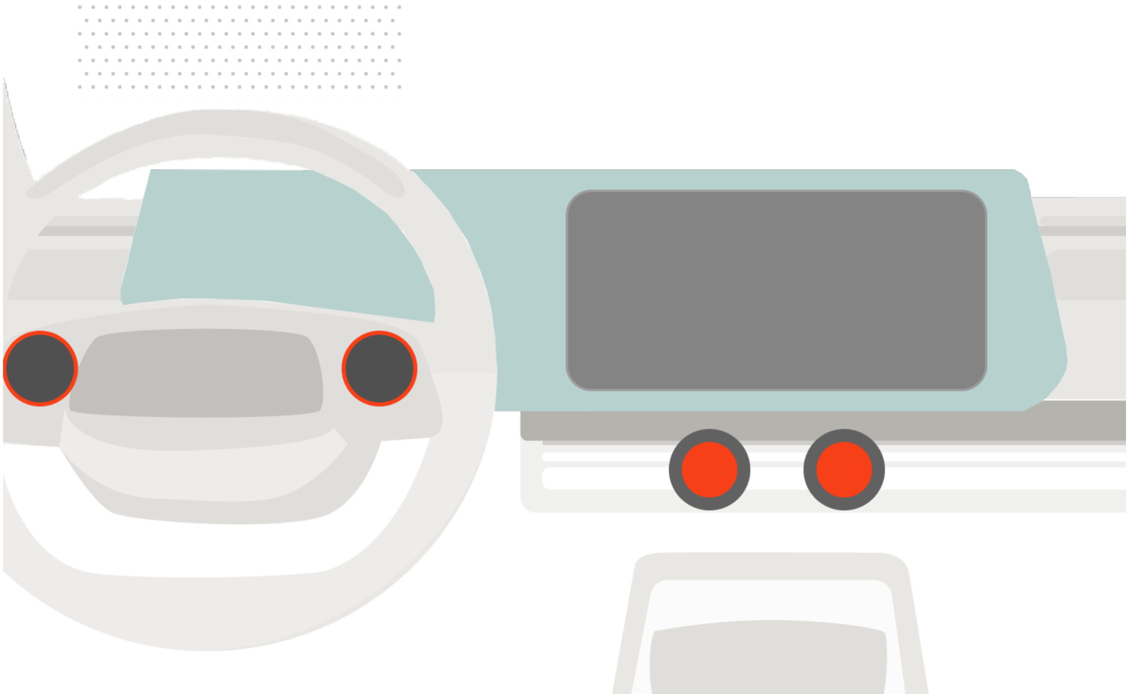

Using vehicle sensor data to enhance in-vehicle multimodal user experiences

How real-time fused media sensor-based data can be used to adapt and enhance multimodal in-vehicle driver and passenger experiences in terms of safety and comfort

Master's thesis in Interaction Design and Technologies

ADITYA GIRIDHAR & GUSTAV SVENSSON

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021
www.chalmers.se

Master's thesis 2021

Using vehicle sensor data to enhance in-vehicle multimodal user experiences

How real-time fused media sensor-based data can be used to adapt and enhance multimodal in-vehicle driver and passenger experiences in terms of safety and comfort

ADITYA GIRIDHAR & GUSTAV SVENSSON

UNIVERSITY OF GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Division of Interaction Design
Chalmers University of Technology | University of Gothenburg
Gothenburg, Sweden 2021

Master Thesis in Interaction Design and Technologies: Thesis Report
Using vehicle sensor data to enhance in-vehicle multimodal user experiences
ADITYA GIRIDHAR & GUSTAV SVENSSON

Supervisor: Morten Fjeld, Department of Computer Science and Engineering
Examiner: Staffan Björk, Department of Computer Science and Engineering

Master's Thesis 2021
Department of Computer Science and Engineering
Division of Interaction Design
Chalmers University of Technology | University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Abstract representative of the interaction design creation space explored by the thesis.

Typeset in LaTeX

Thesis Report
Using vehicle sensor data to enhance in-vehicle multimodal user experiences
ADITYA GIRIDHAR & GUSTAV SVENSSON
Department of Computer Science and Engineering
Chalmers University of Technology | University of Gothenburg

Abstract

The advent of more advanced vehicular sensor platforms, alongside the increase in computing power, has lead to an upswing for companies working in the area called sensor fusion. Proprietary algorithms utilise vast amounts of collected data to make complex predictions otherwise too hard, or impossible, to detect. While these wonderful possibilities exist, they are according to industry leaders currently under-utilised. Among the difficulties of working with novel modality interactions, limitations such as scope, time frame and the at-the-time societal COVID-19 restrictions, are among the challenges faced by this thesis. Two things areas of focus in this thesis, which are two suggested *How-to:s* - How to suitably utilise fused sensor data going forward in automotive interaction design, and: How to measure what is effective about such designs. These suggested How-to:s were the hypotheses, H1 and H2, to the to the research questions. To validate the hypotheses, a design project was held following their suggestions. Parallel to the designing, evaluations (both remote and in-person) measuring task-load and usability indexes were run for potential users, i.e. drivers. Results and discussions revealed that while the first hypothesis, H1, was only validated indicatively, the thesis was successful in validating H2 - the thesis providing a fine-tuned way of measuring positive and negative impacts of such designs from early to late-stage prototypes. Furthermore, the knowledge acquired in the preparation, the learnings from the process and the formative evaluations are potentially valuable for any future work. The thesis also aimed to, and succeeded with, contributing generated ideas and artefacts to the company this thesis was done in collaboration with.

Keywords: Multimodal Interaction, Multimodal Interfaces, Design of Multimodal Interactive Systems, Interaction Design, Automotive Interaction Design, In-vehicle Infotainment, ADAS, IVI, UX Design, UI Design, Thesis.

Acknowledgements

First of we would like to acknowledge the supervision and academic guiding given to us by Morten Fjeld, the supervisor of this thesis, whose helpful advice and counsel have aided in not only the creation of this work, but also any of out future professional en devours.

We would like to thank Aptiv, the company collaborated with in this thesis, and staff for the technical know-how and industry expertise provided throughout the different phases of this thesis. Special thanks goes to William Leeson and Lars-Christian Aadland.

Many thanks goes to the Department of Computer Science and Engineering, specifically the Division of Interaction Design and its faculty, which without, this thesis would not have came to be. The years of lectures, project supervision and academic guidance proved essential for the creation of this thesis.

Acknowledgements also goes to the non-faculty staff at Chalmers, whose facilities and effort the authors have grown used to rely upon in their daily life as graduate students.

Final, collective thanks go to all participants who not only helped in the making of this thesis project, but also in all the minor and major projects conducted before this thesis.

Aditya Giridhar, and Gustav Svensson, Gothenburg, May 2021

Through every phase of writing this thesis I received a ton of assistance and support for which I would like to sincerely thank and acknowledge the Division of Interaction Design, Department of Computer Science and Engineering. Their faculty, whose guidance and lectures equipped me with the empathy and understanding required to be a human-centered designer, something that I will forever treasure and hold in high regard.

I would also like to extend acknowledgements to my friends and peers here, who've been supportive and helpful all along. It was their friendship and warmth which made this Indian lad feel at home in a land far away from his. I genuinely thank you folks for that.

Finally, I couldn't have completed this thesis without the support of my *Amma* and *Appa*, for their sympathetic ear and wise counsel. You were always there for me. I'd like to sincerely thank all of you.

Aditya Giridhar, Gothenburg, May 2021

I would like to acknowledge and thank all the faculty of the Department of Computer Science and Engineering, whose 6 years of lectures and exams provided me with the understanding needed to enroll in a graduate program and gave me the knowledge base of an engineer - something I will continue to build upon in my career.

I would also like to acknowledge the academic and personal support given by associated professor Yuichi Wada and fellow lab member Shinpei Imada, at the Visual Cognition lab at the Human-Social Information Sciences department at GSIS, Tohoku University.

Acknowledgements should also go to the peers I've stood shoulder to shoulder with, whose inspiring interest and generous tips have benefited my academic life immensely.

My final, and greatest, thanks go to the family and close friends, whose support have aided and guided me for many years, and hopefully for many years to come.

Gustav Svensson, Gothenburg, May 2021

List of Acronyms, Abbreviations and Definitions

ACM	–	Association for Computing Machinery (*organisation name*)
ADAS	–	Advanced Driver Assistance Systems
AID	–	Automotive Interaction Design
API	–	Application Programming Interface
ARSCI	–	Adding, Removing, Skewing, Combining, Imbalance
AVD	–	Android Virtual Device
BMW	–	Bayerische Motoren Werke AG (*company name*)
CTH	–	Chalmers Tekniska Högskola (*university name*)
DALI	–	Driving Activity Load Index
EV	–	Electric Vehicle
GB	–	Gigabyte (*unit of data*)
GDPR	–	General Data Protection Regulation (*name of regulation in EU law*)
GM	–	General Motors Company (*company name*)
GPS	–	Global Positioning System (*product name*)
GUI	–	Graphical User Interface
HCI	–	Human Computer Interaction
HMI	–	Human Machine Interface
HUD	–	Heads Up Display
ISO	–	International Organization for Standardization (*organisation name*)
ISP	–	Interior Sensing Platform (*product name*)
IVI	–	In-vehicle Infotainment
LMS	–	Learning Management System (*product name*)
MID	–	Multimodal Interaction Design
ML	–	Machine Learning
MMI	–	Multi Media Interface
NASA	–	National Aeronautics and Space Administration (*governmental agency name*)
NTSB	–	National Transportation Safety Board (*governmental agency name*)
OEM	–	Original Equipment Manufacturer
OS	–	Operating System
PLC	–	Public Limited Company
POV	–	Point-Of-View
PX	–	Pixels
SCAMPER	–	Substitute, Combine, Adapt, Modify, Put to another use, Eliminate, Reverse
SUS	–	System Usability Scale

TLX	–	Task Load Index
TSI	–	Turbocharged Stratified Injection (*product name*)
UI	–	User Interface
URL	–	Uniform Resource Locator
UX	–	User Experience
WHO	–	World Health Organization (*organisation name*)
WPM	–	Words Per Minute

Contents

List of Figures

List of Tables

1
Introduction

The current trend in the automotive industry is to build automation systems to support or assist the human driver in vehicles [1] through aids, alerts and even corrective measures. These advanced embedded systems are collectively referred to as advanced driver assistance systems (ADAS) [2]. On the other hand, in-vehicle infotainment (IVI) systems have developed to a higher level of sophistication; even supporting smartphone-like features in systems such as *Apple CarPlay* [3], *Android Auto* and *Android Automotive* [4]. While both ADAS:s and IVI:s have advanced in their own respective capacities, as Schneider and Nett suggest, the cross-integration between these systems is lacking. Thereby it could be argued that in the case of modern vehicles with ADAS, a conventional human machine interface (HMI) system does not particularly enrich the experience of passengers by virtue of its lack of cross-integration; essentially functioning as two, albeit advanced, isolated systems.

Generally, designers inadvertently or otherwise dictate, or 'inscribe' as Latour (1992) puts it [6], certain ways of using things, which in turn creates new activities, or re-structures well-known activities in particular ways. Activities mould daily experiences with the potential consequence of enabling more or less happiness [7]. In the context of driving or commuting, cars are especially intriguing for two reasons; the first, a substantial amount of time is spent in cars [7]; the second, the car as a space for interactive technology. Vehicles present a highly structured, sensor-rich, personal space with untapped interaction design potential.

In 2006, Heide and Henning proposed the concept of a *"cognitive car"*; defining it as a *"cognitive driving assistance system"*, which leverages the knowledge of mul-

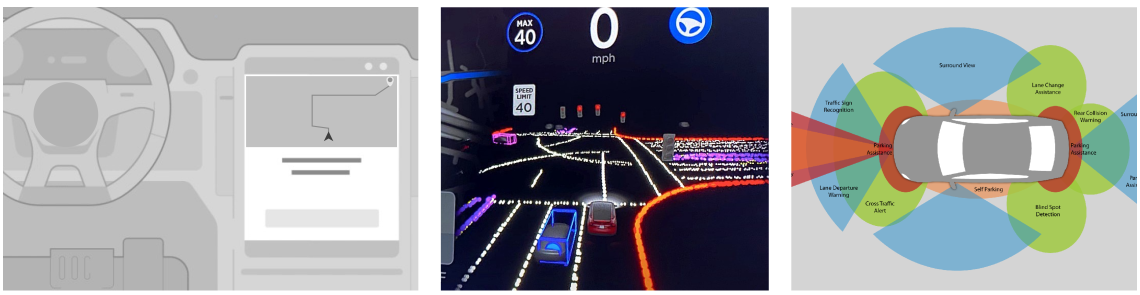

Figure 1.1: Emerging technological trends in AID; Full-size touchscreens (left), driver assistance systems (centre), and the multitude of vehicular sensors (right).

tidisciplinary fields and cognitive sciences to lower the stress or cognitive load placed upon drivers. While this perspective focuses on the drivers' strain; however, this could also lead to better and more comfortable driving experiences as a whole.

According to Wickens, Toplak, and Wiesenthal, the safety aspect of ADAS could be addressed from the outlook of cognitive sciences. A *"cognitive car"* as introduced by Heide and Henning should be capable of monitoring and identifying the errors in human drivers, but also do so in an accurate and transparent manner to avoid accidents and earn user trust. With the rapid evolution of the aforementioned ADAS:s [10] and IVI:s vehicles are gaining more functionality, autonomy and power, shifting the focus of mobility from the core driving activity to other activities. This raises questions about how people will spend their precious time in vehicles and ensure that such novel technologies are organised, integrated and deployed in a way that interactions with them are more enjoyable, safer, meaningful and less stressful.

Scope
Regarding the scope of the thesis, it was determined by external and internal factors, primarily the course requirements of the home universities, the pandemic restrictions of the locations as well as findings from the preparatory work, such as the literary research and interviews with industry experts.

1.1 Stakeholders

The Company - Aptiv
This thesis was executed in collaborations with the automotive company Aptiv (specific personnel as well as the company as a whole), part of Delphi Automotive PLC. The company places itself at the forefront of technological advances in the automotive field, as cited from their about page:

> *"Aptiv holds a leading position as a technology company innovating at the intersection of disruptive trends in the mobility industries. We use our portfolio of technologies to make vehicles safer, greener and more connected and enable the future of mobility."* [11]

Aptiv's current core areas are in the domains of advanced vehicular safety, power electrical systems, autonomous driving and in-cabin UX. The company's future roadmap lies in building trustworthy autonomous driving solutions [12] and developing automotive data connectivity for future V2X (vehicle to everything) applications [13].

Aptiv's interests in this thesis stem from their goal of furthering design knowledge in interior vehicle design, specifically; advancements in in-vehicle infotainment systems leveraging their proprietary ADAS and cabin monitoring system called ISP (Interior Sensing Platform).

Future vehicle occupants, i.e. drivers & passengers

Recipients of this work's design knowledge contribution, as it aims to improve their vehicular user experiences.

Thesis Authors, Home Universities & faculty members
Aditya Giridhar and Gustav Svensson are graduate-level students at Chalmers University of Technology (CTH: Chalmers Tekniska Högskola) - this work represents the graduation project in their academic careers so far.

Design Professionals, specifically Interaction Designers
Any design knowledge created by this work aids and informs them in their craft.

1.2 Research problem

With the advent of improvements to car riding safety, a lot of sensor data from the vehicle's interior and exterior environment is collected. What car manufacturers and designers seem to not know to a sufficient extent yet, is how to best make use of this data to improve other, non-safety aspects of the user experience, namely comfort.

