categorize each stimulus into one of the following categories: "Warnings," "Alerts," "Status Notifications," or "Social Notifications." Following a demographic survey, participants engaged in a brief practice task involving sounds that were not included in the experimental session. Upon sorting the set of 29 sounds, participants were subsequently prompted to select a sound representative of each category, to indicate the reason they grouped sounds the way they did and to give a number between 1 and 100 that would indicate the best urgency level for each category. The results were assessed by examining the percentage of time each sound was categorized within a specific category. Preliminary analysis of the categorization and explanation data revealed that participants did not consistently differentiate between warnings and alerts. Instead, both were perceived as alarming sounds. Consequently, the authors amalgamated the warning and alert categories into a singular "Alarm" category. By employing a blend of backward and stepwise regression analyses to predict membership in the warning group, four primary criteria were established. These criteria accounted for 61% of the variance ($R = .802$, adj. $R^2 = .612$). These criteria, along with their corresponding cutoff values, indicative of the threshold at which sounds are classified as alarming, are outlined in the following table:

| Criterion | Cutoff |
|--------------------------|----------------|
| Peak-to-total-time ratio | ≥ 0.70 |
| Interburst interval | ≤ 125 ms |
| Number of harmonics | ≥ 3 |
| Base frequency | ≥ 1000 Hz |
| Pulse duration | ≥ 200 ms |

Table 14 - Criteria and cutoff from [37][38]

The criteria are arranged in descending order of significance, based on their regression parameters. The "Peak to total time ratio" pertains to a property conceived within the researchers' laboratory to characterize predominantly prolonged onset or offset durations. This ratio is defined as the "peak time," denoting the duration the sound remains at its maximum intensity, divided by the "total time," representing the overall duration the sound is audibly active at any intensity level. According to this definition, a sound devoid of onset or offset duration would yield a ratio of 1, whereas a sound with onset or offset duration encompassing the entire sound length would yield a ratio approaching or reaching 0.

Experiment 1 yielded initial insights into the crucial sound attributes influencing the classification of sounds as urgent and alarming. However, the criteria formulated in this study were derived from the identical dataset used for the classifications themselves. Moreover, only a minimal number of sounds satisfied either all or three out of the four predefined cutoff criteria. Consequently, a supplementary validation test was conducted to scrutinize the efficacy of our criteria in predicting categorization. This validation test involved a broader set of sounds and a fresh set of participants.

Experiment 2 involved the participation of 15 individuals and a total of 52 sounds were used as stimuli. These sounds encompassed those employed in Experiment 1, although they were adjusted to have a combined duration of 1500 milliseconds. Additionally, new stimuli were introduced, including some with added harmonics. The results demonstrated that the criteria and cutoff values derived from Experiment 1 possess a notable degree of efficacy in predicting alarm categorization. This predictability sustains consistency across participants and exhibits enhanced resilience when applied to a more extensive set of sounds. Nonetheless, it's important to recognize that these experiments were conducted under relatively constrained circumstances – devoid of significant context, distractions, and participants were afforded the liberty to listen and re-categorize sounds as frequently as they wished during each session.

Given these considerations, it became imperative to subject our criteria to assessment with individuals engaged in driving scenarios, simultaneously performing a secondary task. This situation demanded participants to categorize sounds while dividing their attention, thus offering a more realistic and challenging environment for evaluation.

Experiment 3 involved 22 participants and the stimuli utilized in this experiment consisted of a subset of 26 sounds extracted from Experiment 2. These particular sounds demonstrated relatively consistent categorization outcomes across subjects and represented a spectrum of alarm, status, and social sounds. During the experiment, participants engaged in simulated driving using a RealTime Technologies desktop driving simulator, which incorporated a Logitech Driving Force GT steering wheel. Ambient sound was delivered through speakers at a sound pressure level (SPL) of 60 dB, while the stimuli were presented through a separate set of speakers at a higher 75 dB SPL. While driving, participants simultaneously performed a visual n-back task displayed on a small touchscreen monitor positioned slightly to their right. They indicated positive or negative responses to each stimulus by pressing buttons located on the right side of the steering wheel. The experimental setup can be seen in the following figure.



Figure 42 - Experimental setup from [37][38]

The outcomes of Experiment 3 reveal that the authors established cutoff criteria effectively forecast alarm categorization within both low and medium driving contexts and levels of simulation fidelity. Of particular significance, the response time data suggests that sounds adhering to all the criteria prompt the quickest categorization decisions. This is evident through shorter response times. This finding potentially implies that sounds conforming to some but not all the criteria might lead to confusion among participants, necessitating more time to arrive at a categorical decision.

In summary, the findings demonstrate the feasibility of defining acoustic attributes for sounds within an in-vehicle setting, enabling the development of clear and intuitive alerting systems. Notable limitations of this study include the limited number of currently employed systems and the vast range of possible acoustic parameters that can be manipulated. Furthermore, our experiments solely focused on low to medium driving scenarios, utilizing a categorization task

instead of real event responses. Future investigations should incorporate more ecologically authentic driving contexts, particularly exploring reactions to actual collision events. However, these three experiments provide compelling evidence for the potential efficacy of this approach in establishing robust and valid criteria for crafting intuitive and unambiguous alerts within in-vehicle information systems.

2.5. Mobile Devices

In recent years, there has been a growing interest in the development of mobile applications designed to alert users about vehicle intentions, provide real-time vehicle status updates, and report on various driving events. These applications serve a crucial role, particularly when employed on a driver's personal mobile device that may be used for NDRTs, functioning as a tool to keep the driver informed about the ongoing activities when the autonomous mode is engaged. This chapter aims to thoroughly examine the interfaces of mobile devices in the context of autonomous driving, exploring their potential impact on user engagement, awareness, and overall user experience.

L. Oliveira et al. [40] investigated the differences among various personal and onboard devices utilized for receiving information about the vehicle. The experiment, carried out at the Urban Development Lab in Coventry (UK), involved a driving simulation featuring a level 4 autonomous vehicle situated in a city center area. Three distinct conditions were analyzed: a mounted tablet with a letterbox overhead display fixed in the vehicle, a mobile phone with a letterbox display, and a standalone mobile phone. The information displayed on both the tablet and mobile phone included the map, the position of the AV, the route, and the expected time of arrival (figure 43).

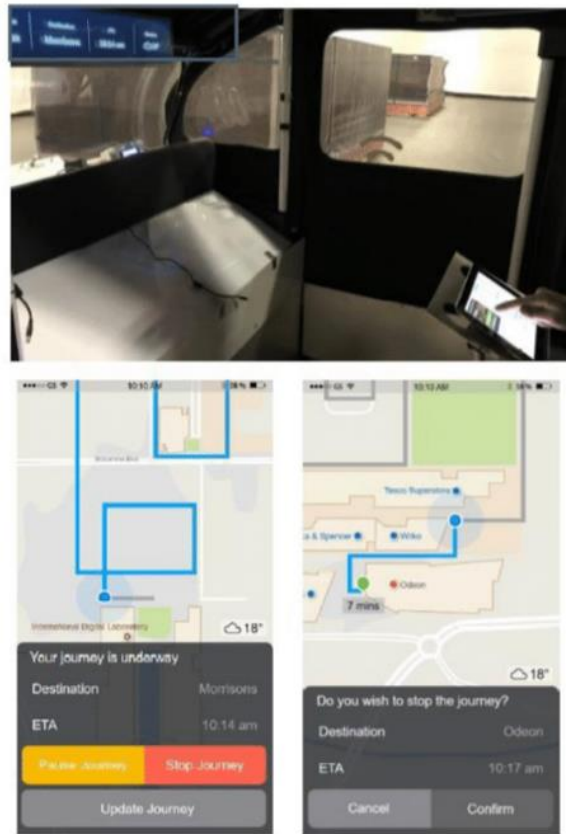


Figure 43 - Setup in [40]. On top the mounted tablet and letterbox. Mobile interface underway displaying the position of the AV on the map.

The twenty participants in the experiment were briefed on the study's objectives and the necessity to record both video and audio, and their consent was obtained. Each participant underwent three simulator journeys, one for each configuration, with varying sequences, each lasting approximately 5 minutes.

The results revealed that participants preferred the integrated tablet combined with an overhead display and regarded devices integrated into the vehicle as more useful and satisfying compared to a mobile phone combined with an overhead display. Furthermore, using a standalone phone without an additional display was found to be more useful than a phone combined with the display. One limitation of the study pertained to the basic information displayed, suggesting that future research could explore more informative interface designs.

Q. Meteir et al. [41] conducted a study to evaluate the impact of providing instructions on the limitations of automation and presenting context-related information through a mobile application on driver situation awareness and takeover performance. Eighty participants, including 54 females, participated in a static screen-based experiment utilizing a fixed-base simulator with the simulated environment projected on a large monitor. The simulation incorporated various events, such as adverse weather conditions and obstacles, each assigned a severity level ranging from 0 to 5, corresponding to distinct modalities of visualization.

The data frame simultaneously transmitted a TOR to the Android app designed for the tablet, aimed at supplying users with relevant information. The app comprised two main sections: the

first section provided information about the presence of obstacles and the status of lane markings, while the second section conveyed information about specific events requiring the driver's attention. The app featured a split-screen design (figure 44). enabling engagement in a NDRT on half of the display. The TOR was communicated through a voice announcement, text, and tablet vibration. Participant feedback regarding the app's concept and design was overwhelmingly positive.

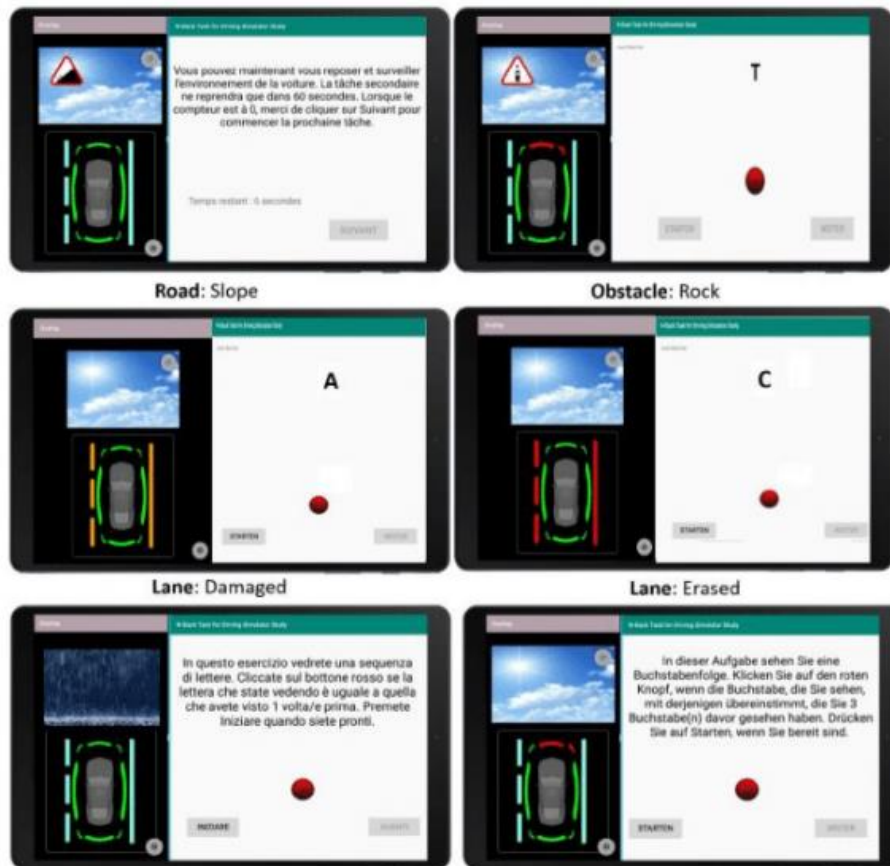


Figure 44 - Android app featuring a split-screen design from [41]

The experiment results indicated that every participant successfully completed the NDRT while also receiving information about the vehicle. All but one participant was able to identify all the events presented. The app was deemed a valuable addition and demonstrated its effectiveness. However, limitations of the study included the relatively small sample size, and future research should incorporate a comparison between conditions with and without the app. Additionally, eye-tracking measures should be considered to assess gaze behavior during the experiment.

A.P. Van Den Beukel et al. [42] proposed a comprehensive framework designed to provide driver support in real-world scenarios. The evaluation encompassed various aspects, including situation awareness, performance in critical scenarios, and the acceptance of the offered support. The framework specifically addressed two types of scenarios: hazardous scenarios that required only the user's attention and critical scenarios that necessitated the user to take control of the vehicle. The experiment was conducted using a dynamic screen-based simulation within a driving simulator at the University of Twente, involving 24 participants who experienced

three hazardous scenarios and three critical scenarios. Simulations were intentionally interrupted immediately after each hazardous or critical scenario. Subsequently, participants were required to complete a subjective questionnaire. Seated within a simulated vehicle, participants could utilize a tablet mounted on the central console to engage in a NDRT. The experimental setup can be seen in the figure below.



Figure 45 - Experimental setup from [42]

Three distinct concepts were employed to deliver information to the driver: Concept A, featuring an audible alert only; Concept B, comprising an audible alert and a text message for each event; and Concept C, incorporating an audible alert, text message, and continuous status information accompanied by graphical warning information (Figure 46).

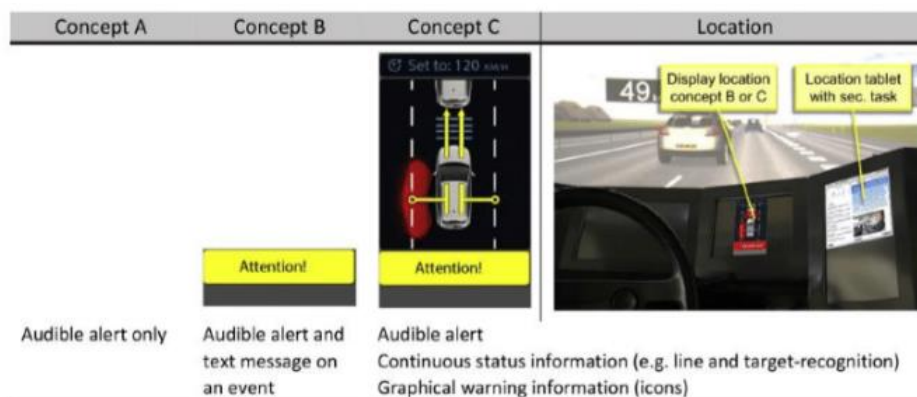


Figure 46 - Concepts design employed by [42]

While small differences were observed between concepts, the framework demonstrated construct validity through consistent measurements of accident avoidance, situation awareness, and perceived usefulness. However, the framework showed limited sensitivity in identifying small differences in optimized support. The authors recommended using the framework's results as a basis for a second round of evaluation based on expert judgment. Despite concerns

about predictive validity, the framework successfully identified the least acceptable form of support, enabling the rapid rejection of inadequate interface support early in the design process. The study concluded that the proposed framework contributes to the efficient development of driver interaction in partially automated vehicles, providing reliable and meaningful assessment data.

