DESIGNING IN-CAR USER INTERFACE FOR LEVEL-3 AUTONOMOUS CAR

A Thesis
Presented to
The Academic Faculty

by

Dheekshana Senthil Murugan

In Partial Fulfillment
of the Requirements for the Degree
Master of Industrial Design in the
School of Industrial Design

Georgia Institute of Technology
April of 2021

DESIGNING IN-CAR USER INTERFACE FOR LEVEL-3 AUTONOMOUS CAR

To be Approved by:

Dr. Wei Wang, Advisor
School of Industrial Design
Georgia Institute of Technology

Dr. Hyun Joo
School of Industrial Design and Interactive Computing
Georgia Institute of Technology

Dr. Bruce N. Walker
School of Psychology and Interactive Computing
Georgia Institute of Technology

Date Approved: April 19, 2021

DESIGNING IN-CAR USER INTERFACE FOR LEVEL-3 AUTONOMOUS CAR

Approved by:

Dr. Wei Wang, Advisor
School of Industrial Design
Georgia Institute of Technology

Dr. Hyun Joo,
School of Industrial Design and
Interactive Computing
Georgia Institute of Technology

Dr.Bruce Walker
School of Psychology and
Interactive Computing
Georgia Institute of Technology

Date Approved: April 19, 2021

ACKNOWLEDGEMENTS

To my advisor and committee; Dr.Wei Wang, Dr.Hyun Joo, and Dr.Bruce Walker for always expressing excitement in learning my progress and providing me with the feedback, support and resources to proceed with my research. They encouraged me throughout and motivated me to surpass my capabilities and go beyond.

A special mention to the members of the sonification lab who helped me in obtaining resources for preparing the user study material.

I am grateful to my fellow classmates, friends and family who provided a constant emotional, financial and educational support system.

Thanks you for all your encouragement!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In this paper the researchers investigate how different instrument cluster designs impact the ability of a person to calibrate their trust to the system, while driving a Level 3, conditional autonomous vehicle. A user study with 15 participants were conducted in a lab environment on a 42” TV screen. The TV displayed videos of three instrument cluster designs responding to a set of pre-recorded simulated roadway driving scenarios. If an alert appears, requesting the driver to take over on the design during the video, the participants were asked to respond by holding onto a steering wheel prop. The recorded reaction times, trust scores and workload scores did not show significant differences between the designs, but certain design elements and layout styles found across the three designs were perceived as beneficial for appropriating user trust and responding faster to take over alerts/requests.

CHAPTER 1. INTRODUCTION

The future mobility will be shaped by the advancements in the autonomous transportation. The US Department of Transportation's National Highway Traffic Safety Administration (NHTSA) defines self-driving cars "are those in which operation of the vehicle occurs without direct driver input to control the steering, acceleration, and braking and are designed so that the driver is not expected to constantly monitor the roadway while operating in self-driving mode". There are plenty of advantages of using autonomous cars. Less traffic accidents and increased personal safety, better use of travel time, reduced fatality rates across all ages, enhanced traffic management are forecasted to be few such benefits [19]. With most driving tasks controlled by vehicle intelligence there is a major change in user experience. Different information needs to be presented to the drivers as per their new mental model of the autonomous driving car [20].

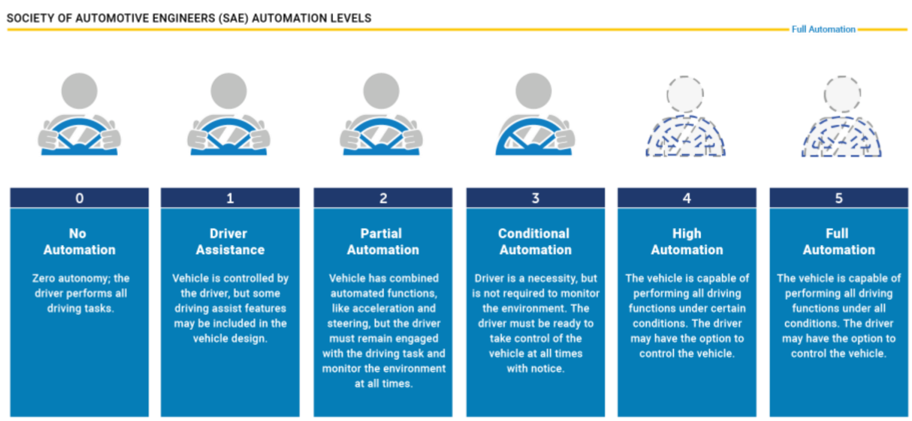


Figure 1.1 SAE Levels of Automation

The information needs vary across the six levels of automation as defined by the SAE (Society of Automotive Engineers, Figure 1.1). Level 3, conditional autonomy poses a unique challenge where the driver must be ready to take over at all times with notice [21]. This notice is famously called as a Takeover Request (TOR). The TOR is a convention used by the AV to communicate to the driver the requirement for a human to take control. Current research on TOR typically assesses the quality of TOR with reaction times [22] [23] and workload [24]. However, the appropriate usage of a system is owed to the calibrated level of trust which is the level of trust that reflects the system's capabilities and performance [25]. Calibrated trust may be the benchmark for safe and efficient design of interaction strategies based on the information provided prior to and during system use [26]. There is limited research that assess the impact on the quality of the TOR displayed based on trust calibration [27]. The objective of this research was to investigate the effect of different visual representations of information during TOR on trust calibration of the user.

The focus in this paper is on the visual modality to communicate TOR based on trust calibration with different designs. The instrument cluster area (dashboard display) is a critical location for visually presenting driving situation and environment related details. Currently, car makers like Tesla and Cadillac explore how the display can be designed for best communication of driving automation status in harmony with other output modalities. Through this research we wish to understand how in-car display interfaces can be designed to visualize the state of the car and alert change of control to establish appropriate trust between the driver and automation. The findings from this research can help understand the factors that would influence success of the new technology with daily users and contribute to the knowledge of car manufacturers venturing into developing systems for advanced AVs. The designs for communication in this research are developed as near future exploration to the technique used to present TOR in current instrument clusters of conditional autonomous cars (Level 3). A user study was conducted with fifteen participants (n = 15) to compare the effectiveness of three designs in calibrating participant's trust with the system. Additionally, the reaction times and workload were assessed to provide a comprehensive evaluation of each instrument cluster design.

CHAPTER 2. BACKGROUND

2.1 The Role of Human in Autonomous Vehicles

With the advancements of automation in cars the amount of driving-related information that drivers receive from the system will be dramatically changed based on different scenarios and driving automation conditions [3]. This introduces complexity in the way in-car displays are designed. The 2019 ADAS/Connected Car report suggests that, in-car experiences are to be designed with empathy meaning that design decisions need to be made in alignment with user needs [31].

There are six levels of autonomy from Level 0 (no automation) to Level 5 (full automation) according to the SAE Standard J3016 [42]. Level 3 conditional automation falls at the center of the spectrum, which is one of the trickiest levels that the human still play a critical role in the system control. In this level of autonomy, the car can perform most aspects of driving tasks and monitoring the environment in well driving conditions but still require human intervention within a reasonable time-frame [4]. Functional system failures like missing lanemarkings, high curvature, or system failures like sensor malfunctions may need manual takeover of the situation [18]. This sort of requirement of human intervention could be a lot for some people since it can influence driver performance by loss of control [40], loss of situational awareness [40], over trust [5] and overconfidence [5]. The usability and acceptance of automation systems depends on the time to successfully complete a take-over [13]. So, the question comes to how can the takeover from the system to the driver be made faster, easier to read and respond to [4]. Additionally, with the lack of experience and the fear of the unknown, the user experience becomes especially important in trust building for Level 3 autonomous drivers [31].

The harmonious communication between system and driver becomes a key factor to the user experience. Humans should be able to understand what the machine can and cannot do, give directions and monitor the machine as well. The information presented by the machine should be useful for decision making, get driver attention to potential risk and provide warnings as per driver-intent [6]. Drivers must also be subject to least mental load when communicated to and the communication behavior of the system should be with respect to current context [6]. Such a design and behavior of the system would make the driver-automation cooperation more transparent, leading to increased trust in the system [7].

2.2 Trust Calibration

Trust is critical in human autonomy interaction. To introduce the concept of informed safety for automation in vehicles, Khastgir et al. [7] pointed out that Trust with the system' means drivers' awareness or attitude towards the limitations of the systems and their subsequent ability to adapt their use of the system to accommodate for the limitations in order to deliver the expected benefit from the system. Lee and See's trust model [35] introduces the appropriateness of trust as a moderator for the relationship between trust evolution and intent formation as well as the relationship between automation and display. The model defines trust calibration as matching of trust capabilities with the trust in the system (Figure 2.1). Hoff and Bashir's model [41] represent dispositional, situational and learned trust. According to their research the design features that influence system performance are appearance, ease of use, communication style, transparency/feedback, and level of control. Li et al. addressed the effect "no risk no trust" that the dynamic learned trust changes as the system performance changes [8][9]. This is how the user's trust changes in line with the system performance during a given interaction.

Mismatching the trust between the user and the system might lead to accidents and other dangers. The mismatch in the driver perception and the capabilities of the system can lead to misuse due to mistrust and disuse due to distrust [7]. Misuse is when the driver uses the system in situations where the system is not designed to perform making it unsafe (Figure 2.1). Disuse is when the user doesn't use the system where the automation is suitable and hence not benefiting from it (Figure 2.1) [7]. Real-time information about the automated system health can bring back drivers "in-the loop". Inaccurate information on the other hand is what can cause over trust or mistrust [7]. In the context of driving, trust was identified as a critical precursor in determining AV attitude [32]. For example, in an experiment to calibrate the trust in the autopark feature of a Tesla, the researchers find that some people intervene often showing more distrust towards the system and absence of intervention showing lack of distrust. They believe design should provide a clear understanding of the process of parking [10]. M.McGruil and B.Sarter find that the status information led to significant improvement in trust calibration as opposed to a command information for pilots [29].

There is some work that explores how the act of interacting and driving an automated vehicle impact trust in automation. Some work talks about the calibrating trust over time [33] and others explore the impact of initial information on user attitude towards automation [34]. Gold and colleagues show that users feel gain in safety, but their perception that the automation allows the driver to perform non-driving tasks reduces after completing a takeover [17]. There is growing research that explores the impact on trust with the presence and absence of an uncertainty

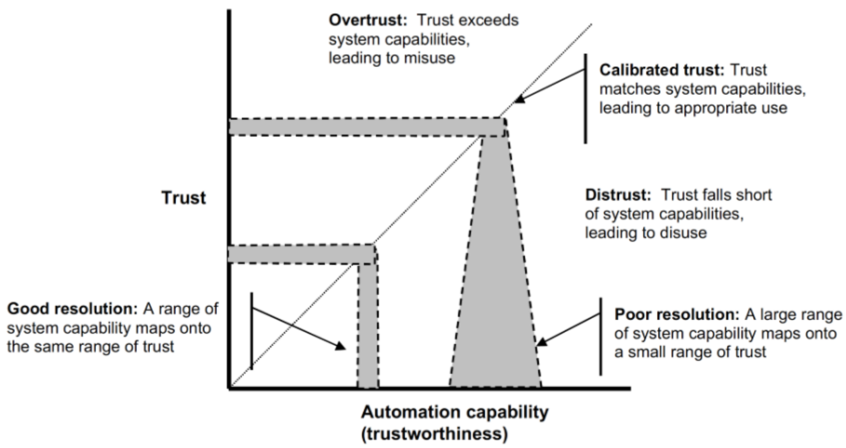


Figure 2.1 Relationship among calibration, resolution and capability by Lee and See [35]

display. Kunze and team perform a driving simulator study [36] where the uncertainty information was conveyed to the user visually. They found that this impacted attention allocation and lead them to monitor the automation state more often than drivers that did not have a display

The current standard for measuring trust is through subjective rating scales and continuous measures. For the purpose of measuring the trust in automation that is context based and dynamic, a method to measure a specific aspect of trust, the situational would be a more nuanced approach as proposed by Hoff and Bashir [11]. Current self-report measurers do not provide the number of measurements without repeated interruptions of interactions. The STS-AD allows for deeper understanding of how experimental manipulations influence specific aspects of trust as opposed to global level of trust and allow for repeated measurements throughout the study [12]. The STS-AD was adopted to measure the situational trust on the designs in this study.

2.3 Designing for Takeover

Information processing for a successful user-system environment interaction can either aim to explain ("understand") certain actions or events, to anticipate (predict) certain outcomes or consequences, or to 'adapt' to changes in the user-system environment relationship [14]. That being said, designing a TOR interface is for when the driver is out of the loop and during transitioning cognitively and physically back to the driving task [15]. A successful environment interaction in such a case should explain, predict and adapt throughout the automated drive and Takeover. The interaction between the autonomous car and the driver could be like two friends on a journey with good team work [20]. For example, simply visualizing the uncertainty in the context of automated driving for a driver can increase the time to collision in cases of automation failure, improve situational awareness, increased acceptance and higher trust ratings [28].

Schmidt and Herrmann reconsider Ben Schneiderman's rules for intervention design [30]. According to their study, designing an autonomous system should:

1. Strive for a dynamic and contextual consistency,

2. Feedback of the automated behavior and intervention must be offered,

3. Design the start of the intervention to be clear and simple,

4. Allow for immediate intervention to avoid unsolicited automated behavior,

5. Allow simple means to reverse the impact of automation actions and impact of interventions,

6. Distribution of control should be communicated,

7. Should not require users to remember a previous system status.

In designing the user interfaces for this study, Rule 1, 2, 3, 6 and 7 proved to be most relevant and useful.

The information that needs to be presented by the car could be sourced from the on-board sensors and in advanced systems even from cooperative perception technology, i.e., information from other road users who have passed that situation [1]. This means that a TOR can be presented immediately or in advance, based on when the information is collected by the car. Gold, Radlymr , Naujoks and Bellam [18] define and classify testing scenarios for the purpose of research. They provide different testing scenarios classified based on urgency, predictability, criticality and driver response. In this study the displays are designed to respond to two such contrasting test scenarios (low urgency and high urgency).

No.	Name	Urgency	Predictability	Criticality	Driver response
1	Sensor failure (Subsystem)	1	1	1	1
2	Sensor failure (Total)	3	1	2–3	1–2
3	End of highway	1	3	1–2	1–2
4	Lane change to deceleration lane not possible (e.g. because of traffic on target lane)	2	2	1	3
5	Lane change from entrance ramp not possible	3	2	2	3
6	Road narrows (known from backend)	1	3	2	2
7	Road narrows (detected by on-board sensors)	3	1	2	2
8	Danger zone/obstacle ahead (known from backend)	1	3	1–3	1–3
9	Danger zone/obstacle ahead (detected by on-board sensors)	3	1	1–3	1–3
10	Loss of reference signals (e.g. lane markings missing)	3	1	2–3	1
11	Predictable loss of reference signals (known from backend)	1	3	2–3	1

Table 2.1 Classifications of Different Design Scenarios by Gold, Redlymr, Noujoks and Bellam [18]

Though the output modalities are varied across publications, the visual display is one of the primary outputs [4]. Nair et al. [9] compared the digital and physical visual indicators to guide user attention in conditional driving automation. The role of instrument clusters become an important mode of visual output for communicating such take over requests (TOR). One way of exploring how to seamlessly switch control is to effectively communicate by taking complex data and presenting it in an easy-to-understand visual format [24]. The challenge is to effectively communicate with the limited display area available to organize the complex data [2]. This study hypothesized that, presenting information on the state of the automation and request to take over in an easy-to-understand visual format can promote proper calibration of trust with the system. We focus on designing the instrument cluster of a Level 3 autonomous vehicle for presenting this information.

CHAPTER 3. METHOD AND PROCESS

Through a process of desk and user research we conceptualized three different designs for the instrument cluster area of a Level-3 autonomous car in near future, which we suppose the vehicle can detect the location and surrounding information of road hazards ahead that require takeover based on its sensors and vehicular networks. The three designs are the independent variables of the study. In all three designs the information presented is the same but the way the information is structured in the instrument cluster area varies. Each design shows the car driving itself (automation mode) and the car requesting the driver to take over (TOR). During a situation that requires human intervention, each design communicates a TOR to the driver.

To create an immersive experience for the users, the designs were made into videos that respond to videos of a simulator driving in roadways. The design videos were then integrated into the simulation videos to make a cohesive video of a design responding to what was happening in the roadways of the simulation video. The purpose of these videos was to test how users calibrated their trust to each design and to understand the workload requirement for each design.

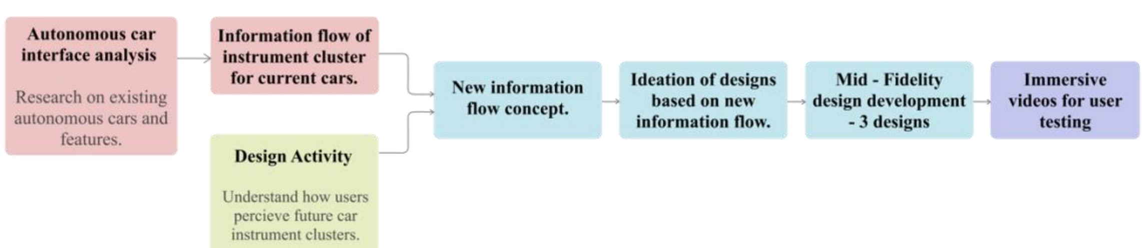


Figure 3.1 Methodology Framework

The framework (figure 3.1) above gives an overlook on the steps that were taken to arrive at the final design that were tested. Each step of the process was taken to make informed decisions on the following steps. The process is similar to the Design process which starts with a research phase and continues into the Design phase. The researcher performed desk research and user research to understand current car displays and then moved into ideating and prototyping the designs. The final product of the process was the immersive videos that were used to evaluate the designs with users.

3.1 Autonomous car interface analysis

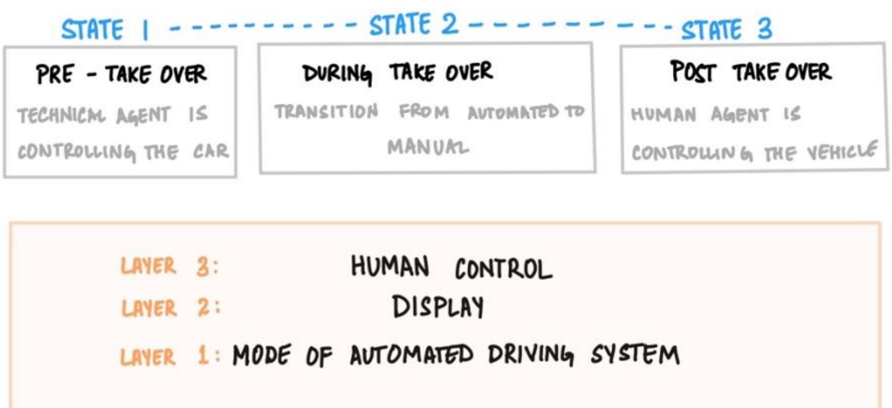


Figure 3.2 Level-3 autonomous car driving model

The model (Figure 3.2) above was framed to understand the different states of the car which would assist us in analyzing the car display interfaces. The State 1 is when the technical agent is in control of the vehicle, State 2 is during the takeover when transitioning from automation to manual and the last state is after the takeover, when the human takes control. We see that there are three layers to communication at any state. The bottom most layer is the status of automated driving system which is predicted using the technical and computational capabilities of the car.

Layer 2: The display is the HMI that communicates to the human through visual cues. Layer 3: Human control, is the perception, interpretation and action of the driver in response to the visual cues. Since the focus of this study is to design the display, we continue to research based on layer 2 and layer 3 for each state.

A few cars in the market today are equipped with SAE Level 2 and Level 3 autonomous features. Self-driving taxis have reached SAE Level 4 autonomy in certain geo-fenced areas. The visual UI of six such cars (Tesla, Waymo, Cruise – General Motors, Audi A8, Cadillac) was analyzed for each state as explained in the model. At each state, the car UIs were asked five questions to help understand better the display and human control levels (Table 3.1). These questions were answered by watching YouTube videos, reading articles and product websites.

Layer 2: Display	What is displayed?
	How is it displayed?
	How is it prioritized?
Layer 3: Human Control	What should human understand?
	What action should human take?

Table 3.1 Car interface analysis questions

This activity gave the researchers an idea of the high-level information drivers expects to receive from autonomous cars. It also reveals how car manufacturers cluster the different information for their drivers. Based on this analysis an information flow (Figure 3.3) was made that highlights the information that is always present on screen and the information that is only occasionally present.

Figure 3.3 Information flow current car interfaces

3.2 Design Activity

The goal for conducting this design activity was to understand how people perceive the instrument cluster design of a Level-3 autonomous car to be like. The activity was conducted with five participants. Two activity sessions were conducted remotely while the other three were conducted in person. The participants were provided with an image of a blank instrument cluster area (Figure 3.4) of the dashboard which was kept stationary on the artboard. Different instruments were color coded and there was a re-sizeable and movable circle next to each instrument. The participants were informed that the car for which they were designing is a level-3 autonomous car that might require human intervention during the introduction presentation. The task for each participant is to pick an instrument and place it on their preferred spot on the blank instrument cluster area (Figure 3.5). They can then resize the circles to indicate how big or small they would like it. The circles were made transparent to allow the participants to overlay the circles if they wished. The participants were asked to think out loud as the make their decision on the position and size of each instrument. The sessions were screen and audio recorded.