On the topic of not knowing; how does a designer know a design decision does, in fact, improve the user experience (UX)? To understand the cause and specifically the effect of a design choice it is vital to evaluate the new design. What methods to use might not be straightforward if no previous work has been done; furthermore, with knowledge of how successful a method has been before, designers (and other stakeholders of the design) can more reliable estimate the added value of any design decisions. In the worst case, untested methods can prove useless, or even misleading, in evaluating a certain design only after it has been taken further, such as being released to the market or delivered to a customer.

With these things in mind, this thesis should therefore not only attempt to find the most UX-improving use of fused sensor data but also attempt to prove through example, a useful methodology for measuring the positive and negative value such a design adds. For the fully articulated research questions R1 and R2, see Section 5.1: **Research Questions** - R1 and R2 was first formulated as such when planing the thesis' practical part, which happened after the background and theoretical concepts were researched and summarized.

1.3 Aim and Planned contributions

Due to the nature of collaborating with a company when writing a thesis, it is quite common that expected results not only include academic contribution but also one or several artefacts valuable for the company - so is the case for this thesis.

For the company
Evaluated ideas and concepts on how to conveniently integrate a proprietary ADAS platform (Aptiv Interior Sensing Platform - ISP) into an existing Android-based

information concept, as well as knowledge on how one might use said technology to, for example, decrease driver distraction, increase driver convenience, and in general give the driver and passengers a better user experience.

For academia
Design knowledge on how user interface (UI) designers might suitably use vehicle occupant (drivers and passengers), interior and exterior sensor data, specifically data related to ADAS, to improve ride experiences, as well as how one might estimate the value added to those experiences. Secondly, the background research conducted and compiled by the authors will act as a primer for the topics covered by the thesis - both to readers and for future works.

2
Background

This chapter covers the main two areas involved in this thesis; Multimodal interaction design and Automotive interaction design. The sections cover terminology and sub fields related to the thesis. The chapter acts as a primer to the main fields touched upon by this thesis.

2.1 Multimodal Interaction Design

Multimodal interaction design (MID), a design field that is becoming increasingly relevant as advances and developments in technology allows for far more interconnected interface systems [7]. Multimodality is the act of either combining multiple input modalities or connecting an input to an output modality, as a way to conceive a new modality, combining the strengths of the parts.

In its early days MID came to be as two alternative input modalities, voice and gesture, were merged to conceive a more natural user modality (Bolt, 1980). While not coined by Bolt, the term *Multimodal Interaction*, sometimes *Multimodal Interface* and many others, is known in academia today - a search on *Association for Computing Machinery's (ACM) Digital library* resulting in nearly five thousand results [15, 16].

2.1.1 Modalities

Any of the body's senses can be viewed as a human input modality; however, the optical and acoustics related ones are the most commonly employed in Human Computer Science [17] due to their higher information transfer rate compared to other modalities. For instance, visual performance is rated at 250 to 300 words per minute (wpm) while aural is rated to be around 150 to 160 wpm [18].

Due to rapid advancements in sensing technology and general improvements in computing performance, the ability to leverage novel modalities has emerged [19]. These advancements also made their way into the automotive industry, resulting in the transformation of the humble car radio into an incredibly complex information aggregation, processing, and delivery system coupled with sensor fusion abetting the use of multiple modalities [20].

2.2 Automotive Interaction Design

Driver distraction still remains a major cause for traffic accidents even in recent years as the Global Status Report of the World Health Organization (WHO) in 2019 reported an estimated 1.25 million yearly deaths [21]. Additionally, modern IVI (In-vehicle Infotainment) systems are complicated to a degree for most people [22].

As Winograd once stated *"Successful interaction design requires a shift from seeing the machinery to seeing the lives of the people using it."* With this he implies that the designer's role in automotive interaction design (AID) is the *"...creation of the "interspace" in which people live, rather than an "interface" with which they interact"*. [23] Part of the design process from the start, is to explore opportunities to make the 'vehicle experience' more human-centred. One important topic for example is the feedback received by the driver from the vehicle. The modern vehicle bombards a user with feedback messages which, for the most part, are in the form of non-descriptive beeps [24].

Arguably the advent of autonomous and semi-autonomous safety and steering features necessitate entirely new approaches for AID [25]. Advanced levels of ADAS, based on new sensors that generate voluminous quantities of data [26], require new approaches to the visualization of insights from said data - especially since the raw form of the data is all but unintelligible to a human. Another aspect to consider is that electric vehicles (EVs) also require novel modes of presenting data and metric such as range awareness, charging station locations and battery status to the driver. These developments have instituted several UX opportunities and trends as a result, from other industries, expanding the design space available to the designer.

2.2.1 ADAS Platform

An ADAS platform is an enclosed, interconnected sensor platform that generates data on vehicle occupant, the interior and exterior environment - a common if not universal presence in modern cars, one such platform being Aptiv's proprietary product Interior Sensing Platform (ISP). This data is typically analysed and processed by software using what is called sensor fusion to predict and estimate complex values such as estimated driver distraction level, child-presence detection, etc. [27]

While the data from the sensors is there, typically what is limiting use cases are algorithms and ML (Machine Learning) powered systems capable of providing estimates based on the data. Companies like Aptiv are developing advanced suites, Aptiv ISP being one, capable of proving such predictions, opening the design space further for automotive developers and interaction designers.

2.2.2 In-vehicle infotainment

In-vehicle infotainment (IVI), or infotainment for short, is the name of the more holistic view of what otherwise is referred to as entertainment system, car stereo, or centre console. As the name implies, it is a piece of technology that provides information to vehicle occupants, such as GPS (Global Positioning System) and climate controls - as well as entertainment, usually in the form of audio. In recent times, a large touch screen and spread-out speakers is the trend for new cars, while older models tend to have far more tactile interfaces with knobs, spinners and buttons. [28]

With the increasing inclusion of modern touch interfaces in vehicles, users expect the same features and UX as with stand-alone touch interface devices such as quotidian smartphones and tablets. This results in harsher demands on innovation on the relatively slower, compared to smartphone and tablet development, automotive industry. Added to this, people tend to upgrade to newer models of tablets/phones far more often than their vehicles. [29, 30]

2.2.3 Design priority: Safety

Traditionally in automotive design development has been, and still is, primarily focused on improving ride safety. Advancements in UX have been at a lower priority for automotive designers, as they probably should be. However, as car safety increases, what is the remaining limiting factor today is mainly the human factor [31]. With causes such as driver distraction or unawareness, the blame can be placed on vehicles' automotive interaction design. An example of such faults could be: a driver is distracted by a frustrating or confusing interface system, as a result not seeing vital road information. To tackle these non-mechanical safety issues, designers in AID have to turn to methods found in conventional interaction design, such as UX and UI design where usability and user satisfaction has been the goal since the fields were introduced.

2.2.4 Design priority: Trust

Another non-mechanical safety issue affecting the success of these systems is the adoption and general reliance on this technology by the average consumer. As Trübswetter and Bengler reports, the results of the survey conducted investigating *'reasons for rejecting the use of driver assistant systems'*, could primarily be chalked up to lack of trust. The reason for this lack of trust in ADAS was reported to be due to *"inappropriate system design"*, *"undesired system feedback"* or most commonly because of the *"lack of perceived usefulness"*. With advanced safety in vehicles emerging as essential, the relatively low trust in ADAS could hamper road safety in general. Therefore, as mentioned earlier and suggested by Trübswetter and Bengler it would be *"...interesting to examine how user experience with ADAS will influence acceptance and willingness to use these systems"*.

2.2.5 Definition of comfort in AID

In the subtitle of this thesis, it is stated that it aims to find uses of vehicular sensor data, both in terms of safety and comfort. The 'safety' part is not the main focus but still an essential part of any design work in AID due to the potential severity of not taking it into consideration. As for the term 'comfort', it can arguably be seen as a highly subjective term, what is comfortable for one user might not be for another. In this thesis the term is understood as the overall comfort across users, where high comfort is equal to high usability and low mental and physical work load under use.

3
Related work

This chapter's sections cover related previous academic work in multimodal interaction design, automotive interaction design as well as estimates of what design knowledge the authors sees as currently lacking in academia.

3.1 Multimodal Interaction Design

As a field, MID hosts designers from all manner of fields - if it has more than one input or output modality, it is included. While not named as such, MID in its current digital form was perhaps first introduced in the work by Bolt (1980). Bolt talks about how new inventions in the field has *"...encouraged the notion that voice and gesture inputs at the graphics interface can converge to provide a concerted, natural user modality"* [14], indicating that with design freedom, multimodal interactions are a not only a possible result but also a perhaps a more natural way of interacting. The following subsections cover academic examples of combined in- and output modalities relevant to both MID and the focus of this thesis, AID.

Input modality combinations
The following section covers related work done in the MID field, with two or more input modalities combined. When no output modality is specified, traditional visual output on the centre console is used.

TACTILE/MECHANICAL (Buttons, dials, knobs, etc.)
This might be the oldest modality available to vehicles, as it originated at a time when vehicles featured limited functionality and thereby offered, a one-to-one mapping from control to function. Even with a growing list of functions and options in the IVIs, until recently the majority of the vehicles still featured a variation on the 'buttons and knobs' layout from the past. The controls however were no longer mapped one-to-one but followed a trend in automotive systems wherein the different functions are combined in a hierarchical menu structure accessed by a set of context-switching physical controls. [28] As an example: before the introduction of touch screens, on-screen content was solely interacted with using buttons and knobs - still today users may forgo the touch interaction for common features such as volume and temperature control, instead use the still present knobs.

GAZE & TOUCH
Prabhakar et al. (2020) explored moving touch interaction from the centre con-

sole to the Heads up display (HUD), while also allowing limited gaze control. The design was evaluated using standardized ISO (International Organization for Standardization) tasks, showing a similar performance in task completion times and an improvement in driving performance compared to just using the standard centre console touch screen. [33]

GAZE, VOICE & GESTURE
In a paper by Roider et al. (2017) it is described how a simulated driving environment was set up, in which the driver could, alongside driving, perform non-driving tasks using a combination of gaze, voice and gesture input [34]. Participants filled in a DALI (Driving Activity Load Index) questionnaire. The test results showed that each modality performed the worst under disturbances in their respective modality field, i.e. optical for gaze and vice versa, showing that a combination is needed to cover weaknesses in each. Individually they react differently to disturbances:

"*"The effect of additional demands on gaze input mainly led to an increase of task completion times, whereas effects on speech input were reflected only in subjective ratings and effects in gesture input were split up between both factors."* [34]"

Pfleging, Schneegass, and Schmidt (2012) explored other input alternatives to the conventional mechanical buttons and knobs, focusing on the then-novel voice and gesture modalities. Their findings were limited by the technologies of the time, showing slower task-completion time and unchanged driving performance. However, using a DALI questionnaire, users expressed that that voice and gesture input had a lower visual demand than the conventional modality. [29]

GESTURE
In the extensive synopsis by Bilius and Vatavu (2020), one of the modalities covered is Gesture. The synopsis highlights the discovered strengths of the modality, primarily being lowered visual demand and distractions upon the driver of the vehicle. [30]

Less than average review was done of this modality as *the company* had, before the review, expressed a desire to avoid this modality; however, by reading the synopsis by Bilius and Vatavu as well as other papers partly involving gesture as a modality, a good enough understanding was achieved by the authors. [30]

Output modality combinations
The role of output modalities is one of feedback to users' interactions, meaning as a response to one or more input modalities. In other cases, it can act based on non-occupant interaction such as relaying vehicle status, warning about road obstacles etc. The following section covers multimodal combinations of at least one output and one input modality.

VISUAL SOURCES
Includes dials, warning lights, screen and most recently augmented sections of the

windscreen: Heads Up Display (HUD) and Audio sources.

Related work targeting UX with these modalities include Knobel et al.'s *CliqueTrip* (2012) wherein cars in a motorcade would point to the leading on the navigation screen (visual source) and when in close proximity a two-way auditory channel would open up enabling cross-vehicle social interactions [35]. Krome et al. (2017) explored in-car music creation building upon the convention of singing in the car, leveraging these modalities to boost in-car UX [36].

HAPTIC & TOUCH
In a study by Vo and Brewster (2020) interaction permanence improvements were measured for a setup where tactile input, either on the steering wheel or the centre console, was combined with mechanically haptic output. Vo and Brewster reported a performance increase of 20%, performance was calculated based on driving performance, task completion performance, glance behaviour and a NASA (National Aeronautics and Space Administration) TLX (Task Load Index) score. [37]

Farooq, Evreinov, and Raisamo (2019) explored non-traditional haptics such as pneumatic vibrations and tangential side-to-side (X-axis) vibrations - both as alternatives to conventional (Z- & Y-axes) haptics that suffer from vibratory noise that is primarily along the Z- and Y-axes when utilised inside a vehicle. Their designs were mainly evaluated using primary (driving) and secondary (text entry, menu navigation, etc.) task performance values. [38]

TEMPERATURE & TOUCH
The pneumatic haptics used by Farooq, Evreinov, and Raisamo (2019) used air redirected from the test vehicle's subwoofers, the air streams feeling slightly cool as a side effect; they did however theorize a system that could more thoroughly utilise the thermo-differential sensitive fingertips humans naturally come equipped with. In their setup, upon touch interaction with the centre console, slightly temperatured blasts of air hit said finger. [38]

MECHANICAL & BUTTONS/DIALS
A common feature in modern cars is adjusting mechanical settings such as seat position indirectly using buttons, forgoing the traditionally used manual pulling of levers to slide the seat. This modern feature effectively decouples the input and output actions into further separate modalities, the input being the button press, and output is the mechanical shifting of physical parts of the interior.

3.2 Frameworks and Guidelines in Multimodal Interaction Design

Guidelines for Multimodal Interface Design [39]
Authored by Reeves et al. (2004), the guidelines state a selection of general rules to follow when designing an interface for multimodal interaction. Some of these apply

less in AID than others, due to environmental restrictions such as driver attention demand. The sections covered are as follows:

DESIGNING MULTIMODAL INPUT AND OUTPUT
"The cognitive science literature on intersensory perception and intermodal coordination has provided a foundation for determining multimodal design principles. To optimize human performance in multi-modal systems, such principles can be used to direct the design of information presented to users, specifically regarding how to integrate multiple modalities or how to support multiple user inputs (for example,voice and gesture). [...] *Designers need to determine how to support intuitive, streamlined interactions based on users' human information processing abilities."* [39, p. 58]

While not common at the time Reeves et al. wrote the guidelines, in current information technology, with the advent of smart voice assistants, it is common, especially in the quotidian devices, to support several in and output modalities as a default. This makes following this aspect of the guideline a non-issue for most projects of this thesis nature.

ADAPTIVITY
"Multimodal interfaces should adapt to the needs and abilities of different users, as well as different contexts of use. Dynamic adaptivity enables the interface to degrade gracefully by leveraging complementary and supplementary modalities according to changes in task and context. Individual differences(for example, age, preferences, skill, sensory or motor impairment) can be captured in a user profile and used to determine interface settings." [39, p. 58]

This guideline can be hard to follow depending on the complexity of the artefact involved; too many variables increase the complexity of the task to make suitable adaptions, too few and adaptions are hard if not impossible to do. Depending on the complexity of the artefacts developed, the scope etc., in works such as this thesis, this aspect of the guideline could or could not be followed fully or at all. With too complex artefacts, only a few adaptions could be catered for within the time frame - leaving many aspects either vague or 'hard coded' for within the project's time frame. With too simple ideas, in terms of complexity, less could be done to them to adapt them as they would have fewer 'moving parts'.