D. Sirkin et al. [43] assessed situation awareness facilitated by real-time event alerts in an AV using the Daze app. This application is modeled after the Waze app, originally designed for mobile devices [44]. In Waze, users can review, confirm, or report new alerts to the community, providing information on current driving conditions or risks by visualizing a map displaying nearby users' locations. For instance, if there's an accident on a highway exit lane, the app proactively notifies approaching cars about the situation. On the display, users will have the option to select from two buttons ('yes' or 'no') to indicate whether an accident or construction has been detected (figure 47). Additionally, an alert will prompt the user to confirm if the disorder or any situation causing traffic is still ongoing. Throughout the journey, the app displays a map of the surroundings, depicting the current position of the vehicle. The alert notifications are displayed on the map accompanied by a text sentence and an audible cue.



Figure 47 - Map centered on the car location (left), option to indicate whether the event has been detected by the user (center), confirmation of the location (right) from [43]

To assess how programs or equipment can impact users' awareness of their surroundings, the authors propose applying a technique like that used in the Daze app to personal navigation, AR, or wearable devices. While understanding a driver's awareness is crucial in the context of AV, this technique can also be employed to convey information and provide insights into future circumstances. For instance, Daze could be utilized to evaluate a driver's anticipatory capabilities when operating in a supervisory role.

3. Technologies

This chapter will provide an in-depth exploration of the technological aspects employed in the thesis, encompassing both software and hardware components.

The primary software utilized for the thesis is Unity3D, serving as the core platform for development. As for hardware, a range of devices is employed to create an immersive simulation experience. These include the HTC Vive for VR, the Forza Motorsport Wheel

Bundle Fanatec for steering controls, and the Atomic 3 Motion Simulator for enhanced motion simulation.

3.1. Unity3D

Unity3D is a freely available, cross-platform integrated authoring tool predominantly used for video games, although it can also be applied for simulations, film and animations, architectural visualizations, VR, AR, and other immersive applications. Within the application, there is a store that enables the download of diverse packages for utilization, including both free and commercially available options. These packages offer additional features, tools, assets, and functionalities that can enhance the development process and the overall quality of the project. Unity offers an user-friendly interface, providing intuitive workflows that cater to developers of all levels of experience. Its visual editor empowers developers to effortlessly create, manipulate, and organize objects, scenes, and components using a convenient drag-and-drop system. The extensive documentation and supportive community make it an accessible and approachable platform. It supports the use of C# as a high-level, object-oriented programming language, facilitating more convenient management of software logic.

3.1.1. CiThrus Traffic Simulator and Windridge City

The CiThrus Traffic Simulator was specifically developed to support the advancement of the CiThrus system. The traffic simulation was developed by extending and modifying the CiThrus package for the project's needs. Constructed using Unity3D and leveraging the AirSim Windridge City Asset, this simulator focuses on capturing 360-degree traffic imagery across different areas of the city. This simulator allows for the creation of videos from multiple cameras positioned throughout the city simultaneously.

CiThrus was incorporated into the project and adapted to serve as the core for overseeing NPCs within the simulation and autonomous driving. Taking charge of this are two manager classes, functioning as script components assigned to the NPCManager. Traffic flow is managed by the VehiclesManager class, while the PedestrianManager class handles the population of the scene with pedestrians and their movement. Vehicles utilize the VehicleAI script component, enabling movement along editable, predefined paths within a waypoint network. Conversely, pedestrians utilize Unity's Navigation Mesh System, with the manager responsible for spawning them, assigning destination objects, and activating their NavMesh components. Information regarding the routes that vehicles can take is stored in files, using sets of three floating-point values to represent the world coordinates of each waypoint. Each waypoint represents a node on the path, and the connection between linked nodes is interpreted as a straight line. Additionally, the maximum speed at which vehicles' VehicleAIControllers can traverse each path segment is stored as supplementary data along the path's edges. For this purpose, a graphical editor was developed to facilitate practical and convenient editing of vehicle routes. This tool was created in previous projects, allowing users to modify paths. Accessible within Unity Editor's play mode, it features a flying camera enabling free movement

around the scene. The editor provides a 3D graphical representation of routes and a toolkit to effortlessly insert or delete nodes, establish and break connections, relocate nodes by dragging, and adjust the maximum speed limit along path segments. It serves to customize the user's vehicle path within the scenario, determine NPC vehicle count and their respective routes, as well as manage the count and positioning of pedestrian NPCs in alignment with the project's requirements.

3.1.2. NWH Vehicle Physics 2

The NWH Vehicle Physics 2, a comprehensive vehicle simulation package specifically designed for Unity, has been used in the project [45]. This package was chosen for its realistic and highly configurable nature, as well as its user-friendly interface. It offers a range of notable features, including:

- **Modular vehicle architecture:** Each aspect of the vehicle's behavior can be enabled or disabled, either manually or using the built-in level of detail (LOD) system. Additionally, runtime adjustments can be made to various components such as suspension, powertrain, friction, and visual effects.
- **External modules:** NWH Vehicle Physics 2 allows for the addition or removal of features as needed. Developing modules is straightforward, and they can significantly influence the vehicle's behavior.
- **WheelController 3D:** This component handles ground detection, offers high configurability, and utilizes a sophisticated friction model. It accurately simulates the behavior of the vehicle's bottom half and wheels.
- **Inertia, torque, and angular velocity:** These characteristics contribute to the accurate behavior of the vehicle in all situations. For example, if the vehicle has an ignition and the clutch engaged, it can be started by rolling it downhill.
- **Engine simulation:** The engine module can replicate fuel consumption, losses, rev limiter, adjustable power curve, and forced induction (supercharger or turbocharger). It also accommodates engine stalling and supports both electric motors and internal combustion engines.
- **Manual or automated clutches:** The package includes adjustable slip torque and engagement for manual or automated clutches.
- **Realistic gear ratios and configurable shifting behavior:** These features make NWH Vehicle Physics 2 suitable for simulating various types of vehicles.
- **Aerodynamic simulations:** The package includes drag and downforce simulations, contributing to a more authentic vehicle behavior.

These features represent just a selection of the capabilities offered by NWH Vehicle Physics 2. Its comprehensiveness, well-documented nature, and the completeness of its physics characteristics make it an invaluable asset for simulating an autonomous vehicle (AV) in this thesis.



Figure 48 - NWH Vehicle Physics 2, Demo Scene

3.2. Virtual Reality

In the last few years, VR has emerged as a game-changing technology, revolutionizing how we interact with and experience digital content. With its ability to push the boundaries of interaction and transport users into captivating realms of immersive experiences, VR has ushered in a new era of digital engagement. By creating virtual worlds that enable users to interact with objects within them, VR has unlocked unprecedented possibilities and opened doors to previously unimagined realms of entertainment, education, and exploration. VR has introduced us to a dynamic and interactive form of content consumption. Through the clever integration of advanced headsets, motion tracking devices, and realistic simulations, VR places users at the center of virtual environments, blurring the lines between the physical and digital realms.

3.2.1. HTC Vive

The HTC Vive was a collaborative creation between Valve and HTC, showcasing an array of impressive features. This innovative device boasts two AMOLED panels, each with a 3.6" diagonal size and a resolution of 1080x1200 pixels per eye, resulting in an overall resolution

of 2160x1200. With a refresh rate of 90 Hz and a generous 110° viewing angle, the visual experience is very immersive.

Beyond its optical capabilities, the HTC Vive transforms the user's surroundings into a dynamic 3D space through the remarkable "room scale" technology. This feature enables users to freely move within their physical environment while interacting with the virtual world. The headset incorporates various sensors, including a video camera that displays the immediate surroundings and identifies potential obstacles within the room. In addition, a range of infrared sensors, a gyroscope, an accelerometer, and a proximity sensor work in unison to precisely track the headset's position in space.

To further enhance the immersive experience, the HTC Vive offers exceptionally accurate head tracking and hand-held trackers, enabling seamless interaction with the virtual environment.



Figure 49 - HTC Vive with accessories

3.3. Driving Simulator

Attaining a sense of realism is crucial for the project's success. Hence, a deliberate choice was made to incorporate a motion platform that faithfully recreates the feeling of being inside an authentic vehicle, complete with a seat, steering wheel, and pedals. The steering wheel serves as an input device that not only governs the vehicle's steering actions but also delivers tactile feedback akin to that of a genuine steering wheel. In real-life scenarios, the steering wheel undergoes movements and encounters forces influenced by the road surface and the level of grip available. For instance, on a well-maintained asphalt road, the steering wheel remains stationary and does not exhibit any movement. The feedback conveyed through the steering wheel is pivotal in driving simulation, as it establishes the driver's confidence in the vehicle by providing insights into the grip limit. Incorporating motors within the steering wheel allows for the simulation of numerous real-world situations. This technology, known as force feedback, can apply resistance against the user's force exertion, modifying the perception of the steering wheel's weight or forcefully moving it, thus restricting the driver's unrestricted control over the steering wheel. Equally significant are the pedals, which contribute to the overall experience. The pedals of the latest generation exhibit precise pedal pressure and a comfortable underfoot

sensation. Additionally, they enable independent control over multiple pedals on distinct axes, permitting concurrent braking and acceleration.

3.3.1. Forza Motorsport Wheel Bundle Fanatec

The driving simulator incorporates a comprehensive input device kit to facilitate seamless interaction with the virtual car. The chosen bundle, provided by Fanatec, includes the Forza Motorsport Wheel, consisting of the following components:

- ClubSport Wheel Base V2.5: This motorized base serves as the central component of the system, offering exceptional accuracy. Powered by a robust brushless servo motor and equipped with two Hall magnetic sensors, it ensures precise and responsive control.
- ClubSport Steering Wheel Universal Hub for Xbox One: This versatile hub enables users to configure their driving position by attaching various Fanatec steering wheels. It features a removable control panel and a compact display that provides real-time information such as speed and gear changes.
- ClubSport Wheel Rim GT Forza Motorsport: This customizable steering wheel adapts to different driving scenarios, whether it be racing or navigating city streets. With powerful motors and a maximum rotation angle of 900°, it delivers a realistic and immersive driving experience.
- ClubSport Pedals V3: The pedal unit, which includes the accelerator, brake, and clutch, offers adjustability and exceptional durability. Equipped with Hall Effect magnetic sensors, it ensures precise and long-lasting performance. Additionally, the accelerator and brake pedals feature a vibration system that provides additional feedback during driving.
- ClubSport Shifter SQ V 1.5: This versatile gearbox can be used in both H configuration, with gears ranging from first to seventh, including reverse, and sequential configuration for quick and intuitive shifting.

By integrating this kit of input devices, the driving simulator provides users with a comprehensive and immersive control system, enhancing the overall realism and enjoyment of the virtual driving experience. [46]



Figure 50 - Forza Motorsport Wheel Bundle Fanatec

3.3.2. Atomic 3 Motion Simulator

The Atomic A3 Motion Simulator offers an unparalleled immersive experience with its exceptional features. The system provides a "big simulator feel" through its impressive 27-degree per axis movement range and speeds of up to 71 degrees per second, surpassing other compact systems in the market. Its powerful feedback and high-fidelity motion, operating at 100Hz, authentically reproduces the sensations of road texture, gear changes, and flight turbulence. With a proven track record of robustness and reliability, the Atomic A3 Motion Simulator is ideal for heavy commercial or professional use. Despite its advanced capabilities, the system remains ultra-compact, weighing less than 50kg and easily collapsible within minutes for convenient transportation in a car boot. The simulator's extensive tool-less adjustability caters to riders of all ages and sizes, offering comfortable rides for children and adults ranging from heights of 1.35m/4'5" to 2.00m/6'7". Furthermore, its versatility shines through the interchangeability between driving, flying, and passenger configurations, while the base system provides ample scope for customization, allowing users to tailor their experience to their preferences [47].



Figure 51 - Atomic A3 Motion Simulator

3.3.3. Actuate Motion

Actuate Motion serves as middleware software specifically developed to facilitate the seamless integration of motion hardware support into games and simulation software.

Designed with extensibility in mind, Actuate Motion incorporates a plugin system that enables effortless compatibility with various hardware options, as well as diverse methods of acquiring telemetry. Simplifying the process of selecting the appropriate input plugin, Actuate Motion incorporates a built-in watcher feature. This watcher actively monitors specified processes and automatically loads the corresponding input plugin upon detecting their execution. With this streamlined approach, end users can effortlessly immerse themselves in their favorite games while Actuate Motion seamlessly delivers comprehensive motion support [48].

3.3.4. Motion Compensation

Motion Compensation (MC) refers to algorithms that mitigate the transfer of certain real-world movements made by the user into the virtual environment. Given that this movement is artificially generated solely for conveying acceleration feedback to the user and does not mirror

an actual movement in a real driving scenario, it would be incoherent with what is happening within the simulation. For instance, replicating the linear frontal acceleration of a car would involve tilting the chair backward, causing the user's entire body to lean in that direction. Within the simulation, however, the car accelerating forward doesn't prompt any adjustment in the relative position or tilt of the user's head compared to the seat, as the head remains fixed in place. Hence, it's necessary to nullify this motion. Broadly, MC aims to nullify any motion induced by the motion simulator that could alter the relative positions and orientation of the VR world's and the real world's frames of reference.

MC is achieved by applying the inverse of the transformation induced by the motion simulator to the virtual world camera (representing the player's gaze in VR), effectively nullifying the motion simulator's impact. The MC applied in the simulation was a specialized software solution crafted within Unity. The initial phase required tracking the chair's movement within the simulator, accomplished by securely mounting a Vive Tracker on the chair's top using a custom 3D-printed mount (Figure 52).



Figure 52 - Vive tracker used for motion compensation

Any changes in the tracker's position and orientation relative to the motion simulator's pivot point must be compensated within the VR environment by applying an inverse rotation and translation to the camera (representing the user's head). The spatial data from the tracker can

be readily accessed in Unity using the OpenXR package. By incorporating a Pose Tracked Driver component into an object and designating the Vive Tracker as the source for updating its Pose data, a game object is created within the scene, containing real-time position and rotation information of the tracker. At the start of the simulation, a second static game object, referred to as "TrackerClone," is instantiated. This object serves as a reference point for the initial position and rotation of the Tracker object. Lastly, a custom C# class executes the MC algorithm, which involves calculating the positional and rotational differences between the Tracker and the TrackerClone. This information is then utilized to modify the player's position and rotation, effectively compensating for the motion generated by the chair.

Given that the player is part of a hierarchy of game objects within the simulation, careful consideration is required to ensure that MC calculations are accurate and coherent. The Player object is positioned and maintained inside the vehicle by another custom software implementation, composed of a base class and a derived class that is used as a C# script component in the Unity scene, called SquaredLimitTracking. The Player is also the transform parent of the XROrigin that holds all VR devices driven object inside its hierarchy. The SquaredLimitTracking mechanism forces the main camera, which is driven by the Vive HMD, to stay in a limited adjustable volume, which in this case is the cabin of the car, blocking the user's movement if he would put its head outside the cabin or penetrate in its interiors. Hence, MC corrections are applied to a higher-level hierarchy, which comprises two empty game objects: a position controller and a rotation controller. By being parented to these controllers, the Player object is adjusted accordingly.

4. Design

This chapter will explore the decisions made regarding the choice of interface for the vehicle, scenario configuration, and the utilization of distinct notifications for critical events.