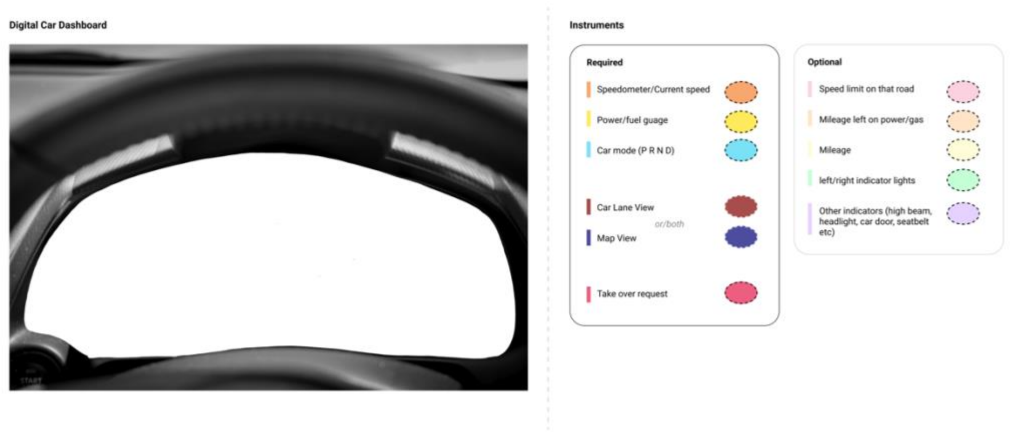


Figure 3.4 Design activity template - Blank instrument cluster area

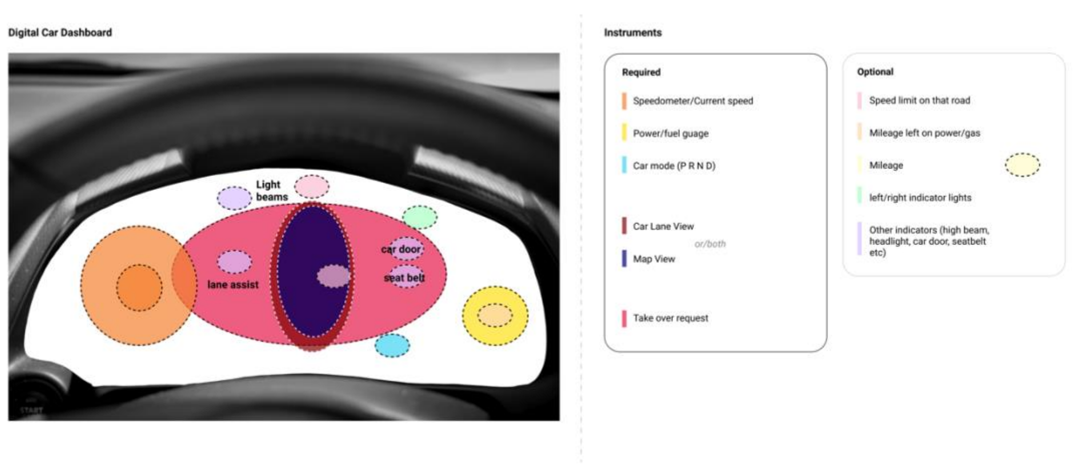


Figure 3.5 Completed design activity example

To analyze the instrument cluster data from the design activity, this study choses one instrument, copied that instrument from each participant's activity and overlayed it onto a new blank instrument cluster area, creating a sort of heat map (Figure 3.6). This gave an idea of position and size of instrument preferred by most participants. The think aloud data was transcribed and the data was grouped by instrument (Figure 3.7).

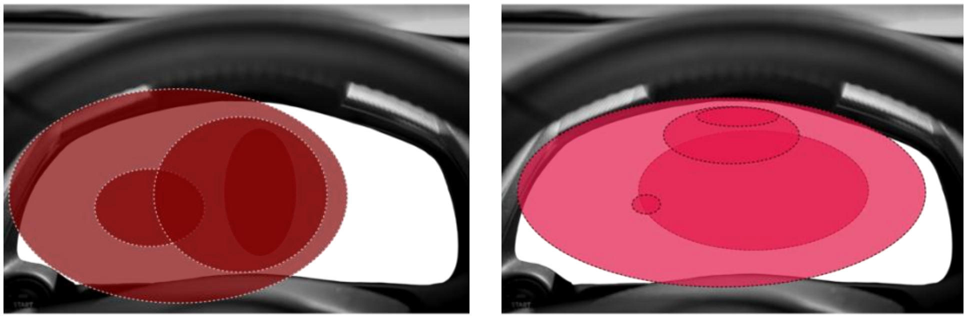


Figure 3.6 Overlay analysis of each instrument

Figure 3.7 Analyzing think aloud data

From the results of the study, we found that users prefer the entire display to change in some way when there is a TOR situation. They also say the change in color of the display to red would be the quickest indication of an emergency. On discussing about the view that they would like to see on the instrument cluster, some preferred just the lane view while the other participants preferred to see both the lane view and map view/navigation. The participants mention that they would like to see the obstacle ahead on the lane view. During autonomous drive the users wished to see which direction the car is moving and the actions the car is about to take (lane change, turns, stop etc.) in prior.

Based on these findings the current information flow (Figure 3.8) that was created from the car interface analysis was modified. The image below highlights the modifications made.

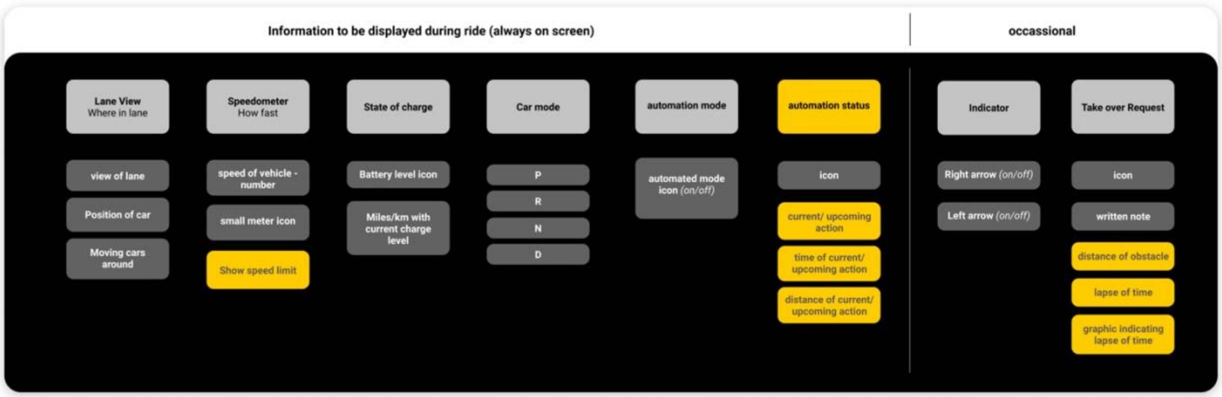


Figure 3.8 Modified Information flow

Some of the most important changes we made was in the automation status and Take over request. The automation status was bought over to be displayed at all times during the automated drive. We added few more information that could be kept constant like the current action of the vehicle. Two important information was added to be presented during a TOR. The distance of the obstacle and lapse of time/time remaining to obstacle in a graphic representation. These changes were made in reflection of the findings from the design activity.

3.3 Ideation and Concept Development

The next step was coming up with design ideas with the help of the modified information flow. Ideas with different layouts and graphical elements were brainstormed. These ideas were discussed in detailed among the researchers and three designs were finalized. The final design concepts were designed and developed on Figma (https://www.figma.com/).

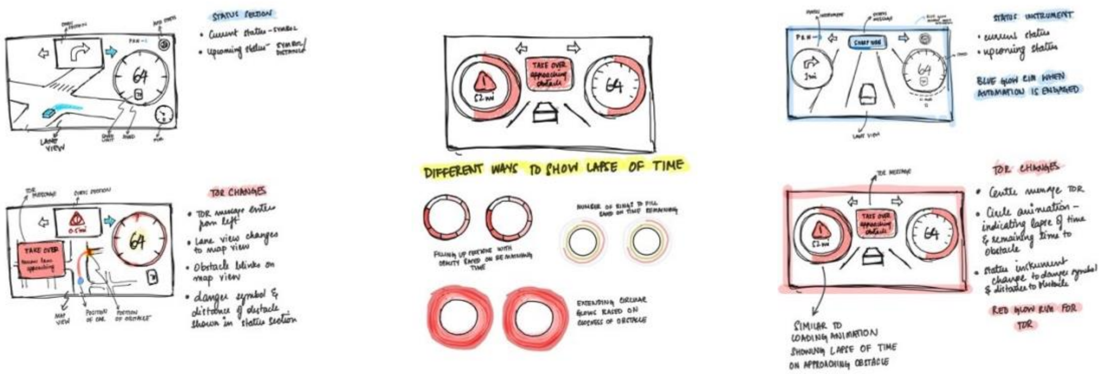


Figure 3.9 Some Ideation sketches

Figure 3.10 State 1 Automated drive designs (Design A, Design B, Design C - left to right)

Figure 3.11 State 2 TOR Designs (Design A, Design B, Design C - left to right)

The Automated drive state (Figure 3.10) and the TOR state (Figure 3.11) was developed for each design as per Rule 2 and 7 of Schmidt and Herrmann rules for intervention design that says feedback must be offered and do not require users to remember any previous states [30]. The design language and color schemes are kept consistent within each design (Rule 1 of Schmidt and Herrmann rules for intervention design [30]). Every TOR design is marked by the change in the overall display colors and change in the information displayed (Rule 3 of Schmidt and Herrmann rules for intervention design [30]). We see some elements that are constant and others that are different across each design. The color schemes, font, font size, symbols, symbol dimensions, gas bar, speed limit and indicators are kept consistent across the designs. The layout, information displayed, road views, graphic elements for time and speedometer of each design is varied. A more detailed description of the differences between each design is presented in table 3.2.

	Design element	**Design A**	**Design B**	**Design C**
1	**Showing lapse of time**	Circular time countdown dial.	Obstacle map view showing car moving towards the obstacle.	Obstacle map view showing car moving towards the obstacle. Two load bars in map and status sections.
2	**TOR status**	On the left, in the circular section in place of the status. Horizontal orientation of time and distance	On the right in the rectangle pop up below the TOR notification in place of status. Vertical orientation of time and distance.	On the right section of the display In the place status is Horizontal orientation of time and distance.
3	**TOR notification**	Top center Rectangle Slides in from top	On the right Rectangle pop up Above status rectangle	On the right section Rectangle Below status symbol
4	**View (during the automated drive)**	Current lane view	Third-person view	Current lane view + map view
5	**Status**	On the left in Circular section	On the right in Rectangle pop up	On the right in Rectangular section
6	**Speedometer**	On the right dial + number	On the left Dial + number	Center top Number

Table 3.2 Differences across each design

3.4 Immersive videos

Research introduces different methods that can be adopted to test HMI for autonomous car applications. Some researchers use advanced automated cars like Tesla [10] others use a driving simulator surrounded with a display [16] or projectors [17]. For the purpose of this study videos of roadway drives were superimposed with each design responding to the roadway scenarios in sync. Participants were made to sit in front of a steering wheel prop as an object for response during a TOR situation. The videos were played on a 42" large TV to bring the users attention to the road and display. The roadway videos had simulated roadway sounds. This user testing session being the first exposure of the designs to users, a low-fidelity method of creating an immersive experience was employed. As per the feedback we received during the user study, the roadway videos were found to mimic real-time scenarios well and the instrument cluster design responding to the roadway situations was easy to discern. The study set up was considered reasonable to test user reactions to each design.

The videos of roadway drives were recorded videos of a person driving a simulator car. The videos were revised through and two specific scenario clips were chosen. The testing scenarios were chosen based on the paper "Testing Scenarios for Human Factors Research in Level 3 Automated Vehicles" [18]. One scenario was chosen to be of high urgency and the other was chosen to be low urgency. Urgency was taken as a factor for scenarios since take over is a time-based challenge for drivers. The low urgency (LU) takeover scenario was a construction site that was known by the system rom backend. This gave users 20 seconds to take over before the construction site arrives. The high urgency (HU) takeover scenario was the presence of a parked car on the driving lane sensed by the on-board sensors. This gives the drivers 6 seconds to take over the driving task. To gauge the innate trust of the users, baseline videos for each design were made. The baseline video did not have any take over scenario.

Once the scenario videos were made, a sequence of screen animations were created in After Effects for each design in response to the scenarios. The scenario videos and design videos were combined on premiere pro (Figure 3.12).

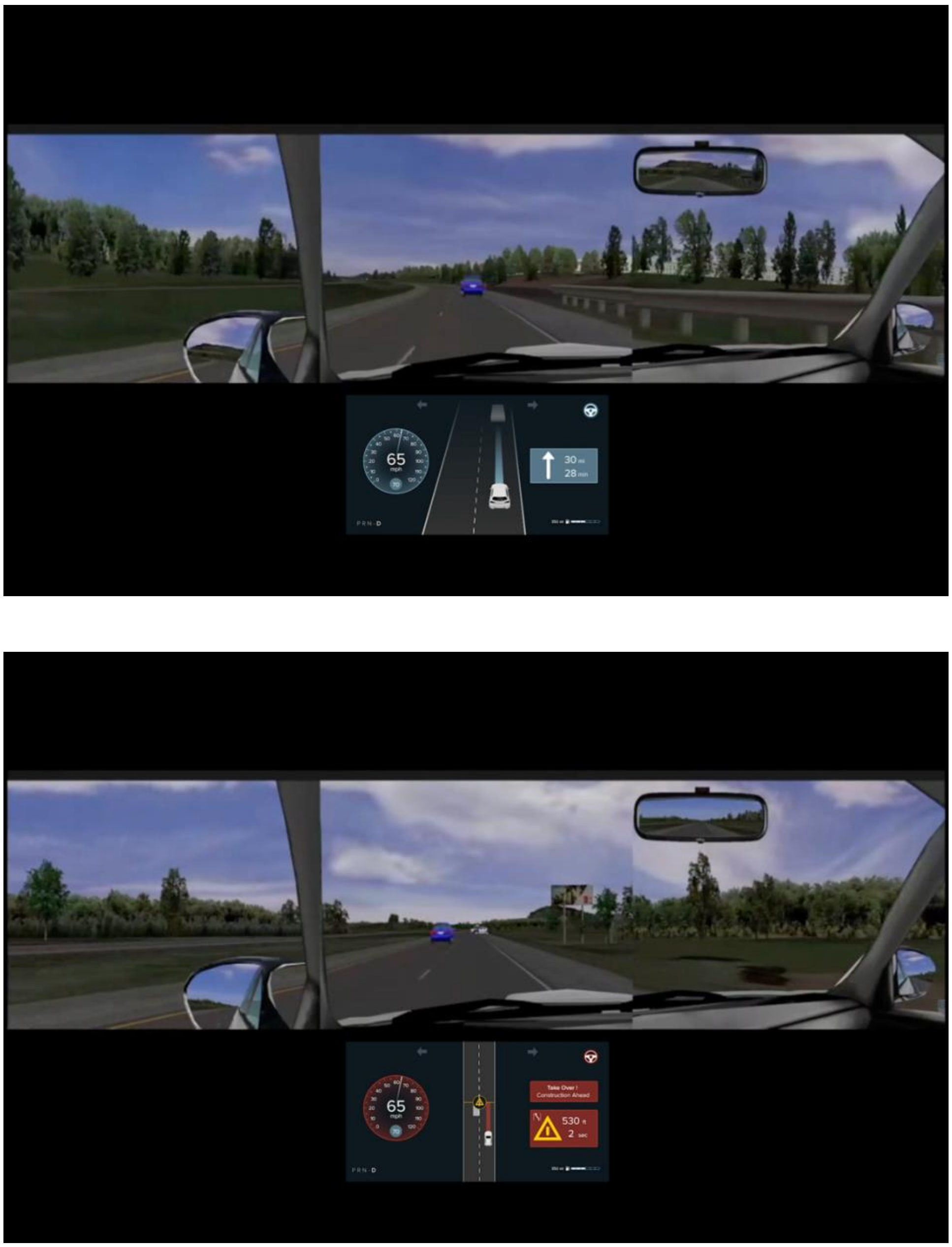


Figure 3.12 Combined scenario and design videos

3.5 User Study

A total of fifteen participants were recruited for the user study. Each participant was given a $15 gift card in compensation for their time and participation. Emails and messages were sent out to public groups and friends inviting them to participate in the user study. 14 out of 15 participants had driven a car in the last three months and one participant in the last six months of the study being conducted. The age of the participants ranged from 20 to 36 years.

The user study took place in three phases (Figure 3.13). Prior to the in-person user study a participant screening survey was sent to collect demographics and learn about their driving experience. This survey included two descriptive questions: 1. How do you think Automated vehicles would impact how people travel in the future? How do you think Automated vehicles would impact how people travel in the future? 2. Why or why not would you trust a highly automated vehicle? These questions were asked to understand their expectations and expected reactions towards autonomous cars in the future.

Phase 2 and Phase 3 of the design study took place in-person. During the Design Evaluation (Phase 1), participants watched each design in response to a high urgency, low urgency and baseline situation (Figure 3.14). Keeping the environment constant, like the experiment location, position of the steering wheel prop, table and the 42” television, the participants watched nine videos one at a time. To avoid selection bias the users were asked to pick a design and scenario from two piles which randomized the order in which each participant watched the videos. The participants were asked to keep their hands free while the autonomous car was driving them in the video. When they felt the instrument cluster display communicated a takeover situation, they were asked to react to the situation by touching some part of the steering wheel model (Figure 3.15). At the moment they made contact with the steering wheel the video was paused and the time on the video was recorded. The place on the steering they held onto was photographed when they reacted. After watching each video, they were asked to complete the STS-AD survey for Trust and the NASA TLX survey for Workload.

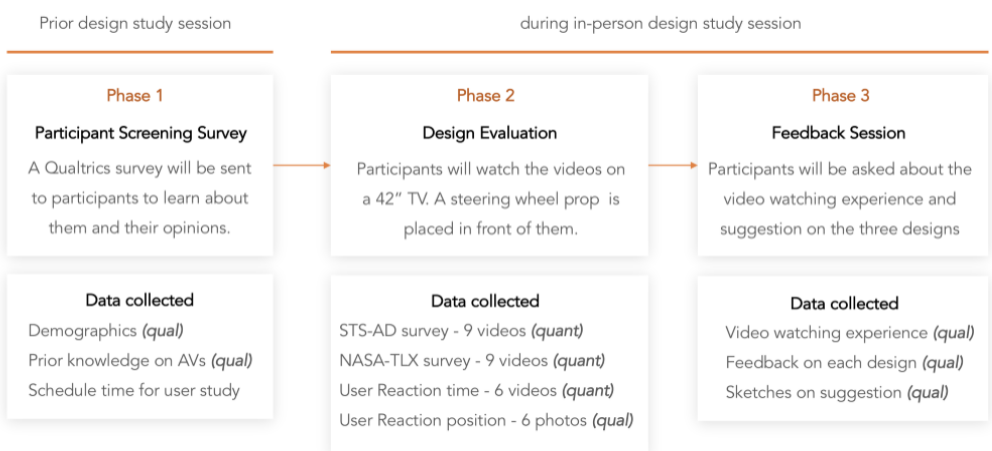


Figure 3.13 Three phase user study

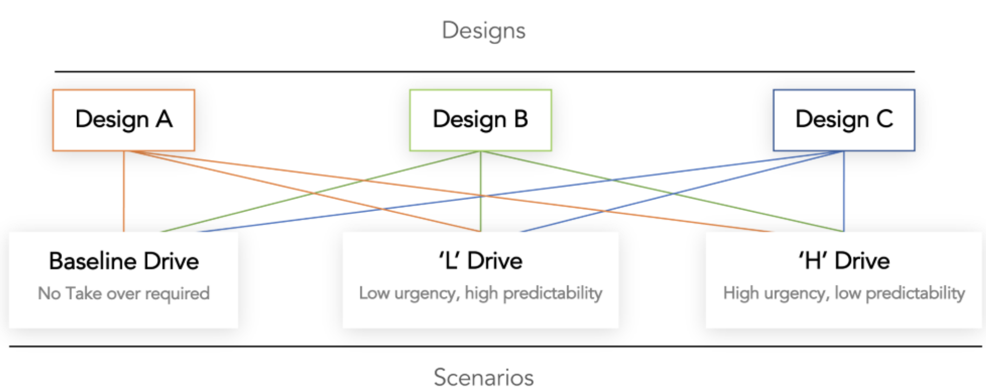


Figure 3.14 Video combination of designs and scenarios

A 10-minute feedback session (Phase 3) was conducted at the end of the design evaluation phase to learn more about the participants critique on each design. This would help us have qualitative explanations for survey results. They were given three sheets of paper that had the images of the designs on each sheet. This helped the participants to recollect the designs they saw on video. The sheets had a blank workspace that encouraged participants to draw out or write down what

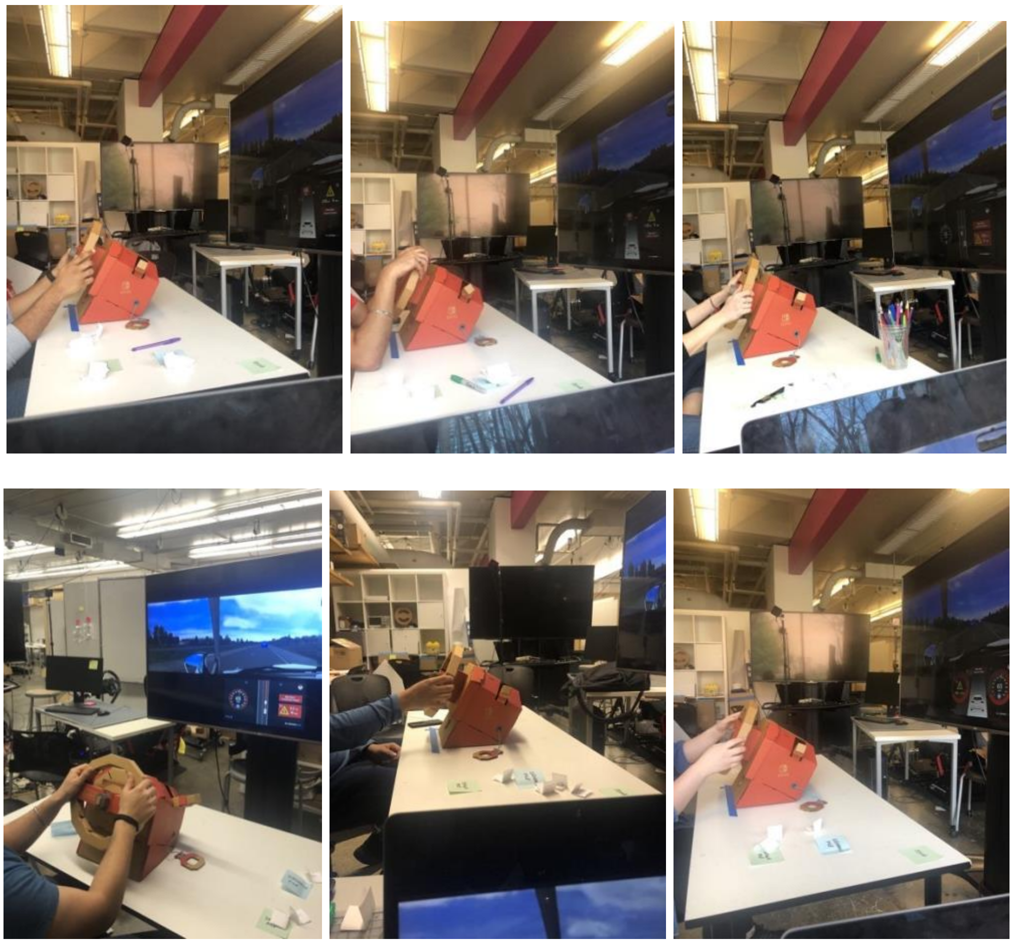


Figure 3.15 Participants contact with steering wheel prop during TOR

they liked and disliked about each design. They were also encouraged to discuss their thoughts out loud which was audio recorded for each participant.