CONSISTENCY
"Presentation and prompts should share common features as much as possible and should refer to a common task including using the same terminology across modalities." [39, p. 58]

Consistency can be easy to achieve, at least on a surface level, when designing for an established modern software environment such as *Android Automotive* - designers are recommended to use existing design guidelines precisely for such a purpose [40]. As for cross-modality consistency, as long as all interactions are possible with all supported modalities this guideline could be seen as fulfilled.

FEEDBACK
"Users should be aware of their current connectivity and know which modalities are available to them. They should be made aware of alternative interaction options without being overloaded by lengthy instructions that distract from the task. [...] *Also, confirm system interpretations of whole user input after fusion has taken place, rather than for each modality in isolation."* [39, pp. 58–59]

One could potentially argue for a clash between this aspect of the guideline and what is the case for the voice input modality in modern devices. On one hand, users are taught about the existence of it and presented it as if it is all-encompassing - something far from the truth as there is no formal requirement for each individual piece of software to support the modality. This means that users might assume the modality as always available - fulfilling the aspect; however, for non-assuming or unknowing users, the modality is suddenly fully invisible. The authors argue for an approach where the target user and platform is taken into account, as it certainly affects whether an artefacts users are assuming or not. In the case of this thesis, the use environment is one where all but the vocal modality is preoccupied with the primary tasks, thereby making it a necessity in almost all newly developed artefacts. For these reasons, this thesis will target users who are aware of the modality, due to this the thesis can take into advantage the fact that the modality is taking up precisely zero interface real estate - in line with the aspect of guideline. As for other modalities present in the artefacts developed, a similar approach will be taken on a case by case basis - estimate user assumptions and develop a guide-following approach for it.

ERROR PREVENTION/HANDLING
"User errors can be minimized and error handling improved by providing clearly marked exits from a task, modality, or the entire system, and by easily allowing users to undo a previous action or command. To further prevent users from guessing at functionality and making mistakes, designers should provide concise and effective help in the form of task-relevant and easily accessible assistance." [39, p. 59]

Interactive aid will be given in the forms of design interface conventions for pliancy, affordance, etc. Following the design guidelines developed by the platform's owner, such as Google when developing for Android Automotive [40], provides designers with an easy way to follow this part of the guideline.

A Pattern-Based Methodology for Multimodal Interaction Design [41]
Ratzka and Wolff (2006) defines and suggest the following framework for addressing a project in MID [41], see Figure 3.1.

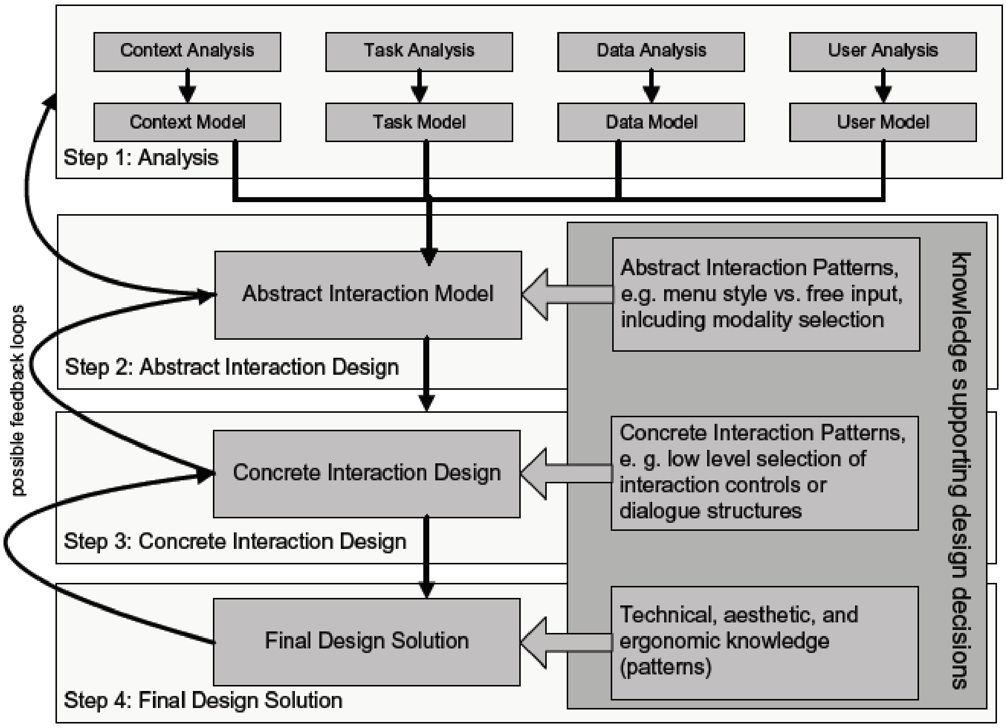

Figure 3.1: The suggested design process for a project in multimodal design, by Ratzka and Wolff [41].

In their paper, they also introduce a selection of design patterns, a common methodology to use in design, created specifically with MID in mind. [41]

From the perspective of a master thesis done in collaboration with a company, in regards to *Step 1: Analysis*, it is common that the company will provide at least a preexisting customer base and perhaps even task and data models. This depends upon if the company has a product they are interested in improving or if they have a solution or service not utilised as of yet. For the latter steps, it is more up to the designers to create new artefacts and knowledge based on informed decisions from guidance given primarily through academia or the company. For improvement or refinement decisions it is more common, however, to employ evaluation methods, sometimes involving real or simulated user group representatives.

Multimodal information presentation: Design guidance and research challenges [42] resp.
Ontology-driven Elicitation of Multimodal User Interface Design Recommendations [43]
Sarter (2006) suggest four basic themes of MID [42]:

- *The selection of modalities*
- *The mapping of modalities to tasks and types of information*
- *The combination, synchronization and integration of modalities*
- *The adaptation of multimodal information presentation to accommodate changing task contexts and circumstances.*

Tourwé et al. (2011) reformulates Sarter's four themes into these three questions [43]:

- *Which input and output modalities are available to the user of an application?*
- *What are the different factors that affect the use of particular input and output modalities?*
- *What are the appropriate (combinations of) modalities to support users in a particular task?*

It goes without saying that any designer tasked with designing an interactive system using multiple modalities should consider following these guidelines and through the design process strive to create a design that, in some way, answer the stated questions.

3.3 Automotive Interaction Design

A steady rise in safety and efficiency features coupled with autonomy, has given rise to a fundamental evolution of driver-vehicle interaction; from the extension of the driver to a distinct entity with which control is shared, as stated by Pettersson and Ju (2017) [44]. These innovations result in new challenges for AID as the system gains more agency, gradually shifting human-vehicle relationships [45]. In the recent past, the acceptance of hedonic characteristics in the field Human Computer Interaction (HCI) has improved [46] resulting in AID researchers also adopting some of these 'UX-focused' features. However, since its genesis, the majority of AID research has typically focused on the design of safety system IVI:s - as a result, the AID research cited in this thesis is primarily drawn from safety-related research.

Multimodal interaction in AID

The primary advantage of employing multimodal interfaces is, among others, in crafting natural user interactions in HMI communications and boosting the robustness of interaction by using redundant or complementary information as suggested by Reeves et al. (2004) [39]. The application and implementation of multimodal interaction in AID, for the most part, has been employed for Driver warning systems (multimodal output) [24] and controlling system navigation [47, 48] (multimodal input). Therefore multimodal interaction in AID as a proponent of primarily UX is the relatively niche design space this thesis aims to contribute to.

Foundational examples of Multimodal interaction in AID at mass-market scale came in the early 2000s with AUDI's MMI (Multi Media Interface) system [49] and BMW's iDrive systems [50]. Both systems featured the pairing of two parts: a single controller and a display. Tertiary purpose controls were often placed in the centre stack. With the addition of the multifunctional steering wheel controls, a few tertiary controls intruded into the domain of secondary devices, for instance, audio controls, phone call buttons, etc., on the steering wheel for quick access to frequently-used functions [28]. The aforementioned systems featured multimodal interaction as the users could choose between touch, knobs and buttons, to perform the same action.

Recently, cutting-edge work in this domain has been focused around a combination of specific modalities such as gesture, voice and gaze [30, 51, 33]. This can partly be attributed to the decreasing cost of computing power and sensors, but mainly thanks to various on-the-road complexities which can necessitate additional demands on the driver that goes beyond the basic demands of driving, i.e. more interactions are being performed and desired within the vehicle environment than initially conceived. Thereby, these factors influence the active development, acceptance and appropriateness of exterior and in-vehicle input modalities to operate secondary tasks in the car.

Research in the domain of cognitive science showed that people have separate pools of attentional resources, referring to the different sensory modalities [52], thereby allowing the use of alternative input modalities to change the proportion of cognitive and perception resources required for engaging with the secondary tasks. Automotive original equipment manufacturers (OEMs) and sensing technology firms like the company associated with this thesis have in recent years worked on advancing the modalities of voice, gestures and gaze [27]. As a rule, Gaze input mainly addresses users' visual resources, gesture input addresses the manual, and speech, the auditory.

In summary
Automotive companies' primary focus of services is foremost automotive safety and have recently started to apply artificial intelligence, machine learning and low-level fusion capabilities to enhance and power novel safety features to minimize the risk of road accidents, irrespective of the cause being internal or external. To increase the safety of cars, automotive companies have for a long time developed solutions to mitigate safety issues present when driving or riding a vehicle. Lately, the focus has gone from mechanical features such as material choice, seat belts and airbags, towards software solutions reacting to, among other things, external spatial information, driver distraction and drowsiness. [53]

3.4 Academic Vacuity

While automotive companies, such as the one connected to this thesis, consistently provide new safety solutions, a hitherto untapped design field remains, and perhaps there is where the last to-be-addressed safety issues are found. The field in question houses the hard, and impossible, to identify issues - what is commonly referred to as *Wicked problems*[54, 55]. Wicked problems by nature cannot be fully solved, using the methodology explained by Gaver (see section Section 4.1) attempts at mitigating them can be made.

Automotive companies have when adding advanced safety features, developed software platforms capable of complex sensor fusion algorithms, providing unique and so-far underutilised modalities - both in the automotive industry and in academia. The following section covers what the authors deem to be missing or lacking in academia; however, as reading every piece of literature is impossible it is possible examples exist but were not found within the scope of the literature review.

ADAS data [2, 27, 53]
With the introduction of ADAS systems, a new form of user interaction emerges, a form where the user does not need to actively, or even knowingly, interact with the system. It includes input modalities such as real-time collected occupant, exterior and interior sensor data in the form of:

PROXIMITY
Proximity as an input modality is an area of immense research work in the recent past, especially in the sub-domains of desktop and smartphone interactions. Seldom explored by prior research efforts, proximity is especially interesting in the domain AID as it enables features such as predictive touch, finger-pointing vector, automatic steering, etc. All of these examples of features promoting passive interaction with the system, i.e. the system responding to the intent of the user, even before the user input has been received - as well as the other form of proximity, the proximity of external entities. These can include other vehicles, pedestrians and curbs - all additional passive live data feeds acting as an input modality to the ADAS platform, allowing predicting the surrounding area's intentions.

POSTURE
Although primarily tracked for determining fatigue and enforcing focus on the primary activity of driving; body posture also is indicative of the mood and intent of the driver and passengers, which could potentially be leveraged for proactive system behaviour.

FINGER POINTING VECTOR & PREDICTIVE TOUCH
Using occupant spatial data it is possible to estimate what vector finger is currently pointing along as well as estimated points of interaction on any surface. This allows for complex intent-based interactions, examples include: what a user is pointing at can be used as a form of input - even if it is outside the vehicle, or: calculations can be made before a finger reaches the screen - resulting in a more responsive interface.

MOOD, DISTRACTION, IDENTIFICATION, OBJECT CLASSIFICATION, ETC.
With the face analysis algorithms developed in the field of computer vision, it is now possible to identify a person by capturing their face digitally, as well as estimate complex user statuses such as mood and distraction. These algorithms are not only limited to humans, some also allow for visual object analysis such as classification.

Manipulation of the Output
As modern interfaces and computing systems replace simpler meters and low-resolution displays, manipulating, an alternative to just adding/replacing, audio and visual output modalities to improve user experiences will be possible.

Guidelines & Frameworks
With these new, unexplored and definitively complex modalities within AID, there

is a clear lack of guidelines and frameworks of use. Until such is made, case examples involving them will be the only guide there is.

4
Methodology

The methodology used was primarily sourced from IDEO's human-centred design toolkit and their resources on the implementation of design thinking. Herein design thinking is termed as a human-centred approach to innovation. The toolkit classifies methods into three categories, *Inspiration*, *Ideation* and *Implementation*, the last involving bringing the design to market - something that is outside the scope of most theses, this one included. [57]

While many of the methods were brought over from the previously mentioned toolkit, suitable adjustments were necessary to adapt them for the project's contexts, and other common methods in the field were sourced as well. The evaluation methods were chosen such that they measure usability as well as the workload for individual tasks as a part of the larger workflows in the driving/commuting activity.

For the most the concepts covered by this chapter, it is described in detail why they were utilised, when and how in chapters 5: **Planning** and 6: **Execution | Process**. As such that information is not covered in this chapter; however, for concepts utilised throughout the thesis, such as general guidelines, why and how they are used is covered in chapter. This stylistic choice is done to increase the readability of the chapters, as well as to maintain the chronological order.

4.1 Research through Design

According to Gaver (2012), what differentiates research through design from traditional scientific methods, primarily found in the natural sciences, is that the research problems it tackles are by definition not falsifiable - making the criterion of falsifiability proposed by Popper in 1963, void [56, 58]. This nature stems from the problems being so called 'wicked' problems - in the sense that there is no simple, factually correct answer or solution to them [54, 55].

Wicked Problems and Research through Design
Rittel and Webber (1973) defined ten characteristics for wicked problems[55], one of them being that they are unique and another that they usually are a symptom of another problem. Gaver comparatively mentions that research through design is defined by the way that what it achieves as it evolves, are the artefacts developed along the way - each a testament of one way of potentially creating successful designs [56]. Gaver argues for the notion that what is created is *'what might be'*

[56], to follow this one could see a preference of working in-so-far untouched areas of design space. Seeing as new designs in already explored areas will most likely yield fewer new insights, the conclusion to prioritize 'gaps' in explored design space can be drawn.

As for the second characteristic, the cause of the wicked problems that research through design tackles, one could formulate it as being the problem of no academic knowledge covers that area. From this it is easily realized that by attempting to solve the wicked problem one actually, at least partly, solves the underlying cause for it - as written by Gaver: *"...one of the features that distinguish design from science is its tendency to make generative statements rather than falsifiable ones"* [56]. Gaver argues that by being generative, research through design achieves the same goal as what Lakatos (1978), as cited by Gaver, states is the strength of scientific research programmes - that over time and attempts, the solution or truth is found [59, 56]. Research through design with its many artefacts has as a goal *"to create theories that are sometimes right"* [56], that *"designing for X can lead to successful outcomes"* [56], this can be understood from what Gaver continues by saying *"Design examples are indispensable to design theory because artifacts embody the myriad choices made by their designers..."* [56] as well as what Collins (1994) said, as cited by Gaver: *"...a designed artifact is a 'theory nexus'..."* [60, 56].

The nature of the research problem being how to best or optimally, leverage the possibilities present, arguably makes it a wicked problem. The design possibilities are endless and what one user thinks is optimal is not a guarantee for the next one asked answering the same. In interaction design and user experience design, subjectivity plays a part, making it impossible to determine what is best - what is often done instead is to compare alternatives and estimate which is the better one for most users. Designers, such as the authors, can only make better interfaces, never the best. The research problem of this thesis is shaped in such a way, that it plays into the second characteristic described above; it targets an area without academic coverage. By attempting to solve the research problem, this thesis will contribute at least by adding some coverage to the explored design space.