4.1. Scenario configuration

4.1.1. Graphic Restyle

The WindridgeCity scenario serves as the backdrop for the simulation, where numerous traffic events occur. A graphic restyling has been applied to the original scenario, involving the addition of new assets to enhance the realism of the simulation and increase user immersion. Some assets were created using Blender, while others were downloaded from the Unity Asset Store.

The road signs originally present in WindridgeCity have been replaced with Italian signs created using Blender, as the test group for the experiment consists of Italian participants.



Figure 53 - Updated road sign assets

Additionally, a roadwork scenario has been introduced near a closed-off road tunnel on the highway, and a separate multi-vehicle accident scenario has been added.



Figure 54 - Roadwork scenario



Figure 55 - Car crash event scenario

To provide context for the first jaywalking event within a suburban area of the map, several bus stops have been incorporated. In terms of NPCs, the simulation now includes a Sedan car model downloaded from the asset store (Figure 57). Furthermore, the pedestrian models have been updated (Figure 56); their design has been modified using Blender.



Figure 56 - Bus stop and updated pedestrian models



Figure 57 - Updated NPC car models

4.1.2. Journey and Events

To comprehend the occurrence of traffic events in AV's simulations, an in-depth analysis of the current state of the art has been conducted. This analysis identified several types of traffic events, that have been classified into the following groups: TORs, maneuvers, degraded driving conditions, critical, domain transitions, and traffic status updates. For the simulation, at least one event of each group has been implemented. The figure below shows the events occurring during the simulation.

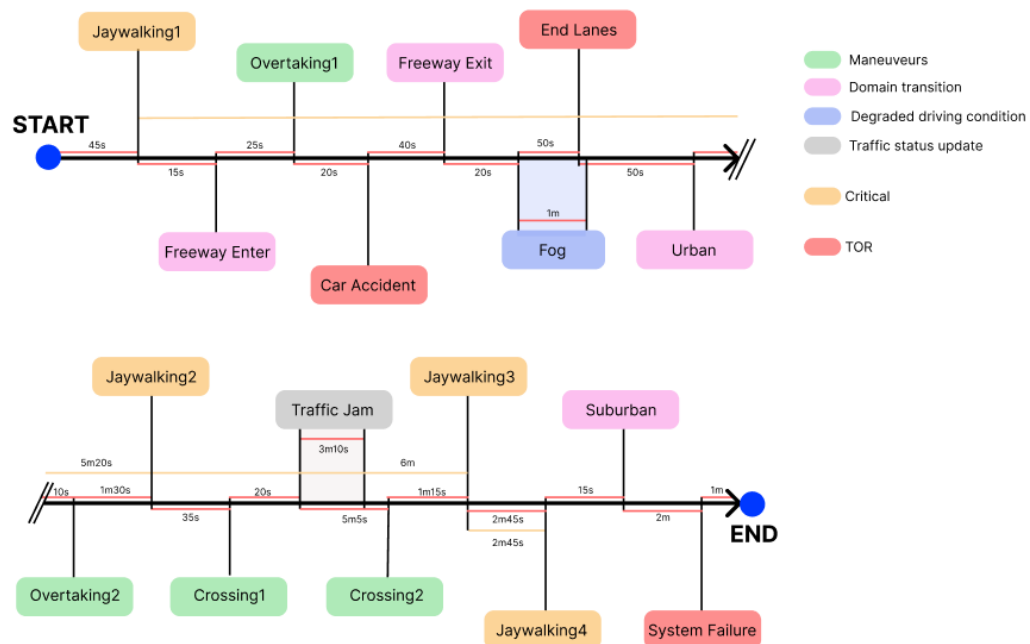


Figure 58 - Traffic events occurring during the simulation

Initially, the vehicle commences its journey from a suburban location, requiring the driver to maneuver the vehicle from a parking slot onto the street before transitioning to autonomous

mode. The first traffic event arises when a pedestrian (referred to as Jaywalker1) crosses the street from one bus stop to another. Subsequently, the car enters a freeway (Freeway Enter) and performs an overtaking maneuver (Overtaking1) (Figure 59).

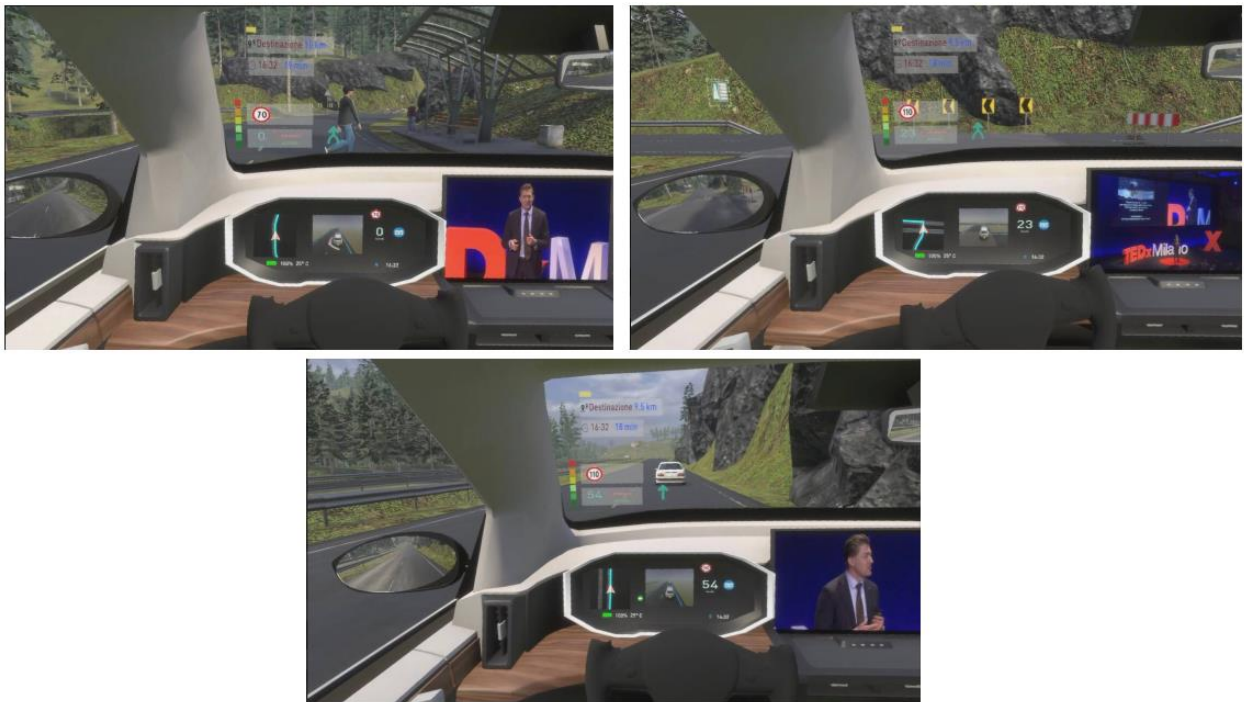


Figure 59 Jaywalking1 event on top left, Freeway enter event on top right, Overtaking1 event on bottom

Along the way, a road sign indicating a car accident prompts the first TOR of the simulation (CarAccident), necessitating the driver to assume control until the accident site is passed. Once past the site, autonomous control can be restored. (Figure 60).



Figure 60 - Car crash event: TOR on the left, manual mode on the right

Upon exiting the freeway (Freeway Exit), the vehicle approaches a suburban area characterized by sloping roads, numerous turns, and dense fog (Fog) suddenly engulfs the surroundings. The fog becomes so thick at one point that the system loses visibility of the street lines, triggering the second TOR (End Lane). As the fog dissipates and the street lanes become visible again, autonomous mode can be reactivated. (Figure 61).



Figure 61 - Freeway exit event on the top left, Fog event on the top right, End lanes event on the bottom

Subsequently, the vehicle enters the city (Urban), where the first traffic event involves another overtaking maneuver of a vehicle (Overtaking2). This is followed by a jaywalker (Jaywalker2) obstructing the vehicle's path while crossing the street. A pedestrian crossing event (CrossingPedestrian1) occurs, this time on a zebra crossing, before encountering a traffic jam (Traffic) as depicted in Figure 62.

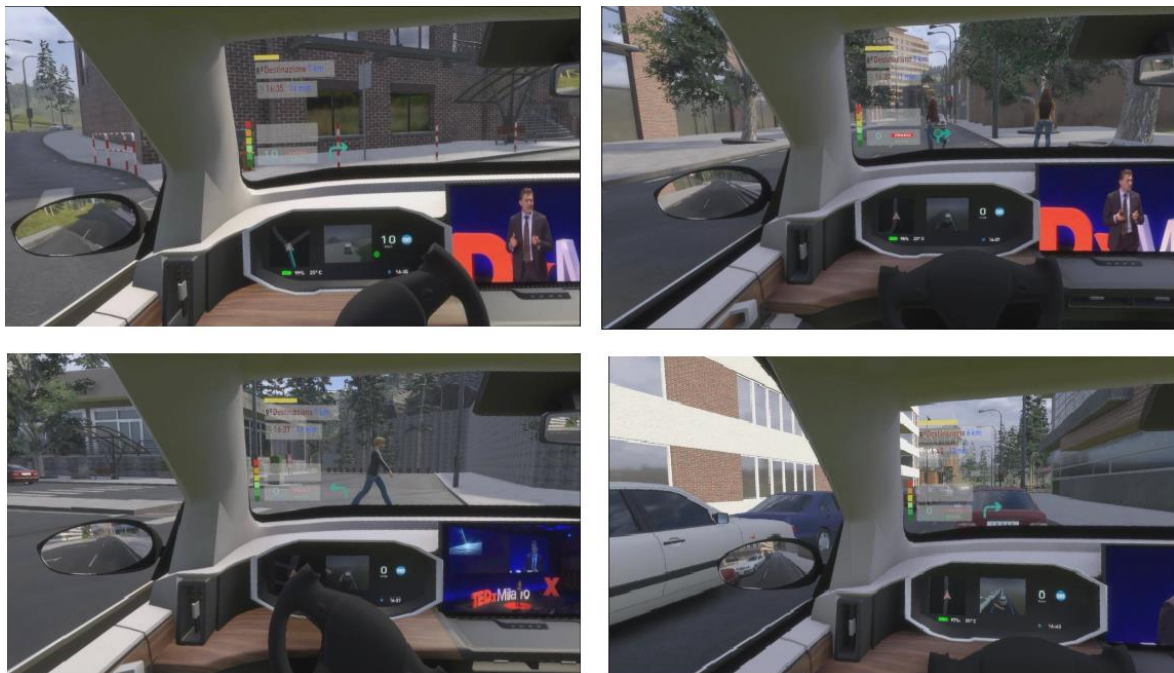


Figure 62 - Urban event on the top left, Jaywalking2 event on the top right, Crossing1 event on the bottom left, Traffic event on the bottom right

Another crossing pedestrian event (CrossingPedestrian2) and two more jaywalking events (Jaywalker 3 and Jaywalker 4) take place within the city limits.

Finally, the vehicle exits the city and enters another suburban area (Suburban) characterized by numerous turns and sloping roads. At a certain point, a system failure disrupts the reliability of the autonomous mode, prompting the third and final TOR (SystemFailure)(Figure 63). The driver assumes control for a short distance until the vehicle reaches its destination.



Figure 63 - Suburban event on the left, System failure event on the right

The simulation's duration is approximately 19 minutes. The following image provides an overview of the map and encountered traffic events.

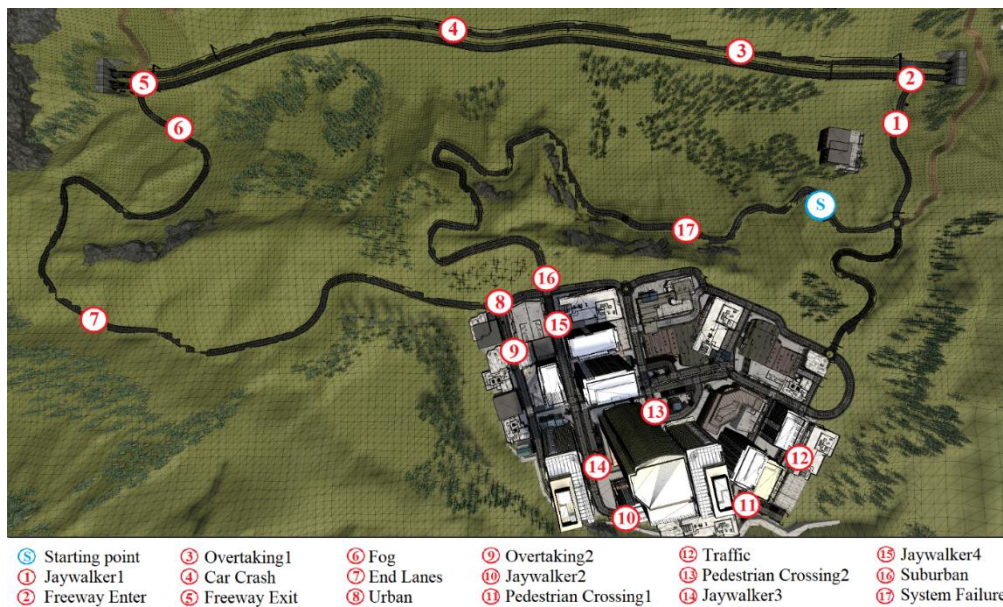


Figure 64 - Roadmap with events

4.2. Autonomous Vehicle Interface

The interface developed for the project consists of both real vehicles inspired components (such as the IC display and touchscreen display) and prototyped interfaces coming from the state of the art of autonomous vehicles (WSD).

The decision to develop the IC display was based on the premise that people are more accustomed to traditional dashboards rather than more advanced technologies. This choice was made with the intention of ensuring user familiarity and ease of use. Furthermore, the IC display retains a level of simplicity that aligns with users' existing mental models of vehicle interfaces. The second component is a touchscreen display that serves as the infotainment system. It is positioned at the central console of the car, ensuring that not only the driver, but the other passengers can enjoy its content. Both the IC display and the touchscreen display are inspired by the Tesla Model X [49]. The third component is an AR-WSD, which combines screen-fixed elements and world-registered AR overlays. From previous studies, it emerges that the support of AR-HUD is preferred over conventional displays and is perceived as more understandable, useful and resulting in higher trust from the driver [21]. The WSD display design was derived from a prototyped version that underwent thorough testing and evaluation, demonstrating a notable enhancement in driver situational awareness [22]. Lastly, the auditory interface provides, using speech and abstract sounds, information and alerts about incoming events.

4.3. Jaywalking Notification

Certain traffic events necessitate sudden braking, potentially causing discomfort and concern for the driver, even though the system can effectively manage these situations and doesn't require the driver to take over control of the vehicle. Alerting the driver about these incidents is critical for building trust in the system by providing pertinent information. However, determining the optimal notification method that enhances trust without causing undue irritation is a significant consideration. Furthermore, it's essential to ensure that the driver isn't misled and to emphasize that they do not need to intervene with vehicle control, as the autonomous system is entirely capable of dealing with such circumstances. Even if the system can handle these events itself, non-communicated autonomous car braking has significant negative consequences for the driver while a voice warning not only mitigate anxiety but also enhance alertness and restore a sense of control for passengers [28]. However, more information does not necessarily lead to more trust and may in fact negatively affect cognitive load [3]. It is imperative to thoroughly explore methods of notifying these types of events to the driver to enhance trust in the system, while also ensuring that the notifications are not overly intrusive or detrimental to trust.