A 10-minute feedback session (Phase 3) was conducted at the end of the design evaluation phase to learn more about the participants critique on each design. This would help us have qualitative explanations for survey results. They were given three sheets of paper that had the images of the designs on each sheet. This helped the participants to recollect the designs they saw on video.

The sheets had a blank workspace that encouraged participants to draw out or write down what they liked and disliked about each design. They were also encouraged to discuss their thoughts out loud which was audio recorded for each participant.

3.6 Data Collection

Qualitative and Quantitative data were collected at different phases of the user study. Figure 3.13 gives an idea of the data that was collected for each participant during and prior to the design study session. From the participant screening survey was sent out to collect demographic information, understand participants driving experience and prior knowledge on AVs.

The quantitative data collected during Design Evaluation phase were the STS-AD and NASA-TLX survey. The STS-AD survey was focused at understanding the participants Situational Trust. It is a six-item scale that is aimed at evaluating the participants perspective of the automated driving context's potential risk and driver's self-efficacy for operating the automated system [12] (Table 3.3). The response for each item is recorded on a 7-point Likert scale (1-completely disagree; 7 – completely agree). Items 2, 4, and 5 are reverse scored and these values were reversed during data analysis.

The NASA-TLX survey was focused at understanding the participant's workload in responding to each automated driving and takeover video. It gives us an overall workload score based on the ratings of six items: Mental Demands, Physical Demands, Temporal Demands, Own Performance, Effort and Frustration [37]. The technique of taking the weighted average of the items was not utilized in this study. The items were utilized in their raw form. The response for each item was recorded on a 7-point Likert scale that was later mapped onto a 100-point scale during data analysis. As mentioned before, the time (seconds passed) on the video at which the

Item	Situational Trust Factor	Item Abbreviation
I trust the automation in this situation.	Type of system, system complexity	Trust
I would have performed better than the automation in this situation. (Reverse scored.)	Self-confidence, subject matter expertise	Performance
In this situation, the automated vehicle performs good enough for me to engage in other activities (such as reading).	Perceived benefits, workload, task difficulty	NDRT (non-driving related task)
The situation was risky. (Reverse scored.)	Perceived risks	Risk
The automated vehicle made an unsafe judgement in this situation. (Reverse scored.)	Perceived risks	Judgement
The automated vehicle reacted appropriately to the environment.	Perceived risks, perceived benefits	Reaction

Table 3.3 Situational Trust Factors related to each STS-AD item

participant makes contact with the steering wheel was recorded. The time at which the TOR was presented on that video was subtracted from the time recorded (time of contact) to obtain the reaction times in seconds.

The qualitative data collected during the design feedback in the form of audio recordings and written/drawn participants notes were transcribed separately for each participant. The designs that were most preferred by each participant were also interpreted from the audio recordings. These transcriptions were much useful in explaining the trends we see in the qualitative data.

Based on previous work, we put forth the following assumptions:

A1: There should be a drop in trust score from the baseline to the TOR situation

A1.1: The drop in trust score from baseline to High Urgency (HU) TOR should be greater than the drop in trust score from baseline to Low Urgency (LU) TOR, i.e. Trust score for LU TOR should be greater than HU TOR

A2: The reaction times for LU TOR can be greater than HU TOR. This can imply how well the design communicated urgency.

A3: The workload should be greater for TOR as compared to baseline

A3.1: The workload score to complete a HU TOR must be greater than the workload score to complete a LU TOR.

CHAPTER 4. RESULTS ANALYSIS

Qualitative and quantitative data was collected from the user study to answer the questions that was raised in the beginning of the research. Trust scores, Workload Scores and Reaction times were the quantitative data collected through the study. After watching each video users were asked to fill out two surveys, the STS-AD for understanding the participants situation trust and the NASA-TLX for understanding the participants required workload. While the user was watching the video, the user reacted to a TOR by holding onto the steering wheel model placed in front of them. At this moment of contact, the video was paused and time on the video was recorded.

Qualitative data was collected for each participant during the feedback session on completion of all Trust and Workload surveys. They were asked their thoughts on each design along with which design they liked best. They were given three sheets of paper with the design prototypes printed on them as worksheets where they could write or sketch out their feedback and suggestions.

4.1 Reaction time data

The reaction times were calculated and recorded on an excel sheet. 3 of 90 (15 * 3 * 2) TOR cases failed to complete the required TOR gesture in the end. The participants mentioned that they reacted much later as they did not understand correctly what needed to be done in the situation. The rest 87 valid reaction times were analyzed for frequency of occurrence (Figure 4.1) in different designs and the averages for different designs (Figure 4.2)

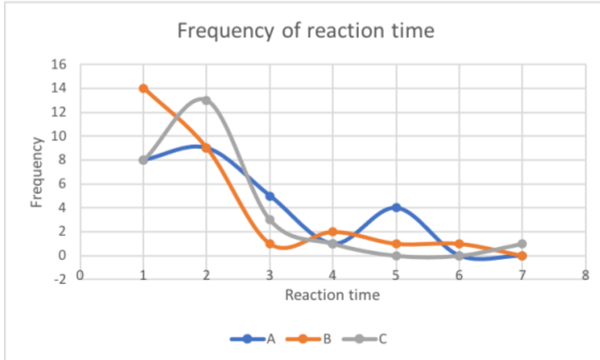


Figure 4.1 Frequency of reaction time occurrence

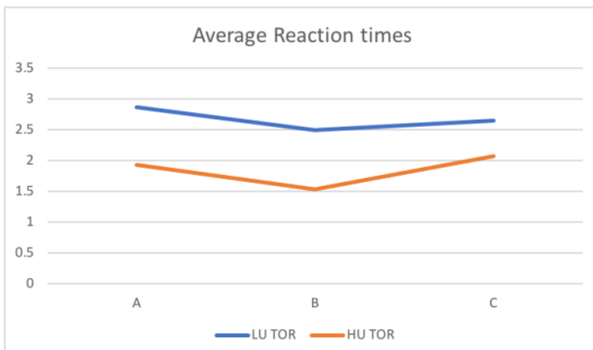


Figure 4.2 Average reaction times

Design B received a better reaction time than others. 14 out 30 TOR reactions had a reaction time of 1s which is the least reaction time. From the average reaction times (Figure 4.2) we can evidently see that Design B has a lower reaction time in both Low and high urgency TOR (LU TOR & HU TOR) as compared to other designs. From the feedback session, participants reported that they found Design B to be more alerting like the view changing from a normal following view (Figure 4.3, left) to a top view (Figure 4.3, right) which is why they might have shorter reaction times as compared to the other two designs. -Additionally, the difference of the average reaction times for the 20 sec and 6 sec TOR in Design B is the most apart. This is also good because it might imply that users get a better sense of different levels of urgency, to react appropriately.

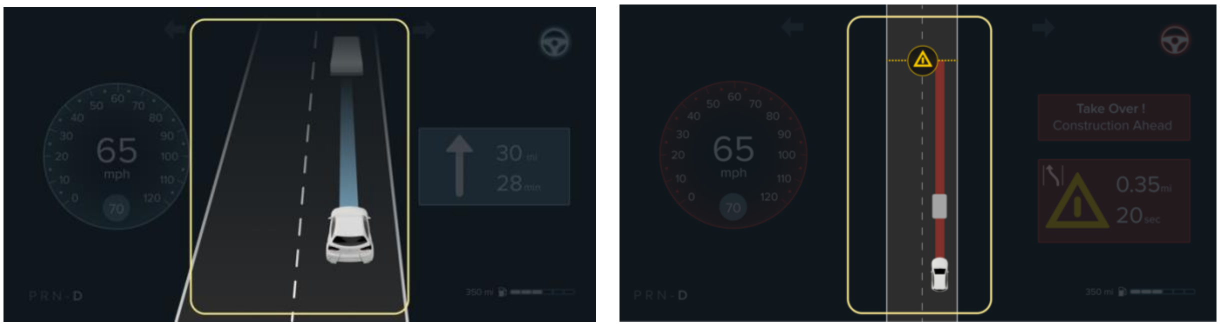


Figure 4.3 Highlighting change of views in Design B during a TOR

Looking at the Frequency chart (Figure 4.1), the second peak around 5s reaction times in Design A is interesting. This could mean that Design A might lead to different responses among participants. For most participants it takes around 2s. But it takes longer to interpret and understand for others. From the average reaction time we can find that participants take longer to respond for LU TOR in Design A. Participants who performed quicker reaction (reaction time = 2s) in Design A mentioned that the lapse of time was best communicated by the countdown element (the two red circle that outline the alert in the left and the speedometer in the right, Figure 4.4) in Design A. Though others found the representation confusing, they felt they could learn this new graphics over time. It could explain why Design A received varied performance.

Figure 4.4 Highlighting countdown element in Design A during LU TOR (left) and HU TOR (right)

From the average reaction times of Design C, we could imply that users reacted slower than the other designs for HU TOR, while it is in between Design B and A for LU TOR. It might be because Design C provide more graphics in different panels than A and B. It is probably fine if users in LU case that have more time to digest, but might be problematic in short time like HU. In the interview, some mentioned that "Design C has too many images of car" (Figure 4.5, right). However, others said "I liked the sectioning of the message and symbol in design C. Makes me feel like they are all telling me one thing".

Figure 4.5 Highlighting automated drive (left) and TOR (right) UI for Design C

4.2 Trust Calibration

Each participant watched nine videos and filled out the STS-AD scale survey after each video. A total of 135 surveys were collected. The survey data is presented in the form of a box plot (Figure 4.6) to understand the distributional characteristics of the trust scores for each video. The average difference of the trust score between scenarios were calculated to understand how much the trust score for each scenario differ within each design (Figure 4.7). Though there is no statistical significance between the overall trust score of the three designs, the box plots and the average of the difference of trust scores between scenarios reveal some interesting facts.

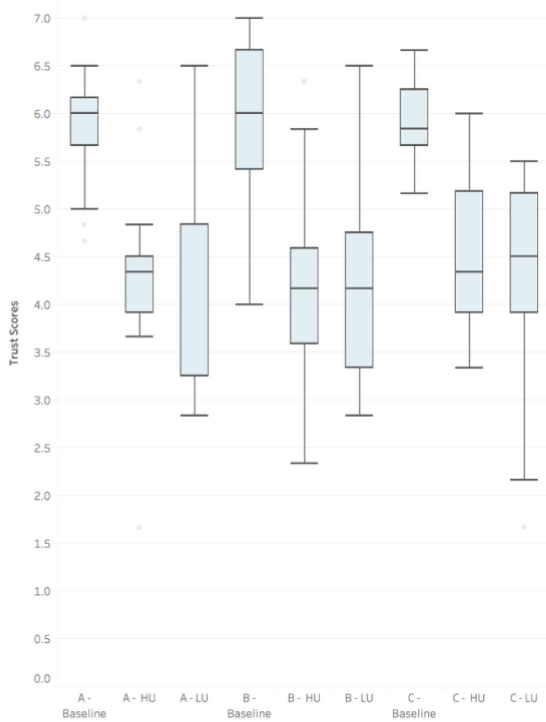


Figure 4.6 Box plot of trust scores

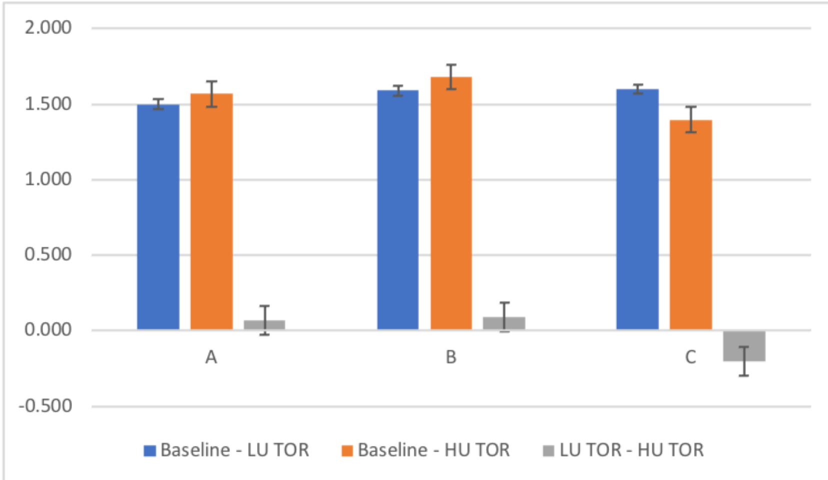


Figure 4.7 Average difference of trust scores between scenarios

The trust scores for both TOR situations are lesser than that of the baseline trust for all designs. Comparing the baselines, we see that the Design B trust scale data is distributed to show higher trust followed by Design A then Design C. Looking at the upper quartile of the boxes for the TORs, Design C seems to have a slightly better trust score (LU TOR = 5.167, HU TOR = 5.183) than Design A (LU TOR =4.833, HU TOR = 4.5) and Design B (LU TOR = 4.75, HU TOR = 4.583). We see an odd behavior of Design C in the average difference chart. The Baseline – LU TOR average (1.5) is greater than the Baseline – HU (1.567) TOR average. This shows that the drop in trust for a low urgency situation is greater than the drop in trust for a high urgency situation which violates our assumption (A1.1). LU TOR – HU TOR being negative also says that the LU TOR trust score is much greater than the HU TOR trust score for Design C.

In Design B the drop from baseline median (median trust score = 6) to TOR 2 median (trust score = 4.167). The drop-in trust difference (1.833) is greater as compared to Design A (1.67) and Design C (1.5). The average difference in trust scores also suggest that the drop from the baselines to the HU TOR is slightly higher for Design B as compared to Design A and Design C, and the drop from the baseline to LU TOR for Design B is slightly greater than Design A. LU TOR-HU TOR is also greatest for Design B showing that low urgency situation generally scored higher trust than high urgency. During the feedback session users did mention that Design B TOR was more "in the face" as compared to the other designs. This might be why Design B might mostly have a greater drop in trust score as compared to the other two designs.

We generally see that there are quite some outliers in the data (grey dots in Figure 4.6). The whisker lengths of some of the cases are longer than their quartile areas. Researchers see that participants tend to have a highly varied perception of trust. From the interview feedback different users had contradicting opinions on each design. This could have led to the highly variable perception of trust among participants.

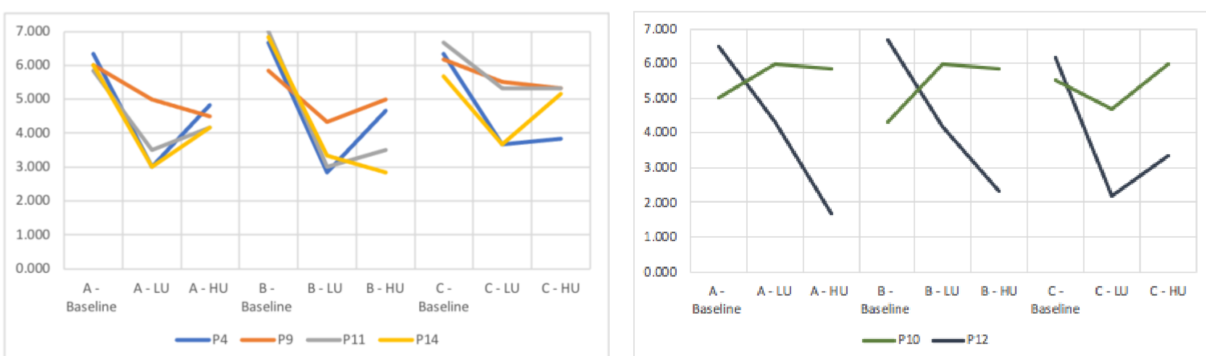


Figure 4.8 Individual Trust Scores for specific participants

Drilling in and taking a closer look at the data I found some interesting patterns in participants whose trust scores violated our initial Assumptions on trust calibration for certain designs. Looking at Figure 4.8, left, we see that for P14 the trust score for HU is greater than LU in Design A and C but for Design B we see that the trust is calibrated as per our assumptions (LU

	RT A	RT B	RT C
P4	1	1.5	1.5
P9	2.5	1	2
P11	1.5	1.5	2
P14	2	4	2.5
P12	1.5	1	1.5
P10	2.5	2	1.5

Table 4.1 Overall Average Reaction Time (RT)

TOR>HU TOR). According to the participant P14's feedback, they strongly dislike Design B and prefer UI elements from Design A and C. P14's average reaction time reflects the participants preference. They have a fastest reaction to Design A. Similarly, In P11 where we see that LU trust score < HU trust scores more so in Design B and A than Design C, the participants expressed dislike for Design C, and the lower reaction time for Design A & B (1.5s) is consistent with their feedback. A similar pattern is seen P4 and P9 where though trust is calibrated as per our assumption in some designs (P4: Assumed calibration in Design C over A and B; P9: Assumed calibration in Design A and C), participants preference (P4: preferred Design A & B; P9: preferred Design B and disliked Design A) and faster reaction times (P4 least RT: Design A, 1s; P9 least RT: Design B, 1s) correspond to designs that have LU trust score < HU trust score.

In Figure 4.6 (right), the patterns revealed by participants P12 and P10 are quite different from the pattern we see earlier. In P12 we see that their trust is calibrated as per our assumption (LU TOR>HU TOR) in Design A and B and not in Design C. In the participants feedback, they mention preferring Design C over the other two designs. But we see that, P12's reaction time is least for Design B. This shows that the reaction time is consistent with the trust calibration as per our assumption. A different pattern is noticed in participant P10. Here, the trust calibration for

Design A and B is not as per our baseline assumption. According to the participants feedback Design B struck to them the most. But, in the reaction times for the same we see the participant react fast for Design C at an average of 1.5s. This shows consistent behavior between reaction time and trust calibration as per our assumptions but is inconsistent with feedback.

4.3 Workload

Each participant filled out the Workload scale survey after watching each video (*n* = 9). A total of 135 surveys were collected. The survey data is presented in the form of a box plot (Figure 4.9) to understand the distributional characteristics of the workload scores for each video. The average of the workload scores for each design are also calculated and presented (Figure 4.10).

We can evidently see from the box plot that the workload for TOR situations is greater than for baseline situations which matched with the fact of urgent events shown in video scenarios. Comparing the baselines, we see that the workloads are lower for Design B followed by Design C then Design A. From the median levels we can also see that the workload requirement for the low urgency situations is lesser than the high urgency situation (A3.1).

Looking at the whiskers for Design A TOR, participants' perception of workload requirement for TOR in Design A is more varied than in Design B and Design C. In LU TOR, Design B provides the lowest workload and more convergent, comparing against Design A which brings quite different cognitive workload among participants. From the feedback it was evident that there

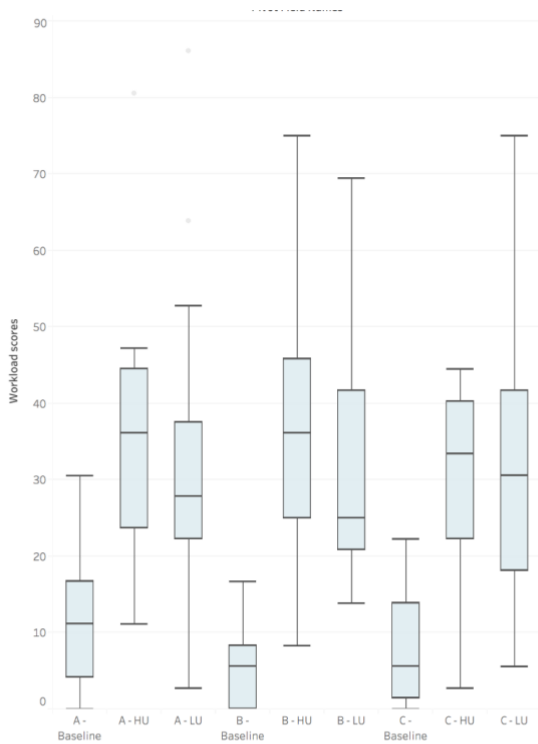


Figure 4.9 Box Plot for workload scores

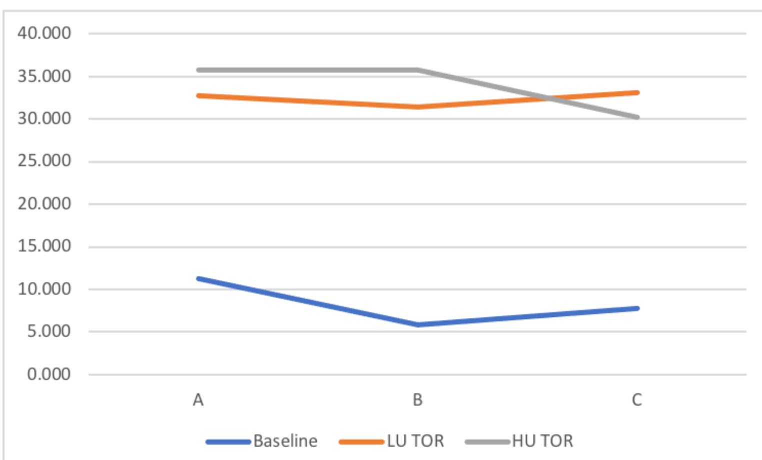


Figure 4.10 Average workload scores

were contradicting opinions on Design A. Some reported that “Design A was redder which I like”. However, others said “In design A the dials (the countdown cycle) are on both sides which makes it hard to look at.”. From the whiskers for Design C, LU scenario we also find the workload scores are more spread for higher scores. The upper quartile of Design C LU TOR is higher than Design C HU TOR, which violates our assumption A3.1.