4.2 Co-design

According to Sanders and Stappers (2008) Co-design as it is exercised and discussed in contemporary design fields assumes quite a few manifestations, based upon the attitudes and expertise of its practitioners. They also clarify that:

> *The terms co-design and co-creation are today often confused and/or treated synonymously with one another.* [...] *Co-creation is a very broad term with applications ranging from the physical to the metaphysical and from the material to the spiritual.* [61]

Sanders and Stappers, for clarification and differentiation from Co-creation, instead define Co-design as follows :

> *We use co-design in a broader sense to refer to the creativity of designers and people not trained in design working together in the design development process.* [61]

Regarding the roles assumed by the non-designers and/or users in this process, Muller and Druin (2002) states that, *"With each role there is a spectrum of user involvement, at differing points in the design of new technology"*. Yamauchi (2009) suggests that the best role for users is as *"peripheral designers"* working with assigned detailed problems rather than whole-system design. In line with this ideology is Light and Luckin's (2008) statement: *"...you can't just 'add users and stir."* and Muller and Druin's remark that *"People's needs differ by work roles and their relationship to the design task, by life stage, by physical or cognitive condition, and by other attributes and dimensions as well"*.

4.3 Literature study

The following steps are a guideline provided by the staff at Chalmers University library [65]:

1. In the initial part of a thesis project, interesting and relevant papers should be selected based on sporadic searches of thematic words found in, for example, the thesis proposal.

2. From these papers, select keywords should be chosen to be used for more thorough literature searches using the digital tool *Scopus*[66].

3. Different combinations of the keywords can then be used to find well-cited older papers as well as more recent ones. This should be done to get both reliable, but older knowledge of the field as a whole as well as design knowledge from more up-to-date design projects. This is especially crucial since the field of automotive interaction design has over recent years gone from the slow and steady design evolution common in the industry towards the rapid, constant changes of the digital interaction design world [29, 30].

4. Throughout the process, the collection of keywords should be expanded upon and evolve as the authors' understanding of the field grows.

5. When a broader base of papers has been covered, further studies should focus on the sub-fields regarded as most relevant to the project. These priorities are based on a multitude of reasons. In this thesis that is primarily what the company the authors works with have an interest in; A) design solutions that do not primarily aim to enhance safety, B) novel to the company-design combinations of different modalities, and C) design solutions utilizing the proprietary ADAS platform's more advanced sensor data.

Regarding A): The company's main expertise is in the safety sector, with this project they desire to expand their non-safety expertise.

Regarding B): Previous thesis projects housed by the company have already involved some of the modalities existing in the field. As a company, they desire to see novel projects, with as little overlap as possible.

Regarding C): The least design knowledge exists in the field when it comes to utilizing the latest sensor data. The company desires to gain this knowledge as it positions them at the head of the industry as a whole.

6. After the study, lacking aspects of the relevant fields should be identified to better shape and target the aim for the academic contributions of the thesis, i.e. to avoid rediscovering the wheel as well as maximizing the new understanding generated by research through a design project.

4.4 Expert Interview

To design for actual needs and problems in a field, getting expert know-how is invaluable for a designer, and can sometimes be a better source of inspiration than the actual end-user of the design - an example would be asking a customer instead of a chef for ideas on a new menu item at a restaurant. Experts in the field that are not designers, such as manufacturers or engineers, have knowledge on what the latest innovations are, known failures as well as a systems-level view of the subject area. Designers seldom have the technical expertise of the field they design in, interviewing experts following the guide found at IDEO's toolkit may help in such scenarios [67]:

1. *"Determine what kind of expert you need."*
2. *"When recruiting your experts, give them a preview of the kinds of questions you'll be asking and let them know how much of their time you'll need."*
3. *"Choose experts with varying points of view. You don't want the same opinions over and over."*
4. *"Ask smart, researched questions. Though you should come prepared with an idea of what you'd like to learn, make sure your game plan is flexible enough to allow you to pursue unexpected lines of inquiry."*
5. *"Record your Interview with whatever tools you have. A pen and paper work fine."*

4.5 ARSCI

Abbreviated ARSCI (Adding, Removing, Skewing, Combining, Imbalance) by the authors, this unnamed structured creativity method was mentioned in a CTH course by Mafalda Samuelsson-Gamboa, referring to a method mentioned by Staffan Björk. It is possible that ARSCI is based on the less structured SCAMPER (Substitute, Combine, Adapt, Modify, Put-to-another-use, Eliminate, Reverse) method [68], both being evolutions from the methodology first introduced by Osborn (1953) in chapter XXIV of the book *"Applied Imagination: Principles and Procedures of Creative Thinking"* [69]. The method is structured as follows:

1. State the main problem [X] that the ideas should aim to solve.
2. Identify all design elements [E_1-E_n] defining the design space.
3. State a personal goal [Y] for the ideas to fulfill.
4. Enter [E_1-E_n] into the following templates and write an idea that is based on the filled-in template and fulfills [Y].
   (a) Solve X by adding more of E_a.
   (b) Solve X by removing E_a.
   (c) Solve X by skewing E_a.
   (d) Solve X by combining E_a and E_b.
   (e) Solve X by creating an imbalance between E_a and E_b.
5. Roughly cull undesired ideas, based on possibility, skill, cost, etc.
6. When few enough 'good' ideas remain, continue with the design process.

The method's goal is to generate many ideas that fulfil a design goal, not all good and not necessarily ideas with every permutation of the templates. This design goal can be based on a wicked problem or a desire to improve the user experience. The method prompts designers to look upon and consider every permutation, thus covering otherwise missed permutations in the ideation phase. As a result of this thorough process, 'unideal' permutations, either through possibility or suitability, will also be considered by designers. It is not necessary to come up with good ideas on such permutations as the end goal is to find all good/suitable ideas, not all possible ideas.

The method is adapted for use in multimodal interaction in that template d) is the primary template used, as it allows paring two or more modalities for ideation.

4.6 Theory of Change

IDEO's *Theory of Change* is a method sourced from the aforementioned IDEO Design Kit [70]. The method comprises of 5 steps:

1. Reviewing the list of desired outcomes, which were defined originally when framing the design challenge.

2. Next, using the *Theory of Change* worksheet provided by IDEO, each one of the issues and the corresponding ideas/concepts is listed out using a grid structure on the workspace.

3. From the grid generated, a priority is established regarding the most critical problems in the design space and the related ideas. This worksheet then serves a guide to redirect efforts towards the higher priority issues and their corre-

sponding solutions.

4. This process of establishing priorities incites articulation of theories, or rationales, for how said solutions will bring about change and achieve the desired outcomes (from Step 1). Interrogation of this emerging theory of change questions the logic, assumptions or risks of each solution.

5. Documenting the theory of change involves using the *Impact Ladder* in IDEO's activity guide to rapidly capture the output of this method in the first instance of its use. Future iteration could make use of the *Logic Model* activity instead to acquire a more detailed and clearly structured visualization of the solution model.

As a rule, using the method does not require one to follow all the steps, specifically the steps suggesting a follow-up method - instead designers can apply a different method to the output of the method.

4.7 Prototyping

Prototyping is a vital and effective way to give ideas a tangible form, to seek insights through crafting said prototypes, and to acquire important feedback from the target user groups. Prototypes are classified according to the level of their fidelity to match the level of the evaluation they're in and the maturity of the idea being represented.

Low-fidelity (Early Stage)
The low fidelity level of prototypes supports a rapid generation workflow for receiving quick feedback from evaluators, such as stakeholders or user groups, on the potential of early ideas/concepts. This also allows for test participants to focus on the critical aspects of the idea rather than on the aesthetics of the execution. [71]

High-fidelity (Later Stage)
The ideas and concepts which exhibit potential and receive positive feedback from prior tests are 'promoted' to a stage wherein they would be reconstructed at a higher fidelity level for further refinements and, eventually, final evaluations. These prototypes would also potentially be interactive enabling evaluation for user-interaction performance (e.g., time to complete a task), and in demonstrating actual solutions to clients, management, etc. (for the Company) and others. [72]

4.7.1 Scenarios

Scenarios serve as a foundation, a 'work context', that allows for prototypes of ideas to be defined, created and evaluated against. In this regard Nardi defined scenarios as:

> *"...a description of a set of users, a work context, and a set of tasks that users perform or want to perform. A scenario sketches future technolo-*

> *gies that will help users do the things they want to do. A scenario blends a carefully researched description of some set of real ongoing activities with an imaginative futuristic look at how technology could support those activities better.”* [73]

For the purposes of some theses, it is necessary to create deliberate scenarios such that they could accommodate and/or necessitate the application of ideas and concepts developed in a meaningful context. For example, although traditionally used in the evaluation of the developed artefacts or products, Scenarios could also be employed as aids to craft user-flows around the proposed ideas of this project when prototyping.

4.7.2 Storyboard

Although a vital part of human-centred design is communicating with the end-users themselves; however, in some unprecedented circumstances, this is potentially not an option. Therefore, to better understand and interpret the users' physical, cognitive, and emotional needs, in such situations, storyboards would be created around the aforementioned scenarios to serve the following needs as enumerated by McQuaid, Goel, and McManus:

1. Direct-experience story-boarding with narratives to understand customers' frustrating and pleasurable experiences with the artefacts in question.
2. Communicate the user experience enhancements to stakeholders and to serve as an aid in empathizing with end-users.
3. Potentially leverages the storyboards and narratives to further develop design recommendations. [74]

Computing has become progressively integrated into the environment and as a result has blurred the lines between the actual system and the context around, strongly necessitating the use of storyboards to depict context explicitly as implied by Truong, Hayes, and Abowd.

4.7.3 Wizard of Oz

First defined as a method by Kelley in the 1980s [76], the Wizard of Oz method, originally *the OZ Paradigm* [76, 77], it was initially only meant for use within the *Natural Language* field - something that is apparent from the definition given by Kelley:

> *“...experimental participants are given the impression that they are interacting with a program that understands English as well as another human would. In fact, at least in the earlier stages of development, the program is limping along, only partly implemented. The experimenter, acting as "Wizard", surreptitiously intercepts communications between participant and program, supplying answers and new inputs as needed.”* [76]

Since the 80s the method has been applied in many design fields, as what Kelley defines as 'intercepts communication' can be reinterpreted to mean any human interaction with an artefact. Instead of relying on messages manually behind a curtain, designers today manually supply, hidden or not hidden, features yet to exist to artefacts for a participant to experience. Extreme examples would be architectural house sketches and beta versions of a software feature. In the first, most if not all of the interaction is faked, the observer of the sketch can imagine how the house will look and then base their thoughts on that. In the other case, as little 'wizarding' as possible is used, and the interaction is almost fully real, only slight adjusting or resetting is done by the designers to keep the interaction working. As covered previously, the participant in a Wizard of Oz setting can be oblivious of the mock part of the interaction, but is not always so, as seen in the house sketch example. Of course, the more oblivious the wizard can make the participant the more realistic, reliable and representative data can be observed from the interaction.

4.7.4 Video prototyping

According to Karras et al., as well as the many sources cited by them in their 2017 paper, prototypes in video format do a better job at relying on the complex nature of designed interactions [78]. Karras et al., as well as Mackay and Fayard, suggest an approach where a written scenario, showcasing an intended interaction, is accompanied by a video showcasing the scenario, by using techniques such as analogue (paper, plastic films, etc.) or digital mock-ups interacted with in a Wizard of Oz setting [78, 79]. The use of Wizard of Oz in video prototyping was actually suggested already back in 1983 by Kelley in his original description of how to use the Wizard of Oz method.

Videos resulting from the use of the method can be seen and evaluated by participants knowing or unknowing of the 'fake' nature of the interaction - as it is primarily meant to reinforce and strengthen the relaying of the underlying idea; however, if done well, the interaction seen in the video can be indistinguishable from recorded interaction with the 'real thing'. Mackay and Fayard suggest that video prototypes done in this way be evaluated using the method *Walkthrough-based evaluation* - see Section 4.8.4.

4.8 Evaluation methods

The two primary qualities of interests in this thesis are comfort and safety, and any design created in this thesis needs to be, at some point, evaluated based on those criteria. One way to evaluate safety is to look at how big the subjective workload is on users, as the bigger it is the more distracted users will be - a common cause of vehicular accidents. To evaluate comfort, a common method is to utilise standardised usability scales, as they provide clear aspects to analyse for most if not all usable systems. For the first criteria, safety, this project will utilise the DALI, for the second, comfort, the system usability scale (SUS).

4.8.1 NASA TLX and DALI

The NASA TLX is a subjective workload assessment tool, developed in the 1980s at NASA's Ames Research Center (ARC), to evaluate pilots' cognitive and physical workload in aviation and spaceflight systems [80]. It enables its utilisers to perform subjective workload assessments on test participants working with various human-machine interface systems. Since its inception, however, it has been repurposed for use in computer science, healthcare and other complex socio-technical domains. [80]

A modified version of the NASA TLX has been proposed by Pauzié (2008) in order to adapt it to driving tasks, known as Driving Activity Load Index (DALI) [81]. The basic principle of the DALI is the same as the NASA TLX, using a scale rating procedure across six pre-defined factors, followed by a weighing procedure in order to combine the six individual scales into an overall score. The main difference lies in the choice of the main factors composing the workload score. [81]

4.8.2 System Usability Scale

The System Usability Scale (SUS) was first introduced by John Brooke in 1986, and its use cases are described by usability.gov (2021) as *"...it allows you to evaluate a wide variety of products and services, including hardware, software, mobile devices, websites and applications"* [82]. It is primarily used to differentiate usable and unusable systems from each other and is reliable even with smaller participant numbers. Participants in a SUS usability tests evaluates a system and scores it for each of the following statements [82]:

1. *I think that I would like to use this system frequently.*
2. *I found the system unnecessarily complex.*
3. *I thought the system was easy to use.*
4. *I think that I would need the support of a technical person to be able to use this system.*
5. *I found the various functions in this system were well integrated.*
6. *I thought there was too much inconsistency in this system.*
7. *I would imagine that most people would learn to use this system very quickly.*
8. *I found the system very cumbersome to use.*
9. *I felt very confident using the system.*
10. *I needed to learn a lot of things before I could get going with this system.*

The scoring is done using a Likert-scale [83] from 0-4, 0 being *Strongly Disagree*, 4 being *Strongly Agree*. The scores from the test are added up and multiplied by 2.5, making the possible scores go from 0-100 instead of 0-40. According to usability.gov and Sauro (2011), an average system would score 68 points, making anything below it worse than average usability wise. [82, 84] Sauro states that translating the numeric score, based on how a teacher grades on a normal distribution curve, into a more readable A-F grading, better conveys the percentile meaning of the score, see Figure 4.1. For example, a score of 70 should not be misunderstood as better than 70% of all interfaces, but instead only better than 50%. Sauro provides a rough guide to how to do this translation: A for the top 10% of scores (>80.3), B for scores over 68, C for scores over 60, and F for scores lower than 51. Sauro also adds that in the top ten percentile, i.e. a grade of A and over 80.3 score-wise, is where the interfaces or systems which users most often tend to recommend to others are located. [84]