As mentioned before, a voice warning can enhance alertness, reduce anxiety and restore a sense of control by giving an explanation about what is happening to the driver.

Q. Zhang et al. [50]. presented a review of the literature to better understand what has been done and what should be done regarding AV explanations.

| SAE Level Name | Time Period | | | | | |
|---------------------------------|-------------|------|------|------|------|---------------|
| | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 |
| Level 0: No Automation | | | | | | |
| Level 1: Driver Assistance | | | | | | |
| Level 2: Partial Automation | [4] | [5] | | | | |
| Level 3: Conditional Automation | | | [3] | [6] | [2] | |
| Level 4: High Automation | | | | | [1] | |
| Level 5: Full Automation | | | | [7] | [8] | [9] [10] [11] |

Table 15 - Summary of literature review by automation level (SAE) from [45]

They considered two explanation attributes: content and timing. Content refers to the message conveyed by the system during an event. It can be a message providing an explanation about what the vehicle is going to do or just did (what) or about the reason behind its action (why). This explanation can be provided alone or as a combination of the two (what + why). The timing refers to the moment when the content explanation is provided. It can be either provided before the vehicle action (before action) or after (after action).

| Explanation attribute | Category | Description | References |
|-----------------------|---------------|--|------------------------------|
| Content | What-only | What-only explanation provides descriptions of the vehicle action itself (i.e., what will/did the vehicle do?) | [4], [9] |
| | Why-only | Why-only explanation describes the reasoning for action (i.e., why did the vehicle perform that action?) | [4], [5], [9] |
| | What + Why | What + Why explanation describes the vehicle action itself and the reasoning for the action. | [1], [3], [4], [6], [8] [10] |
| Timing | Before action | The time to provide explanations is before the AV acts. | [1], [3], [4], [5], [6] |
| | After action | The time to provide explanations is after the AV acts. | [1], [2], [9], [10] |

Table 16 - Summary of literature review by content and timing from [45]

The review highlights how the provision of an explanation prior to the autonomous vehicle's actions is closely tied to increased positive attitudes (like trust, anthropomorphism, acceptance, preference, situational awareness, sense of control, and alertness) and decreased negative emotions. On the other hand, the studies providing explanation after the event have mixed results. About the content explanation attribute, results show that the “what-only” explanation led to the worst performance and lowest acceptance compared to the “what + why”, “why-only” explanation and the “no-explanation” condition. About the “why-only” and the “what + why” explanations, results show how both have shown positive results in trust and acceptance. For the thesis experiment, the “what + why” message has been chosen as one of the configurations (Speech) for notifying jaywalker events. While also why-only configuration provided positive results, they were tested in a static simulator with a SAE lv2 autonomous vehicle [8] while the results for the “what + why” configuration were related to motion simulation study with a SAE lv3 autonomous vehicle [31], more related to the thesis experiment. As underlined by the authors, none of the studies examined the effectiveness of a particular modality over another modality. Prior research in vehicle display design has identified variations in the efficacy of presenting signals through visual versus auditory means. For instance, auditory signals have generally been proven more effective than visual signals in conveying hazard alerts quickly and indicating the magnitude of potential hazards [51] [52]. However, it's worth noting that the auditory modality has also been linked to heightened levels

of annoyance when compared to the visual modality [53]. Hence, in the second configuration, a visual notification is provided to jaywalkers. This involves a 3D AR bounding box displayed via the WSD, which highlights the jaywalker throughout the crossing period (Visual). To explore the combined use of visual and auditory modalities, a third arrangement was devised. In this configuration, the jaywalking event is alerted through a fusion of the previous two configurations (Visual + Speech). Auditory cues outside of speech can be designed to provide an alert for an approaching event through the manipulation of acoustic variables. From previous studies, there are acoustic characteristics that contribute to a sound being unambiguously perceived as an urgent alarm within a vehicle context [40][52]. As the AV is capable of autonomously managing this event, the abstract sound is intentionally crafted to avoid being interpreted as an urgent alarm by ensuring its acoustic variables remain below a specified threshold. In the fourth configuration, the jaywalking event is alerted using a combination of visual feedback and specially crafted abstract sound (Visual + Abstract). The fifth and final configuration serves as a baseline, with no feedback provided for this event (Baseline).

5. Implementation

5.1. Timeline event system

Unity's Timeline was utilized as the logical framework for managing all events within the simulation. Known for its applications in gaming sequences, cinematic content, audio sequences, and complex particle effects, Timeline proved to be a versatile tool. By adding timeline tracks and associating them with GameObjects, specific methods from scripts owned by those game objects could be called at predetermined times using timeline event clips. This streamlined approach facilitated the efficient management and timing of events, as well as the behavior of various game objects. For instance, the InterfaceManager GameObject housed dedicated methods for each event and assumed responsibility for handling the feedback generated by different interface components.

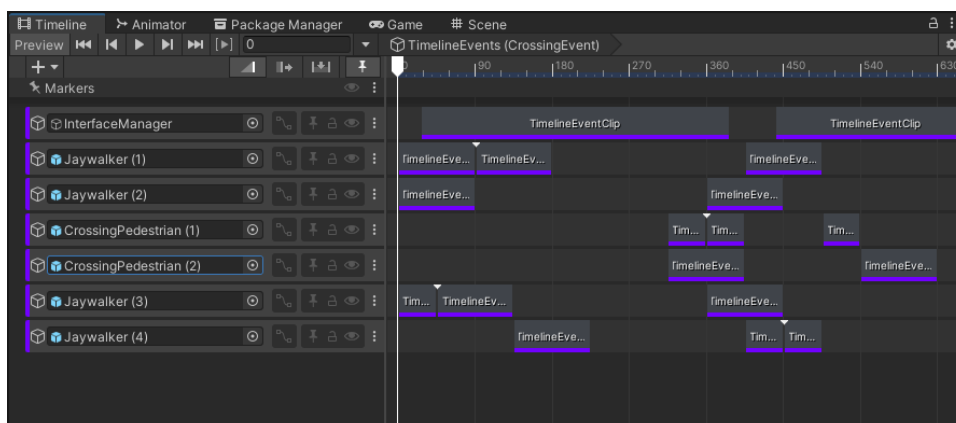


Figure 65 - Timeline for the Pedestrian Crossing events

5.2. Interface

5.2.1. Instrument cluster

The instrument cluster, created using Blender, presents diverse types of information by using Unity Canvas.



Figure 66 – User vehicle's dashboard

The bottom part of the display features the vehicle's state of charge, depicted through a percentage value and a battery icon, alongside additional information such as time (taken from Unity struct System.DateTime) and temperature. The script CurrentBattery handles battery depletion, decreasing both the percentage (created with a TextMeshProGUI) and reducing the green bar (a 2D image) inside the battery icon (created with a 2D sprite) by scaling the image dimension over time. Considering the lifespan of current electric vehicles, the decision was made to assign a battery duration of 10 hours to the simulated vehicle, serving as the basis for determining the rate of battery depletion. The gear position indicator (PRND) is positioned at the bottom right of the display, where a TextMeshProGUI dynamically presents the corresponding letter based on the current gear engaged. While the autonomous mode is active, the letter will be displayed in blue color. Information about the current speed of the vehicle and the current speed limit is also displayed in the IC. The current speed is a TextMeshProGUI which changes according to the vehicle speed, the speed limit is a 2D Sprite corresponding to the actual speed limit which dynamically updates throughout the simulation as the vehicle transitions between different domains (e.g., from suburban area to urban area) by the script InterfaceManager. At the center of the IC, the world in miniature is displayed, showing the car and surrounding vehicles, as well as the lanes of the street. As an additional visual cue, the lanes undergo a color transformation from white to blue when the autonomous mode is engaged. This distinctive color change serves as a clear indication of the active autonomous functionality. Furthermore, a proximity sensor is employed to detect the proximity of vehicles, pedestrians, or obstacles in relation to the vehicle. When the distance between the vehicle and another object falls below a predetermined threshold, the ProximityInMiniature script promptly activates a corresponding indicator to provide a clear visual alert. Importantly, the indicator's

activation is specifically tailored to the direction from which the object is detected, ensuring that the driver is promptly informed of the potential hazard's location. To represent the world in miniature, a camera with an appropriate culling mask that only renders the target objects (by specifying the layers) is used, and its view is rendered on a texture used for the portion of the IC destined to this view. On the left of the IC, the navigation information is displayed. Following the same principle of the world in miniature, a camera with an appropriate culling mask showcases a white and orange triangle icon (representing the car, placed on top of it), and the path is highlighted with blue color over the road and renders its view on a texture. In order to communicate the availability of the autonomous mode, an icon depicting a steering wheel in gray and white is displayed on the central-right section of the instrument cluster (IC), positioned next to the current speed information. Once the user activates the autonomous mode, the color of this icon transitions to blue, serving as a visual indication that the autonomous mode is now engaged. In situations where the system necessitates user intervention, a relevant text message will be presented, prompting the user to assume control of the vehicle.

5.2.2. Touchscreen display

The touchscreen display serves as a primary source of entertainment within the vehicle. When the autonomous mode is engaged, a video is played on the screen to enhance the user experience. This is achieved by utilizing a game object equipped with the VideoPlayer component, enabling the playback of video files directly onto a texture assigned to a plane GameObject.



Figure 67 - Touchscreen display playing a video

When an event causing a TOR occurs, the video playback is paused, and a circular progress bar, displayed on a black background, along with a text message, is utilized to alert the user to assume control of the vehicle. The script TouchscreenTOR is responsible for managing this process. During a TOR event, the TouchscreenTOR script initiates a gradual transition of the progress bar's fill amount using interpolation (lerping). Additionally, as the progress bar

reaches its halfway point, the color of the progress bar and the accompanying text message change from orange to red. [12]. While the manual is active, a pause icon is displayed over the paused video. The vehicle will stop if the driver doesn't take manual control before the bar reach the end.



Figure 68 - TOR on the touchscreen display

5.2.3. Windshield display

The windshield display provides both screen-fixed elements and world-registered AR overlays. The screen-fixed elements are presented on a Canvas located behind the windshield and divided into various sections, as shown in the figure below:



Figure 69 - Windshield display screen fixed information and AR detection box

The driving panel is located on the bottom left corner of the windshield. It provides information about the current speed and uses two icons (2D sprites) that highlight when the vehicle is accelerating or braking. The AccelBrakeBinary script computes the average of the most recent 50 samples of the z component of acceleration, specifically the frontal component. The acceleration is calculated by the script Accelerometer by taking the time derivative of the vehicle speed, obtained from the "RigidBody.Velocity" property. When the average acceleration is above zero, the corresponding icon will increase its alpha value, indicating that the vehicle is accelerating. Conversely, if the acceleration is negative, the braking icon will raise its alpha value so that the driver knows that the vehicle is slowing down. To avoid rapid toggling of the binary icons due to the acceleration value fluctuating around the threshold, a timer is employed, allowing each icon to be switched on or off at most once every 0.5 seconds. The driving panel information is made more visible to the driver by utilizing a gray background.

The traffic regulation panel is positioned directly above the driving panel, utilizing a gray background. It serves the purpose of presenting crucial information such as the current speed limit and displaying traffic signs to enhance the user's awareness of the vehicle's understanding of the surrounding environment. For each traffic sign type, a timeline event has been established, and a trigger is placed corresponding to each sign. This arrangement allows the script InterfaceManager to activate the respective 2D sprite icon on the panel when the vehicle reaches a certain distance from the sign. In the case where multiple signs are active simultaneously, the position of the 2D sprites is adjusted to handle such scenarios effectively. Furthermore, a 2D sprite informs the user when the vehicle is slowing down due to a hazard.

The time and destination panel, located on the top left of the windshield, displays crucial information including the current time, remaining time, and distance to the destination. The current time is obtained in the same manner as the one provided by the instrument cluster. The remaining time and distance to the destination are constantly updated by the InterfaceManager script using the events placed along the route, serving as checkpoints to refresh these variables. In addition to the aforementioned information, the time and destination panel also features a clock icon and a location icon. Furthermore, a gray background is utilized to provide a visually cohesive and consistent presentation for the panel.

The navigation panel, located on the right of the driving panel, presents information regarding the upcoming turn directions through the utilization of distinct 2D sprites for each direction. This functionality is achieved through the timeline event system, wherein a trigger is positioned after each turn to effectively convey the information about the subsequent turn.

The traffic light panel, positioned to the right of the navigation panel, employs a traffic light icon (2D sprite) to effectively communicate the currently active color light to the user. The script "TrafficLightProgression" manages this functionality by utilizing two types of triggers. These triggers are responsible for displaying or hiding the traffic light icon before and after each traffic light, while also updating the source from which the information about the currently active color light is obtained.

The automation confidence bar is situated on the left side of the driving panel and provides a visual representation of the autonomous system's confidence level, which is divided into six distinct levels. This bar is exclusively displayed during the autonomous mode. Throughout the simulation, as a result of dense fog, the confidence level gradually diminishes until a TOR is initiated due to the system's inability to accurately recognize the road lanes.

Positioned at the bottom-central part of the windshield, the high-priority panel showcases a blinking 2D sprite whenever a TOR takes place. This sprite serves as a visual indication to the driver, providing information regarding the occurrence of the TOR and the underlying reason for it.

The AR overlays are positioned in the world scene. Bounding boxes (2D sprites) are used to highlight both the encountered traffic signs and traffic lights. The implementation follows a similar approach as the traffic sign icons in the traffic information panel, utilizing timeline event triggers positioned near each road sign and specific traffic light detection used to show the traffic light icon in the traffic light panel. As the vehicle approaches an intersection without a traffic light, the oncoming traffic indicator notifies the user about the presence and direction of other vehicles. This feature enables the driver to be aware of approaching cars without the need to lean or turn their head.



Figure 70 - Oncoming traffic indicator notifying a car approaching from the right

5.2.4. Auditory Interface

The auditory interface incorporates both verbal and abstract sound feedback elements. To convey specific events to the driver, a natural-sounding female voice generated through text-to-speech technology is employed. This voice is reproduced using an AudioSource that contains an AudioClip, played by the InterfaceManager when a traffic event is initiated via a timeline event trigger. To ensure that the speech feedback is clearly heard, the volume of the NDRT is temporarily lowered while the speech clip plays, returning to normal afterwards. Additionally, for TORs, the speech feedback is complemented by an alarm-like abstract sound, intensifying the perceived urgency of the event and emphasizing the driver's imperative to assume control of the vehicle, following guidelines of [37] [38].