Looking at the mean scores of the workload we can see that Design C has a lower workload in the HU situation as compared to the LU situation. This consistently shows that Design violates assumption 3.1. In interview, different participants addressed different aspects of Design C that might make them feel less workload. For example, one said “I like the route in Design C”. another said “I like Design C view where I can see how far I am from the obstacle” and also “The loading symbol in design C made me clear how much time was remaining”. But others also reported that “In Design C, I like the loading bar on top but it is not that obvious” and “In Design C there is too much information on one side”. We feel further investigation is needed to explain Design C’s workload scores.

	A	**B**	**C**
20 sec TOR	-0.2562309	-0.610	-0.364
6 sec TOR	-0.3879689	-0.551	-0.502

Table 4.2 Correlation coefficients between Trust and Workload scores

From the correlation coefficients (Table 4.2) we can see that Trust is negatively correlated to workload in all cases which means Workload reduces with increased Trust on the system. It also could mean if we want users to have better trust calibration with the system their perceived workload should be calibrated inversely. There is a larger negative correlation for both TOR in Design B and HU TOR in Design C, which means some of the same design in Design B and Design C like the top obstacle top view of the obstacle provide a better effect to reduce Trust while increasing Workload. And it was also mentioned in interview that participants who like B or C said the top view of the obstacle is the key for them.

CHAPTER 5. DISCUSSIONS

5.1 Designing for Appropriate Trust

Reviewing the feedback sessions that were conducted at the end of the user study. It became clear that the most preferred design was Design B. Eight participants preferred Design B. Four participants preferred Design A and three participants preferred Design C. The quantitative data of Reaction Time collected during both TORs also suggest that Design B might be slightly better than other Designs in takeover performance. It is interesting to see that Design B has a greater drop from its baseline scores to the TOR trust scores. From this we can interpret that Design B comminute more urgent information to the user. As Lee and See proposed [35], designing for appropriate trust in automation, the calibration of trust in Design B according to different urgent situation is as per our assumptions, since drivers do need a reduced trust level to raise their risk awareness and to appropriately react to a situation that the system cannot handle. Similarly, the workload data suggest that Design B required lower workload for a LU TOR and higher for the HU TOR.

One participant mentions that the information in Design B struck to him making it easier to react. The emphasis on concrete realistic representation of the lanes could have resulted in increased level of baseline trust in Design B [35]. Participants found the birds eye view that was presented in the non-TOR situation (Figure 5.1, left) more contextual than any other view presented giving them a better idea of lanes. Another participant also mentioned they liked the overall layout of Design B with the lanes in the center and the speed on the side.

We can see from the difference of trust scores (Figure 4.2) that Design A has the second highest drop from Baselines to the HU TOR's (Design A average differences = 1.5, 1.567). One participant liked the design since they felt it was more red. Another participant mentioned that they liked the natural layout of Design A, that mimics the current speedometer, tachometer layout. This could be the effect of trust in familiarity as the users might have previously seen how such a system works. The participants did not enjoy the car view represented in Design A (Figure 5.1, right), since they felt it was not contextual compared with Design B. It might disclose why Design A received a higher baseline Trust and longer Reaction Time in both TORs.

Figure 5.1 Design B (right) and Design A (left) during automated drive

In Design C we see that trust calibration is not as per our assumptions (A1.1 & A3.1). Though the drop in trust from baseline is greater for Design B LU situations as compared to other designs, people tend to trust the system more during a HU TOR versus a LU TOR which is not a corrected trust calibration we intent to. This suggests an over trust in the system when it needs to reduce the trust and raise the risk awareness. Over trust is risky in HU TOR. On a first glance, as per our assumptions on trust calibration we may think of it as misuse which does not maximize the potential benefit of the automation [11]. But here we question if our assumptions correctly reflect the people's perception of trust and its impact on their reaction to the TOR. Figure 4.6 makes an important finding that though participants may not show trust calibration as per our assumptions, their reaction times and preferences may say a different story. This makes us question whether a generalized assumptions as made by the researchers in this study can perfectly fit to everyone's trust calibration needs in order to respond appropriately. We might have to consider trust calibration in a more individualistic manner as opposed to a general assumption.

In this study we compare the consistency of the results between our assumptions, participant feedback and reaction times. We see in some cases there is consistency with the user feedback and the reaction time but not our assumptions, and in some cases, there is consistency with our assumptions and reaction timbes but the user preferences say differently. Previous work suggests that appropriate trust in automation is a safety critical requirement [11]. Thinking about criticality of the situation the driver is in, designers should make it a priority that there is consistency between the trust calibration and reaction times rather than focusing on user preferences for reasons of safety.

In the case our assumptions are a valid assessment of trust calibration, we might say that the dense information in Design C, both the map in the left with a clear obstacle location and the driving information in the right (Figure 5.2) can give the user an impression that the risky situation is still “under the control” by the autonomous system. It suggests us if we want to let the driver have a trust calibration as per our assumption in high urgency, the design might not provide “over” information but focus on reporting the next action the driver should do.

Other detailed elements of Design C were discussed to be beneficial. Design C baseline has a lower distributed workload data as compared to Design B. One participant mentioned that they liked how the design sectioned out the information. Most participants liked seeing the route map

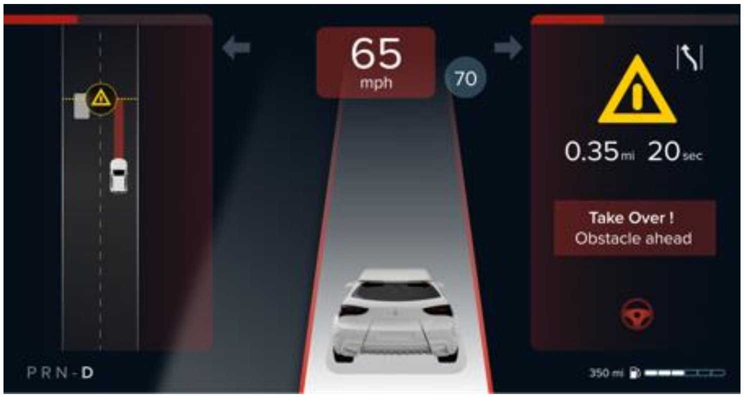


Figure 5.2 Design C TOR UI

view on the display. Design C baseline yielded lower distributed trust data as compared to other baselines. One participant mentioned that there was too much information on one side. Another participant mentioned that there are too many images of cars in the TOR view.

5.2 Influence of Design Details

A general observation is that participants tend to take over in the first few seconds (less than 7s) of the TOR being displayed irrespective whether it is a low urgency or high urgency TOR which match with literature (1.4s - 6.7s) [23].

Participants tend to react to Design B faster (Figure 4.4 above). The top obstacle view presented during TOR was well received by users as it showed where the obstacle was and communicated the lapse of time. One participant mentioned that the distance and time represented in a vertical orientation (shown in Figure 5.3 left) makes it easier to read as opposed to numbers in a horizontal orientation as in Design A and C (Figure 5.3 right). Another participant mentioned they liked having the notification below the distance and time alert as the numbers is faster to interpret than words.

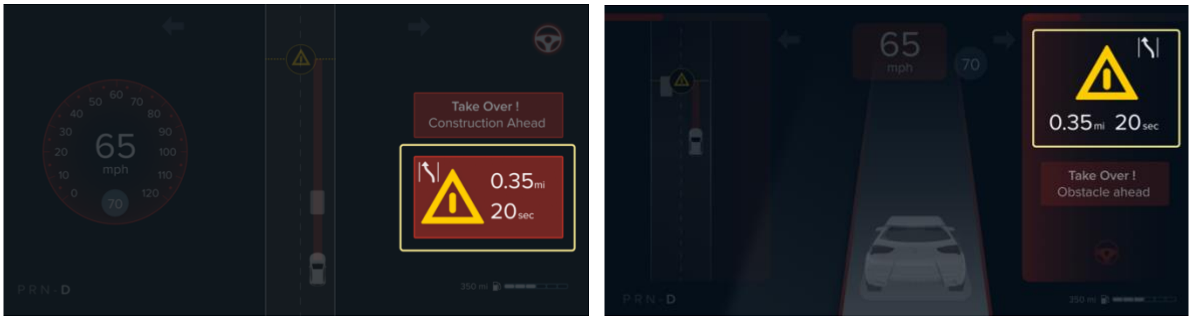


Figure 5.3 Design B alert (left), Design C alert (right)

The lapse of time (Figure 5.4, left) according to a lot of participants was best communicated by the countdown element in Design A. Though some found the representation confusing they felt they could learn it over time. It can explain the two waves of Design A's reaction frequency (Figure 4.3). The same design details might mean different depending on user's background. It also suggests that an adaption process needs to be taken into account with novel graphics for first time users. The time bars that are seen in Design C to communicate lapse of time in the top (Figure 5.4, right) was liked by some participants, but they felt the graphic was too small.

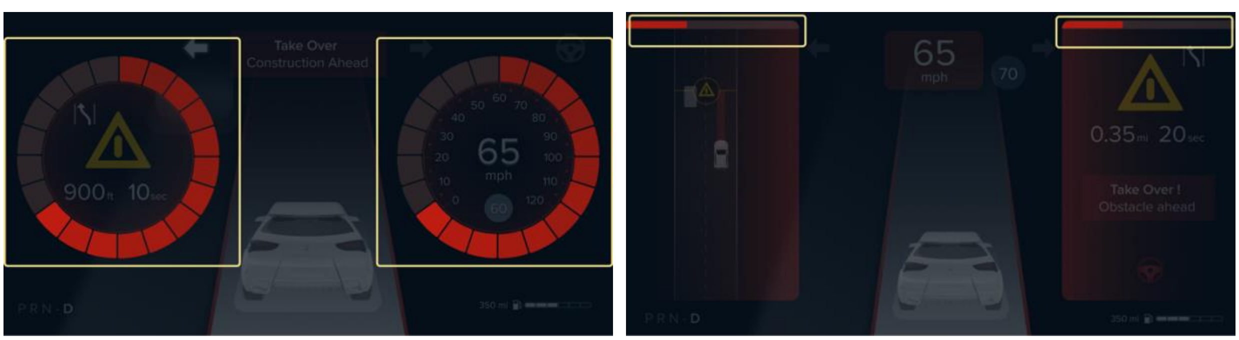


Figure 5.4 Design A countdown element (left), Design C time bar (right)

Participants did find displaying the time (like 4 sec in Figure 5.4) and distance (like 530 ft in Figure 5.4) at the same time to be confusing sometimes. They felt distance remaining to obstacle was difficult to interpret, and would confuse with other distance/mileage information. They felt time remaining was a quicker judgement to make. This suggests that the time should be prioritized over distance in such TOR cases when presented as a number on the display. It was also supported from post-hoc interview. For example, one participant mentioned that the distance and time remaining was not that obvious and he wished the display just showed him the time. This reflects Brittany E. Noah findings that quantitative display or qualitative displays that are direct or slight abstractions of numeric information is appropriate for trust calibration [38].

The pop-up notification ("Take Over!" in Figure 5.5) also plays an important role in participants' reaction to TOR. Koo et al. disclosed that a message reporting *Why* information maintains a good driving performance [39]. In this study, the position of such an alerting notification in the overall layout was preferred in two ways: the notification separately showing in the top center like in Design A (Figure 5.5, leftmost) or all TOR related information integrated on one side like in Design B and C (Figure 5.5, center and rightmost). It received polarized feedback. Participants.

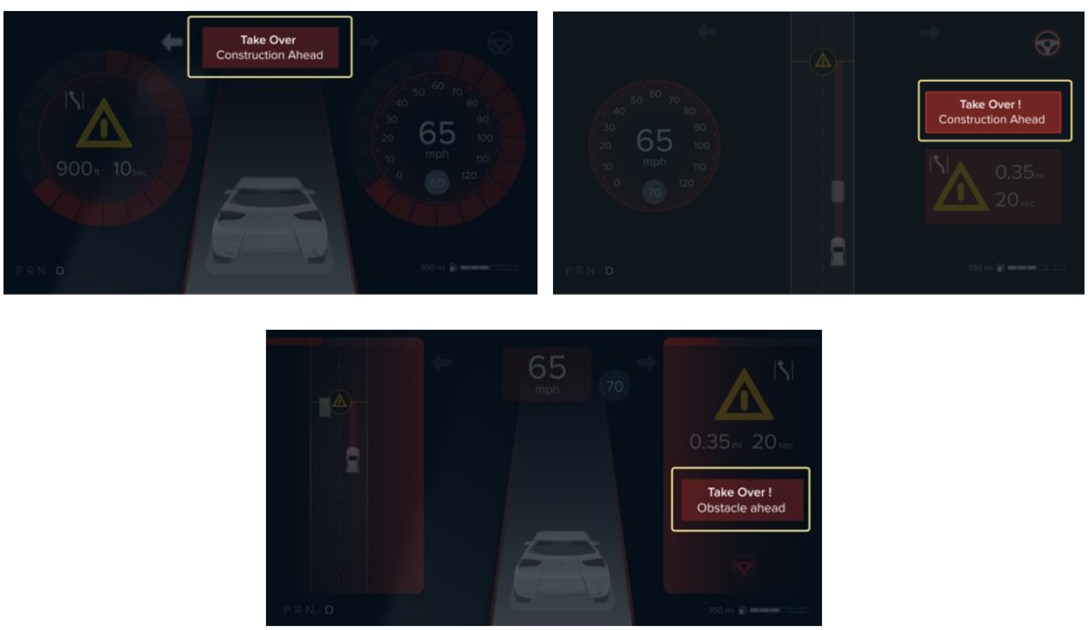


Figure 5.5 TOR notification Design A (top left), Design B (top right), Design C (bottom)

who have the behavior that seek the information between the real road condition and the display prefer Design A. They mentioned that the center position request lesser eye movement for them when looking down at the display and quickly looking back-to the road. On the other hand, other participants looking for more *Why* information from display felt it was quicker to look at all information on one side rather than spending time gathering information from different parts of the screen

5.3 Other Design Implications

Even though all three designs received satisfying overall results, the two main drawbacks in the three designs that were discovered in the feedback sessions were: 1) the lack of obvious communication of urgency and 2) need to communicate how the driver should take action (eg. change lanes, slow down etc). For example, using the change in hues and colors should communicate if the emergency situation is immediate or is low urgency. Arrows superimposed on the lane views should communicate what the driver need to do to react appropriately, like change lanes, slow down or stop.

Participants pointed out that they definitely would need another mode of communication to convey the TOR, like integrated audio and haptic feedback. One participant suggested that a conversational dialogue-based interface could be highly beneficial. The added output modalities would have an effect on the way the visual output would be perceived. Such interactions should be taken care of when designing for multi-modality.

Throughout the study the researchers see that participants' reactions and opinions are contradicting among Design A, B and C. As discussed above, some design features received controversial feedbacks. Though most people liked Design B, two participants expressed dislike for the design. One participant said that the design was too sudden while the other mentioned the change in views was too disorienting. In Design C, while one participant enjoyed the way information was sectioned, another participant mentioned that there was too much information on one side which made her feel imbalanced. The same in other design details. For example, there was an almost even distribution of participants who preferred a number representation of speed (digital style, in Figure 5.6, right) and the traditional dial representation of speed (analog style, in Figure 5.6, left). While seven participants liked the ring countdown in Design A, three participants felt it was too far apart, too scary or confusing. It indicates that there might not be a magic design that can fit all well. One design implication is, providing multiply visual representation with redundancy for the fundamental information, like the traditional speedometer dial design (Figure 5.6, left) with the number in the center of the dial could better accommodate different preferences. Designers should also support options to let users customize some display layout and graphics.

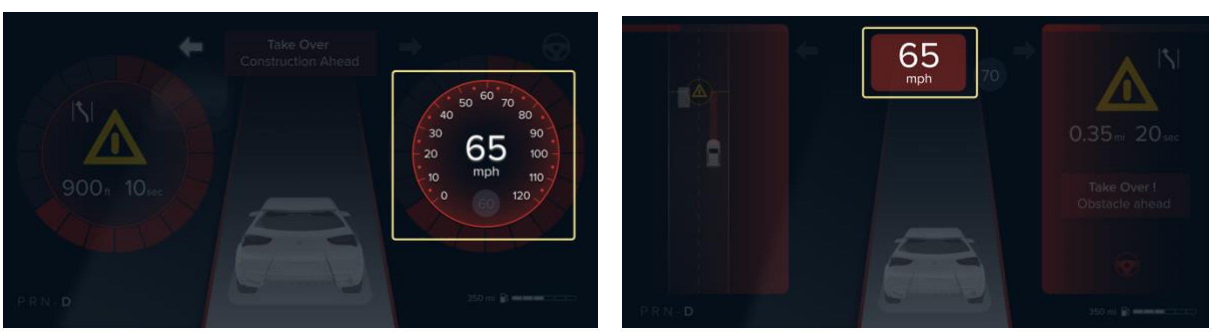


Figure 5.6 Speedometer Design A dial (left), Design C digital (right)

CHAPTER 6. CONCLUSION

At the beginning of the research, we aimed to understand how in-car display interfaces can be designed to visualize the state of the car and alert change of control to establish appropriate trust between the driver and automation. This study proofed that presenting information on the state of the automation and request to take over in an easy-to-understand visual format can promote proper calibration of trust with the system and TOR reaction. To understand the impact of different visual representations of information, we conducted a user study with participants who were instructed to watch videos of three different designs (A, B, and C) responding to three fixed driving scenarios (baseline, Low Urgency TOR, High Urgency TOR). While some elements remained constant, Design A, B and C explore different layout, information displayed, road views, graphic elements for time and speedometer design. We assessed the situational trust and workload of each participant at different points during the driving. After reviewing the results, we find Design B to be slightly better than the other Designs according to our assumptions. With the autonomous driving information communicated by realistic and contextual roadway view, that TOR information communicated by top view (TOR obstacle view) and distinctly visible time and distance to obstacle, Design B had better trust calibration and lower reaction times. Design A was also found to communicate TOR time well with the dial countdown graphic. Though trust calibration is not as per our assumption in Design C, it yields a lower workload. The layout sectioning and the time bar graphic in Design C were mentioned to be beneficial. The trust calibration seen in Design C could be owed to the over information presented in the Design. Though we make these conclusions from our initial assumption, some participant data tells us a different story. With inconsistencies between our assumptions and reaction times, we come to question if our assumptions on trust calibration can be a generalized for all. We conclude by saying that according to our initial assumptions Design B seems to have the best trust calibration among the three designs. But further research needs to be done on the subjective nature of trust calibration to make solid conclusions on design details.

CHAPTER 7. LIMITATIONS & FUTURE WORK

Though the testing conditions and environment were maintained as constant for all participants, the set-up is far from the actual environment and conditions during an automated drive as such. Researchers might find different results when participants are subjected to more realistic environments such as simulators with a responsive system in place, similar to a video game. One participant mentioned that her immediate response to the TOR was to hit the brake pedals. Not having all possible interactive options available to a driver in a car might have forced participants to react only by holding on to the steering wheel, which could be an unnatural reaction to such a scenario for them.

When asked to provide trust scores, the scenarios and driving style seemed to distract participants perception of trust. The participants focus might have shifted to evaluating the scenarios rather than the designs in such situations, which might have influenced the trust scores.

With the quantitative data not being statistically significant, it is clear that the sample size is not large enough to show pronounced patterns in the data. Maybe evaluating the designs with a larger number of participants can yield statistically significant results. Better equipment and real-time evaluation could also yield better results.

Considering the results of this experiment as the earlier part of the design process, we would like to make design iterations and improvements based on these findings. This would help in narrowing down the key visual factors and information that greatly impact trust during a takeover scenario. Constraining and limiting design variability across the designs to be evaluated could also yield in more specific results. One important area that might lack research in current day is “How should drivers calibrate their trust to the system for the best and quickest response?". From this study we see that there cannot be a single assumption that correctly assess everyone's perception of trust. This also brings up the question, "how should designers interpret the calibration of trust in different people and translate those findings into visual design that meets the users trust requirements?"