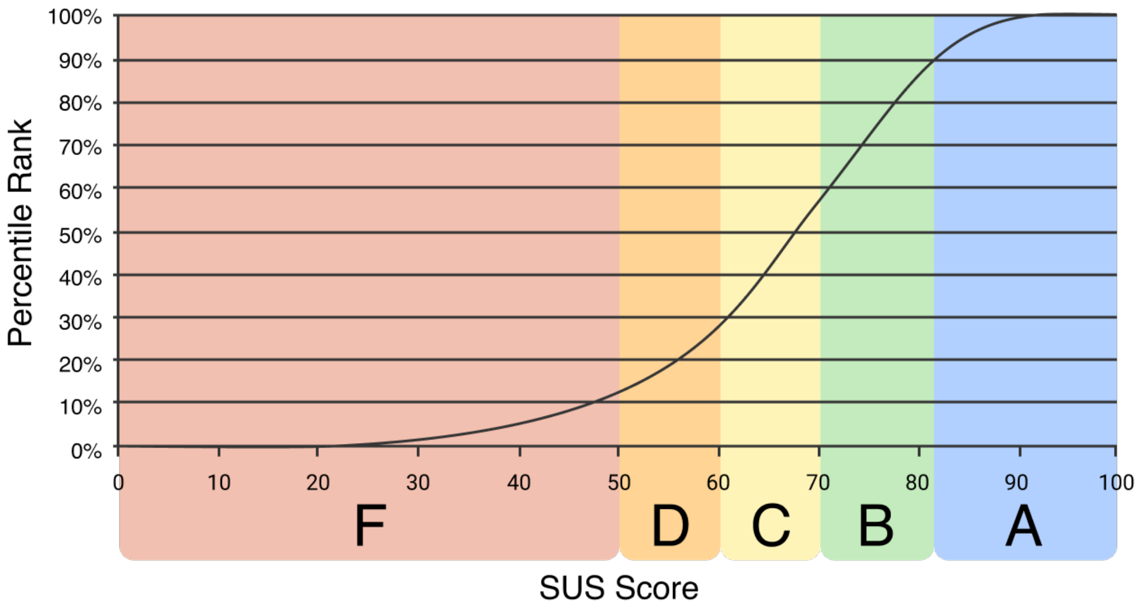

Figure 4.1: How one might translate a SUS score into the more representative A-F grading scale, based on Sauro's parable for grading UI's on a curve. [84]

4.8.3 t-tests

To detect any statistically significant patterns in collections of participant data, two types of t-tests can be conducted, the types being one and two-tailed - refereeing to if the test covers one or both tails of a normal distribution curve. One-tailed ones are better used to detect statistically relevant patterns as it, for a given significance alpha, covers one big region compared to two-tailed ones, that look at two, smaller sized regions instead to detect inverse relationships. With inverse relationships is meant that while a test might be conducted to see if proposition P is true, one-tailed might detect it easier, but two-tailed might detect that actually its the inverse of P that is probable - one-tailed tests do not take into account the difference between the proposition and its inverse.

For each t-test, there are two hypotheses, H_0 or *null hypothesis* which is usually *'the two population means are equal'*, and H_1 *'the two population means are not equal'*, i.e. there is a difference between them. t-tests generate a p-value, the lower the value is the stronger the statistical significance is for H_1 to be true and for the null hypothesis to be rejected. Conventionally, p-values equal to or lower than .05 are considered low enough to indicate statistical significance. [85]

When utilizing t-tests and stating that a p-value or alpha, lower than .05 is indicative of strong statistical significance, certain assumptions are made. These include that the data is normally distributed, that the population (user group) is randomly sampled, as well that the sample group is adequately sized. To test whether the sample size is big enough, one could perform a *Power Analysis*[86] on the data. Power analysis takes into account the means of the different populations, number of subjects, standard deviation as well as the p-value, or alpha. The analysis outputs a measure of power, beta, from 0 to 1 - indicating the probability of *"...recovering a statistically significant result"* [86] from the p-value. Conventionally, a beta of at least 0.8 (often written as 80%) is desired. [86]

It is important to decide upon a significance threshold before calculating the actual value(s), as otherwise, authors might introduce what is commonly referred to as *"..error of first kind."* [87], where an author believes their result to be significant - no matter if it is or not.

4.8.4 Walkthrough-based evaluation

In this type of evaluation, participants are walked through a designed artefact, as compared to free exploration or performing a task by interacting with it. The first alternative can be too unreliable if the designers want the entire design to be seen by the participant, as free exploration might lead to missing certain parts of it. The second, requiring interaction with the design, is sometimes not an option, such as when the design's fidelity is too low to interact with or when external factors such as physical access, legal regulations, etc. prohibit interaction.

To allow a sense of possible interactions, however, designers utilise certain mediums and strategies when creating observable artefacts. For example, while digital prototyping supports creating interactive prototypes with responsive controls mapped to mockups, the process of capturing the events of these controls to obtain an understanding of the interaction is crucial. Videos for example, which demonstrate interaction sequences, can serve as additional support for the pre-crafted scenarios and aforementioned digital prototypes. Mackay and Fayard used videos in their design process, which were evaluated using walkthrough-based evaluations, to demonstrate scenarios of using the system under development. Different variants of scenarios can emerge by modifying the captured event sequences and mockups [78], thus showing different interaction paths. According to Karras et al.'s evaluation, such a generated video allows for a faster understanding of an artefact's possible interactions when compared to static mockups - making the walkthrough, and the attached evaluation, more effective and reliable.

4.8.5 Task-based evaluation

Task-based evaluation, also called *Task based usability testing*, have many benefits, McCloskey (2014) of the Nielsen Norman Group states that:

> *"The most effective way of understanding what works and what doesn't in an interface is to watch people use it.* [...] *you gain qualitative insights into what is causing users to have trouble. These insights help you determine how to improve the design.* [...] *In order to observe participants you need to give them something to do. These assignments are frequently referred to as tasks."* [88]

McCloskey suggest creating realistic tasks and scenarios for the participants to act out, as a more convincing environment and activity allows the participants to better emerge themselves in the fictional (or actual, in case it is an existing product) reality they are pretending to take part in - resulting in more reliable and representative data.

The second note provided, is to make the task actionable, meaning to give them an action to do, and not to ask them how to do an action, e.g. *'Do X'*, not *'How would you do X'* or *'How did you do X'*. This suggestion stems from the fact that with 'How' questions participants tend to reply with words and not with observable actions - thus providing less natural interaction data. If a participant starts explaining what they would like to do, it can be a sign that the design is not actionable enough, they have nothing to do so they say what they would like to do instead.

The last guideline provided is to avoid, as much as possible, to give clues on what to do - task should allow for freedom of interaction choice while at the same time not be too vague in what to do. As an example, the task could be to *'Make an appointment for next Tuesday at 10 am with your dentist, Dr Petersen.'* instead of *'Make an appointment with your dentist.'*, which is too vague, or *'Go to the website, sign in, and tell me where you would click to make an appointment.'* which clues the participant too much into pre-selected interactions. [88]

Task-based evaluation necessitates an artefact that participants can interact with, or at least pretend to interact with, such as a paper prototype, or an existing product. Simply showing a design to a participant, such as when using the *Walkthrough-based evaluation* described in Section 4.8.4, allows for no interaction at all, making it impossible for a participant to perform a task the way suggested by McCloskey.

4.9 Project structure

This section of the methodology covers primarily the background of, and alternatives to, a design process defined by *d.school*. It follows it up by explaining in short the communication and documentation tools utilised in this thesis.

4.9.1 Five-step process

A popular design process structure is to follow the five stages, or a derivative of them, defined by *The Hasso Plattner Institute of Design at Stanford (aka the d.school)*, as cited by Interaction Design Foundation (2021); *Empathize*, *Define*, *Ideate*, *Prototype* and *Test* [89]. The structure has similarities with other alternatives, such as the related one found at the IDEO Human-Centered Design Toolkit (2021), which is based on the principles of Design Thinking [57]. Both of these might have been inspired by the double diamond first introduced by the Design Council (2004), that consists of: *Discover*, *Define*, *Develop* and *Deliver* [90].

In the double diamond, *Discover* is the acquisition of knowledge and definition of the problem. While its *Define* phase is the convergence to *Discover*'s divergent thinking, refracting and redefining the problem into a challenge. This first half of the double diamond can be seen as the *Empathize* and *Define* phases defined by d.school, first, a gathering of knowledge(usually about the users and their needs) followed by an establishment of the design space (the specifics about the users' needs and problems). The third part and second divergent part of the double diamond is the *Develop* phase, coming up with answers to the design challenge [89]. This, in d.school's stages, is represented first by *Ideate*, the creation of ideas, and then by *Prototype*, creation of solutions to the problems. As the fidelity increases and *Prototype* bleeds into the last, and final convergent part of the double diamond: *Deliver* - where solutions are tested out, rejected or improved, by d.school this phase is called *Test* [89, 90]. It is important to note, that when using an agile process, with several iterations, the phases are performed several times, with each Prototype phase followed by a Test phase. Depending on how one sees it, evaluating ideas in the Ideate phase could be seen as another occurrence of the Test phase, albeit at a different scale and thoroughness due to the difference in the quantity of the ideas in a project compared to the prototypes in it.

In the first phases, Empathize and Define, it is common to primarily involve the end-user stakeholder, in the research where the result is primarily academic and secondary an artefact, the end-users is the academic(i.e. **Design Professionals**) and industrial stakeholders(i.e. **The Company**). In that case, the needs of the end-users can be researched through related work written by **Design Professionals** as well as through industry expertise about such artefacts. As understanding is built, a clearer picture can be constructed about the needs and problems these stakeholders face.

For the following phase, Ideation, if the thesis is done at a company, it is more suitable to involve **The Company** stakeholder, as that stakeholder's know-how of the field enables suitability checks of ideas before going into the following phase. In Prototype and Test, understanding of design and evaluation methods drawn from **Design Professionals** is first used, followed by utilizing technology form **The Company** in case a final, non-prototype, artefact is realized.

4.9.2 Meetings and channels of communication

The main communication channel for this project was Skype [91], for its ease of use and support for cross-platform messaging capability. Teams [92] was used primarily for written communication with the company officials as it was the preferred business communication channel internally. Zoom [93] was also used occasionally for remote work sessions, certain meetings and distance-based participant sessions. Canvas [94], the LMS (Learning Management System) was, at the time of writing, the official channel for thesis-work submissions and for seeking authorizations with the University.

4.9.3 Documentation

Day-to-day documentation throughout the project was primarily done using Notion.so's student version of its online documentation suite [95]. For conventional documents such as presentations, pdfs, questionnaires and pure text documents Google Drive's Presentations, Forms and Docs was used due to the low latency in cooperative work as well due to the authors' preexisting knowledge of how to use it. For the thesis' planning and final reports, Overleaf's online LaTeX editor [96] was used.

4.9.4 Research diary

Gray-Grant (2018) describes why one should keep a research diary and what should go into it:

> *"Your diary should include a description of what you've done each day—the people you met and what they said, books or papers that you read, lectures or conferences that you attended, notes from discussions or conversations, and ideas you want to remember to follow up. But, even more important, it should also include your personal views and opinions of all that you have learned and your analysis of any problems you've noticed. Questions, hunches, thoughts and plans for future actions also belong in this free-flowing document. Finally, make a diary entry even if you do nothing else that day. It will help keep you connected with your project and maintain your motivation." [97]*

The arguments discussed by Gray-Grant are partly quoted from a video by Gray (2014), bringing up four points as mentioned by Gray [97, 98]:

- *It helps you keep a detailed history of your research as it unfolds*
- *It provides a reference point for when your thoughts changed/matured during the process*
- *It allows you to trace the development of your research skills*
- *It gives you a place to reflect on your research*

Gray-Grant adds two additional points [97]:

- *A research diary will require you to declare your opinions about what you are reading*

- *A research diary will allow you to maintain your writing habit*

In this project several separate channels were used to chronicle project activity, primarily to separate meeting notes from summaries of general progress and motivations: meeting notes for each meeting as well as a weekly report. These reports were utilised as a reminder when writing chapter 5: **Planning**, 6: **Execution | Process**, and 8: **Discussion**.

5
Planning

In any project, it is common to have a preconceived plan of action, based on preliminary information, preparation and experience from prior project work. Since the main purpose of a research project is to acquire new knowledge along the way, the plans and proposals presented in this chapter are subject to change, withdrawal and modifications as newer and better ways of design, research, development and evaluation are understood and/or discovered.

The following chapter presents, in chronological order, the research questions, details the overall project plan based on the research questions, which in this thesis is the lion's share of the hypotheses. More precisely, it introduces the cardinal phases of the project including the evaluation plans and concludes with the planned contributions of the project. The final decisions with regards to the plans in this chapter, motivations, comments and detailed step by step sub-results are instead covered in Chapter 6.

5.1 Research questions (RQ1 and RQ2)

With the theoretical and methodological background explained in the previous chapters in mind, the authors observe that examples providing answers to the following research questions are needed in the field, and thus are what the thesis aims to answer. The first part of the research problem being a wicked problem necessitated a rephrasing of the terms *best use* and *most UX-improving* into the less absolute terms *better use* and *improved user experience*.

RQ1: How can designers make better use of vehicle sensor data to create improved vehicular user experiences?

RQ2: How can designers estimate the value, resulting from UX designs relying on vehicle sensor data, added to the vehicular user experience?

5.2 Overall plan

The following section covers the overall plan for answering the research questions (**RQ1** and **RQ2**), which by being *How*-questions means their respective hypotheses should be suggested *How-to* instructions. The section first covers hypotheses one -

a suggested design process (**H1**), hypothesis two - a suggested evaluation strategy (**H2**), together they make up the thesis' overall design project plan, used in tandem for the purpose stated in the hypotheses: creating good vehicular user experiences using vehicle sensor data (**RQ1**), as well as evaluate their UX value contribution (**RQ2**). After presenting the hypotheses, the section explains how this overall plan aims to achieve the planned contributions introduced in Section 1.3: **Aim and Planned contributions**.

It is important to note when reading the hypotheses, that domain-specific knowledge related to the research questions, covered primarily in Chapter 3, is utilised to derive them. For example, the term multimodal interaction design is not mentioned in the research questions themselves, but is mentioned in the hypotheses due to it being part of what makes up automotive interactions design, which in turn is what design the *vehicular user experience* mentioned in the research questions.

The chapter after this, 6: **Execution | Process**, covers the process of following this suggested design project plan, as an attempt to validate whether the hypotheses, the How-to suggestions, are valid or not. For the more practical project plan, i.e. what should be done and when, the initially suggested timeplan for the thesis can be found in Appendix C.

5.2.1 Hypothesis 1: the suggested design process (H1)

This subsection covers the authors' hypothesised answer to the first research question (**RQ1**), which by being an answer to a How-question is a *How-to*, i.e. a suggested plan of action, to achieve designs that uses fused vehicle sensor data to add value to vehicular user experiences.

A design project, one that aims at making *...better use of vehicle sensor data to create improved vehicular user experiences* - to quote the first research question (**RQ1**), should aim to follow a conventional design approach of first acquiring an understanding of the field, then, using an iterative process, produce and evaluate one or several artefacts - successively compiling further knowledge of both how to follow this suggested plan, i.e. the hypothesised answers to research question one (**RQ1**), as well as to what degree that suggestions were valid, i.e. how successful they were. This last instruction serves to possibly further validate the hypothesis as more and more projects have utilised it.

d.school's five stages
To the intended effect, the popular design process structure called *Five Stage Process*, defined originally by *The Hasso Plattner Institute of Design at Stanford (aka the d.school)*, is the suggested base structure to follow. Another process structure that could also be considered, for it's similarity in approach, is the double diamond structure, first introduced by Design Council (2004), that consists of: *Discover*, *Define*, *Develop* and *Deliver*. As cited by Interaction Design Foundation (2021); the *Five Stage Process* structure unfolds in the following stages, namely : *Empathize*,

Define, *Ideate*, *Prototype* and *Test* [89]. The following paragraphs cover the suggested adaptation of the *Five Stage Process* with regards to the project:

EMPATHIZE
The first stage of the process should be to gain an empathetic understanding of the design space set by the research questions. In a thesis, the first stage should involve engaging in literature studies and interviewing domain experts from the company to obtain a clear view of the area of concern and design space. This can potentially aid in understanding the merits of prior work, user experiences and stakeholder motivations in the automotive industry.