5.3. Traffic Events

- For the crossing and jaywalking events, a timeline event is triggered when the vehicle enters a designated area, which activates the pedestrian's walking animation. The pedestrian continues to walk until they reach their intended destination. Using raycast, the vehicle identifies the presence of a pedestrian and engages the braking system. Once the pedestrian has safely cleared the street, the vehicle can resume its course.
- For overtaking events, a distinct pathway has been allocated for the overtaken vehicle, with a reduced speed set to ensure a seamless overtaking maneuver. This enables the overtaking to occur naturally as the vehicles adhere to their predefined paths. Additionally, a specific trigger ensures precise timing for the overtaking event to take place.
- When a TOR occurs, a first trigger initiates the TOR, activating the transition mode. During this mode, the driver is informed about the TOR and its reason. They must take control of the vehicle while it gradually slows down until it comes to a complete stop, by reaching a second trigger. Once control is assumed, the manual mode remains active until another trigger marks the area where autonomous mode is disallowed due to the TOR-causing event. Once the vehicle exits this area, autonomous mode becomes available again.
- The implementation of fog involves using a Unity Volume, which creates a visual effect to simulate the presence of fog in the environment. It allows the customization of different properties such as fog attenuation distance, base and maximum height and more.
- The traffic jam scenario is implemented by adjusting the duration of traffic lights. This causes a buildup of cars near a specific traffic light until the user's vehicle arrives in a specified location, mimicking a traffic jam situation. The duration of red and green

signals has been modified to create a situation where vehicles remain stacked in that area for a few minutes.

6. Experiment

This chapter will discuss the experimental procedure, the metrics utilized, and the results.

6.1. Procedure

23 participants (15males, 8 females), ranging from age 21 to age 58, with a valid driving license took part to the experiment. After being welcomed, they completed a brief demographic questionnaire, which gathered general information as well as inquired about their driving experience, experience with VR, motion simulators, and their opinions regarding AVs. Subsequently, the experiment was comprehensively explained to the subjects. To acquaint them with the simulator, they engaged in a demonstration where they had the opportunity to drive the simulation's vehicle, activate and deactivate the autonomous mode, familiarize themselves with the interface and observe how a TOR is communicated. The demo lasted approximately 5 minutes. Afterward, the subjects performed a total of 5 runs, one for each notification type for the jaywalker event. Every run took roughly 19 minutes. To eliminate any bias due to exposition 'order, a Latin square design was employed. Both objective and subjective measurements have been gathered during the experiment. Finally, participants were thanked and rewarded with biscuits at the end of the experiment. The experimental session lasted approximately 2 hours and 30 minutes.

6.2. Metrics

After each run, the participants answered questionnaires regarding both the entire simulation, as well as specific questionnaires about jaywalker notifications and TOR.

In relation to the entire simulation run, the dependent variables encompass the subjective attitudes of the participants. These attitudinal dimensions include trust, mental workload, and anxiety as measured factors.

Trust was measured using 7-point Likert scales with six dimensions [6]: competence, predictability, dependability, responsibility, reliability and faith (1: not at all; 7: extremely well) which consist of an adapted version of the Muir questionnaire [54].

Anxiety comprised the four subscales describing the AV experience: fearful, afraid, anxious and uneasy measured using 7-point Likert scales of [6]. In conclusion, to assess mental workload, the raw NASA-TLX, a self-reported mental workload assessment tool, with a 0 to 20 scale was employed [55]. The raw NASA-TLX comprises six subscales that gauge mental, physical, and temporal demands, as well as frustration, effort, and performance.

Regarding the jaywalker notifications, the first part consists of a section of the Extended User Experience Questionnaire (UEQ+) [56], on the categories of Trustworthiness of content, Quality of content and Value. The overall score is computed using the authors' provided

analytical tool. Each category is assessed based on four pairs of characteristics. A score of 1 indicates alignment with the first adjective, while 7 corresponds to alignment with the second adjective. Scores falling in between are interpolated accordingly. In each category, the user is also prompted to assign a score on a consistent scale, reflecting the level of importance they assign to that category. A rating of 1 implies irrelevant while a rating of 7 denotes extremely important. The scores attributed to individual characteristics are combined to calculate the category score. Subsequently, each category score is subjected to a weighted mean based on its importance, resulting in the overall UEQ score.

Furthermore, with an ad-hoc questionnaire, participants also rated the perceived urgency (PU), perceived annoyance (PA), alerting effectiveness, and message clarity of each notification using a 7-point Likert scale adapted from [57]. Finally, the intrusiveness of the notification is measured with a seven-point scale with 7 dimensions: distracting, disturbing, forced, interfering, intrusive, invasive, obtrusive, with a rate ranging from 1 (strongly agree) to 7 (strongly disagree) [58]

The same PU, PA, alerting effectiveness, message clarity and intrusiveness scales are used for the TOR notification.

At the end of the 5 runs, participants also answered other questionnaires. First, the simulator sickness questionnaire (SSQ) [59] to assess if any symptoms of simulator sickness occurred. After that, participants had to explicitly rank the 5 configurations for jaywalking notification for the following metrics: trust, message clarity, PU, PA, alertness effectiveness, intrusiveness, and overall preference. Furthermore, the immersion and presence questionnaire (IPQ) was used to evaluate the sense of immersion and presence inside the virtual environment. It has four subscales, i.e. general presence, spatial presence, involvement, and realism [60]. Lastly, to evaluate the autonomous vehicle interface, a section of the UEQ+, on the categories of Efficiency, Usefulness, Novelty, Clarity, Intuitive use, Aesthetics, Trustworthiness of Content, Quality of Content and Value. During each run, the eye gaze behavior was tracked as an objective metric.

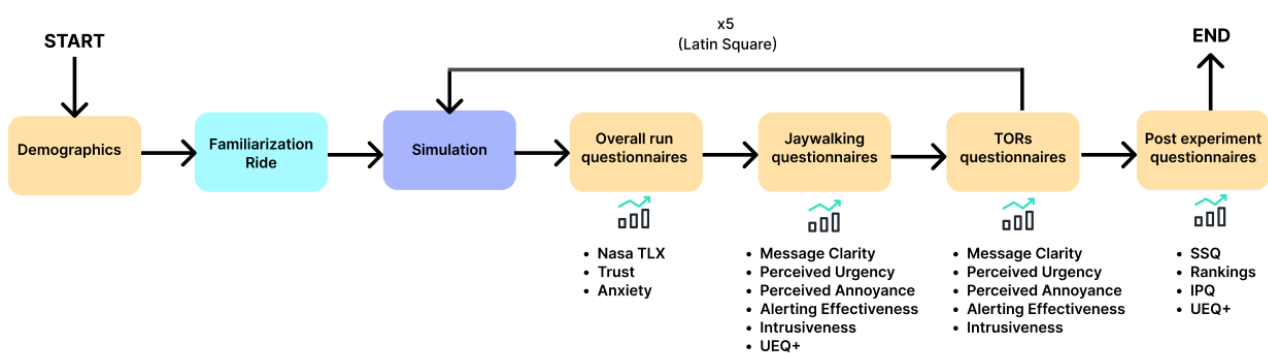


Figure 71 - Experimental Protocol

6.3. Results

All the collected data gathered for each configuration (so everything excluding the SSQ, the IPQ and the UEQ+ regarding the interface made at the end of the experiment) undergoes statistical evaluation using Friedman’s Test [61], with a predetermined alpha level of 0.05. Confrontations resulting in a p-value greater than 0.05 are deemed not-statistically significant. The same conditions regarding p-values apply for pairwise comparisons between categories, involving data aggregated from two combined categories across all experiment runs. The interpretation of effect-size values for each pairwise comparison was as follows: an absolute value ranging from 0.25 to 0.40 represents a small effect, from 0.40 to 0.80 signifies a medium effect, and anything exceeding 0.75 in absolute value indicates a large effect.

6.3.1. Trust, Anxiety and Mental Workload

About the seven dimensions of the trust survey, competence yields statistically significant difference in the group with a p -value=.116. The measurement in the Baseline condition reports statistically significant differences with respect to the Visual + Speech condition with a small effect in favor of the latter. The degree of faith yields statistically significant difference in the group with a p -value=.040. The measurement in the Baseline condition reports statistically significant differences with respect to the Visual + Speech condition with a medium effect in favor of the latter. The predictability yields statistically significant difference in the group with a p -value=.005. The measurement in the Baseline condition reports statistically significant differences with respect to the Visual + Speech condition with a medium effect in favor of the latter. The measurement in the Visual condition also reports statistically significant differences with respect to the Visual + Speech condition with a medium effect in favor of the Visual + Speech condition. Dependability yields statistically significant difference in the group with a p -value of .096. The measurement in the Baseline condition reports statistically significant differences with respect to the Visual + Speech condition with a medium effect in favor of the latter. The overall score yields statistically significant difference in the group with a p -value=.096. The measurement in the Baseline condition reports statistically significant differences with respect to the Visual + Speech condition with a medium effect in favor of the latter. No statistically significant differences were found for the other pairs in the pairwise analysis.

| # | Item | AVG (C.I.95%) | | | | | p-value (effect size) | | | | | | | | | | |
|---|-----------------------|---------------|-------------|-------------|-------------|-------------|-----------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------------|--------------|-------------|
| | | B | V | S | VS | VA | Friedman | B-V | B-V | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA |
| 1 | Competence | 6.13 (0.3) | 6.26 (0.29) | 6.26 (0.29) | 6.34 (0.27) | 6.21 (0.29) | .116 | .233 (-0.18) | .298 (-0.18) | .036 (-0.32) | .571 (-0.12) | .637 (0.00) | .345 (-0.13) | .772 (0.06) | .345 (-0.13) | .772 (0.06) | .371 (0.19) |
| 2 | Responsability | 6.17 (0.33) | 6.34 (0.21) | 6.3 (0.24) | 6.39 (0.21) | 6.3 (0.24) | .100 | .203 (-0.26) | .233 (-0.19) | .088 (-0.33) | .049 (-0.19) | 1.00 (0.08) | 1.00 (-0.08) | 1.00 (0.08) | 0.345 (-0.16) | 0.637 (0.00) | 1.00 (0.16) |
| 3 | Reliability over time | 6.3 (0.33) | 6.3 (0.33) | 6.43 (0.31) | 6.47 (0.31) | 6.3 (0.35) | .158 | .907 (0.00) | .233 (-0.17) | .071 (-0.23) | .887 (0.00) | .371 (-0.17) | .173 (-0.23) | .841 (0.00) | 1.00 (-0.05) | .148 (0.16) | .173 (0.22) |
| 4 | Degree of faith | 5.78 (0.34) | 5.91 (0.25) | 6 (0.26) | 6.04 (0.27) | 6 (0.26) | .040 | .350 (-0.18) | .072 (-0.30) | .041 (-0.36) | .072 (-0.30) | .345 (-0.14) | .344 (-0.21) | .345 (-0.14) | 1.00 (-0.07) | 1.00 (0.00) | 1.00 (0.07) |
| 5 | Predictability | 5.39 (0.38) | 5.52 (0.36) | 5.69 (0.35) | 5.78 (0.36) | 5.69 (0.35) | .005 | .407 (-0.15) | .064 (-0.35) | .017 (-0.44) | .064 (-0.35) | .071 (-0.20) | .047 (-0.30) | .071 (-0.20) | 1.00 (-0.10) | 1.00 (0.00) | 1.00 (0.10) |
| 6 | Dependability | 5.95 (0.24) | 6 (0.29) | 6.13 (0.23) | 6.17 (0.24) | 6.04 (0.27) | .096 | .789 (-0.07) | .129 (-0.31) | .036 (-0.38) | .587 (-0.14) | .148 (-0.21) | .071 (-0.27) | .765 (-0.06) | 1.00 (-0.07) | .345 (0.14) | .371 (0.21) |
| 7 | Overall trust | 5.82 (0.3) | 5.95 (0.2) | 6 (0.22) | 6.08 (0.22) | 5.95 (0.24) | .051 | .298 (-0.21) | .240 (-0.27) | .047 (-0.41) | .482 (-0.2) | .772 (-0.08) | .148 (-0.26) | .841 (0.00) | .345 (-0.16) | .772 (0.08) | .371 (0.24) |

Table 17 - Trust survey results

About the four dimensions of the anxiety survey, no statistical significance in the group has been found. Nevertheless, the anxiety levels remained consistently low across all configurations. The results are shown in the table below.

| # | Item | AVG (C.I.95%) | | | | | p-value (effect size) | | | | | | | | | | |
|---|---------|---------------|-------------|-------------|-------------|-------------|-----------------------|--------------|-------------|-------------|-------------|-------------|--------------|-------------|--------------|--------------|--------------|
| | | B | V | S | VS | VA | Friedman | B-V | B-S | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA |
| 1 | Anxious | 2.08 (0.55) | 2.13 (0.62) | 2.08 (0.5) | 2.21 (0.55) | 2.04 (0.52) | .967 | .887 (-0.03) | .900 (0.00) | .903 (-0.1) | .705 (0.03) | .885 (0.03) | .935 (-0.06) | .571 (0.06) | .719 (-0.1) | .890 (0.03) | .516 (0.13) |
| 2 | Afraid | 1.78 (0.52) | 1.78 (0.63) | 1.56 (0.38) | 1.6 (0.36) | 1.6 (0.44) | .090 | .789 (0.00) | .088 (0.20) | .33 (0.16) | .203 (0.15) | .173 (0.17) | .58 (0.14) | .173 (0.13) | 1.00 (-0.05) | 1.00 (-0.04) | .637 (0.00) |
| 3 | Fearful | 1.69 (0.44) | 1.65 (0.59) | 1.52 (0.38) | 1.56 (0.36) | 1.6 (0.44) | .406 | .853 (0.03) | .173 (0.18) | .49 (0.13) | .571 (0.08) | 1.00 (0.11) | 1.00 (0.07) | 1.00 (0.03) | 1.00 (-0.04) | .345 (-0.08) | 1.00 (-0.04) |
| 4 | Uneasy | 2 (0.47) | 2.04 (0.61) | 1.78 (0.39) | 1.95 (0.46) | 1.86 (0.51) | .756 | .785 (-0.03) | .167 (0.21) | .886 (0.04) | .429 (0.11) | .197 (0.21) | .0832 (0.06) | .269 (0.13) | .409 (-0.17) | .423 (-0.08) | .678 (0.07) |

Table 18 - Anxiety survey results

About the six dimensions of the NASA TLX, no statistical significance in the group has been found. The results are shown in the table below.

| # | Item | AVG (C.I.95%) | | | | | p-value (effect size) | | | | | | | | | | |
|---|-----------------|---------------|--------------|--------------|--------------|--------------|-----------------------|-------------|---------------|--------------|--------------|---------------|--------------|--------------|--------------|--------------|--------------|
| | | B | V | S | VS | VA | Friedman | B-V | B-S | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA |
| 1 | Mental Demand | 6.04 (1.77) | 5.95 (1.77) | 6.13 (1.87) | 5.86 (1.65) | 5.95 (1.5) | .667 | .830 (0.02) | -.797 (-0.02) | .68 (0.04) | .55 (0.02) | .44 (-0.04) | .888 (0.02) | .833 (0.00) | .511 (0.06) | .370 (0.04) | 1.00 (-0.02) |
| 2 | Physical Demand | 5.13 (1.67) | 5.13 (1.6) | 5.21 (1.82) | 5.43 (1.71) | 5.17 (1.65) | .856 | .785 (0.00) | -.586 (-0.02) | .269 (-0.07) | 1.00 (-0.01) | -.588 (-0.02) | .300 (-0.07) | .782 (-0.01) | .586 (-0.05) | 1.00 (0.01) | .461 (0.06) |
| 3 | Temporal Demand | 5.95 (2.06) | 5.78 (1.87) | 5.86 (1.93) | 6.04 (1.86) | 5.86 (1.89) | .890 | .577 (0.03) | -.712 (0.01) | .916 (-0.01) | .916 (0.01) | -.571 (-0.01) | .546 (-0.06) | .931 (-0.01) | .786 (-0.03) | 1.00 (0.00) | .671 (0.03) |
| 4 | Performance | 16.73 (1.52) | 16.73 (1.71) | 16.69 (1.69) | 16.69 (1.53) | 16.73 (1.71) | .781 | .891 (0.00) | .916 (0.01) | .582 (0.01) | 1.00 (0.00) | .853 (0.01) | .855 (0.01) | 1.00 (0.00) | .855 (0.00) | .785 (-0.01) | .786 (-0.01) |
| 5 | Effort | 6.78 (1.9) | 6.78 (1.9) | 6.73 (1.9) | 7 (1.88) | 6.95 (1.83) | .995 | 1.00 (0.00) | 1.00 (0.00) | 1.00 (-0.04) | .586 (-0.04) | 1.00 (0.00) | 1.00 (-0.04) | .586 (-0.04) | .850 (-0.05) | .422 (-0.05) | .892 (0.01) |
| 6 | Frustration | 6.43 (2.17) | 6.47 (2.19) | 6.43 (2.27) | 6.56 (2.14) | 6.26 (2.2) | .630 | .850 (0.00) | 1.00 (0.00) | .730 (-0.02) | .407 (0.03) | 1.00 (0.00) | .832 (-0.01) | .461 (0.04) | 1.00 (-0.02) | .586 (0.03) | .850 (0.06) |

Table 19 - Nasa TLX results

6.3.2. Jaywalking notifications

6.3.2.1. Ad-hoc questionnaire

Regarding jaywalker notifications, the message clarity score yields statistically significant difference in the group with a p -value<.001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other notification conditions with a large effect favoring the other conditions. Between the other jaywalker notifications, the confrontations that show statistically significant results are Visual in respect of Speech, Visual in respect of Visual + Speech with a large effect favoring the second condition over the Visual condition and Visual in respect of Visual + Abstract with a small effect favoring the latter, Speech in respect of Visual + Speech with a large effect favoring the latter, Visual + Speech in respect of Visual + Abstract with a large effect favoring the former condition. No statistically significant differences were found regarding the Speech in respect of the Visual + Abstract condition.