APPENDIX A. CAR INTERFACE ANALYSIS

Vehicle		Tesla	Waymo	Cruise - General Motors	Audi A8 Traffic Jam Pilot	Cadillac - super cruise (Adaptive advanced cruise control
SAE Level of automation		**Level 3**	**Level 4**	**Level 4**	**Level 3**	**Level 2**
						
Links referred						https://www.cadillac.com/ownership/vehicle-technology/super-cruise?ppc=GOOGLE_700000001297222_71700000025000383_58700003076487421_p24180568874&ds_rl=1253750&d_src=313715&d_adsrc=3876787&d_campaign=71700000025000383&d_site=GOOGLE&d_adgroup=58700003076487421&d_keyword=cadillac+super+cruise&gclick=CjwKCAjwgZuDBhBTEiwAXNofRARWUyCAEv8avyiDk3JGVA6tV-_MzmmrXc45U_fPkNYtxexLEqHDVxoCr4sQAvD_BwE&ds_rl=1253750&gclid=CjwKCAjwgZuDBhBTEiwAXNofRARWUyCAEv8avyiDk3JGVA6tV-_MzmmrXc45U_fPkNYtxexLEqHDVxoCr4sQAvD_BwE&gclsrc=aw.ds
		https://www.youtube.com/watch?v=zd6bnhgJMCI&t=786s	https://www.youtube.com/watch?v=__EoOvVkEMo&t=32s	https://www.youtube.com/watch?v=sliYTyRpRB8	https://www.youtube.com/watch?v=2TPloNOFvfk	https://www.youtube.com/watch?v=zq43DQ9yVJY&t=195s
		https://www.youtube.com/watch?v=Uvy4uZvQr1A&t=20s	https://www.youtube.com/watch?v=B8R148hFxPw&t=11s	https://www.slashgear.com/epic-games-unreal-engine-used-in-gm-hummer-digital-cockpit-07641511/	https://www.youtube.com/watch?v=oV4ee17Nf44	https://www.youtube.com/watch?v=R_WcDjoeEyE

		https://www.youtube.com/watch?v=ckib1ABJ_sM	https://www.theturnsignalblog.com/blog/waymo-design/	https://www.gm.com/content/dam/company/docs/us/en/gmcom/gmsafetyreport.pdf	https://www.youtube.com/watch?v=WsiUwq_M8lE	https://www.cadillac.com/ownership/vehicle-technology/super-cruise
Description of communication system		The vehicle is attached with a 17 inch touchscreen in the center of the front dashboard. There is a display behind the wheel (where there is the usual speedometer)	The vehicle is attached with an interface and there is a mobile application users interact with.	The vehicle does not have a steering wheel. The digital communication system can be found in the center dashboard of the vehicle.	screen behind the wheel, screen in the center of the dashboard, AI activation button on the bottom right (below the center infotainment screen.	The car has a digital display on the center and behind the wheel. There are car steering controls
State 1: Pre-Takeover	**What is displayed**	Steering screen - The directions on map to destination, current path, The lane the car is in, speed limit by law, maximum speed the car will take, the current speed of car, the average power usage, the battery charge, distance on map, temperature, time and car mode (P R N D), car/human mode (steering wheel image on screen), Other notifications. Center screen (during drive) - Map, rear camera view (can be customized to any other view)	The in car display - map - roads, route, vehicles, crosswalk, traffic lights, traffic signs, pedestrians. Unplanned events - construction work, emergency vehicles. Why the car has stopped. 2D status layer - detailed information on the decision the vehicle has taken.	The screen is said to display details about the trip, allow riders to request stops. Driving mode, path, map view displayed.	The main view is the car on the lane with a teal light on the back indication movement. There is a speed bar on bottom of the screen with a limit of 60. This bar is dynamic based on the car speed. It also shows the exact speed right next to the bar on the right side. It shows the speed limit on that road. There is a green Ai icon on the bottom right and green AI button activated on the top bar of the screen.	There is a green light on the steering wheel. Here there is a steering wheel icon on the left bottom display and on the top mid right of the display. The green light on the steerign wheel changes through three modes during autonomous cruise. Green mode, green mode blinking, blue mode flashing.
	How is it displayed	Steering screen - three informations (map, current path, power used) is split into three portions on the screen. The other information is placed below or above the screen. The steering wheel glows in blue when in car mode. The lane glows in blue lines. Other notifications is displayed as pop up from bottom of the screen. Center screen (during drive) - The map is placed on the top	In an abstract animated visual display. The display is real time and changing based on camera and sensor feedback.	*I am not sure if the picture shows what is on the in-car display or is it what is visible to remote drivers.* *The visual is a top view map of how the car is driving. Cruise os showed in white and the other cars in blue, pedestrians in green and cyclists in purple. The cars path is projected with a green light. The length of the gren light depends on how fast the car is*	Mostly the display is pictorial or graphical way representating. Except the speed is represented in two ways.	The steering wheel is a graphical icon on the display. The lights are displayed as an arc on top of the steering wheel and there are three dots lights on either side of the light arc.

		and any other on the bottom.		*movin in a straight direction . The other cars also have a blue line in fromt of them showing how they are moving. There is a small tile on the top left corner of the screen that shows driving mode (Autonomous), The speed in mph and a steering wheel image that shows real time movement based on how the AV is steering. There is also a double arrow icon that blinks wne the vehicle moves left or right. Thso could be the indicator icons.* *A caution symbol appears and the steering turns caution yellow on the left top tile when the AV is not aware what needs to be done and requires assistance from the remote driver controlling the AV*		
	How is it prioritized	Steering screen - The three information is what takes most screen real estate showing that It is prioritized. The current path information is prioritized by keeping that information in the center of the screen. The other information on the top and bottom of the screen is prioritized by font size and placement (eg, current speed vs max speed the car will go). Notifications are high priority	Prioritized Information that is the most important is displayed in high contrast color and with a 3d layer of information. Route reason of vehicle action (2d layer) pedestrians emergency/construction	The top map view takes the majority of the screen. Since the AV on the screen is white attention is maintained on that. Attention is also maintained on the green pathway since it is constantly changing in the scene. The tile on top is less prioritized than the map view.	The priority is the car movement staying in lane. The second priority is speed. And the rest is least priority.	Th elights on the steering wheel are of top priority. It immedialtely grabs the drivers attention.

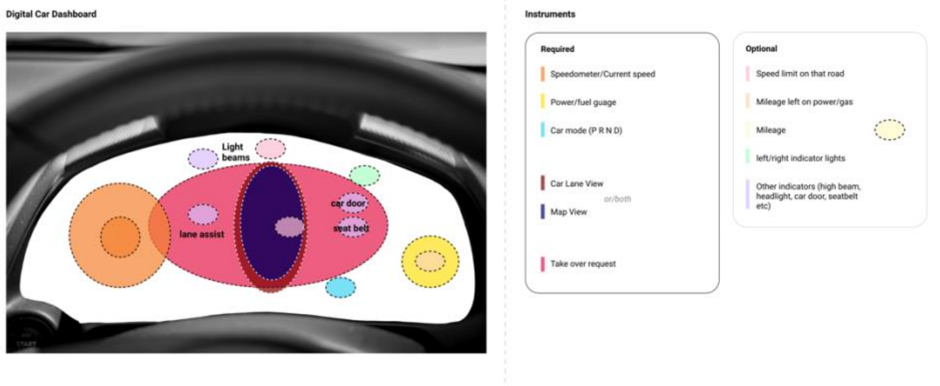
		when they appear Center screen (during drive) - Placement of information (top or bottom)				
	What should human understand ?	Steering screen - They should immediately understand the path the vehicle is taking, what lane it is on, the situation the car is at in any given time (surrounding cars, is there anything too close to the vehicle), The directions the car is about to take. Notifications on what the car is about to do and what the human needs to do in return.	Human must understand what the vehicle sees, what the vehicle is doing and what the vehicle intends to do next. The graphics of the surrounds shows what the car sees, The current location of the car shows what the car is doing (stationary or moving. The highlighted route of shows what path the vehicle is goign to take.	Most times human must understand the path the vehicle is taking. Is it talking optimized routes and lanes given the larger context. The user definitely understands the larger context of the vehicle compared to some other views. The user must understand what driving mode the vehicle is in and what is the speed of the vehicle in comparison to context. The remote driver must understand when the AV needs assistance.	The human must understand that the car is driving in lane and at the required speed to move and stop. He must understand the car is engaged in autonomous mode.	In green mode the driver must understand that the car is doing well by itself. If the green light is blinking the human must understand that the user must pay attention to the road (the car is tracking human eye movement). Lastly the blue mode says that it would be good for the driver to pay more attention or take control, but the car is still doing fine and don't need immediate taking over by a human.
	What action should human take?	There are scenarios in which human must confirm a cars action to follow with a button on left side clicker.	*Waymo does not require drivers to take control. If the driver touches the steering wheel the vehicle comes to a halt.*	Usually the human in the car has to take no action. He only views the display. The remote driver on the other hand must be aware when the AV needs assistance.	The human can relax and occasionally monitor	The human must pay attention to the road when the green light blinks. The human may take control when the light is blue.
State 2: During Takeover	**What is displayed**	*Usually change happens when driver realizes the car has understood the situation incorrectly, or other scnarios such as turns etc.)* *Everything on pre is still displayed.* *The car slows down when it does not undersatand it's surroundings and also gives a written*	*Waymo does not require drivers to take control. If the driver touches the steering wheel the vehicle comes to a halt.*	The AV stops and a yellow caution symbol appears on the top left of the dashboard along with the steering symbol turning yellow.	**Human to car change :** During change (post hitting on the button), both the displays pull up a notification for few seconds and they disappear. The screen then shifts to the car mode display. Car to Human change : There is a sound. . The UI changes to red. Red flashing lights on either side of behind the wheel display. Then there is text	**Human to car change** : There is now a green light on the steering wheel. There is written notification on the right side display of the behind the wheel display. There is also what seems like green bar that comes on top of the green icon on the left side of the display below the lane line view. Car to Human change : blinking

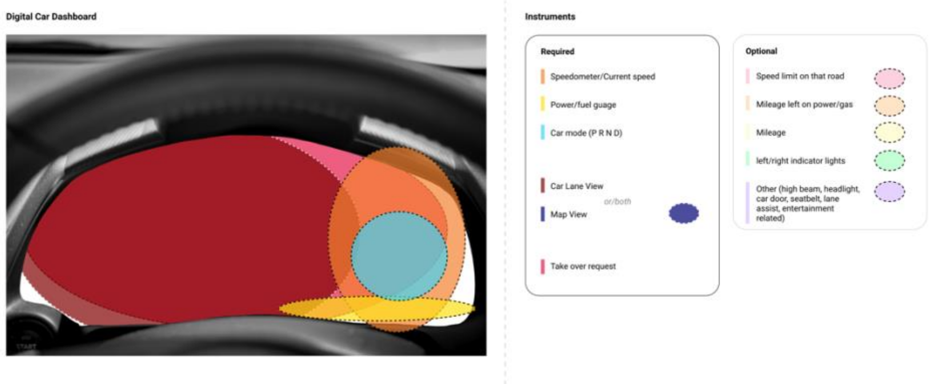
		notification on the bottom of the screen. Disengagement - There is a sound indication disengagement.			indicating human must resume control.	red light and vibrating seat. The steering wheel on the display changes red and the icon shows a human hand holding onto the steering. There is also a text note on the right side of the behind the wheel display.
	How is it displayed	In a the form of text notification. There might be a white arc around the car graphic on screen. Disengagement - There is a sound indication disengagement.		Displayed on the top left corner tile of the display	**Human to car change :** It takes up the whole of both the screens for a few seconds. Car to Human change : The red UI is there all over and the text appears a little on the top of the behind the wheel screen.	**Human to car change** : The information is sort of split onto different sections of the behind the wheel display. The middle section seems constant where as some change is noticed in the left and right displays either in grapic elements or written text Car to Human change : The light is displayed as the same arc light. The red wheels are icons. The vibrating seat is in the back of the seat the driver is sitting on.
	How is it prioritized	The notification that appears, enters from bottom of the screen. It is on a dark grey tile with white text. Disengagement - The sound is of mid level priotity		It is prioritized with change in color to draw user attention.	**Human to car change** : The notification is prioritized before it transitions into the car control UI. Car to Human change : The sound is of high priority. The red color is prioritized then the text.	**Human to car change :** The higher priority is the light on the steering wheel. It is immediately visible to the user. Car to Human change : Here the haptic feedback would be priority then the red blinking light and then iconography and text.
	What should human understand ?	The Humans must understand that there is something off from normal. Disengagement - must understand the car is now in manual mode.		The in car user must understand that the AV is confused but there is a human operator in-charge. The human operator must understand what the matter	**Human to car change** : The human must understand there are some terms as the control is transitioning to the car. Car to Human	**Human to car change :** The human must understand that the vehicle can now handle itself. But also that the human must pay attention.

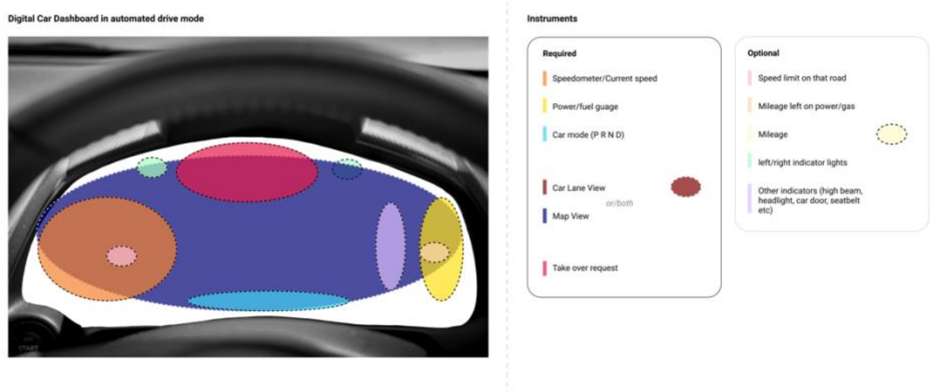
				is based on the camera view and top map view.	change : The human must understand that the human must take immediate action and take control of the car.	Car to Human change : The human must immediately understand that the car needs a human to take over.
	What action should human take?	Hit on the engagement clicker on the left of steering. Regain control by holding the wheel.		In car user must take no action, the human operator must create a new path for the vehicle.	**Human to car change** : No action Car to Human change : Human must place hand on wheel and legs on the required pedal to control the car.	**Human to car change** : No action required expect keep attention on road. Car to Human change : Human must place hand on the steering wheel and legs onto the required pedal and take control of the car.
State 3: Post-Takeover	**What is displayed**	*Everything on pre is still displayed.*	*Similar to pre*	*Similar to pre*	In normal mode both screens show a map view. The behind the wheel dashboard shows the speedometer and the tachometer. When there is opportunity to go autonomous the behind the wheel display has two border lights that blink in white and a text and AI icon appears on that screen indicating the driver has the opportunity to engage in the driverless traffic jam pilot system.	The display has the speedometer, tachometer and the fuel gauges. The digital display in the center shows the speed. And the digital display on the left shows the path of the car (car and the lane lines). There is also a PRN_D mode When there is opportunity for the human to engage the car in cruise a green wheel appears on the top mid right of the display.
	How is it displayed	The car lanes don't glow and the steering wheel image don't glow.			The speedometer and the tachometer are on the right and the left, while the center screen is taken up by the map view in the behind the wheel display. The lights appear on the left and right rims of the behind the wheel screen. The text appears on center screen with a perspective lane view. The AI	The gauges are meter based. The speed is a number. The center screen is one of the camera view overlayed with the car path. (what the car sees). The steering wheel appears as a symbol (no words) on the display.

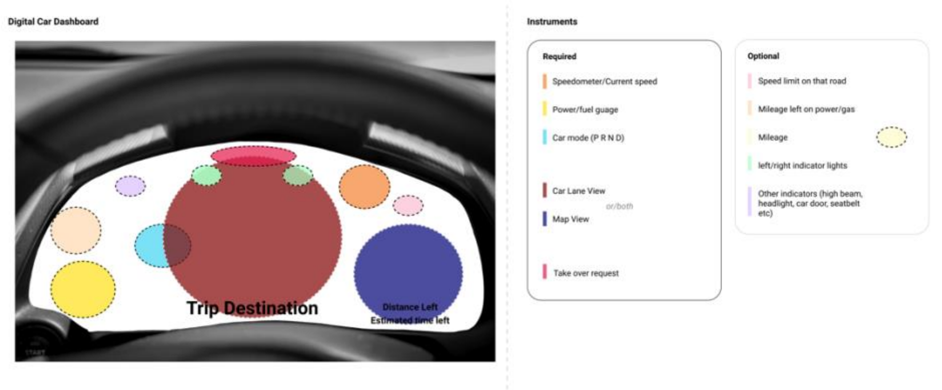
					icon appears on the bottom of the screen.	
	How is it prioritized	*Everything on pre is still displayed.*			The blinking lights are the highest priority it immediately grabs the attention of the driver. The next being the text that says takeover is now possible in the current situation. And the next being the icon in the bottom right.	The prioritization at this mode is the speed of the vehicle. Which is at the center of the behind the wheel display. The steering wheel may not be a priority symbol. It is pretty small in the top right.
	What should human understand ?	With the absence of glowing the human must understand he is on engaged mode.			The human in normal driving conditions just understand the speed they are travelling in and possibly the route they are taking from the screens. When the screen communicates there is possibility for automation at that point, the human must understand there is possibility for him to free his control of the vehicle by confirming/reject the transition with some interaction.	The human in normal driving conditions just understand the speed they are travelling in and possibly the route they are taking from the screens. The human now knows that the road is clear to engage in cruise drive mode.
	What action should human take?	The human must continue to resume regular driving activities and behaviours			The person must ignore the prompt if he does not want to transition control. If he wants to allow the car to take control he must press the button on the lower right side of the driver.	The user can either activate the cruise mode using a button or series of buttons on the steering wheel of the vehicle.

APPENDIX B. PARTICIPATORY DESIGN RESULTS

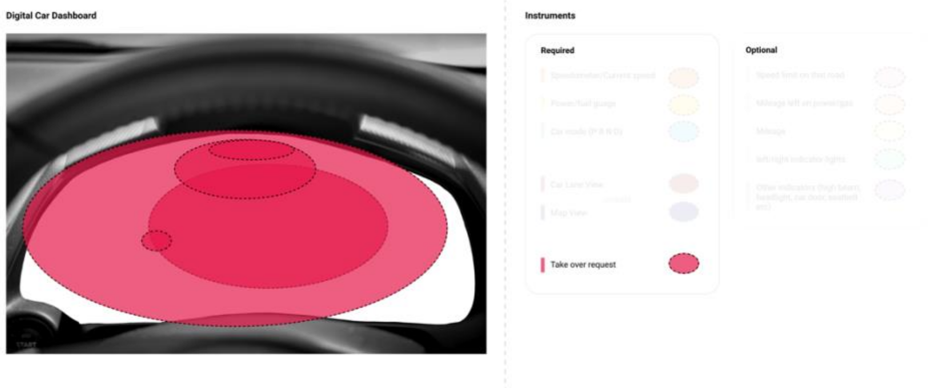


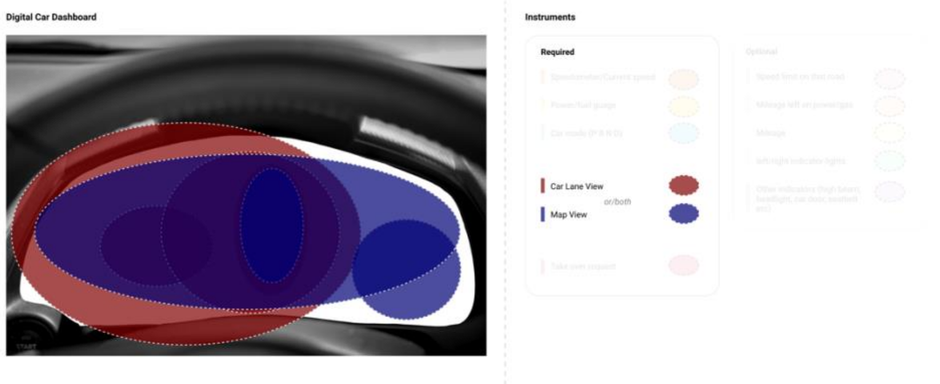


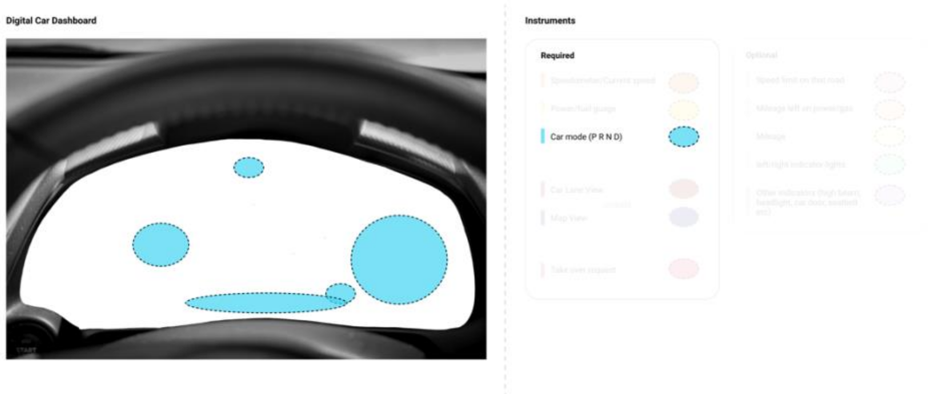

Digital Car Dashboard in automated drive mode
Instruments
Required
Speedometer/Current speed
Power/fuel guage
Car mode (P R N D)
Car Lane View
or/both
Map View
Take over request
Optional
Speed limit on that road
Mileage left on power/gas
Mileage
left/right indicator lights
Other indicators (high beam, headlight, car door, seatbelt etc)

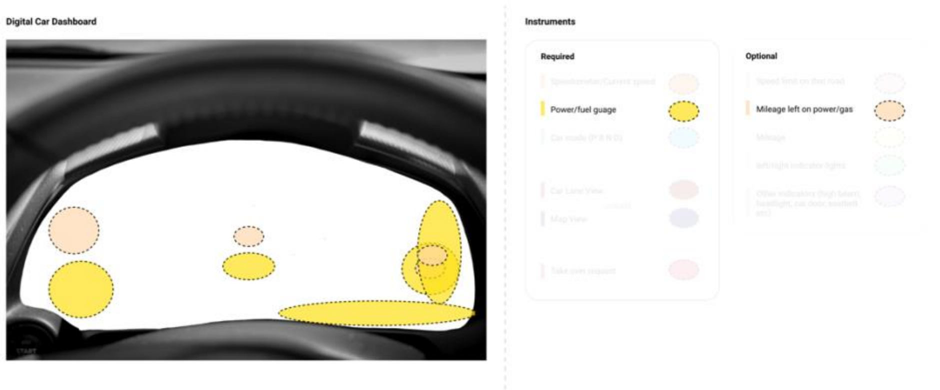

Digital Car Dashboard
Trip Destination
Distance Left
Estimated time left
Instruments
Required
Speedometer/Current speed
Power/fuel guage
Car mode (P R N D)
Car Lane View
or/both
Map View
Take over request
Optional
Speed limit on that road
Mileage left on power/gas
Mileage
left/right indicator lights
Other indicators (high beam, headlight, car door, seatbelt etc)

APPENDIX C. PARTICIPATORY DESIGN RESULTS OVERLAY ANALYSIS



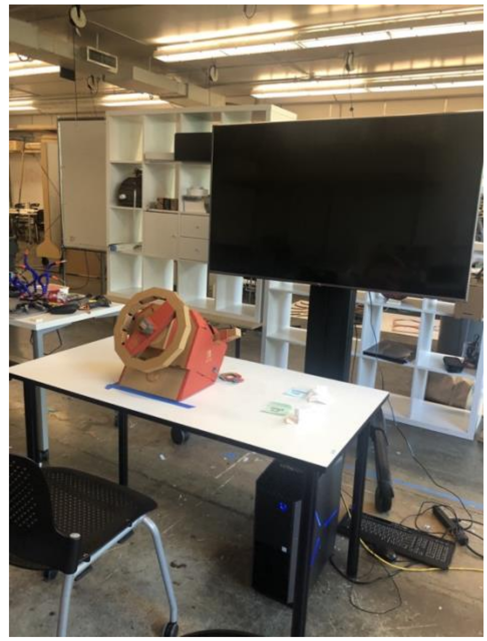



Digital Car Dashboard
Instruments
Required
Car mode (P R N D)


Digital Car Dashboard
Instruments
Required
Power/fuel guage
Optional
Mileage left on power/gas


Digital Car Dashboard
Instruments
Optional
left/right indicator lights

APPENDIX D. VIDEO SETUP AND ENVIRONMENT



The experiment setup consists of a High Definition 42" television that is placed at a fixed distance from the table. The steering wheel prop is also placed at a fixed spot on the table marked by the blue tape. The steering wheel is a dummy prop placed to record user reactions and to ensure an immersive environment. Interaction with the wheel is not responded with any active feedback, it is non-responsive. The participants are seated in front of the table. During the user study the immersive videos prepared for each design and scenario are played to the participant on the television.