DEFINE
During the Define stage, the plan should be to collate and structure the information derived from the literature studies and expert interviews gathered during the Empathize stage. This is where an analysis of observations should take place, if possible, in order to define a design space with respect to the available modalities and core issues targeted by the research questions up to that point.

IDEATE
With a foundational set of literature in mind, along with insights from the domain of AID, the identification of new solutions can begin. These shall be based on the varied combination of the previously identified modalities, potentially leading up to solutions leveraging multimodal interactions to address the research questions. It will be vital to develop as many ideas or solutions as possible at the beginning of this stage to ensure that all probable ideas are available for filtering out the rest towards the end and to find the 'best', i.e. the most suitable, ones.

PROTOTYPE
This stage begins with a number of cost-effective, lower-fidelity prototypes of the artefact(s), developed specifically to investigate the feasibility of the few selected multimodal solutions ideated in the previous stage. The aim of this stage would be to attempt to develop the most suitable solution, that addresses the problems identified during the first three stages. While being an experimental and non-permanent phase, it should enable the authors to gain more permanent expertise in HMI development.

TEST
Herein the designs, that are implemented within the prototypes, are investigated and either approved, improved and re-examined or rejected on the basis of the feedback from evaluations with, both or either, experts and representatives of the target group, i.e. vehicular occupants (see Section 1.1), as participants. This is the final stage of the five stage-model, but as elaborated further later on in this chapter, the plan is to enter an iterative process loop, wherein the outcome, or results, generated during the testing phase would be used to analyse the findings and reinterpret the designed artefact.

Iterative process

Working in an iterative fashion, starting from Ideate, will have the following layout: first plan the iteration, then execute the plan, and lastly to analyse any resultant output - turning it into the outcome of the iteration, which is brought into the next iteration. The output will take different forms depending on what was executed, for example: in the case of phases including an evaluation of a prototype, it might take the form of any collected feedback, task load and usability scores. Analysing and discussing these outputs will lead to new ideas, alteration plans or reaffirmation of the current course of action. In cases where the phase is extensive, in either time or work performed, the mid-phase output might be required to, in detail, plan the rest of the phase. Examples of this include the prototyping and test iterations later in the project, where, depending on what transpired during the prototyping phase, the plan for the evaluation will have different possibilities and limitations. For a more holistic perspective, or overview, of both the previously described stages, but also this iterative process see Figure 5.1 for a graphic representation.

Figure 5.1: A graphical overview of the first hypothesis - the suggested design process (**H1**).

5.2.2 Hypothesis 2: the suggested evaluation strategy (H2)

This subsection covers the authors' hypothesised answer to the second research question (**RQ2**), which by being an answer to a How-question is a *How-to*, i.e. a suggested plan of action, to be able to estimate any added value of designs that full fill R1's requirements.

Throughout a design project evaluations of the designs should be held, mainly due to two reasons; the first being to be able to separate suitable ideas from unsuitable ones, i.e. which to keep and which to discard, the other to estimate the value added to the vehicular user experience - as stated in research question two (**RQ2**). To achieve a balance between time spend and value-added, the methods employed should, as the number of designs decreases, go from rapid and quantitative, towards slow and qualitative. In the initial phases, the method *Expert Interview*, or an alternative, should be employed as a way to utilise the expertise of any company stakeholder or recruited expert. This should be done as an aid in defining the design space, in the define phase, for the project. Their know-how could, for example, help steer away from less promising avenues not already known by the designers. Later, in the Ideate phase, the vast amounts of created ideas should be rapidly graded by the similar experts using a questionnaire version of the method - enabling early and fast exploration of a large design space.

Following the design of all iterations' prototypes, SUS scores should be calculated due to the low cost and fidelity requirements of doing such as compared to the benefit. The scores calculated shall then act as a guideline both when selecting ideas for continuation as well as what aspects of the idea needs addressing in the next iteration. Calculating a SUS score for each iteration also adds the possibility to compare how the score changes with alterations and with increases of fidelity across the iterations - 'Was the improvement effective?', 'Is the idea still good?', etc., are examples of questions that can easier be answered with the help of these SUS scores.

For the last iteration's evaluation, done upon the final artefact(s), to properly verify whether the used methodology was valid or not, i.e. whether the created artefact(s) is 'good' or not, a *Task based evaluation* will need to be held - this to allow the calculation of a DALI score for the final artefact(s). A DALI score will enable verification whether the resultant artefact(s) are perceived and experienced as intended, to quote R1: *"How can designers make* ***better*** *use of vehicle sensor data to create* ***improved*** *vehicular user experiences?"* After raw numerical data have been collected, deeper analysis might be possible, depending on size of participant group, whether demographic data was measured or it is of interest to the design project. Such analysis can be made using statistical analytical tools such as *t-test* and *Power Analysis*. For a holistic overview of H2, see Figure 5.2.

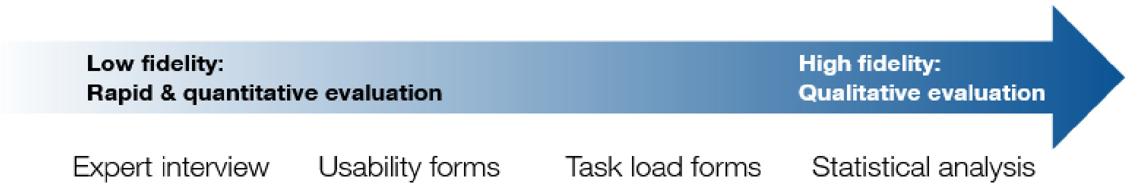

Figure 5.2: A graphical overview of the second hypothesis - the suggested evaluation strategy (**H2**).

5.2.3 Contributions

The following paragraphs go into detail how the project plan presented in this chapter aims to achieve the possible contributions described in Section 1.3: **Aim and Planned contributions**

For the company
To achieve the aim of creating 'evaluated ideas and concepts of how to incorporate an ADAS platform into an Android context', as covered in the introduction of the thesis, the planned process will result in one or several final artefacts. They will go from simple ideas in the Ideate phase, into successively higher fidelities - each version evaluated and analysed for possible improvements. The final iteration will aim to produce Android Automotive applications that in some way incorporate the company's proprietary ADAS platform.

Prototypes from previous iterations, including discarded ideas and earlier versions of the final artefact(s), will also serve as contributions - either as examples of how to do or not do, but also act as inspiration for future projects.

The expertise gained with any tools used could also be seen as contributing to the aim of 'how to' develop concepts such as the final and other prototypes - which is also of use to the company that might transition into using it themselves for future work.

For academia
The thesis aims to academically contribute by creating two hypotheses, i.e. answers, to the research questions. For the first one that is the process, and artefact, suggested and created in this and the following chapter, for the second research question that is the suggested evaluation plan introduced earlier in this chapter.

Chapter 6: **Execution | Process** covers by example, how one might follow and combine the hypotheses (**H1** and **H2**) as well as how one could motivate the use of, and how to use, the methodology, methods and strategies incorporated throughout the process.

Throughout the process, the evaluations made will not only act as an answer to the second research question (**RQ2**), i.e. the evaluation strategy (**H2**), but also verify the validity of the hypothesis to the first research question, i.e. the design process (**H1**). Post-process analysis, made after each iteration as well as after the entire process, of evaluations made, will aim to verify the validity of the evaluation strategy (**H2**).

6

Execution | Process

The following chapter covers, in chronological order, the execution of the proposed plan, see Chapter 5: **Planning**. It covers d.school's five stages in the following phases: *Empathize & Define*, *Ideate* and three iteration loops, each covering *Prototype* followed by *Test*. For each phase there were two major stages, first the planning of the phase - this included reacting to the outcome of the previous phase as well as method choices, thus including motivations behind the later stage: the execute stage. As a result of this, the sections in this chapter, that each cover a process phase, are split into the parts **Planning of phase** and **Execution of phase**, see Figure 6.1, the first consisting of preparative work such as planning and strategizing, the second, of both or either generative work (designing, constructing, etc.) and evaluative work.

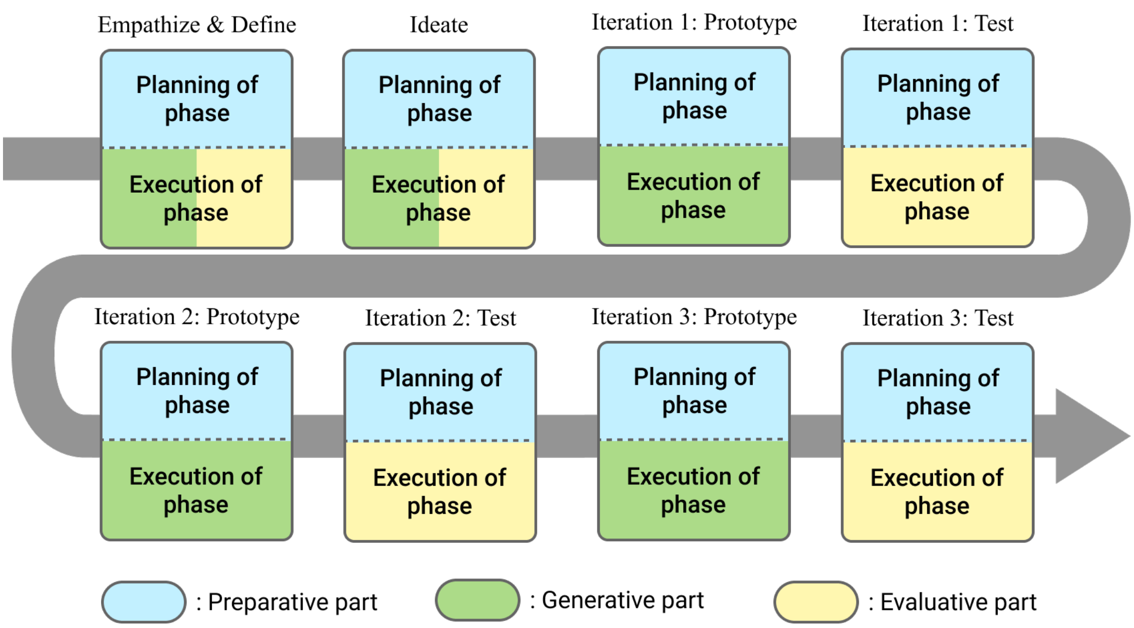

Figure 6.1: A graphical overview of the different phases of the process, and their internal steps.

6.1 Empathize and Define

The initial phase of the process consisted of planning and executing the first steps necessary to initiate the project. This involved reading up on literature, first in an unstructured than structured manner, interviewing field experts, and finally defining a concrete design space suitable in regards to the research questions (**RQ1** and **RQ2**). The result from the phase was such a design space, in the form of a set of modality combinations not found (by the authors) in either academia or the industry.

The findings for this phase is covered under Chapter 2: **Background** and Chapter 3: **Related work**, and the process more in detail under Section 4.3: **Literature study**.

Planning of phase

The overall plan for the empathize & define phase was decided to be; initially - reading papers to get a better understanding of the field, followed by defining the design space based on the overlap of the expertise and interest of the company Aptiv, as well as parts of the academic field identified as lacking or not covered by related work. The interest of the company was partly formed of what no other company, at the time, was providing, i.e. concepts not covered by the industry.

A clearly defined design space was desired, as it would allow for a structured approach to the following phase, Ideate. To properly define such a space, terminology and background understanding would need to be acquired. Therefore, a study of academic works, i.e. a literature study, should be conducted. To achieve this, a meeting should be scheduled and held with staff from the Chalmers Tekniska Högskola - both as a way to better understand what academic resources were available, and also to acquire expert recommendations on methodology for further literature study. From the notes taken during such a meeting, a more effective strategy could be formulated - resulting in a larger and more relevant collection of related works to take into account in the project.

To properly harness their expertise and to understand the interest of the company, IDEO's Inspiration method *Expert Interview* was planned to be utilised, preferably with representatives from different departments of the company and thereby different technical fields. The method could possibly be conducted in a less formal setting than what is conventional with the method, partly due to the fact that the experts being recruited, would be from the company connected to the thesis, and partly due to the difficulty of preparing specific questions for fields where the authors had little to no knowledge.

Execution of phase

Departing from the plan, informal interviews with industry experts from the company was conducted to gather enough of an understanding to start looking for papers. Following this, an unstructured approach was taken and sporadic, well-cited literature was collected and read to further the base knowledge gained during the interviews, as well as to get the academic perspective of the discussed fields. With this, fairly solid, footing in relevant fields for the thesis, a more structured literature study was initialized.

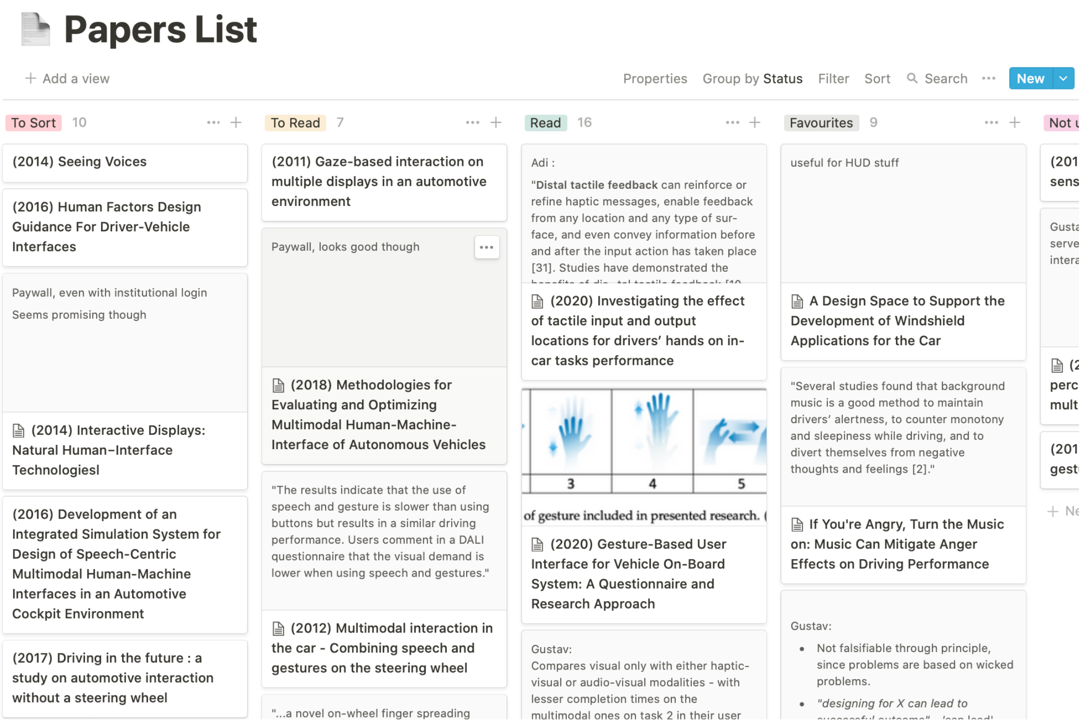

Figure 6.2: A screenshot of the authors use of the online work space organizational tool Notion, papers and other academic works being organised into three main columns: To sort, To read and Read.

6.1.1 Literature study

The authors held an extensive meeting with staff from the Chalmers Tekniska Högskola, and from the meeting, a step by step process was formulated, see Section 4.3: **Literature Study**, and then executed - the outcome being a large collection of papers, both well-cited, foundational works, and more recent and relevant papers. This collection of papers were organised using the online tool Notion, see Figure 6.2. The use of tables on Notion allowed the authors to quickly insert, sort and tag collected material, allowing for the rapid aggregation of sources as well as lending an overview of any comments made on them. Notion's filter feature allowed the authors to effortlessly look for papers in the collection covering a certain topic, e.g. filtering on 'gaze' would only display the papers mentioning 'gaze'.