The PU score yields statistically significant difference in the group with a p -value<.001. The measurement in the Baseline condition reports statistically significant differences in the respect of the Speech, Visual + Speech, Visual + Abstract conditions with a large effect disfavoring the Baseline (having lower score, meaning it was perceived as less urgent). No statistically significant differences were found in respect of the Baseline in respect of the Visual condition. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of Speech with a medium effect favoring the latter, Visual in respect of Visual + Speech and Visual in respect of Visual + Abstract with a large effect favoring the second condition over the Visual condition, Speech in respect of Visual + Speech with a small effect favoring the latter. No statistically significant differences were found

regarding the Speech in respect of Visual + Abstract and Visual+ Speech in respect of Visual + Abstract conditions.

The PA score yields statistically significant difference in the group with a p -value=.002. The measurement in the Baseline condition reports statistically significant difference in respect of the Visual + Abstract conditions with a medium effect disfavoring the Baseline (so the former was perceived less annoying). The measurement in the Visual condition reports statistically significant differences in the respect of the Visual + Speech and Visual + Abstract conditions with a medium effect disfavoring the Visual, which was perceived as less annoying. No statistically significant differences were found regarding the other pairwise comparisons.

The alerting effectiveness score yields statistically significant difference in the group with a p -value<.001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions with a large effect disfavoring the Baseline. The measurement in the Visual condition reports statistically significant differences in the respect of the Speech, Visual + Speech, Visual + Abstract conditions with a large effect disfavoring the Visual condition. Furthermore, measurement regarding Speech condition in respect of the Visual + Speech condition reports statistically significant difference with a large effect, favoring the latter. The measurement in the Visual + Speech condition in respect of the Visual + Abstract condition reports statistically significant difference with a medium effect, favoring the former. No statistically significant differences were found regarding the Speech in respect of Visual + Abstract conditions.

| # | Item | AVG (C.I.95%) | | | | | p-value (effect size) | | | | | | | | | | |
|---|------------------------|---------------|-------------|-------------|-------------|-------------|-----------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | | B | V | S | VS | VA | Friedman | B-V | B-S | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA |
| 1 | Message Clarity | 2.6 (0.54) | 4.86 (0.49) | 6 (0.41) | 6.78 (0.22) | 5.52 (0.46) | <.001 | <.001(-1.87) | <.001(-3.02) | <.001(-4.3) | <.001(-2.46) | .002 (-1.07) | <.001(-2.15) | .031 (-0.58) | <.001(-1.01) | .175 (0.46) | <.001(1.48) |
| 2 | Perceived Urgency | 2.82 (0.89) | 3.6 (0.79) | 4.56 (0.63) | 5.17 (0.66) | 5.08 (0.7) | <.001 | .099 (-0.4) | .002 (-0.97) | .001 (-1.29) | .001 (-1.21) | .017 (-0.57) | .001 (-0.92) | .003 (-0.85) | .021 (-0.4) | .247 (-0.33) | .517 (0.05) |
| 3 | Perceived Annoyance | 2.3 (0.7) | 2.26 (0.58) | 2.86 (0.71) | 3.08 (0.65) | 3.47 (0.79) | .002 | .818 (0.02) | .123 (-0.34) | .055 (-0.49) | .016 (-0.67) | .079 (-0.40) | .020 (-0.57) | .005 (-0.75) | .255 (-0.13) | .093 (-0.34) | .209 (-0.23) |
| 4 | Alerting Effectiveness | 2.26 (0.71) | 4.43 (0.64) | 5.82 (0.38) | 6.47 (0.31) | 5.86 (0.39) | <.001 | <.001(-1.37) | <.001(-2.68) | <.001(-3.29) | <.001(-2.69) | <.001(-1.12) | <.001(-1.72) | <.001(-1.15) | .001 (-0.8) | .952 (-0.04) | .01 (0.73) |

Table 20 - Ad-hoc questionnaire results for jaywalking notification

6.3.2.2. Intrusiveness questionnaire

About the intrusiveness questionnaire, all the seven dimensions yield statistically significant difference in the group with a p -value<.001.

Regarding the score of the “distracting” dimension, the measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a medium with a large effect favoring the Baseline (which was so considered less distracting). Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Visual + Speech and Visual in respect of the Visual + Abstract condition with a large effect disfavoring the Visual condition. Finally, Speech in respect of the Visual + Speech condition with a medium effect favoring the latter. No statistically significant differences were found regarding the Speech in respect of Visual + Abstract condition and the Visual + Speech in respect of the Visual + Abstract condition.

Regarding the score of the “disturbing” dimension, the measurement in the Baseline condition reports statistically significant differences in the respect of the Visual + Abstract with a medium effect and in respect of the Visual + Speech with a large effect favoring the Baseline. No statistically significant differences have been found in respect of the Visual condition. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech with a medium effect, the Visual in respect of the Visual + Speech and in respect of the Visual + Abstract conditions with a large effect favoring the former, Speech in respect of the Visual + Speech and in respect of the Visual + Abstract conditions with a medium effect favoring the former. No statistically significant differences were found regarding the Visual + Speech in respect of the Visual + Abstract condition.

Regarding the score of the “forced” dimension, the measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a medium effect in respect of the Visual and in respect of the Speech condition, and with a large effect respect of the Visual + Speech and in respect of the Visual + Abstract condition favoring the Baseline. No statistically significant differences were found regarding the other pairwise comparisons.

Regarding the score of the “interfering” dimension, the measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a medium effect in respect of the Visual and with a large effect respect of the Speech, in respect of the Visual + Speech and in respect of the Visual + Abstract condition favoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech with a medium effect favoring the former, Visual in respect of the Visual + Speech and in respect of the Visual + Abstract conditions with a large effect favoring the former, Speech in respect of the Visual + Speech with a small effect favoring the former. No statistically significant differences were found regarding the Speech and the Visual + Speech conditions in respect of the Visual + Abstract condition.

Regarding the score of the “intrusive” dimension, the measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a large effect respect of all the other conditions, favoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech with a medium effect favoring the former, Visual in respect of the Visual + Speech and in respect of the Visual + Abstract conditions with a large effect favoring the former, Speech in respect of the Visual + Speech with a medium effect favoring the former. No statistically significant differences were found regarding the Speech and the Visual + Speech conditions in respect of the Visual + Abstract condition.

Regarding the score of the “invasive” dimension, the measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a large effect respect of all the other conditions, favoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Visual + Speech with a large effect favoring the former, Visual in respect of the

Visual + Abstract and Speech in respect of the Visual + Speech with a medium effect favoring the former. No statistically significant differences were found regarding the Visual in respect of the Speech condition, Speech in respect of the Visual + Speech conditions and Speech in respect of the Visual + Abstract conditions.

Regarding the score of the “obtrusive” dimension, the measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a large effect respect of all the other conditions, favoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Visual + Speech, Visual in respect of the Visual + Abstract with a medium effect favoring the Visual condition and Speech in respect of the Visual + Speech with a medium effect favoring the former. No statistically significant differences were found regarding the Speech and the Visual + Speech conditions in respect of the Visual + Abstract condition and regarding the Visual condition in respect of the Speech condition.

The following table shows the results of the intrusiveness questionnaire analysis.

| # | Item | AVG (C.I.95%) | | | | | p-value (effect size) | | | | | | | | | | | |
|---|-------------|---------------|-------------|-------------|-------------|-------------|-----------------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|--------------|--|
| | | B | V | S | VS | VA | Friedman | B-V | B-S | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA | |
| 1 | Distracting | 6.21 (0.66) | 5 (0.59) | 4.21 (0.73) | 3.39 (0.76) | 3.56 (0.67) | <.001 | .003 (0.83) | <.001(1.23) | <.001(1.70) | <.001(1.71) | .053 (0.50) | <.001(1.01) | .001 (0.97) | .002 (0.47) | .07 (0.39) | .620 (-0.10) | |
| 2 | Disturbing | 5.86 (0.81) | 5.69 (0.55) | 5.17 (0.53) | 4.47 (0.67) | 4.52 (0.68) | <.001 | .612 (0.10) | .057 (0.43) | .018 (0.80) | .026 (0.77) | .034 (0.41) | <.001(0.84) | .001 (0.90) | .005 (0.49) | .023 (0.45) | .872 (-0.02) | |
| 3 | Forced | 6.32 (0.4) | 5.91 (0.46) | 5.78 (0.43) | 5.65 (0.48) | 5.47 (0.67) | <.001 | .021 (0.59) | .008 (0.75) | .006 (0.84) | .010 (0.80) | .492 (0.12) | .142 (0.23) | .212 (0.32) | .148 (0.12) | .202 (0.23) | .586 (0.12) | |
| 4 | Interfering | 6.34 (0.6) | 5.43 (0.53) | 4.56 (0.66) | 4.04 (0.67) | 4.21 (0.61) | <.001 | .007 (0.69) | <.001(1.21) | <.001(1.55) | <.001(1.51) | .004 (0.62) | <.001(0.99) | <.001(0.91) | .019 (0.33) | .24 (0.23) | .488 (-0.11) | |
| 5 | Intrusive | 6.52 (0.38) | 5.65 (0.46) | 4.78 (0.67) | 4.34 (0.65) | 4.39 (0.68) | <.001 | .002 (0.88) | <.001(1.36) | <.001(1.73) | <.001(1.65) | .015 (0.64) | <.001(0.98) | <.001(0.93) | .037 (0.28) | .355 (0.24) | .739 (-0.02) | |
| 6 | Invasive | 6.65 (0.24) | 5.52 (0.58) | 4.91 (0.67) | 4.26 (0.7) | 4.56 (0.67) | <.001 | <.001(1.09) | <.001(1.47) | <.001(1.95) | <.001(1.77) | .084 (0.41) | .001 (0.84) | .002 (0.65) | .009 (0.40) | .312 (0.22) | .319 (-0.19) | |
| 7 | Obtrusive | 6.65 (0.24) | 5.65 (0.54) | 5.34 (0.64) | 4.56 (0.72) | 4.86 (0.66) | <.001 | <.001(1.01) | <.001(1.15) | <.001(1.66) | <.001(1.52) | 0.339 (0.21) | .002 (0.73) | .005 (0.55) | .011 (0.49) | .139 (0.31) | .372 (-0.18) | |

Table 21 - Intrusiveness questionnaire results for jaywalking notifications

6.3.2.3. UEQ+

The trustworthiness of content core yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other notifications with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of Speech with a small effect favoring the latter, Visual in respect of Visual + Speech with a large effect favoring the latter, Visual in respect of Visual + Abstract with a medium effect favoring the latter, Speech in respect of Visual + Speech with a medium effect favoring the latter, Visual + Speech in respect of Visual + Abstract with a medium effect favoring the former. No statistically significant differences were found regarding the Speech in respect of the Visual + Abstract notification.

The quality of content score yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other notifications with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of Visual + Speech with a medium effect favoring the latter, Visual in respect of Visual + Abstract with a small effect favoring the latter, Speech in respect of Visual + Speech with a medium effect favoring the latter, Visual + Speech in respect of Visual + Abstract with a small effect favoring the former. No statistically

significant differences were found regarding the Speech in respect of the Visual notification and regarding the Speech in respect of the Visual + Abstract notification.

The value score yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other notifications with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of Visual + Speech with a large effect favoring the latter, Speech in respect of Visual + Speech with a medium effect favoring the latter, Visual + Speech in respect of Visual + Abstract with a medium effect favoring the former. No statistically significant differences were found regarding the other pairwise comparisons.

| # | Item | AVG (CI.95%) | | | | | p-value (effect size) | | | | | | | | | | |
|---|----------------------------|--------------|-------------|-------------|-------------|-------------|-----------------------|----------------|----------------|----------------|----------------|---------------|----------------|---------------|----------------|---------------|--------------|
| | | B | V | S | VS | VA | Friedman | B-V | B-S | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA |
| 1 | Trustworthiness of content | 1.22 (0.22) | 1.88 (0.2) | 2.05 (0.16) | 2.31 (0.14) | 2.15 (0.14) | < .001 | < .001 (-1.30) | < .001 (-1.79) | < .001 (-2.48) | < .001 (-2.11) | 0.019 (-0.38) | < .001 (-1.02) | 0.002 (-0.65) | < .001 (-0.70) | 0.104 (-0.28) | 0.005 (0.45) |
| 2 | Quality of content | 1.24 (0.14) | 1.84 (0.15) | 1.89 (0.15) | 2.11 (0.17) | 1.99 (0.21) | < .001 | < .001 (-1.69) | < .001 (-1.88) | < .001 (-2.37) | < .001 (-1.77) | 0.354 (-0.15) | < .001 (-0.72) | 0.008 (-0.35) | < .001 (-0.58) | 0.096 (-0.23) | 0.007 (0.26) |
| 3 | Value | 1.19 (0.13) | 1.67 (0.13) | 1.76 (0.13) | 1.92 (0.12) | 1.76 (0.18) | < .001 | < .001 (-1.67) | < .001 (-1.84) | < .001 (-2.35) | < .001 (-1.55) | 0.262 (-0.29) | 0.001 (-0.87) | 0.089 (-0.25) | < .001 (-0.53) | 0.643 (0.01) | 0.021 (0.43) |

6.3.3. TORs

6.3.3.1. Ad-hoc questionnaire

Regarding the TORs, no statistical significance was found within the group for measurements of perceived annoyance and alerting effectiveness.