APPENDIX E. RECRUITMENT SCRIPTS

Online Recruitment Script (Email, Whatsapp, Slack)

The recruitment message contained a survey link that collected participant screening information. The survey also had a link for participants to schedule a time for the in-person user study.

Hi all,

I am designing the in-car displays of autonomous cars for my Master's Thesis. I would love for you to take part in my user study. Your valuable feedback would heavily influence future design and iterations.

If you are an enthusiastic car and automation person, this would be an enjoyable experience. If you are interested, please fill out the three-minute survey attached to this message.

Hope to see you soon, until then drive safe. 😁

Note: I will be giving out **Amazon gift cards** as my gratitude for your time and participation in the study.

https://gatech.co1.qualtrics.com/jfe/form/SV_0HsLuKuPxxmIEnQ

User study confirmation email

Once participants filled out the participant screening survey and scheduled a time for the in-person user study an email was sent out to them with further details on the study.

Hi *participant name*,

Hope you are having a great day. 😋

Thank You so much for scheduling a time with me for an **in-person user study on in-car displays for Level-3 autonomous cars.**

About the study

In this study,

- you will be asked to **watch a series of videos**
- videos will be displayed on a **42" screen** with different display designs
- You will then evaluate the designs using **surveys** *(on an iPad)* and a **design feedback session** *(a small chat/on paper)***.**
- The session will include only you *(participant)* and me *(designer)*
- My thesis advisor *(Prof. Wei Wang)* might visit us while the session is in progress.

Please note that **the study will be conducted in person *(location details attached in this mail)*.**

Feel free to **reach out to me regarding any concerns** you have in advance or during the design study. (email: dmurugan6@gatech.edu)

Pre-study survey

If time permits and you have not already filled out the pre-study survey, **please find the survey link attached here**:

Pre-study survey link

Confirmation of in-person user study

Address:

Room 150, 1st Floor,

School of Industrial Design

247 4th St NW. Atlanta, GA 30332

google map location

Scheduled Time :

March 14, Sunday,

4:00pm -4:45pm (45 min)

I have attached the [Calendly link here](#) if you wish to make any changes to the schedule. Please notify me in case any changes are made.

--

Please make sure to stay masked and safe while you arrive for the study. If you or anyone you have been in close contact with is feeling unwell, please stay home and take care.

Thank You in advance for your time and participation. **You will receive a $15 amazon gift card for your contribution to the research.** 😄

Thanks & Regards

Dheekshana Senthil Murugan

--

Image of location attached

APPENDIX F. PARTICIPANT SCREENING SURVEY

Hi there,

I am Dheekshana, an Industrial Design Student at Georgia Tech. I am **Designing the in-car User Experience for a Level-3 automated car** with Prof.Wei Wang of the Industrial Design department.

I invite you to participate in a user study that will take place early this month at the School of Industrial Design, Georgia Tech campus. A maximum of three people (participant, student and professor) will be present at the testing area at any given time. Strict sanitation protocols will be conducted following every participant study.

This 3 min survey is to briefly share with us about you, your driving experience and opinions.

Q1 Please enter your Name

____

Q2 Please enter your Age

____

Q3 To which Gender identity do you most identify

○ Male

○ Female

○ Transgender Female

○ Transgender Male

○ Gender Variant/Non Conforming

○ Not Listed ____

○ Prefer not to say

Q4 Please enter your preferred email address

____

Q5 Where do you currently stay? (City, State)

____

Q6 This study will be conducted at :

247 4th St NW, Atlanta, GA 30332
School of Industrial Design,
Georgia Institute of Technology.

Will you have any trouble commuting to the above mentioned address?

○ Yes

○ No

Q7 This study will include watching a video on a 42 inch Television. Do you have any vision related issues or discomfort we should be aware of?

◯ Yes

◯ No

Q8 Do you currently or previously own a valid drivers license?

◯ Yes

◯ No

Q9 When was the last time you drove a car?

◯ In the last 3 months

◯ In the last 6 months

◯ In the last year

◯ Few years ago

Q10 Select the automated features that you have used while driving a car? (Multiple Select)

▢ Cruise Control

- [ ] Adaptive Cruise Control
- [ ] Automatic Emergency Braking
- [ ] Lane Keep Assist
- [ ] Traffic Jam Assist
- [ ] Automatic Parking
- [ ] Auto Pilot
- [ ] Other ____

A short Introduction to Autonomous Cars

Autonomous cars are predominantly driverless and is one that is able to operate itself and perform necessary decisions and functions without any human intervention, through sensing it's surroundings.

Attached video on autonomous vehicles

Please click on the following links for more real world examples.

The Waymo Experience
Tesla Self-Driving

Q11 How do you think Automated vehicles would impact how people travel in the future?

____

____

____

____

____

Q12 Why or why not would you trust a highly automated vehicle?

____

____

____

____

____

Thank You for sharing your interest to participate in our user study. The user study will take place in person at the following location:

Room 150
School of Industrial Design
247 4th St NW, Atlanta, GA 30332

The following is a link to schedule a meeting :D. Please choose a time that best suits you for the study.

https://calendly.com/dmurugan6/autonomous-car-display-user-study?month=2021-03

Contact Information:

Dheekshana Senthil Murugan,
dheekshana@gatech.edu

Prof.Wei Wang
wei.wang@design.gatech.edu

Thank You so much for your interest in participating. Hope to see you soon. Until then, have a safe drive. :)

Note: *Please take a screenshot* of this page to save our contact and location information.

APPENDIX G. STS-AD AND NASA-TLX SURVEYS DURING STUDY

This module is filled by the moderator

Participant Number (1,2, etc.)

____

Design Number

○ A

○ B

○ C

Scenario Number

○ Baseline

○ 37 sec drive (low urgency)

○ 29 sec drive (high urgency)

SITUATIONAL TRUST SCALE FOR AUTOMATED DRIVING

The evaluation you're about to perform is a technique that has been developed to measure situational trust in automated vehicles. Read through the six items to make sure you understand what each item means. If you have any questions, please ask the administrator. You'll now be presented with a series of rating scales. For each of the six scales evaluate the task you recently performed by choosing a number on the scale that matches your experience. Each item scale has two endpoint descriptors that describe the scale.

Read each item and rate them on a scale of 1 (Fully Disagree) to 7 (Fully Agree)

	Fully Disagree **1**	2	3	4	5	6	**Fully Agree** **7**
Trust : I trust the automated vehicle in this situation.	○	○	○	○	○	○	○
Performance : I would have performed better than the automated vehicle in this situation.	○	○	○	○	○	○	○
Non Driving Related Tasks : In this situation, the automated vehicle performs well enough for me to engage in other activities (such as reading)	○	○	○	○	○	○	○

Risk : The situation was risky	○	○	○	○	○	○	○
Judgement : The automated vehicle made an unsafe judgement in this situation	○	○	○	○	○	○	○
Reaction : The automated vehicle reacted appropriately to the environment	○	○	○	○	○	○	○

NASA TASK LOAD INDEX

The evaluation you're about to perform is a technique that has been developed by NASA to assess the relative importance of six factors in determining how much workload you experienced while performing a task that you recently completed. These six factors are defined on the following page. Read through them to make sure you understand what each factor means. If you have any questions, please ask the administrator.

Mental Load (low/high)

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

Physical Demand (low/high)

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

Temporal Demand (low/high)

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

Performance (good/poor)

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

Effort (low/high)

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

Frustration Level (low/high)

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

You'll now be presented with a series of rating scales. For each of the six scales evaluate the task you recently performed by dragging the slider to the location that matches your experience. Each scale has two endpoint descriptors that describe the scale. Consider your responses carefully in distinguishing among the different task conditions, and consider each scale individually.

Read each item and rate them on a scale of **1 (Low) to 7 (High)**

	Low **1**	2	3	4	5	6	**High** **7**
Mental Load : How much mental and perceptu	○	○	○	○	○	○	○

al activity did you spend for this task							
Physical Demand : How much Physical activity did you spend for this task?	○	○	○	○	○	○	○
Temporal Demand: How much time pressure did you feel in order to complete the task?	○	○	○	○	○	○	○

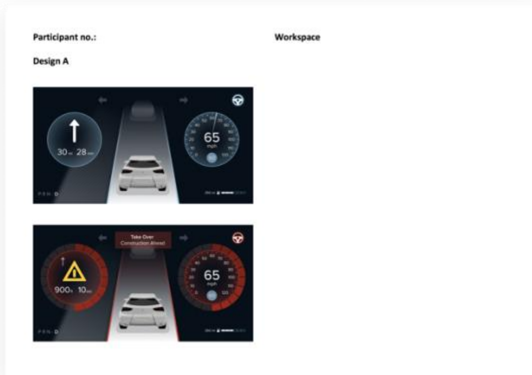
Read each item and rate them on a scale of **1 (Good) to 7 (Poor)**

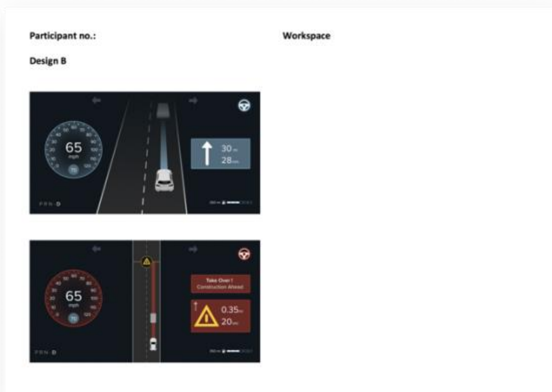
	Good **1**	2	3	4	5	6	**Poor** **7**
Performance: How successful do you think you were in accomplishing the goals of the task?	○	○	○	○	○	○	○

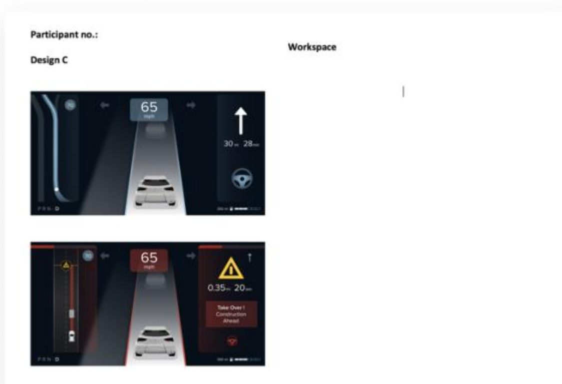
Read each item and rate them on a scale of **1 (Low) to 7 (High)**

	Low **1**	2	3	4	5	6	**High** **7**
Effort: How hard did you have to work to accomplish your level of performance?	○	○	○	○	○	○	○
Frustration: How insecure, discouraged, irritated, stressed and annoyed were you during the task?	○	○	○	○	○	○	○

APPENDIX H. TEMPLATE FOR DESIGN FEEDBACK







APPENDIX I. STS-AD AND NASA-TLX RESULTS

Participant Number (1,2, etc.)	Design Number	Scenario Number	() to 7 () - Trust : I trust the automated vehicle in this situation.	I would have performed better than the automated vehicle in this situation.	Reserse scored (E1)	In this situation, the automated vehicle performs well enough for me to engage in other activities (such as reading)	The situation was risky	Reverse Scored (H1)	The automated vehicle made an unsafe judgement in this situation	Reverse Scored (J1)	The automated vehicle reacted appropriately to the environment	How much mental and perceptual activity did you spend for this task	How much Physical activity did you spend for this task?	How much time pressure did you feel in order to complete the task?	How successful do you think you were in accomplishin g the goals of the task?	How hard did you have to work to accomplish your level of performance ?	How insecure, discouraged, irritated, stressed and annoyed were you during the task?
BASELINE																	
1	A	Baseline	5.00	7	1	6.00	1	7	1	7	7.00	1.00	1.00	1.00	6.00	2.00	1.00
2	A	Baseline	6.00	3.00	5	6.00	1	7	2.00	6	6.00	1.00	1.00	1.00	1.00	1.00	1.00
3	A	Baseline	6.00	3.00	5	6.00	1	7	1	7	7.00	1.00	1.00	1.00	7.00	1.00	1.00
4	A	Baseline	6.00	1.00	7	4.00	1	7	1	7	7.00	2.00	1.00	1.00	1.00	1.00	1.00
5	A	Baseline	7.00	2.00	6	4.00	2	6	2.00	6	7.00	1.00	1.00	1.00	2.00	1.00	2.00
6	A	Baseline	5.00	4.00	4	1.00	2	6	2.00	6	6.00	3.00	1.00	1.00	2.00	1.00	2.00
7	A	Baseline	5.00	2	6	4.00	1	7	1	7	7.00	1.00	1.00	1.00	2.00	1.00	3.00
8	A	Baseline	7.00	2	6	6.00	2	6	1	7	4.00	4.00	1.00	4.00	3.00	1.00	4.00
9	A	Baseline	6.00	3	5	6.00	2	6	1	7	6.00	3.00	2.00	1.00	2.00	2.00	2.00
10	A	Baseline	4.00	5	3	6.00	3	5	2	6	6.00	2.00	1.00	2.00	2.00	1.00	1.00
11	A	Baseline	7.00	1	7	3.00	2	6	2	6	6.00	1.00	1.00	1.00	1.00	2.00	4.00
12	A	Baseline	7.00	3	5	7.00	2	6	1	7	7.00	1.00	1.00	1.00	7.00	1.00	1.00
13	A	Baseline	7.00	1	7	7.00	1	7	1	7	7.00	1.00	1.00	1.00	1.00	1.00	1.00
14	A	Baseline	7.00	3	5	3.00	1	7	1	7	7.00	1.00	1.00	1.00	1.00	1.00	1.00
15	A	Baseline	5.00	4	4	5.00	3	5	3	5	5.00	4.00	2.00	2.00	2.00	2.00	3.00
LOW URGENCY TOR																	
1	A	37 sec drive (urgency)	6.00	7	1	5.00	4	4	1	7	6.00	2.00	2.00	5.00	6.00	2.00	1.00
2	A	37 sec drive (urgency)	5.00	4	4	4.00	3	5	2	6	5.00	2.00	2.00	3.00	1.00	1.00	2.00
3	A	37 sec drive (urgency)	7.00	3	5	7.00	1	7	2	6	7.00	2.00	1.00	1.00	7.00	2.00	1.00
4	A	37 sec drive (urgency)	4.00	6	2	3.00	5	3	4	4	2.00	3.00	2.00	3.00	1.00	2.00	2.00
5	A	37 sec drive (urgency)	5.00	4	4	2.00	4	4	2	6	4.00	2.00	2.00	3.00	3.00	2.00	2.00
6	A	37 sec drive (urgency)	6.00	3	5	1.00	4	4	1	7	6.00	4.00	1.00	4.00	3.00	3.00	1.00
7	A	37 sec drive (urgency)	4.00	3	5	4.00	5	3	1	7	6.00	2.00	1.00	5.00	2.00	4.00	4.00
8	A	37 sec drive (urgency)	4.00	2	6	3.00	4	4	1	7	5.00	2.00	3.00	4.00	1.00	2.00	4.00
9	A	37 sec drive (urgency)	5.00	3	5	4.00	3	5	2	6	5.00	5.00	3.00	5.00	4.00	3.00	5.00
10	A	37 sec drive (urgency)	6.00	3	5	5.00	2	6	1	7	7.00	2.00	3.00	4.00	2.00	2.00	1.00
11	A	37 sec drive (urgency)	3.00	5	3	1.00	2	6	4	4	4.00	7.00	6.00	7.00	3.00	7.00	7.00
12	A	37 sec drive (urgency)	5.00	6	2	5.00	6	2	2	6	6.00	3.00	1.00	2.00	6.00	1.00	1.00
13	A	37 sec drive (urgency)	2.00	5	3	2.00	4	4	5	3	3.00	5.00	5.00	5.00	4.00	5.00	5.00
14	A	37 sec drive (urgency)	5.00	7	1	1.00	5	3	4	4	4.00	1.00	1.00	2.00	1.00	1.00	1.00
15	A	37 sec drive (urgency)	3.00	5	3	3.00	6	2	4	4	3.00	5.00	3.00	3.00	3.00	2.00	5.00
HIGH URGENCY TOR																	
1	A	29 sec drive (urgency)	6.00	7.00	1	3.00	5.00	3	1.00	7	6.00	2.00	1.00	7.00	6.00	2.00	1.00
2	A	29 sec drive (urgency)	4.00	4.00	4	3.00	6.00	2	3.00	5	4.00	3.00	2.00	6.00	3.00	2.00	4.00
3	A	29 sec drive (urgency)	7.00	2.00	6	5.00	2.00	6	1	7	7.00	3.00	2.00	1.00	7.00	2.00	1.00
4	A	29 sec drive (urgency)	5.00	3.00	5	4.00	5.00	3	2	6	6.00	3.00	2.00	4.00	1.00	2.00	1.00
5	A	29 sec drive (urgency)	3.00	5.00	3	2.00	4.00	4	3	5	5.00	3.00	3.00	5.00	4.00	3.00	4.00
6	A	29 sec drive (urgency)	5.00	4.00	4	1.00	3.00	5	2	6	6.00	5.00	1.00	4.00	5.00	2.00	2.00
7	A	29 sec drive (urgency)	4.00	5	3	2.00	5.00	3	1	7	5.00	2.00	2.00	2.00	2.00	2.00	4.00
8	A	29 sec drive (urgency)	4.00	3	5	2.00	4	4	2	6	6.00	5.00	3.00	4.00	1.00	3.00	4.00
9	A	29 sec drive (urgency)	5.00	4	4	5.00	4	4	2	6	3.00	5.00	3.00	4.00	3.00	4.00	3.00
10	A	29 sec drive (urgency)	6.00	3	5	6.00	2	6	2	6	6.00	2.00	2.00	3.00	1.00	2.00	2.00
11	A	29 sec drive (urgency)	2.00	4	4	2.00	1	7	3	5	5.00	7.00	5.00	7.00	3.00	6.00	7.00
12	A	29 sec drive (urgency)	2.00	7	1	1.00	7	1	6	2	3.00	3.00	2.00	6.00	4.00	5.00	3.00
13	A	29 sec drive (urgency)	4.00	3	5	5.00	3	5	3	5	3.00	3.00	2.00	2.00	3.00	3.00	2.00
14	A	29 sec drive (urgency)	4.00	2	6	1.00	3	5	3	5	4.00	2.00	1.00	2.00	1.00	1.00	3.00
15	A	29 sec drive (urgency)	4.00	5	3	4.00	6	2	3	5	5.00	6.00	3.00	3.00	4.00	3.00	4.00