After collecting and reading more academic literature, specific modalities were identified or defined, such as *Gaze input*, *Voice input* and *Haptic output* - most of these were directly extracted from papers. In the multimodal interaction field, two or more modalities are combined in a design, therefore to find combinations lacking in the academic field, as initially evidenced from the collected literature and later verified through external source libraries, a combinatory grid was constructed. The grid was filled in with references to examples of an academic source implementing and evaluating said combination. When examples were not found in the collection (which was the fastest way to search), the external sources Scopus and ACM was utilised. This resulted in a partially filled-in grid, wherein the empty sections represent what design areas possibly have yet to be covered by academia or the industry. This process was conducted to clearly identify an area where the thesis has the highest potential to contribute to the field. This area was chosen to represent the outer reaches of the design space for the project.

6.1.2 Expert Interview

Following the literature study, the method *Expert Interview* was applied to further narrow down the design space - based on: what parts of the initial design space they could provide insight and expertise on, what parts they were interested in, and what parts they thought there's a demand and/or potential in. As planned, the interview took a less formal and structured approach than recommended by IDEO, due to the reasons mentioned previously under Chapter 5 **Planning**.

The planned approach was conducted without issue and resulted in a well-defined design space, see white sections of Table 6.1. Most of the modalities within were extracted from academic sources, others were coined by the authors in case no pre-existing academic work had already coined a term.

The rows **Body Tracking** and **Occupant profile** were constructed and their names coined by the authors, and each represents a collection of modalities connected to the proprietary ADAS platform utilised in the thesis. The term occupant referred to one of the stakeholders, the vehicular occupants, which includes both drivers and passengers of vehicles.

The columns *Visual Source manipulation* and *Audio Source manipulation* were also coined by the authors - as a way to differentiate visual/auditory output/feedback that purely relays information as traditionally intended or the second and arguably more complex one: the cases were the distortion or shift, of whatsoever info is currently being relayed in the visual/auditory feed, is the output/feedback. Two examples would be the act of lowering music volume or adjusting the brightness of the screen as an output modality - the content played/display is not affected per se, merely *manipulated* in its presentation.

Input \ Output	Visual Source	Audio Source	*Visual Source manipulation*	*Audio Source manipulation*	Haptic
Touch	YES	YES			YES
Voice	YES	YES			YES
Gesture	YES	YES			YES
Gaze	YES	YES			YES (VR)
Body Tracking					
Occupant profile					

Input \ Input	Touch	Voice	Gesture	Gaze	**Body Tracking**
Occupant profile					
Body Tracking					
Gaze	YES	YES	YES		
Gesture	YES	YES			
Voice	YES				

Output \ Output	Visual Source	Audio Source	*Visual Source manipulation*	*Audio Source manipulation*
Haptic	YES	YES		
Audio Source manipulation				
Visual Source manipulation				
Audio Source	YES			

Table 6.1: The initial table of input and output modality combinations after being filled in. White sections indicate combinations not found in either academia or the industry.

The input modality groups **Body Tracking** and **Occupant profile** were coined at a stage where the authors did not yet know fully know what exact modalities were available for use through the utilised ADAS platform. As this was made apparent, the authors fully expanded the groups into the following list of modalities:

Body Tracking

- Body proximity to parts of interior
- Finger pointing vector
- Predictive touch

Occupant profile

- Distraction level
- Emotion/Mood
- Activity estimation

Overlap **Body Tracking** & **Occupant profile**

- Identification - who is where

Other
- Object classification
- Left-behind object
- Cabin snapshot (video call)

Most of these new and (at least to the extent of the authors' knowledge) untapped modalities had a common component: as input modalities, they require little or no user intent/focus to use - a very useful detail in a setting where user distraction/frustration directly results in fatalities. Information about the vehicle occupants is, by the ADAS platform's sensors and sensor fusion algorithms, automatically(from the user's viewpoint) entered into the IVI system similarly to how a button press or voice commands is.

6.2 Ideate

With the definition of a concrete design space, ideation could commence. The choice of ideation method, *ARSCI*, was selected based on the difficulty of combining unfamiliar technologies and interactions, and rapid, but effective, evaluation methods, *Expert Interview* and *Theory of Change*, were utilised to comb through a vast selection of bad and good ideas in a short time frame. The planned process proved practical, as an unforeseen reiteration needed to be performed as new information about the design space emerged. The phase resulted in 37 ideas formulated form 275 possible combinations, culled down to a final set of six ideas. These six ideas were the ones brought with into the next phase of the project, prototyping.

Planning of phase
Multimodal interactions combine two or more modalities, some of these combinations are more challenging to discover as combinations are essentially based on the designers inherent associations, this can cause ideas based on un-thought of combinations to never emerge or be considered. Designers in the field should be mindful of this human fault in the ideation process. To address this potential weakness, the plan to utilise and adapt the structured creativity ARSCI method was made.

The plan was to, after generating a set of ideas, using ARSCI, a modified version of IDEO's *Expert Interview* would be used to further cull these ideas. The culling should be done looking at the desirability, feasibility and viability of ideas - all based on expert know-how recruited from the company. To remain cost-effective (time wise) this could be achieved by doing it through a streamlined questionnaire with pointed multiple-choice questions on each idea. The aim of this method choice was to reduce the design options as well as improve the average quality and suitability of the possible avenues for the rest of the thesis, i.e. what ideas remained after the culling.

Theory of Change [70] would be employed as a post-ideation method to refine concepts that emerge from the core ideation sessions. Specific concepts, albeit well thought-out, lack a clear motivation as to which issues within the design they address and how they do so. As such, to craft a clear mapping between ideas/concepts and corresponding issues within the design space, this method would be put into practice. Theory of Change is a method sourced from the aforementioned IDEO Design Kit[57].

Execution of phase
In line with the initial plan for Ideation, using the *ARSCI method*, a vast selection of ideas were generated - touching upon most valid combinations of modalities deemed to be inside the design space created in the *Define* phase. This was followed by the culling of ideas using a questionnaire version of *Expert Interview*. However, after following the plan and receiving feedback from the company stakeholder, a new set of modalities were introduced. The rapid nature of the chosen ideation methodology allowed the authors to reiterate both the ARSCI method as well as the questionnaire. After the unplanned for reiteration, the planned methodology could continue, with the method *Theory of Change* culling the ideas further - leaving a selection of six ideas to bring into the next design phase - *Prototyping*.

6.2.1 ARSCI

The *ARSCI* method (*Adding, Removing, Skewing, Combining, Imbalance* - see Section 4.5) was applied as intended onto the possible modality combinations - resulting in 275 possible combinations. From these combinations, specific ideas (some combinations yielding no apparent ideas) were documented - also through Notion. These ideas of varying quality and suitability were subjectively culled by the authors based on knowledge gained so far in the project, though literature and interviews, resulting in a final set of 28 ideas.

However, ideas outside the design space, but still believed to be intriguing, were kept and later readdressed. Most of the ideas were created without a method and written down as they spontaneously emerged. This was an impromptu decision made by the authors, as the goal of the Ideation phase was not to only create ideas with a method, rather, the method was used to aid in idea creation - it was thus seen as logical that any ideas were valid, no matter how they were created.

For the 'good', but outside-the-design-space ideas, the *Adapt* segment from SCAMPER was later utilised to transform and transfer said ideas and concepts from outside to the inside of the design space. This was done to ensure that interesting concepts weren't discarded solely because they existed outside the design space when they originated.

6.2.2 Expert Interview - questionnaire

The ideas generated and retained so far, were entered into a questionnaire with the two simple, follow-up questions: *Do you like the idea?*, and: *If the idea cannot be*

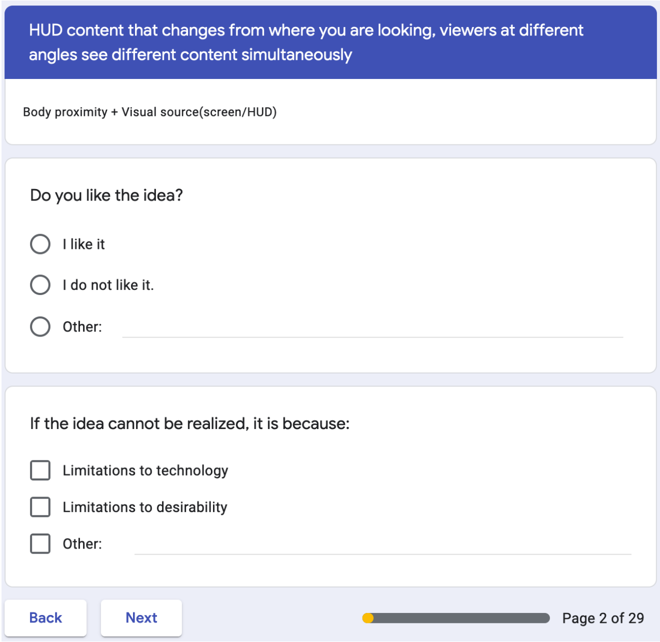

Figure 6.3: An example section taken from the *Expert Interview* (questionnaire), showcasing question setup and possible answers.

realized, it is because:, as can be seen in Figure 6.3. The questions were kept to a minimum to streamline the process of utilizing expert know-how on as many ideas as possible. The 'Other:' options allowed the respondents to enter specific opinions if any on the ideas especially interesting to them. The questionnaire was constructed using *Google Forms* due to the authors' familiarity with the tool.

Responses from experts were collected and analysed using *Google Sheets*, initially auto-generated from the questionnaire. To assist in the analysis process, coloured formatting was applied, see fig. 6.4. Ideas that received no 'I do not like it.' replies were selected for further analysis, in which the second question's answers and, if any, the text replies under 'Other:' was looked at more in detail. This process was selected as a way to rapidly comb through and narrow down the still relatively large number of possible avenues for the design process. Based on limitations revealed and extra comments, if any, ideas suitable for further development were selected and extracted into a separate collection for use in the next phase of the project.

Timestamp	Do you like the idea?	If the idea cannot be realized, it is because:	Do you like the idea?	If the idea cannot be realized, it is because:	Do you like the idea?	If the idea cannot be realized, it is because:
2/8/2021 16:48:03	I like it	Limitations to technology	I do not like it.	Limitations to desirability	I'm not sure I understand. So it is used to "teach" the driver a good posture? Then it's not disirable. The HUD projection must at all time adjust to driver posture, angle and gaze.	Limitations to desirability
2/9/2021 15:47:00	I like it		I like it	Limitations to technology	I like it	

Figure 6.4: A screenshot taken from the colour coded replies to the *Expert Interview* questionnaire.

6.2.3 ARSCI: revisited

Following the analysis of the expert responses to the first questionnaire, feedback from the company advised the authors to also include any possible modalities born from the exterior sensor systems as well. Another *Expert Interview* was set up with a representative for that industry, similar to how it had been done at the start of the project, see Section 6.1.2.

Based on the interview, a new set of unexplored modalities were identified and coined:

- 360° Awareness (Fused Modality)
- Camera (Black/White or Limited Colour)
- Path and Heading Estimation
- Blind-spot Detection / Crossing info

Again the ARSCI method was applied, using the previous modalities as well, generating a final set of 9 ideas.

6.2.4 Expert Interview - questionnaire: revisited

Similarly to how it had been done before, the ideas from the latest iteration of ARSCI was entered into a *Expert Interview* questionnaire.

6.2.5 Theory of Change

As intended the method *Theory of Change* was applied upon the final ideas as finally culled using the two *Expert Interview* questionnaires. In fig. 6.5 one can observe how the execution of the method was prepared according to the recommended worksheet.

The ideas were grouped under umbrella terms according to their constituent modalities, then individually colour coded based on the recommended legend: green for *Addresses this well*, and blue for *Addresses this somewhat*, see fig. 6.6.

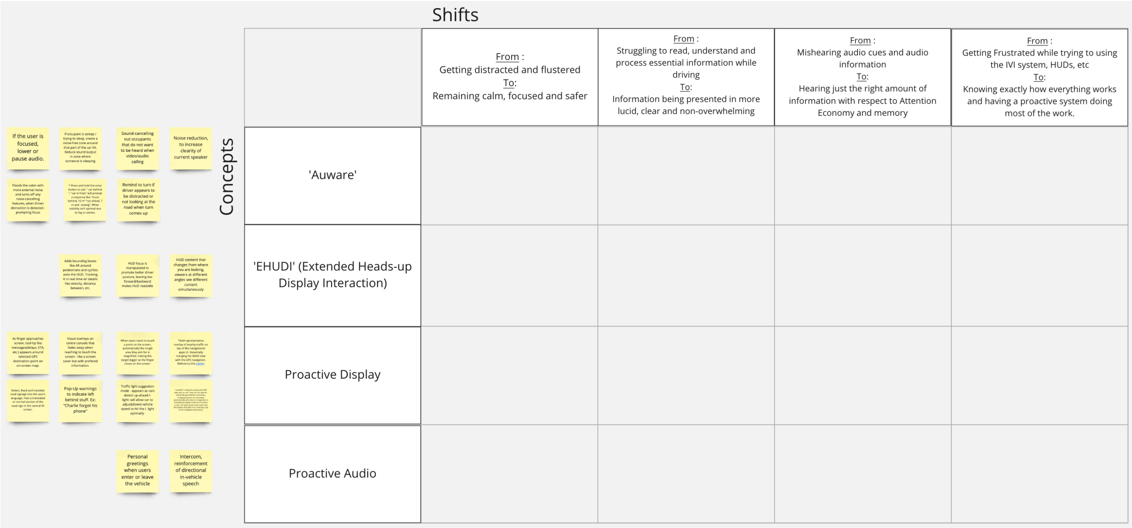

Figure 6.5: The prepared work space for the method *Theory of Change*, implementing the suggested IDEO worksheet in the online tool Miro.

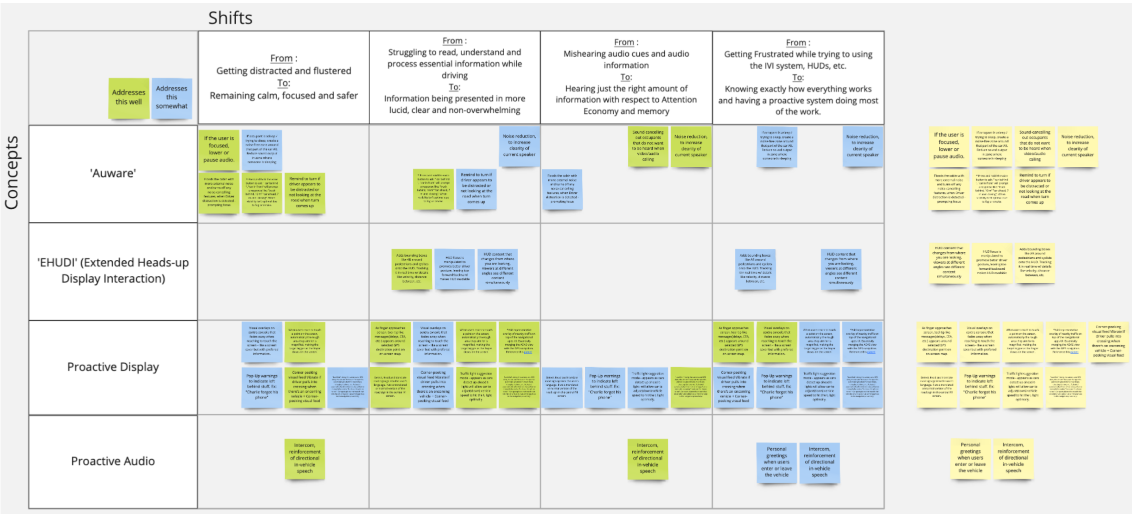

Figure 6.6: The *Theory of Change* worksheet after applying the method.