Message clarity score yields statistically significant difference in the group with a p -value = .041. The measurement in the Baseline condition reports statistically significant differences in the respect of the Visual + Speech condition with a small effect favoring the Baseline. No statistically significant differences were found regarding the other pairwise comparisons.

PU score yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of the Speech condition and in respect to the Visual + Abstract condition with a small effect favoring the former with respect to the Visual + Speech condition with a medium effect favoring the Baseline. of the Visual + Speech and the Visual + Abstract conditions with a small effect favoring the Visual condition. No statistically significant differences were found regarding the other pairwise comparisons.

| # | Item | Classifica | | | | | p-value (effect size) | | | | | | | | | | |
|---|------------------------|-------------|-------------|-------------|-------------|-------------|-----------------------|-------------|--------------------|--------------------|--------------------|--------------|--------------------|--------------------|-------------|-------------|-------------|
| | | B | V | S | VS | VA | Friedman | B-V | B-S | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA |
| 1 | Message Clarity | 6.65 (0.25) | 6.57 (0.26) | 6.52 (0.26) | 6.39 (0.34) | 6.39 (0.36) | .041 | .345 (0.14) | .148 (0.22) | .047 (0.38) | .094 (0.36) | .772 (0.07) | .203 (0.25) | .203 (0.23) | .148 (0.18) | .371 (0.17) | .841 (0.00) |
| 2 | Perceived Urgency | 6.00 (0.57) | 5.87 (0.46) | 5.57 (0.55) | 5.43 (0.53) | 5.39 (0.66) | < .001 | .49 (0.10) | .019 (0.33) | .007 (0.44) | .007 (0.42) | .064 (0.26) | .010 (0.37) | .010 (0.36) | .371 (0.10) | .407 (0.12) | .890 (0.03) |
| 3 | Perceived Annoyance | 3.09 (0.79) | 2.91 (0.69) | 3.00 (0.73) | 2.83 (0.72) | 2.83 (0.70) | .235 | .173 (0.10) | .571 (0.04) | .188 (0.14) | .169 (0.15) | .423 (-0.05) | .484 (0.05) | .423 (0.05) | .173 (0.10) | .173 (0.10) | .637 (0.00) |
| 4 | Alerting Effectiveness | 6.22 (0.29) | 6.22 (0.29) | 6.26 (0.30) | 6.17 (0.31) | 6.09 (0.45) | .539 | 1.00 (0.00) | 1.00 (-0.06) | 1.00 (0.06) | .586 (0.14) | 1.00 (-0.06) | 1.00 (0.06) | .586 (0.14) | .345 (0.12) | .371 (0.19) | .850 (0.09) |

Table 22 - Ad-hoc questionnaire results for TORs

6.3.3.2. Intrusiveness questionnaire

Regarding the seven dimensions of the intrusiveness questionnaire, only the “distracting” dimension yields statistically significant difference in the group with a p -value=.031. Statistically significant difference has been found in the Baseline measurement in respect of the Visual + Speech condition with a small effect disfavoring the former. No statistically significant differences were found regarding the other pairwise comparisons.

| # | Item | AVG (C.I.95%) | | | | | p-value (effect size) | | | | | | | | | | |
|---|-------------|---------------|-------------|-------------|-------------|-------------|-----------------------|--------------|---------------|---------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | | B | V | S | VS | VA | Friedman | B-V | B-V | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA |
| 1 | Distracting | 3 (0.95) | 3.04 (0.94) | 3.08 (0.93) | 3.26 (0.91) | 3.08 (0.89) | .031 | 1.00 (-0.01) | .345 (-0.03) | .047 (-0.12) | .423 (-0.04) | .772 (-0.01) | .088 (-0.10) | .772 (-0.02) | .071 (-0.08) | .637 (0.00) | .071 (0.08) |
| 2 | Disturbing | 3.86 (0.91) | 3.82 (0.87) | 4 (0.83) | 4.04 (0.82) | 4.08 (0.8) | .064 | 1.00 (0.02) | .233 (-0.06) | .33 (-0.08) | .34 (-0.10) | .071 (-0.08) | .088 (-0.11) | .094 (-0.13) | .772 (-0.02) | .586 (-0.04) | .850 (-0.02) |
| 3 | Forced | 5.39 (0.59) | 5.39 (0.59) | 5.43 (0.59) | 5.47 (0.59) | 5.47 (0.6) | .222 | 1.00 (0.00) | 1.00 (-0.03) | .345 (-0.06) | .345 (-0.06) | 1.00 (-0.03) | .345 (-0.06) | .345 (-0.06) | 1.00 (-0.03) | 1.00 (-0.03) | .637 (0.00) |
| 4 | Interfering | 3.13 (0.77) | 3.21 (0.74) | 3.26 (0.74) | 3.43 (0.76) | 3.39 (0.75) | .007 | 1.00 (-0.04) | .371 (-0.07) | .053 (-0.17) | .088 (-0.14) | 1.00 (-0.02) | .088 (-0.12) | .071 (-0.09) | .071 (-0.09) | .148 (-0.07) | .772 (0.02) |
| 5 | Intrusive | 3.6 (0.88) | 3.56 (0.86) | 3.65 (0.86) | 3.73 (0.88) | 3.73 (0.85) | .245 | 1.00 (0.02) | .0772 (-0.02) | .49 (-0.06) | .49 (-0.06) | .345 (-0.04) | .203 (-0.08) | .203 (-0.08) | .423 (-0.04) | .586 (-0.04) | .785 (0.00) |
| 6 | Invasive | 3.78 (0.77) | 3.78 (0.77) | 3.86 (0.75) | 3.91 (0.73) | 3.82 (0.73) | .620 | 1.00 (0.00) | .345 (-0.04) | .374 (-0.07) | .765 (-0.02) | .345 (-0.04) | .374 (-0.07) | .765 (-0.02) | .765 (-0.02) | .772 (0.02) | .345 (0.05) |
| 7 | Obstrusive | 3.39 (0.91) | 3.47 (0.9) | 3.65 (0.9) | 3.65 (0.88) | 3.56 (0.88) | .033 | 1.00 (-0.04) | .173 (-0.12) | .094 (-0.12) | .173 (-0.08) | .173 (-0.08) | .173 (-0.08) | .345 (-0.04) | .785 (0.00) | 1.00 (0.04) | .345 (0.04) |

Table 23 - Intrusiveness questionnaire results for TORs

6.3.4. Ranks

The trust rank yields statistically significant difference in the group with a p -value< .001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech, Visual in respect of the Visual + Speech, Visual in respect of the Visual + Abstract with a large effect disfavoring the Visual condition, the Speech in respect of the Visual + Speech with a large effect favoring the latter and the Visual + Speech in respect of the Visual + Abstract with a large effect favoring the former. No statistically significant differences were found regarding the Speech and the Visual + Abstract condition.

The message clarity rank yields statistically significant difference in the group with a p -value< .001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech, Visual in respect of the Visual + Speech, Visual in respect of the Visual + Abstract with a large effect disfavoring the Visual condition, the Speech in respect of the Visual + Speech with a large effect favoring the latter and the Visual + Speech in respect of the Visual + Abstract with a large effect favoring the former. No statistically significant differences were found regarding the Speech and the Visual + Abstract condition.

The perceived urgency rank yields statistically significant difference in the group with a p -value< .001. The measurement in the Baseline condition reports statistically significant differences in the respect of the Speech condition, in the respect of the Visual + Speech condition and in the respect of the Visual + Abstract condition with a large effect disfavoring the Baseline. No statistically significant differences were found in respect of the Visual

condition. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech, Visual in respect of the Visual + Speech, Visual in respect of the Visual + Abstract with a large effect disfavoring the Visual condition, the Speech in respect of the Visual + Speech and in respect of the Visual + Abstract with a medium effect disfavoring the former. No statistically significant differences were found regarding the Visual + Speech and the Visual + Abstract conditions.

The perceived annoyance rank yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of the Visual + Abstract condition with a large effect disfavoring the Baseline. No statistically significant differences were found in respect of the other conditions. Between the other jaywalker notifications, the confrontation that reported statistically significant results is Visual in respect of the Visual + Speech with a large effect disfavoring the former and in regards of the Visual + Abstract with a large effect disfavoring the Visual condition. No statistically significant differences were found regarding the other pairwise comparisons.

The alerting effectiveness rank yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions, with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech, Visual in respect of the Visual + Speech, Visual in respect of the Visual + Abstract with a large effect disfavoring the Visual condition, the Speech in respect of the Visual + Speech with a large effect favoring the latter and the Visual + Speech in respect of the Visual + Abstract with a large effect favoring the former. No statistically significant differences were found regarding the Speech and the Visual + Abstract condition.

The intrusiveness rank yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of the Visual condition with a medium effect, and in respect of the Speech, Visual + Speech, Visual + Abstract conditions with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech, Visual in respect of the Visual + Speech, Visual in respect of the Visual + Abstract with a large effect disfavoring the Visual condition with a large effect favoring the latter and Speech in respect of the Visual + Abstract with a medium effect favoring the former. No statistically significant differences were found between the Speech and the Visual + Speech conditions and between the Visual + Speech and Visual + Abstract conditions.

The overall rank yields statistically significant difference in the group with a p -value < .001. The measurement in the Baseline condition reports statistically significant differences in the respect of all the other conditions with a large effect disfavoring the Baseline. Between the other jaywalker notifications, the confrontations that reported statistically significant results are Visual in respect of the Speech, Visual in respect of the Visual + Speech with a large effect disfavoring the Visual condition, the Speech in respect of the Visual + Speech with a large

effect favoring the latter. No statistically significant differences were found between the Visual and the Visual + Abstract conditions and between the Speech and the Visual + Abstract conditions.

| # | Item | Rank | | | | | p-value (effect size) | | | | | | | | | | | | |
|---|------------------------|------|---|---|----|----|-----------------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|--|--|
| | | B | V | S | VS | VA | Friedman | B-V | B-S | B-VS | B-VA | V-S | V-VS | V-VA | S-VS | S-VA | VS-VA | | |
| 1 | Trust | 5 | 4 | 2 | 1 | 2 | <.001 | <.001(2.53) | <.001(4.82) | <.001(12.15) | <.001(2.98) | .001 (1.44) | <.001(4.2) | .011 (1.04) | <.001(2.66) | .912 (-0.04) | <.001(-1.83) | | |
| 2 | Message Clarity | 5 | 4 | 2 | 1 | 3 | <.001 | <.001(2.20) | <.001(5.26) | <.001(12.15) | <.001(3.52) | <.001(1.90) | <.001(4.28) | .003 (1.23) | <.001(2.25) | .298 (-0.45) | <.001(-2.33) | | |
| 3 | Perceived Urgency | 5 | 4 | 3 | 1 | 1 | <.001 | .065 (0.52) | <.001(1.42) | <.001(1.82) | <.001(1.83) | .004 (1.08) | .001 (1.56) | <.001(1.57) | .030 (0.74) | .042 (0.71) | .611 (-0.07) | | |
| 4 | Perceived Annoyance | 5 | 4 | 1 | 1 | 1 | .003 | .389 (0.15) | .161 (0.50) | .064 (0.7) | .012 (1.05) | .124 (0.52) | .040 (0.77) | .002 (1.23) | .305 (0.29) | .054 (0.73) | .232 (0.42) | | |
| 5 | Alerting Effectiveness | 5 | 4 | 2 | 1 | 2 | <.001 | <.001(2.24) | <.001(4.28) | <.001(8.14) | <.001(3.30) | <.001(1.56) | <.001(3.38) | .002 (1.33) | <.001(1.56) | .826 (0.00) | .002 (-1.29) | | |
| 6 | Intrusiveness | 5 | 4 | 3 | 1 | 1 | <.001 | .029 (0.54) | .005 (1.23) | .002 (1.54) | <.001(1.86) | .014 (0.99) | .002 (1.38) | <.001(1.82) | .130 (0.45) | .035 (0.76) | .365 (0.25) | | |
| 7 | Overall | 5 | 4 | 2 | 1 | 2 | <.001 | .015 (0.80) | <.001(2.01) | <.001(3.25) | <.001(1.56) | .002 (1.25) | <.001(2.51) | .027 (0.90) | .003 (1.16) | .816 (-0.11) | .006 (-1.01) | | |

Table 24 - Rank results

6.3.5. SSQ, IPQ and UEQ+

The simulator sickness questionnaire revealed that the sickness scores were consistently low for all participants.

About the three subscales of IPQ, with a score ranging from 1 to 5, spatial presence was rated with an average of 4.55 (C.I._{95%} = 0.44), the involvement with an average of 3.62 (C.I._{95%} = 0.43) and experience realism with an average of 3.39 (C.I._{95%} = 0.39), while the general additional item, sense of being there, was rated with an average of 4.22 (C.I._{95%} = 0.38).

The UEQ+ for the interface reveals a KPI score (ranging from -3.00 to +3.00 representing the overall user experience impression) of 2.00, showing interface provided a positive experience. For the single scales results are: Efficiency rated with an average of 0.72 (C.I._{95%} = 0.04), Usefulness rated with an average of 0.72 (C.I._{95%} = 0.04), Novelty rated with an average of 0.56 (C.I._{95%} = 0.44), Clarity rated with an average of 0.76 (C.I._{95%} = 0.05), Intuitive use rated with an average of 0.69 (C.I._{95%} = 0.03), Aesthetics rated with an average of 0.56 (C.I._{95%} = 0.4), Trustworthiness of Content rated with an average of 0.74 (C.I._{95%} = 0.05), Quality of Content rated with an average of 0.69 (C.I._{95%} = 0.04) and Value rated with an average of 0.57 (C.I._{95%} = 0.05)

7. Conclusions and future works

7.1. Conclusions

This section will summarize the conclusions of the thesis.

7.1.1. Trust, Anxiety, Mental Workload

The initial questionnaires aim to evaluate the overall experience and investigate whether varying jaywalker notifications can influence the overall levels of trust, anxiety, and mental workload within the entire simulation.

The trust results highlight how trust levels remained consistently high across all configurations of jaywalking notifications. In five items out of seven, a statistically significant difference was

found between the Baseline and the Visual + Speech configuration, in one of the items (predictability) between Visual and Visual + Speech configurations. While the jaywalking events consisted only of a small part of the simulation, the finding suggests that the Visual + Speech notification could enhance the overall trust compared to not having a notification.

Regarding anxiety, there were no statistically significant differences found between the configurations. However, it's worth noting that the testers reported very low levels of anxiety across all configurations.

Similarly, there were no statistically significant differences observed among the configurations in terms of mental workload. Specifically, performance was consistently high across all configurations, while mental demand, physical demand, time demand, effort, and frustration levels remained low.