Participant Number (1,2, etc.)	Design Number	Scenario Number	() to 7 () - Trust : I trust the automated vehicle in this situation.	I would have performed better than the automated vehicle in this situation.	Reserse scored (E1)	In this situation, the automated vehicle performs well enough for me to engage in other activities (such as reading)	The situation was risky	Reverse Scored (H1)	The automated vehicle made an unsafe judgement in this situation	Reverse Scored (J1)	The automated vehicle reacted appropriately to the environment	How much mental and perceptual activity did you spend for this task	How much Physical activity did you spend for this task?	How much time pressure did you feel in order to complete the task?	How successful do you think you were in accomplishing the goals of the task?	How hard did you have to work to accomplish your level of performance ?	How insecure, discouraged, irritated, stressed and annoyed were you during the task?
BASELINE																	
1	B	Baseline	2.00	7.00	1.00	3.00	3.00	5.00	1.00	7.00	6.00	3.00	1.00	1.00	1.00	1.00	2.00
2	B	Baseline	6.00	3.00	5.00	5.00	1.00	7.00	2.00	6.00	6.00	1.00	1.00	1.00	1.00	1.00	1.00
3	B	Baseline	6.00	3.00	5.00	7.00	2.00	6.00	2.00	6.00	7.00	2.00	1.00	1.00	7.00	2.00	7.00
4	B	Baseline	6.00	1.00	7.00	6.00	1.00	7.00	1.00	7.00	7.00	2.00	1.00	1.00	1.00	1.00	1.00
5	B	Baseline	4.00	2.00	6.00	4.00	1.00	7.00	1.00	7.00	6.00	1.00	1.00	1.00	3.00	1.00	1.00
6	B	Baseline	6.00	2.00	6.00	7.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	2.00	2.00	1.00
7	B	Baseline	6.00	2.00	6.00	4.00	1.00	7.00	7.00	1.00	7.00	1.00	1.00	1.00	1.00	1.00	1.00
8	B	Baseline	6.00	2.00	6.00	6.00	2.00	6.00	1.00	7.00	7.00	2.00	1.00	2.00	3.00	1.00	2.00
9	B	Baseline	6.00	5.00	3.00	6.00	1.00	7.00	1.00	7.00	6.00	2.00	2.00	1.00	2.00	1.00	1.00
10	B	Baseline	5.00	6.00	2.00	6.00	2.00	6.00	2.00	6.00	1.00	2.00	1.00	1.00	2.00	2.00	1.00
11	B	Baseline	7.00	1.00	7.00	7.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	1.00	1.00	1.00
12	B	Baseline	6.00	2.00	6.00	7.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	7.00	1.00	1.00
13	B	Baseline	7.00	7.00	1.00	7.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	1.00	1.00	1.00
14	B	Baseline	7.00	2.00	6.00	7.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	1.00	1.00	1.00
15	B	Baseline	6.00	4.00	4.00	5.00	3.00	5.00	5.00	3.00	5.00	3.00	2.00	2.00	2.00	2.00	4.00
LOW URGENCY TOR																	
1	B	37 sec drive (urgency)	3.00	7.00	1.00	2.00	6.00	2.00	1.00	7.00	6.00	5.00	4.00	5.00	5.00	3.00	2.00
2	B	37 sec drive (urgency)	4.00	4	4.00	5.00	3.00	5.00	2.00	6.00	5.00	2.00	1.00	3.00	2.00	1.00	2.00
3	B	37 sec drive (urgency)	7.00	3	5.00	6.00	1.00	7.00	1.00	7.00	7.00	2.00	1.00	2.00	7.00	2.00	1.00
4	B	37 sec drive (urgency)	3.00	6	2.00	3.00	5.00	3.00	4	4.00	2.00	4.00	2.00	3.00	1.00	2.00	3.00
5	B	37 sec drive (urgency)	3.00	4	4.00	2.00	4	4.00	2	6.00	5.00	2.00	2.00	4.00	3.00	2.00	2.00
6	B	37 sec drive (urgency)	5.00	4	4.00	1.00	2	6.00	2	6.00	6.00	6.00	1.00	4.00	5.00	2.00	2.00
7	B	37 sec drive (urgency)	4.00	5	3.00	1.00	2	6.00	1.00	7.00	6.00	2.00	1.00	6.00	1.00	1.00	1.00
8	B	37 sec drive (urgency)	5.00	2	6.00	4.00	3	5.00	1.00	7.00	7.00	2.00	2.00	2.00	1.00	2.00	2.00
9	B	37 sec drive (urgency)	4.00	4	4.00	3.00	3	5.00	1.00	7.00	3.00	3.00	2.00	3.00	5.00	2.00	2.00
10	B	37 sec drive (urgency)	6.00	3	5.00	6.00	2	6.00	2	6.00	7.00	3.00	1.00	2.00	2.00	1.00	2.00
11	B	37 sec drive (urgency)	3.00	6	2.00	1.00	5	3.00	3	5.00	4.00	6.00	6.00	6.00	2.00	5.00	6.00
12	B	37 sec drive (urgency)	6.00	5	3.00	5.00	6	2.00	3	5.00	4.00	4.00	2.00	3.00	6.00	2.00	1.00
13	B	37 sec drive (urgency)	3.00	5.00	3.00	3.00	5	3.00	4	4.00	4.00	3.00	3.00	3.00	4.00	4.00	5.00
14	B	37 sec drive (urgency)	3.00	5.00	3.00	1.00	2	6.00	4	4.00	3.00	2.00	1.00	5.00	1.00	1.00	5.00
15	B	37 sec drive (urgency)	4.00	5.00	3.00	3.00	6	2.00	4	4.00	4.00	6.00	3.00	3.00	3.00	3.00	5.00
HIGH URGENCY TOR																	
1	B	29 sec drive (urgency)	5.00	7.00	1.00	2.00	6.00	2.00	2	6.00	6.00	5.00	4.00	4.00	6.00	4.00	1.00
2	B	29 sec drive (urgency)	4.00	3	5.00	3.00	4.00	4.00	3	5.00	5.00	3.00	1.00	4.00	2.00	1.00	3.00
3	B	29 sec drive (urgency)	6.00	3	5.00	7.00	2.00	6.00	1.00	7.00	7.00	2.00	1.00	3.00	7.00	2.00	1.00
4	B	29 sec drive (urgency)	5.00	4	4.00	4.00	5	3.00	2	6.00	6.00	3.00	2.00	3.00	1.00	2.00	1.00
5	B	29 sec drive (urgency)	4.00	5	3.00	2.00	6	2.00	3	5.00	6.00	2.00	2.00	5.00	4.00	3.00	3.00
6	B	29 sec drive (urgency)	5.00	4	4.00	1.00	5	3.00	2	6.00	6.00	5.00	2.00	5.00	2.00	3.00	2.00
7	B	29 sec drive (urgency)	5.00	5	3.00	1.00	2	6.00	1.00	7.00	4.00	3.00	1.00	7.00	3.00	3.00	5.00
8	B	29 sec drive (urgency)	5.00	4	4.00	2.00	4	4.00	1.00	7.00	5.00	3.00	3.00	4.00	1.00	3.00	2.00
9	B	29 sec drive (urgency)	5.00	4	4.00	5.00	3	5.00	2	6.00	5.00	4.00	2.00	2.00	5.00	3.00	2.00
10	B	29 sec drive (urgency)	5.00	5	3.00	7.00	2	6.00	1.00	7.00	7.00	2.00	1.00	1.00	2.00	2.00	1.00
11	B	29 sec drive (urgency)	5.00	4	4.00	5.00	6	2.00	5	3.00	2.00	6.00	5.00	6.00	4.00	6.00	6.00
12	B	29 sec drive (urgency)	3.00	6	2.00	2.00	7.00	1.00	5	3.00	3.00	5.00	3.00	5.00	5.00	4.00	3.00
13	B	29 sec drive (urgency)	4.00	5	3.00	4.00	3	5.00	3	5.00	4.00	4.00	4.00	3.00	3.00	3.00	2.00
14	B	29 sec drive (urgency)	2.00	5	3.00	1.00	6	2.00	3	5.00	4.00	2.00	1.00	4.00	1.00	1.00	5.00
15	B	29 sec drive (urgency)	4.00	5	3.00	3.00	6	2.00	5	3.00	5.00	6.00	3.00	3.00	3.00	3.00	5.00

Participant Number (1,2, etc.)	Design Number	Scenario Number	() to 7 () - Trust : I trust the automated vehicle in this situation.	I would have performed better than the automated vehicle in this situation.	Reserse scored (E1)	In this situation, the automated vehicle performs well enough for me to engage in other activities (such as reading)	The situation was risky	Reverse Scored (H1)	The automated vehicle made an unsafe judgement in this situation	Reverse Scored (J1)	The automated vehicle reacted appropriately to the environment	How much mental and perceptual activity did you spend for this task	How much Physical activity did you spend for this task?	How much time pressure did you feel in order to complete the task?	How successful do you think you were in accomplishing the goals of the task?	How hard did you have to work to accomplish your level of performance ?	How insecure, discouraged, irritated, stressed and annoyed were you during the task?
BASELINE																	
1	C	Baseline	5.00	7.00	1.00	7.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	3.00	2.00	2.00	1.00
2	C	Baseline	6.00	3.00	5.00	6.00	2	6.00	2	6.00	6.00	1.00	1.00	1.00	1.00	1.00	1.00
3	C	Baseline	6.00	3.00	5.00	6.00	1.00	7.00	1.00	7.00	7.00	3.00	2.00	2.00	2.00	2.00	1.00
4	C	Baseline	6.00	1.00	7.00	4.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	1.00	1.00	1.00
5	C	Baseline	6.00	4.00	4.00	5.00	1.00	7.00	2	6.00	6.00	1.00	1.00	2.00	3.00	2.00	2.00
6	C	Baseline	6.00	4.00	4.00	2.00	1.00	7.00	1.00	7.00	6.00	4.00	1.00	2.00	2.00	1.00	2.00
7	C	Baseline	5.00	2.00	6.00	2.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	2.00	1.00	2.00
8	C	Baseline	5.00	2.00	6.00	6.00	2	6.00	1.00	7.00	5.00	1.00	1.00	1.00	2.00	1.00	1.00
9	C	Baseline	6.00	2.00	6.00	6.00	2	6.00	1.00	7.00	6.00	2.00	2.00	1.00	2.00	2.00	2.00
10	C	Baseline	6.00	4.00	4.00	6.00	3	5.00	2	6.00	6.00	2.00	1.00	1.00	1.00	1.00	2.00
11	C	Baseline	7.00	1.00	7.00	7.00	2	6.00	2	6.00	7.00	1.00	1.00	1.00	2.00	1.00	1.00
12	C	Baseline	6.00	3.00	5.00	6.00	2	6.00	1.00	7.00	7.00	1.00	1.00	3.00	1.00	1.00	1.00
13	C	Baseline	7.00	1.00	7.00	7.00	1.00	7.00	1.00	7.00	4.00	1.00	1.00	1.00	1.00	1.00	1.00
14	C	Baseline	7.00	3.00	5.00	1.00	1.00	7.00	1.00	7.00	7.00	1.00	1.00	1.00	1.00	1.00	1.00
15	C	Baseline	6.00	4.00	4.00	5.00	3	5.00	3.00	5.00	6.00	3.00	2.00	2.00	2.00	2.00	3.00
LOW URGENCY TOR																	
1	C	37 sec drive (urgency)	3.00	7.00	1.00	4.00	3	5.00	1.00	7.00	6.00	3.00	2.00	2.00	6.00	2.00	1.00
2	C	37 sec drive (urgency)	5.00	2	6.00	4.00	3	5.00	2	6.00	6.00	4.00	3.00	4.00	1.00	2.00	2.00
3	C	37 sec drive (urgency)	7.00	4	4.00	5.00	4	4.00	2	6.00	6.00	4.00	3.00	6.00	7.00	5.00	2.00
4	C	37 sec drive (urgency)	4.00	6	2.00	3.00	1.00	7.00	4	4.00	2.00	3.00	2.00	3.00	1.00	2.00	2.00
5	C	37 sec drive (urgency)	5.00	4.00	4.00	5.00	3	5.00	2	6.00	5.00	2.00	1.00	3.00	3.00	1.00	2.00
6	C	37 sec drive (urgency)	5.00	4.00	4.00	1.00	3	5.00	2	6.00	6.00	5.00	1.00	4.00	3.00	2.00	4.00
7	C	37 sec drive (urgency)	4.00	5.00	3.00	2.00	5	3.00	1.00	7.00	6.00	2.00	1.00	6.00	3.00	2.00	4.00
8	C	37 sec drive (urgency)	5.00	2.00	6.00	1.00	3	5.00	1.00	7.00	6.00	2.00	2.00	2.00	1.00	1.00	2.00
9	C	37 sec drive (urgency)	6.00	3.00	5.00	6.00	3	5.00	2	6.00	5.00	4.00	2.00	4.00	2.00	3.00	2.00
10	C	37 sec drive (urgency)	6.00	5.00	3.00	6.00	2	6.00	2	6.00	1.00	2.00	2.00	1.00	2.00	2.00	2.00
11	C	37 sec drive (urgency)	6.00	4.00	4.00	6.00	4	4.00	2	6.00	6.00	3.00	3.00	3.00	2.00	5.00	5.00
12	C	37 sec drive (urgency)	3.00	7.00	1.00	2.00	5	3.00	6	2.00	2.00	3.00	2.00	5.00	6.00	3.00	2.00
13	C	37 sec drive (urgency)	1.00	7.00	1.00	1.00	5	3.00	6.00	2.00	2.00	7.00	5.00	6.00	4.00	4.00	7.00
14	C	37 sec drive (urgency)	5.00	5.00	3.00	1.00	4	4.00	3.00	5.00	4.00	1.00	1.00	2.00	1.00	1.00	2.00
15	C	37 sec drive (urgency)	5.00	4.00	4.00	3.00	5	3.00	2.00	6.00	4.00	4.00	4.00	6.00	6.00	3.00	4.00
HIGH URGENCY TOR																	
1	C	29 sec drive (urgency)	5.00	7.00	1.00	2.00	4	4.00	2.00	6.00	6.00	3.00	2.00	3.00	6.00	6.00	2.00
2	C	29 sec drive (urgency)	5.00	4.00	4.00	4.00	3	5.00	3.00	5.00	5.00	3.00	2.00	5.00	2.00	2.00	3.00
3	C	29 sec drive (urgency)	6.00	4.00	4.00	6.00	4	4.00	2.00	6.00		5.00	3.00	3.00	7.00	3.00	1.00
4	C	29 sec drive (urgency)	5.00	5.00	3.00	1.00	5	3.00	2.00	6.00	5.00	4.00	3.00	4.00	2.00	3.00	2.00
5	C	29 sec drive (urgency)	5.00	6.00	2.00	3.00	5	3.00	2.00	6.00	4.00	2.00	2.00	3.00	3.00	2.00	2.00
6	C	29 sec drive (urgency)	5.00	4.00	4.00	1.00	4	4.00	2.00	6.00	6.00	6.00	1.00	4.00	3.00	4.00	2.00
7	C	29 sec drive (urgency)	4.00	3.00	5.00	1.00	4	4.00	1.00	7.00	4.00	3.00	1.00	2.00	3.00	1.00	4.00
8	C	29 sec drive (urgency)	4.00	2.00	6.00	2.00	4	4.00	1.00	7.00	5.00	3.00	3.00	5.00	1.00	2.00	3.00
9	C	29 sec drive (urgency)	6.00	3.00	5.00	5.00	3	5.00	2.00	6.00	5.00	4.00	3.00	3.00	2.00	3.00	3.00
10	C	29 sec drive (urgency)	6.00	3.00	5.00	6.00	2	6.00	1.00	7.00	6.00	3.00	1.00	2.00	2.00	1.00	2.00
11	C	29 sec drive (urgency)	6.00	6.00	2.00	6.00	2	6.00	2.00	6.00	6.00	1.00	2.00	2.00	1.00	2.00	3.00
12	C	29 sec drive (urgency)	4.00	6.00	2.00	3.00	4	4.00	4.00	4.00	3.00	2.00	2.00	5.00	6.00	2.00	2.00
13	C	29 sec drive (urgency)	5.00	6.00	2.00	3.00	4.00	4.00	5.00	3.00	4.00	4.00	3.00	5.00	3.00	4.00	2.00
14	C	29 sec drive (urgency)	5.00	2.00	6.00	1.00	2.00	6.00	1.00	7.00	6.00	1.00	1.00	2.00	1.00	1.00	1.00
15	C	29 sec drive (urgency)	4.00	4.00	4.00	4.00	6.00	2.00	3.00	5.00	6.00	6.00	3.00	3.00	3.00	3.00	4.00

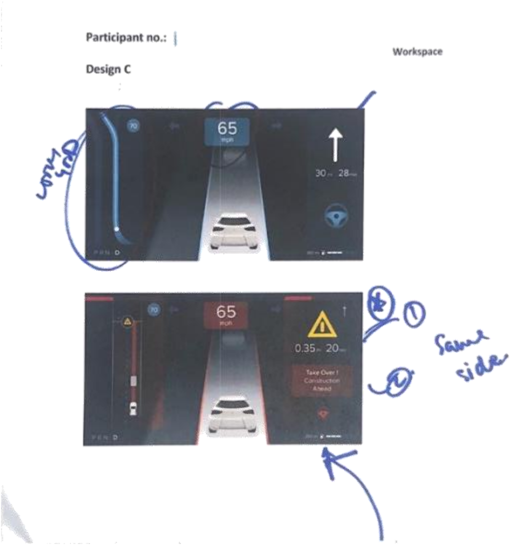
APPENDIX J. REACTION TIME RESULTS

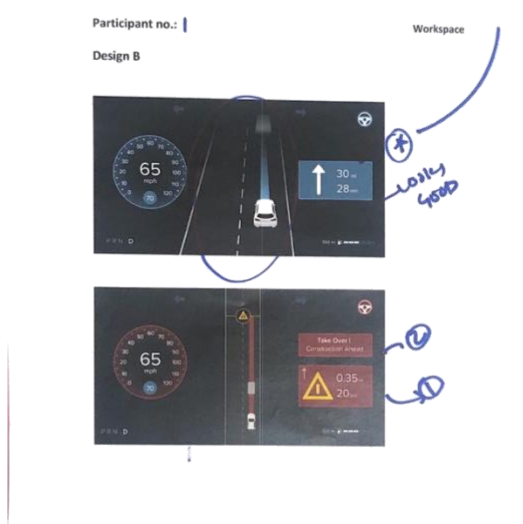
Participant number	Video Order	Video	Time on video	TOR begin	Reaction Time
1	**5**	**DA-LU**	17	11	6
	6	**DA-HU**	24	21	3
	1	**DB-LU**	12	11	1
	2	**DB-HU**	25	21	4
	4	**DC-LU**	16	11	5
	3	**DC-HU**	25	21	4
2	**3**	**DA-LU**	13	11	2
	4	**DA-HU**	24	21	3
	5	**DB-LU**	13	11	2
	6	**DB-HU**	22	21	1
	1	**DC-LU**	16	11	5
	2	**DC-HU**	24	21	3
3	**5**	**DA-LU**	16	11	5
	6	**DA-HU**	22	21	1
	3	**DB-LU**	12	11	1
	4	**DB-HU**	23	21	2
	1	**DC-LU**	13	11	2
	2	**DC-HU**	22	21	1
4	**6**	**DA-LU**	12	11	1
	5	**DA-HU**	22	21	1
	4	**DB-LU**	13	11	2
	3	**DB-HU**	22	21	1
	2	**DC-LU**	12	11	1
	1	**DC-HU**	23	21	2
5	**2**	**DA-LU**	16	11	5
	1	**DA-HU**	26	21	5
	3	**DB-LU**	16	11	5
	4	**DB-HU**	23	21	2
	6	**DC-LU**	14	11	3
	5	**DC-HU**	23	21	2
6	**4**	**DA-LU**	15	11	4
	3	**DA-HU**	26	21	5
	1	**DB-LU**	15	11	4
	2	**DB-HU**	23	21	2
	6	**DC-LU**	18	11	7
	5	**DC-HU**	25	21	4
7	**5**	**DA-LU**	13	11	2

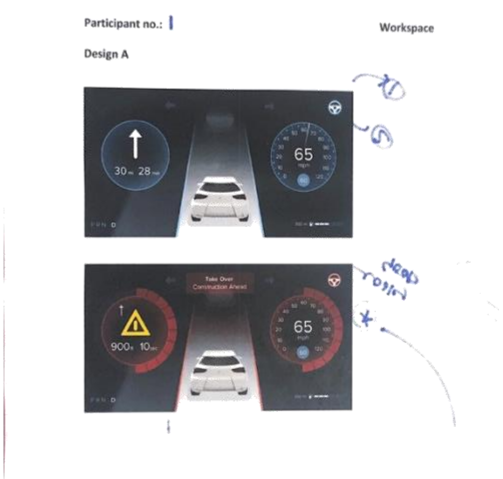
	6	**DA-HU**	22	21	1
	1	**DB-LU**	12	11	1
	2	**DB-HU**	22	21	1
	3	**DC-LU**	12	11	1
	4	**DC-HU**	23	21	2
8	**2**	**DA-LU**	14	11	3
	1	**DA-HU**	24	21	3
	3	**DB-LU**	15	11	4
	4	**DB-HU**	22	21	1
	6	**DC-LU**	13	11	2
	5	**DC-HU**	23	21	2
9	**4**	**DA-LU**	14	11	3
	3	**DA-HU**	23	21	2
	1	**DB-LU**	28	11	17
	2	**DB-HU**	22	21	1
	6	**DC-LU**	24	11	13
	5	**DC-HU**	23	21	2
10	**1**	**DA-LU**	14	11	3
	2	**DA-HU**	23	21	2
	4	**DB-LU**	13	11	2
	3	**DB-HU**	23	21	2
	6	**DC-LU**	13	11	2
	5	**DC-HU**	22	21	1
11	**4**	**DA-LU**	13	11	2
	3	**DA-HU**	22	21	1
	1	**DB-LU**	13	11	2
	2	**DB-HU**	22	21	1
	5	**DC-LU**	13	11	2
	6	**DC-HU**	23	21	2
12	**3**	**DA-LU**	13	11	2
	4	**DA-HU**	22	21	1
	5	**DB-LU**	12	11	1
	6	**DB-HU**	22	21	1
	1	**DC-LU**	12	11	1
	2	**DC-HU**	23	21	2
13	**6**	**DA-LU**	13	11	2
	5	**DA-HU**	19	21	-2
	4	**DB-LU**	14	11	3
	3	**DB-HU**	22	21	1
	2	**DC-LU**	12	11	1
	1	**DC-HU**	22	21	1
14	**4**	**DA-LU**	13	11	2
	3	**DA-HU**	23	21	2
	1	**DB-LU**	17	11	6