7.1.2. Jaywalking Notifications

From the ad-hoc questionnaire, Visual + Speech results as the best notification for jaywalking events in terms of trust, message clarity, alerting effectiveness and overall preference showing statistically significant differences with respect of all the other notifications. Speech and Visual + Abstract result as the second-best options with no statistically significant differences in those metrics between them. The Visual notification turned out to be worse than the Speech notification. The baseline, however, was by far the worst. Visual + Speech notification results also as the one being perceived more urgent together with the Visual + Abstract one with no statistically significant difference between them but yielding statistically significant differences between the ones that consisted of only one modality (Visual and Speech) or none (Baseline), proving that the use of multiple modalities can improve the urgency of a notification. The Visual notification turned out to be perceived as less urgent than the Speech one. About the annoyance, the Visual + Abstract notification was rated as the most annoying one however with no statistically significant differences between the Speech and the Visual + Speech notifications. Since no statistically significant difference were found between the Speech and the Visual + Speech it appears that the use of the visual feedback together with the speech notification can increase trust, message clarity, alerting effectiveness in respect of the use of speech alone without impacting much on the annoyance of the notification.

The intrusiveness questionnaire reveals how the Visual + Speech and Visual + Abstract notifications resulted has the more intrusive ones for all the items except from the “forced” one, with no statistically significant differences between them. The forced item reveal that no statistically significant differences were found between the notifications (aside from the baseline) revealing how, even if those two configurations were perceived as more intrusive and annoying, they are not considered forced compared to the other single modality channel ones. The UEQ+ reveals as well how the Visual + Speech configuration was rated as the best in all the items (trustworthiness of content, quality of content, value) with statistically significant differences in respect of all the other configurations. The second-best configurations are Visual + Abstract and Speech with no statistically significant differences between them.

7.1.3. TORs

The results indicate that, while the TORs notification remained consistent across configurations, various notifications for jaywalking events can influence the perception of

TORs. Notably, statistically significant differences were observed in terms of PU, showing that a more urgent jaywalker notification led to a less perceived urgent TOR. The intrusiveness questionnaire also identified a statistically significant difference between the baseline and the Visual + Speech configuration regarding the interfering and regarding the distracting items.

The findings highlight the importance of considering how an event is notified, as it can have a significant impact not only on that specific notification but also on how other notifications are perceived and the design of this alerts should be careful to ensure a consistent and effective user experience.

7.2. Future works

While the primary focus of this thesis centered on the comprehensive analysis of jaywalking events in the ODD, there remains a significant avenue for forthcoming research to explore optimal notification methodologies applicable to various critical events. Notably, this includes scenarios involving abrupt vehicular actions, such as sudden lane changes or unexpected maneuvers.

The significance of this pursuit lies in the recognition that different critical events demand tailored notification strategies. Just as this study has shed light on the impact of modifying jaywalking notifications on altering perception of TORs, it becomes imperative to extend this inquiry to encompass a broader spectrum of critical events. However, the scope of prospective studies could encompass a more expansive exploration into whether the variations in TORs notifications exert a similar influence on the perception and response to other alerts and notifications.

By bringing the investigation beyond jaywalking incidents to encompass diverse critical events, future research endeavors could unravel more information about perception and response dynamics. Understanding how differing notification approaches impact the perception of various critical events will not only enrich the comprehension of driver behavior but also lay the groundwork for the development of more advanced and context-aware notification systems within the driving environment.

References

1. Ma RH, Morris A, Herriotts P, Birrell S. Investigating what level of visual information inspires trust in a user of a highly automated vehicle. *Applied Ergonomics*. 2021;90:103272.
2. Ekman F, Johansson M, Sochor J. Creating appropriate trust in automated vehicle systems: A framework for HMI design. *IEEE Transactions on Human-Machine Systems*. 2017;48(1):95–101.
3. Mackay A, Fortes I, Santos C, Machado D, Barbosa P, Boas VV, Ferreira JP, Costa N, Silva C, Sousa E. The impact of autonomous vehicles' active feedback on trust. In: *Advances in Safety Management and Human Factors: Proceedings of the AHFE 2019 International Conference on Safety Management and Human Factors, July 24-28, 2019, Washington DC, USA 10*. Springer; 2020. p. 342–52.
4. Stanton NA, Young MS. Vehicle automation and driving performance. *Ergonomics*. luglio 1998;41(7):1014–28.
5. Wiese EE, Lee JD. Attention grounding: a new approach to in-vehicle information system implementation. *Theoretical Issues in Ergonomics Science*. 2007;8(3):255–76.
6. Du N, Haspiel J, Zhang Q, Tilbury D, Pradhan AK, Yang XJ, Robert Jr LP. Look who's talking now: Implications of AV's explanations on driver's trust, AV preference, anxiety and mental workload. *Transportation research part C: emerging technologies*. 2019;104:428–42.
7. Körber M, Prasch L, Bengler K. Why do I have to drive now? Post hoc explanations of takeover requests. *Human factors*. 2018;60(3):305–23.
8. Koo J, Kwac J, Ju W, Steinert M, Leifer L, Nass C. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)*. 2015;9:269–75.
9. Lee S, Hong J, Jeon G, Jo J, Boo S, Kim H, Jung S, Park J, Choi I, Kim S. Investigating effects of multimodal explanations using multiple In-vehicle displays for takeover request in conditionally automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour*. 2023;96:1–22.
10. Häuslschmid R, von Buelow M, Pfleging B, Butz A. Supporting Trust in autonomous driving. In: *Proceedings of the 22nd international conference on intelligent user interfaces*. 2017. p. 319–29.
11. McKeown D, Isherwood S. Mapping candidate within-vehicle auditory displays to their referents. *Human factors*. 2007;49(3):417–28.
12. Forster Y, Naujoks F, Neukum A. Your turn or my turn? Design of a human-machine interface for conditional automation. In: *Proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications*. 2016. p. 253–60.
13. Stadler S, Cornet H, Huang D, Frenkler F. Designing Tomorrow's Human-Machine Interfaces in Autonomous Vehicles: An Exploratory Study in Virtual Reality. *Augmented Reality and Virtual Reality: Changing Realities in a Dynamic World*. 2020;151–60.
14. Drezet H, Colombel S, Avenel ML. Human-man interface concept for autonomous car. In: *2019 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE; 2019. p. 1–5.

15. Sandhaus H, Hornecker E. A WOZ study of feedforward information on an ambient display in autonomous cars. In: Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. 2018. p. 90–2.
16. Löcken A, Heuten W, Boll S. AutoAmbiCar: Using ambient light to inform drivers about intentions of their automated cars. In: Adjunct proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications. 2016. p. 57–62.
17. Jose R, Lee GA, Billinghamurst M. A comparative study of simulated augmented reality displays for vehicle navigation. In: Proceedings of the 28th Australian conference on computer-human interaction. 2016. p. 40–8.
18. Doshi A, Cheng SY, Trivedi MM. A novel active heads-up display for driver assistance. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*. 2008;39(1):85–93.
19. Medenica Z, Kun AL, Paek T, Palinko O. Augmented reality vs. street views: a driving simulator study comparing two emerging navigation aids. In: Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services. 2011. p. 265–74.
20. Schömig N, Wiedemann K, Naujoks F, Neukum A, Leuchtenberg B, Vöhringer-Kuhnt T. An augmented reality display for conditionally automated driving. In: Adjunct proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications. 2018. p. 137–41.
21. Feierle A, Beller D, Bengler K. Head-up displays in urban partially automated driving: Effects of using augmented reality. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC). IEEE; 2019. p. 1877–82.
22. Pfanmüller L, Kramer M, Senner B, Bengler K. A comparison of display concepts for a navigation system in an automotive contact analog head-up display. *Procedia Manufacturing*. 2015;3:2722–9.
23. Haeuslschmid R, Shou Y, O'Donovan J, Burnett G, Butz A. First Steps towards a View Management Concept for Large-sized Head-up Displays with Continuous Depth. In: Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications [Internet]. Ann Arbor MI USA: ACM; 2016 [citato 14 ottobre 2023]. p. 1–8. Disponibile su: <https://dl.acm.org/doi/10.1145/3003715.3005418>
24. Lindemann P, Lee TY, Rigoll G. Supporting driver situation awareness for autonomous urban driving with an augmented-reality windshield display. In: 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). IEEE; 2018. p. 358–63.
25. Diels C, Bos JE. User interface considerations to prevent self-driving carsickness. In: Adjunct proceedings of the 7th international conference on automotive user interfaces and interactive vehicular applications. 2015. p. 14–9.
26. Morra L, Lamberti F, Praticó FG, La Rosa S, Montuschi P. Building trust in autonomous vehicles: Role of virtual reality driving simulators in HMI design. *IEEE Transactions on Vehicular Technology*. 2019;68(10):9438–50.
27. Bolton A, Burnett G, Large DR. An investigation of augmented reality presentations of landmark-based navigation using a head-up display. In: Proceedings of the 7th

- International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2015. p. 56–63.
28. Koo J, Shin D, Steinert M, Leifer L. Understanding driver responses to voice alerts of autonomous car operations. *International journal of vehicle design*. 2016;70(4):377–92.
 29. Waytz A, Heafner J, Epley N. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of experimental social psychology*. 2014;52:113–7.
 30. Choi JK, Ji YG. Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*. 2015;31(10):692–702.
 31. Forster Y, Naujoks F, Neukum A. Increasing anthropomorphism and trust in automated driving functions by adding speech output. In: 2017 IEEE intelligent vehicles symposium (IV). IEEE; 2017. p. 365–72.
 32. Hock P, Kraus J, Walch M, Lang N, Baumann M. Elaborating feedback strategies for maintaining automation in highly automated driving. In: Proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications. 2016. p. 105–12.
 33. S. Alpers B, Cornn K, E. Feitzinger L, Khaliq U, Park SY, Beigi B, Joseph Hills-Bunnell D, Hyman T, Deshpande K, Yajima R. Capturing passenger experience in a ride-sharing autonomous vehicle: the role of digital assistants in user interface design. In: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2020. p. 83–93.
 34. Lee S, Ratan R, Park T. The voice makes the car: Enhancing autonomous vehicle perceptions and adoption intention through voice agent gender and style. *Multimodal Technologies and Interaction*. 2019;3(1):20.
 35. Naujoks F, Forster Y, Wiedemann K, Neukum A. Improving usefulness of automated driving by lowering primary task interference through HMI design. *Journal of Advanced Transportation*. 2017;2017.
 36. Forster Y, Naujoks F, Neukum A, Huestegge L. Driver compliance to take-over requests with different auditory outputs in conditional automation. *Accident Analysis & Prevention*. 2017;109:18–28.
 37. Lewis BA, Eisert JL, Roberts DM, Baldwin CL. Designing unambiguous auditory crash warning systems. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications Sage CA: Los Angeles, CA; 2014. p. 2078–82.
 38. Lewis BA, Eisert JL, Baldwin CL. Validation of essential acoustic parameters for highly urgent in-vehicle collision warnings. *Human factors*. 2018;60(2):248–61.
 39. Hellier E, Edworthy J. On using psychophysical techniques to achieve urgency mapping in auditory warnings. *Applied Ergonomics*. 1999;30(2):167–71.
 40. Oliveira L, Luton J, Iyer S, Burns C, Mouzakitis A, Jennings P, Birrell S. Evaluating how interfaces influence the user interaction with fully autonomous vehicles. In: Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2018. p. 320–31.
 41. Meteier Q, Capallera M, De Salis E, Sonderegger A, Angelini L, Carrino S, Abou Khaled O, Mugellini E. The effect of instructions and context-related information about limitations

- of conditionally automated vehicles on situation awareness. In: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. 2020. p. 241–51.
42. van den Beukel AP, van der Voort MC. How to assess driver's interaction with partially automated driving systems—A framework for early concept assessment. *Applied ergonomics*. 2017;59:302–12.
 43. Sirkin D, Martelaro N, Johns M, Ju W. Toward measurement of situation awareness in autonomous vehicles. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017. p. 405–15.
 44. Silva TH, De Melo POV, Viana AC, Almeida JM, Salles J, Loureiro AA. Traffic condition is more than colored lines on a map: characterization of waze alerts. In: *Social Informatics: 5th International Conference, SocInfo 2013, Kyoto, Japan, November 25-27, 2013, Proceedings 5*. Springer; 2013. p. 309–18.
 45. NWH Vehicle Physics 2 | Physics | Unity Asset Store [Internet]. [citato 23 ottobre 2023]. Disponibile su: <https://assetstore.unity.com/packages/tools/physics/nwh-vehicle-physics-2-166252>
 46. ClubSport Steering Wheel GT Forza Motorsport V2 for Xbox | Fanatec [Internet]. [citato 23 ottobre 2023]. Disponibile su: <https://fanatec.com/eu-en/steering-wheels/clubsport-steering-wheel-gt-forza-motorsport-v2-for-xbox>
 47. Atomic A3 Motion Simulators - compact and robust for commercial venues [Internet]. Atomic Motion Systems. [citato 23 ottobre 2023]. Disponibile su: <https://www.atomicmotionsystems.com/atomic-a3-motion-simulators/>
 48. Actuate Motion [Internet]. [citato 23 ottobre 2023]. Disponibile su: <https://actuate-motion.com/>
 49. Model X Owner's Manual [Internet]. Tesla. [citato 23 ottobre 2023]. Disponibile su: https://www.tesla.com/ownersmanual/modelx/en_us/GUID-29A7E205-A689-41D1-B69C-3AE821CB70E7.html#CONCEPT_WRS_4X5_T4B
 50. Zhang Q, Yang XJ, Robert LP. What and when to explain? A survey of the impact of explanation on attitudes toward adopting automated vehicles. *IEEE Access*. 2021;9:159533–40.
 51. Cao Y, Mahr A, Castronovo S, Theune M, Stahl C, Müller CA. Local danger warnings for drivers: The effect of modality and level of assistance on driver reaction. In: *Proceedings of the 15th international conference on Intelligent user interfaces*. 2010. p. 239–48.
 52. Davenport WG. Vigilance for simultaneous auditory and vibrotactile signals. *Australian Journal of Psychology*. 1969;21(2):159–65.
 53. Nass C, Jonsson IM, Harris H, Reaves B, Endo J, Brave S, Takayama L. Improving automotive safety by pairing driver emotion and car voice emotion. In: *CHI'05 extended abstracts on Human factors in computing systems*. 2005. p. 1973–6.
 54. Muir BM. Trust between humans and machines, and the design of decision aids. *International journal of man-machine studies*. 1987;27(5–6):527–39.
 55. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: *Advances in psychology*. Elsevier; 1988. p. 139–83.

56. Laugwitz B, Schubert U, Ilmberger W, Tamm N, Held T, Schrepp M. Subjektive Benutzerzufriedenheit quantitativ erfassen: Erfahrungen mit dem User Experience Questionnaire UEQ. Tagungsband UP09. 2009;
57. Baldwin CL. Verbal collision avoidance messages during simulated driving: perceived urgency, alerting effectiveness and annoyance. *Ergonomics*. 2011;54(4):328–37.
58. Li H, Edwards SM, Lee JH. Measuring the intrusiveness of advertisements: Scale development and validation. *Journal of advertising*. 2002;31(2):37–47.
59. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*. 1993;3(3):203–20.
60. igroup presence questionnaire (IPQ) overview | igroup.org – project consortium [Internet]. [citato 13 novembre 2023]. Disponibile su: <https://www.igroup.org/pq/ipq/index.php>
61. Sheldon MR, Fillyaw MJ, Thompson WD. The use and interpretation of the Friedman test in the analysis of ordinal-scale data in repeated measures designs. *Physiotherapy Research International*. 1996;1(4):221–8.