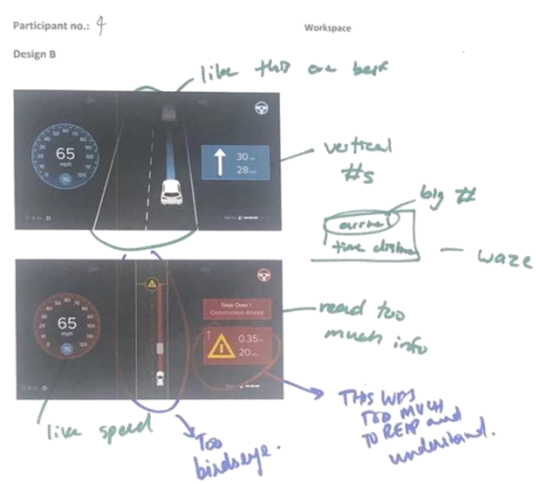
	2	**DB-HU**	23	21	2
	6	**DC-LU**	14	11	3
	5	**DC-HU**	23	21	2
15	**4**	**DA-LU**	12	11	1
	3	**DA-HU**	22	21	1
	5	**DB-LU**	12	11	1
	6	**DB-HU**	22	21	1
	1	**DC-LU**	13	11	2
	2	**DC-HU**	22	21	1

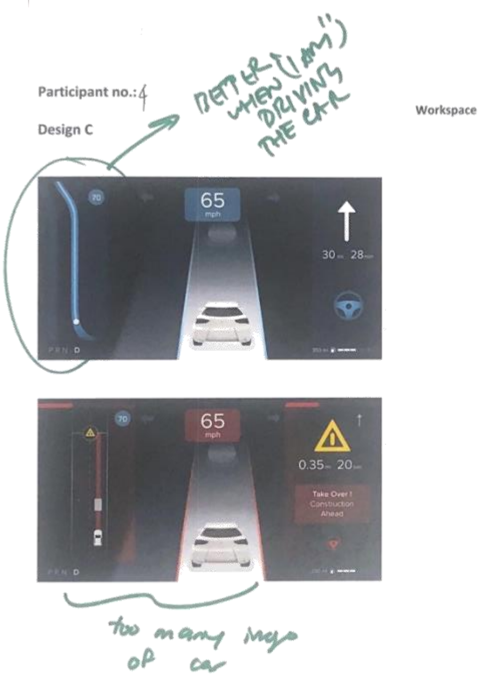
APPENDIX H. PARTICIPANT SKETCHES

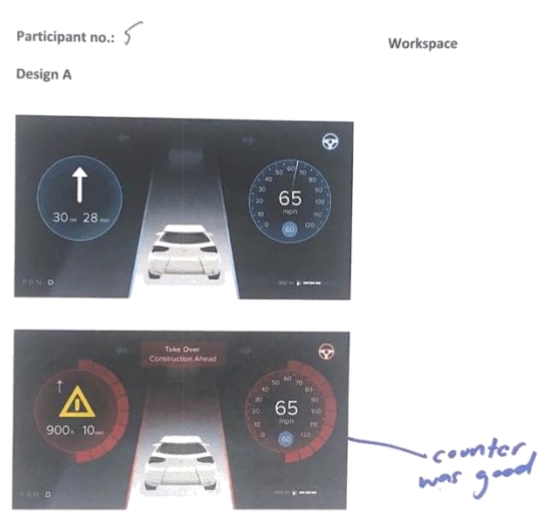












Participant no.: 3

Design C

Workspace

MORE OBVIOUS CAUSE

MINIMAL

©

Participant no.: 4

Design A

Workspace

I LIKED THIS THE BEST CAUSE OF THE TIMELAPSE MORE RE

✓ center

like 2 circles (symmetry)

like this

more relevant

reversed

Participant no.: 4

Design B

Workspace

like this one best

vertical #s

big #

current time distance

— waze

read too much info

like speed

Too birdseye.

THIS WAS TOO MUCH TO READ and understand.

Participant no.: ✓

Design A

Workspace

ok

0

Participant no.: 2

Design C

also agree

Workspace

Participant no.: 3

Design A

Workspace

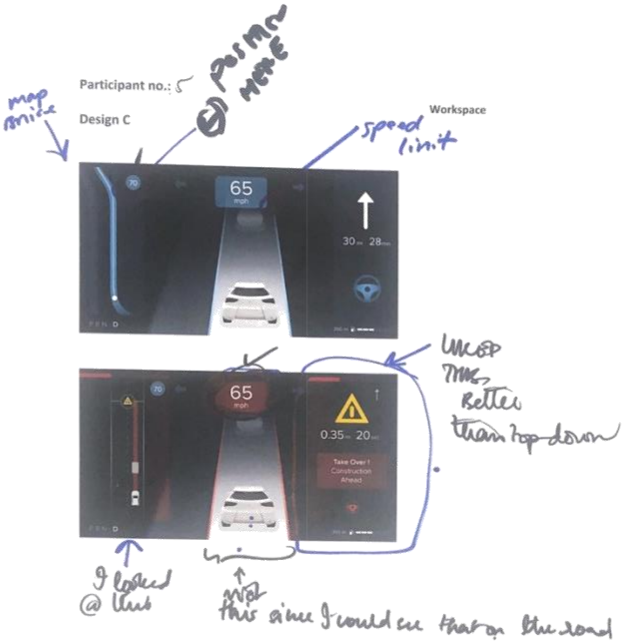
B/W A&B

COUNTDOWN BETTER WHEN USED TO

Participant no.: 5

Design C

Workspace



Participant no.: 5

Design B

Workspace

Participant no.:

Design A

Workspace

Participant no.: 6

Design B

Workspace

Participant no.: 6

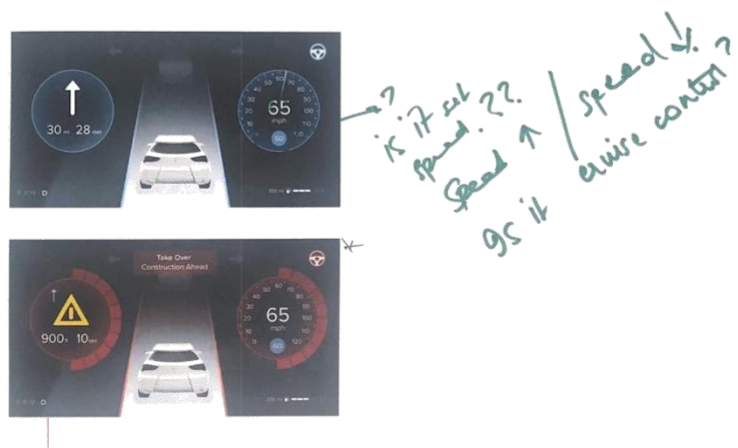
Design C

Workspace

Participant no.: 7

Design A

Workspace



Participant no.: 7

Workspace

Design B

?

Signals what?

Participant no.: 7

Design C

Workspace

3

Participant no.: 8

Workspace

Design A

· did not move w/ car

picture

✓ middle

→ · continuous countdown.

↑ watch car

1

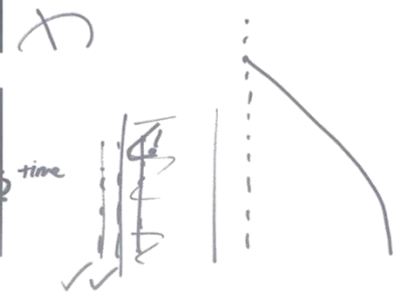
Participant no.: 8

Workspace

Design B

time

↑↑↑



2

Participant no.: 8

Workspace

Design C

OBVIOUS ?

few limited icons

message

How would this look in 10 km

how you're doing speed

Participant no.

Design A

Workspace

Cannot see the whole driving lane. ~~Time~~

→ RING WAS SCARY

Participant no.:

Design B

Workspace

① have show the situation more clearly like which lane and what happend. and the reaction time.

Participant no.:

Design C

Workspace

→ MORE REACTION TIME

→ what happen ✓

Lane ✓

Participant no.: 16

Design C

Workspace

→ how big symbol was clear.

→ I could what's the distance / how far

→ This was a good point

it better view vehicle

Participant no.: 10

Design B

Workspace

→ what does this stands for? Take over limit or collision.

→ Took my first attention

→ 3 Information

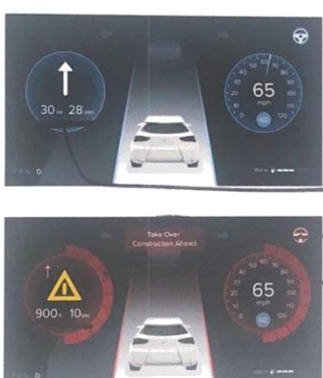
→ Felt something dangerous.

→ Distance in the design view

Participant no.: 10

Design A

Workspace



→ what does this means

→ Information was well placed.

→ why speedo became red?

→ Am I overspeeding?

→ Oh that's the distance b/w vehicles

Participant no.: 11

Design A

Workspace

Too big & unneccessarily big

like concept of countdown.

would prefer different color schemes based on time left.

Participant no.: 11

Design C

Workspace

much better representation.

gives better info about speed limit.

I do not like it here since it's too much to digest in a panicking sitn.

Too much info on one side of screen.

Participant no.: 11

Design B

Workspace

Better than Design C because it's visually more appealing.

may be change the icon to represent meaning better

Much better than Design C, - feel it is more alerting.

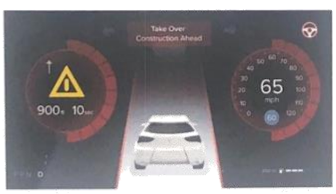
Better than other design
- Lanes
- Obstruction well visible.

Participant no.: 12

Design A

Workspace

I personally prefer if the speedometer were on the right.



Participant no.: 12

Design B

Workspace

This also gives details on which lane I'm in. So, I prefer this over other two.

Wasn't big enough to grab my attention.

Participant no.: 12

Design C

Workspace

I would want this on my side display. Not on the driver's display

I liked this among the three. Would have been better if this alert came in the middle of the screen. Also, initially I would like to see a big warning covering the entire screen.

Participant no.: 13

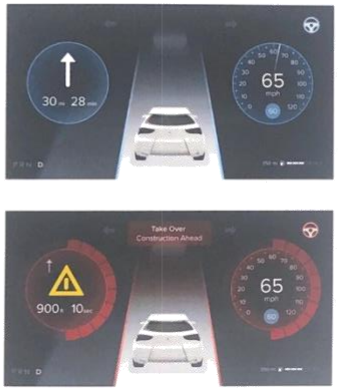
Design B

Workspace

Participant no.: 13

Workspace

Design A



Participant no.: 13

Design C

Workspace

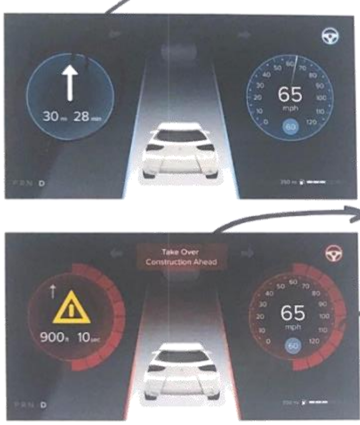
Participant no.:

Workspace

Design A

WHAT IS THIS FOR

↑ arrow just unnecessary



notification should be in centre

like the countdown

Participant no.:

Workspace

Design C

liked it but too small

like how it shows where in relation to me

Participant no.:

Workspace

Design B

2-lane view

Don't like

Sudden

Participant no.:

Design A

SEPARATING DISTANCE & TIME TO DESTINATION.

Workspace

LIKED COLOR CHANGE.
STANDARDIZED SYMBOLS.

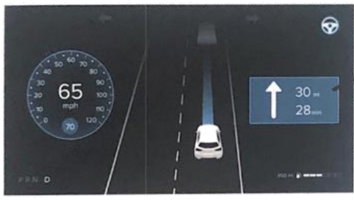
WHAT DOES IT EXACTLY CONVEY?

DOES IT SHOW TIME & DISTANCE OF OBSTACLE?

Participant no.: 15

Workspace

Design B



CLEARER.

LIKE THE SEPARATION.

CLEARER THAN PREVIOUS DESIGNS.

LIKE THE POSITION & INFO.

Participant no.: 15

Design C

Workspace

LABELING WOULD BE BETTER

LIKED THE MAP PREVIEW.

TRANSITION FROM DEST^n. TO OBSTACLE WAS CONFUSING. COULD BE GRAPHICAL.

LIKE THIS VERY MUCH.

REFERENCES

[1] Naujoks, Frederik, Yannick Forster, Katharina Wiedemann, and Alexandra Neukum. "A human-machine interface for cooperative highly automated driving." In *Advances in Human Aspects of Transportation*, pp. 585-595. Springer, Cham, 2017.

[2] Bellotti, Francesco, Alessandro De Gloria, Andrea Poggi, Luisa Andreone, S. Damiani, and P. Knoll. "Designing configurable automotive dashboards on liquid crystal displays." *Cognition, Technology & Work* 6, no. 4 (2004): 247-265.

[3] Wang, Wei, Feng Zhou, Wayne Li, and Jim Budd. "Designing the product-service system for autonomous vehicles." *IT Professional* 20, no. 6 (2018): 62-69.

[4] Mirnig, Alexander G., Magdalena Gärtner, Arno Laminger, Alexander Meschtscherjakov, Sandra Trösterer, Manfred Tscheligi, Rod McCall, and Fintan McGee. "Control transition interfaces in semiautonomous vehicles: A categorization framework and literature analysis." In *Proceedings of the 9th international conference on automotive user interfaces and interactive vehicular applications*, pp. 209-220. 2017.

[5] Inagaki, Toshiyuki. "Smart collaboration between humans and machines based on mutual understanding." *Annual reviews in control* 32, no. 2 (2008): 253-261.

[6] Neßelrath, Robert, and Michael Feld. "Towards a cognitive load ready multimodal dialogue system for in-vehicle human-machine interaction." In *Adjunct Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 49-52. 2013.

[7] Khastgir, Siddartha, Stewart Birrell, Gunwant Dhadyalla, and Paul Jennings. "Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles." *Transportation research part C: emerging technologies* 96 (2018): 290-303.

[8] Holthusans – Bruce paper,

//Trust in Automation: Designing for Appropriate Reliance

[9] Li, Mengyao, Brittany E. Holthausen, Rachel E. Stuck, and Bruce N. Walker. "No risk no trust: Investigating perceived risk in highly automated driving." In *Proceedings of the 11th*

International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 177-185. 2019.

[10] Tenhundfeld, Nathan L., Ewart J. de Visser, Kerstin S. Haring, Anthony J. Ries, Victor S. Finomore, and Chad C. Tossell. "Calibrating trust in automation through familiarity with the autoparking feature of a Tesla Model X." *Journal of cognitive engineering and decision making* 13, no. 4 (2019): 279-294.

[11] Trust in Automated Vehicles, Brittany E. Holthausen, Rachel E. Stuck, Bruce N. Walker Georgia Institute of Technology, School of Psychology

[12] Holthausen, Brittany E., Philipp Wintersberger, Bruce N. Walker, and Andreas Riener. "Situational Trust Scale for Automated Driving (STS-AD): Development and Initial Validation." In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 40-47. 2020.

[13] Gold, Christian, Daniel Damböck, Lutz Lorenz, and Klaus Bengler. "“Take over!” How long does it take to get the driver back into the loop?." In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 57, no. 1, pp. 1938-1942. Sage CA: Los Angeles, CA: Sage Publications, 2013.

[14] Mirnig, Alexander G., Philipp Wintersberger, Christine Sutter, and Jürgen Ziegler. "A framework for analyzing and calibrating trust in automated vehicles." In *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 33-38. 2016.

[15] Debernard, Serge, C. Chauvin, Raïssa Pokam, and Sabine Langlois. "Designing human-machine interface for autonomous vehicles." *IFAC-PapersOnLine* 49, no. 19 (2016): 609-614.

[16] Larsson, Pontus, Emma Johansson, Mikael Söderman, and Deborah Thompson. "Interaction design for communicating system state and capabilities during automated highway driving." *Procedia Manufacturing* 3 (2015): 2784-2791.

[17] Gold, Christian, Moritz Körber, Christoph Hohenberger, David Lechner, and Klaus Bengler. "Trust in automation–before and after the experience of take-over scenarios in a highly automated vehicle." *Procedia Manufacturing* 3 (2015): 3025-3032.

[18] Gold, Christian, Frederik Naujoks, Jonas Radlmayr, Hanna Bellem, and Oliver Jarosch. "Testing scenarios for human factors research in level 3 automated vehicles." In *International conference on applied human factors and ergonomics*, pp. 551-559. Springer, Cham, 2017.

[19] Thomopoulos, Nikolas, and Moshe Givoni. "The autonomous car—a blessing or a curse for the future of low carbon mobility? An exploration of likely vs. desirable outcomes." *European Journal of Futures Research* 3, no. 1 (2015): 1-14.

[20] Gowda, Nikhil, Wendy Ju, and Kirstin Kohler. "Dashboard design for an autonomous car." In *Adjunct proceedings of the 6th international conference on automotive user interfaces and interactive vehicular applications*, pp. 1-4. 2014.

[21] https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety

[22] Naujoks, Frederik, Christoph Mai, and Alexandra Neukum. "The effect of urgency of take-over requests during highly automated driving under distraction conditions." *Advances in human aspects of transportation* 7, no. Part I (2014): 431.

[23] Melcher, Vivien, Stefan Rauh, Frederik Diederichs, Harald Widlroither, and Wilhelm Bauer. "Take-over requests for automated driving." *Procedia Manufacturing* 3 (2015): 2867-2873.

[24] Borojeni, Shadan Sadeghian, Lewis Chuang, Wilko Heuten, and Susanne Boll. "Assisting drivers with ambient take-over requests in highly automated driving." In *Proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications*, pp. 237-244. 2016.

[25] Muir, Bonnie M. "Trust between humans and machines, and the design of decision aids." *International journal of man-machine studies* 27, no. 5-6 (1987): 527-539.

[26] Kraus, Johannes, David Scholz, Dina Stiegemeier, and Martin Baumann. "The more you know: trust dynamics and calibration in highly automated driving and the effects of take-overs, system malfunction, and system transparency." *Human factors* 62, no. 5 (2020): 718-736.

[27] Helldin, Tove, Göran Falkman, Maria Riveiro, and Staffan Davidsson. "Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving." In *Proceedings of the 5th international conference on automotive user interfaces and interactive vehicular applications*, pp. 210-217. 2013.

[28] J. Beller, M. Heesen, and M. Vollrath, "Improving the Driver–Automation Interaction: An Approach Using Automation Uncertainty," Human Factors: The Journal of the Human Factors and Ergonomics Society, 2013.

[29] McGuirl, John M., and Nadine B. Sarter. "Supporting trust calibration and the effective use of decision aids by presenting dynamic system confidence information." *Human factors* 48, no. 4 (2006): 656-665.

[30] Schmidt, Albrecht, and Thomas Herrmann. "Intervention user interfaces: a new interaction paradigm for automated systems." *Interactions* 24, no. 5 (2017): 40-45.

[31] SPECIAL REPORT: ADAS/CONNECTED CAR

[32] Zhang, Tingru, Da Tao, Xingda Qu, Xiaoyan Zhang, Rui Lin, and Wei Zhang. "The roles of initial trust and perceived risk in public’s acceptance of automated vehicles." *Transportation research part C: emerging technologies* 98 (2019): 207-220.

[33] Walker, Francesco, Anika Boelhouwer, Tom Alkim, Willem B. Verwey, and Marieke H. Martens. "Changes in trust after driving level 2 automated cars." *Journal of advanced transportation* 2018 (2018).

[34] Zhang, Tingru, Da Tao, Xingda Qu, Xiaoyan Zhang, Rui Lin, and Wei Zhang. "The roles of initial trust and perceived risk in public’s acceptance of automated vehicles." *Transportation research part C: emerging technologies* 98 (2019): 207-220.

[35] Lee, John D., and Katrina A. See. "Trust in automation: Designing for appropriate reliance." *Human factors* 46, no. 1 (2004): 50-80.

[36] Kunze, Alexander, Stephen J. Summerskill, Russell Marshall, and Ashleigh J. Filtness. "Automation transparency: implications of uncertainty communication for human-automation interaction and interfaces." *Ergonomics* 62, no. 3 (2019): 345-360.

[37] NASA TASK LOAD INDEL (TLX), Human Performance Research Group , NASA Ames Research Center , Moffet Field, Claifornia

[38] Noah, Brittany Elise. "Understanding automation handoff impacts on workload and trust when mitigated by reliability displays." PhD diss., Georgia Institute of Technology, 2018.

[39] Koo, Jeamin, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. "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)* 9, no. 4 (2015): 269-275.

[40] Endsley, Mica R., and Esin O. Kiris. "The out-of-the-loop performance problem and level of control in automation." *Human factors* 37, no. 2 (1995): 381-394.

[41] Hoff, Kevin Anthony, and Masooda Bashir. "Trust in automation: Integrating empirical evidence on factors that influence trust." *Human factors* 57, no. 3 (2015): 407-434.

[42] https://www.sae.org/standards/content/j3016_201806