Following the method, instead of following the recommended follow-up method *Impact Ladder*, the visual end-status of the worksheet was used as documentation of the results - the number of green and blue notes counted for each idea.

This approach was adopted as the culling aspect of *Theory of Change* was the primary motivation of using it, not the more lofty goal of *"Articulate and interrogate your assumptions about how your solution will create positive change."*, stated on the IDEO website [70]. As the societal changing aspects of the method, and its successive follow-up method, was deemed to be off-topic for the thesis' scope, which the ethical and societal discussions made under Section 8.4 already covers in suitable detail for the size of this study.

The six ideas that fit best into the *shifts* defined in Theory of Change were selected for continuation in the next phase of the project. These were:

- *Add to route* - v. command to adjusts the route to include a new destination
- *Car behind* - driver can ask if there is a car behind or in front
- *Intercom* - speech volume is boosted for clarity of in-car conversations
- *Left-behind object* - vehicular occupants are told if 'X'/someone leaves 'Y'/something in the car
- *Magnification* - aimed for part of screen is magnified as finger nears
- *Tooltip* - pointing to an app opens temporary info windows

6.3 Iteration 1: Prototype

In the first, out of three, prototyping iterations the ideas procured as a result of the previous phase were moved from the intangible into the tangible using the scenarios and storyboards - turning the previous text-based ideas into situational graphics more useful for showcasing not only the interactive aspects but also the beginnings of imagined graphical and general UI components.

Planning of phase

As stated in Section 4.7.1, scenarios are an amalgamation between carefully researched descriptions of a set of real ongoing activities with an imaginative futuristic look at how technology could support those activities better [73]. While scenarios were initially reserved strictly for the purposes of evaluation, it was found by Nardi to also be a vital aid in the creation of low-fidelity prototypes and storyboards. As suggested by [73]: *"... showing the actual circumstances under which people work, scenarios provide guidelines on how technology should perform. The very act of making things explicit and clear also helps in making design choices"* [73]. Therefore scenarios would indurate, both as an aid for prototyping and as a tool for assisting artefact evaluation. In short, due to these reasons, *Scenarios* were selected as a method to accompany the follow-up method *Storyboard*.

Storyboards are brief graphical depictions of scenarios or narratives [75]. As for this thesis, wherein the focus lies upon designing for novel and advanced technologies, storyboards could serve to showcase the context of use, including the environment of use, physical embodiment/placement of a system, and user interactions with and reactions to system elements to illustrate an envisioned scenario of how an application feature would work. Thus, storyboards would, in a fidelity suitable for initial prototypes, serve to depict not only a user's interaction with computing technology but also higher-level concepts/factors as well as surrounding user motivations and emotions during system use [75] in this, initial, iteration.

Execution of phase

As planned, the selected ideas were further developed using scenarios, where simple text scripts of example use-cases were written as a support for creating storyboards. The resultant storyboard prototypes can be seen in Figure 6.7 - 6.9, for full versions see Appendix A.1.

The storyboards were made using an iPad and the iOS application *Procreate* which allows for digital sketching of storyboards. The sketches were imported into Google Presentations for positioning, cropping and insertion of other digital components. This choice was done as a way to allow for easy sharing, to parallelize the finishing work and due to the relative ease of text insertion, Google Slides has over Procreate (according to the authors).

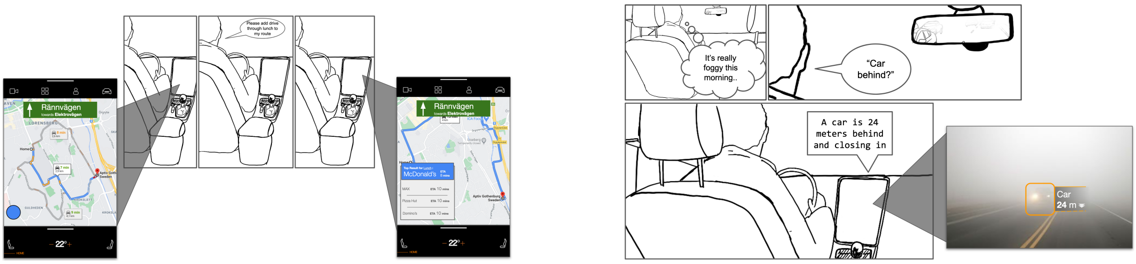

Figure 6.7: The storyboards for the scenarios *Add to route* (left) and *Car behind* (right), for larger versions see Figure A.2 and A.1 in Appendix A.1.

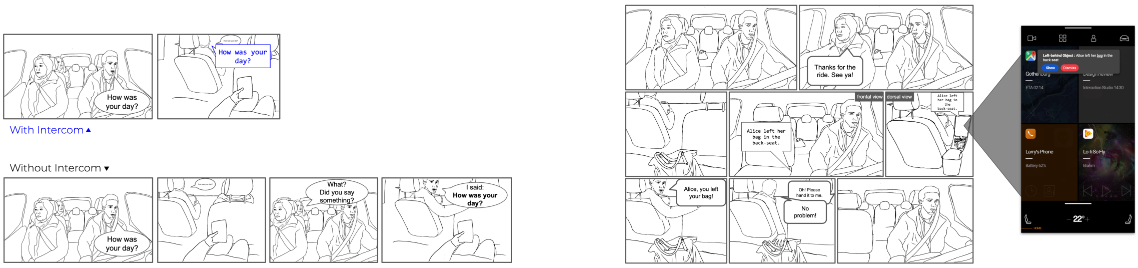

Figure 6.8: The storyboards for the scenarios *Intercom* (left) and *Left-behind object* (right), for larger versions see Figure A.3 and A.4 in Appendix A.1.

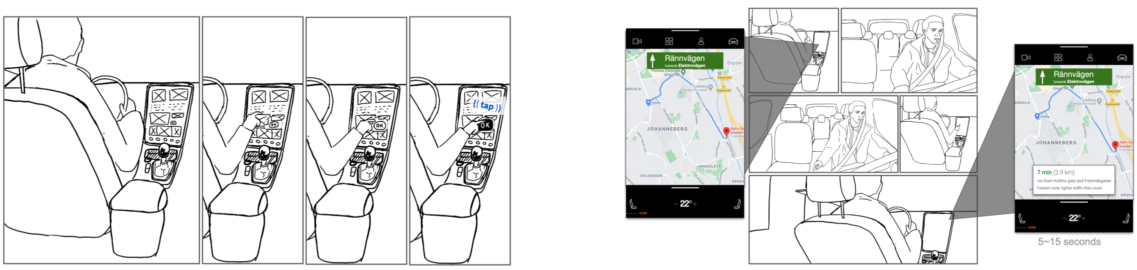

Figure 6.9: The storyboards for the scenarios *Magnification* (left) and *Tooltip* (right), for larger versions see Figure A.5 and A.6 in Appendix A.1.

6.4 Iteration 1: Test

Given the storyboards from the prototype phase, a quick, but not so dirty [84], evaluation approach was employed in the form of internally generated SUS scores for the prototypes. With these scores, as well as knowledge from the Ideate phase, in mind, another culling of ideas was conducted - resulting in a smaller selection of ideas to bring into the second iteration.

Planning of phase
To evaluate the first set of prototypes, the storyboards, it was decided to calculate SUS scores based on the average of the author's individual responses to a SUS questionnaire. This would be done as a way to differentiate the ideas based on the authors' estimations on the standardised measure points in SUS. Due to the novelty of the technology involved, lengthy explanations of each idea and its involved technology would be necessary for priming an impartial respondent enough to fully answer the scale. Due to the time scope of the project phase, the authors deemed it too inefficient, and early, to apply at this point of the design process - instead, parties already versed in the project should be utilised in these evaluations.

Based on the resulting SUS scores, some ideas could then be excluded from future iterations of the design process.

Execution of phase
A questionnaire covering the SUS statements mentioned in Section 4.8.2 for each of the six storyboards was constructed using Google Forms, then answered individually by the authors. The SUS score was calculated for each resulting in the following scores (out of 100), see Figure 6.10 for where the scores placed on the SUS diagram:

- *Intercom*: 95
- *Magnification*: 93.75
- *Left-behind object*: 86.25
- *Car behind*: 85
- *Add to route*: 78.75
- *Tooltip*: 57.5

The fact that some ideas had almost a perfect score was not surprising, as the ideas selected from Theory of Change fulfilled user experience shifts towards the positive, something that ought to generate high SUS scores in general. What was surprising, however, was the low scores given to the bottom two ideas, *Add to route* and *Tooltip*. The reason these ideas passed through Theory of Change and still got a low SUS score might be due to 'good UX' intentions of the ideas and poor realizations in the form of scenarios and/or storyboards. It might be the case that another way of realizing the ideas could have gotten a higher SUS score; however, the case could be that the ideas looked better than they actually were, and the prototyping revealed hidden faults.

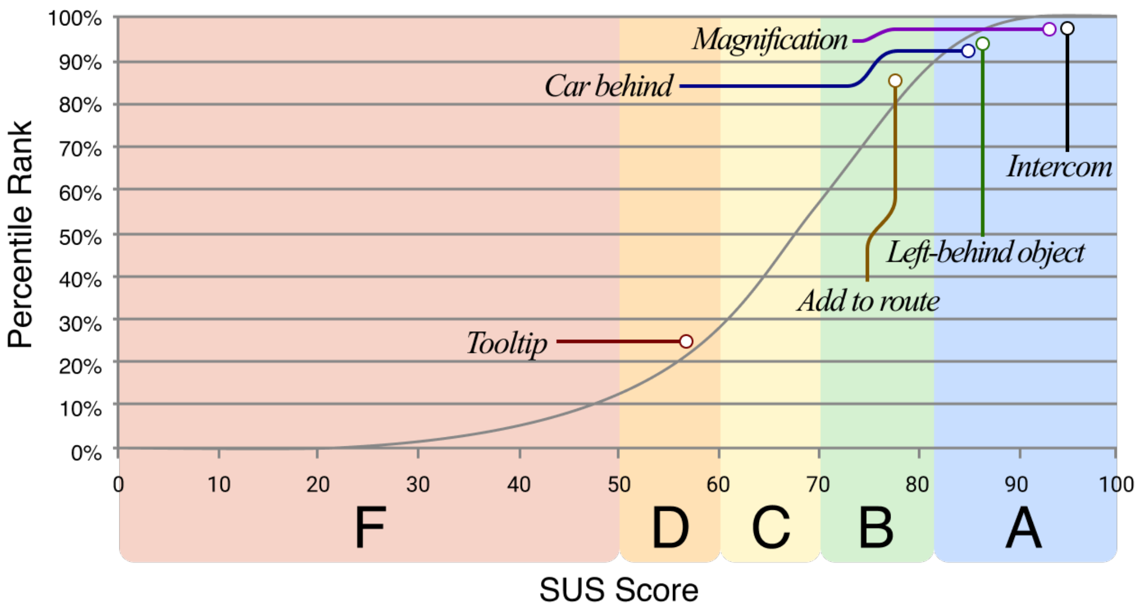

Figure 6.10: How the first iteration's prototypes placed in the SUS diagram (explained in Section 4.8.2) after the Test phase.

In any case, the low scores were deemed argument enough to not further those ideas - this left four ideas for the next iteration.

6.5 Iteration 2: Prototype

To increase the level of fidelity of the prototypes, as well as move towards the aim for fully interactive application, a plan was first made, then executed, to create an interactive mock-up, then record live-action video prototypes based on the scenarios created in the previous iteration. These videos were recorded in an acquired car, using the Figma Mirror app running on an attached iPad acting as the IVI display. GUI (graphical user interface) aspects of the mock-up involved were developed using official guidelines for the Android Automotive operating system.

Planning of phase

The second iteration of prototyping should increase the level of fidelity, as it allows more accurate evaluations of the designed interactions of the prototypes within the intended use-case environment i.e. a car. It was decided that the storyboards from the previous iteration should be recreated as video prototypes, the GUI parts in higher detail using *Figma* and the videos should be shot in either a fake or real car using a modern smartphone camera. To finalize the prototypes, video editing software should be used.

The choice of prototype type (video) was based on several reasons, the two main ones being; the tool Figma was already employed in the previous iteration to create the visible GUI parts. The second reason was that Figma's companion app, Figma Mirror, would allow for some forms of interaction through a touch screen directly inside the Figma environment. This would make a recorded interaction sequence, based on the previously designed scenarios, look highly convincing as no physical interaction with the designs (i.e. taps, drags, pinches, etc.) would need to be faked. One could argue for allowing participants to themselves interact with these Figma designs; however, at this point of the project, it was deemed unnecessary to take this, more temporally costly approach.

Execution of phase
Details of the GUI in the four storyboarded scenarios from the previous iteration were recreated in higher fidelity on Figma, Inc, the tool allowing for animated transitions between created frames for a dynamic GUI appearance, as compared to the static frames illustrated in the storyboards. This choice was made as it would make for a more convincing prototype in the Wizard of Oz setting planned for in the video prototyping, with smooth transitions and semi-interactive interfaces.

6.5.1 GUI Design

For designing the interactive high-fidelity GUI, which was consistent, intuitive and highly functional, the use of a design system was deemed to be necessary. Since the research, development and testing of a novel design system was outside the timeline and scope of the thesis, the decision to adopt a pre-existing design system was made.

The attributes which the design systems were required to embody were as follows:
1. A familiar design language, featuring iconography and navigational paradigms familiar to users, thereby being perceived as intuitive.
2. Featuring well-researched design patterns and guidelines suitable for automotive use and deployment.

After some comparison and online research, the design system by Google called *Material* [102] was adapted. This decision was made, for it best fit the requirements of having a familiar design language, thanks to its widespread use on Android and on Google's online services and its well-defined guidelines for AID (automotive interaction design) [40]. The use of *Material*, [102] allowed the artefact's design to inherit the rich usability and functionality of Google's design research, thereby alleviating major design problems.

The company's own branding guidelines [103] were also accommodated into the design by leveraging the *Theming* [104] aspect of this design system.

6.5.2 Development of Video Prototypes

A fairly modern car was used as a prop (2018 Volkswagen Golf Estate 1.0 TSI) to recreate scenes illustrated in the storyboards. The graphical interfaces created in Figma, Inc were displayed on an iPad attached on top of the centre console, to mimic a large IVI system display, using the Figma, Inc companion app *Figma Mirror* - allowing both direct touch interaction with the interfaces as well as remote control of transitions.

The shot video clips were edited together using the editing software *Final Cut Pro*, due to the authors' preexisting skill set in that tool. Other tools were not considered due to the relatively minor effect the choice of editing tool was planned to have on the prototype. The prototypes, when done, was shared with the company for feedback before evaluating them more thoroughly using the planned method. No immediate changes were requested by the company stakeholder and thus the process continued as intended. Here's a list of *YouTube* URL's to the four video prototypes (valid as of June 2, 2021):

- *Magnification*: https://youtu.be/Svot8XpoiII
- *Left-behind object*: https://youtu.be/LOyjVYbRK80
- *Car behind*: https://youtu.be/T8UPZLM6dsA
- *Intercom*: https://youtu.be/qFmEQBYIm3c

In Figure 6.11 and 6.12 one can see a screenshot of the designs for Left-behind object, Car behind and Magnification. As Intercom has no UI, no screenshot was possible. For more screenshots, see Appendix A.2.

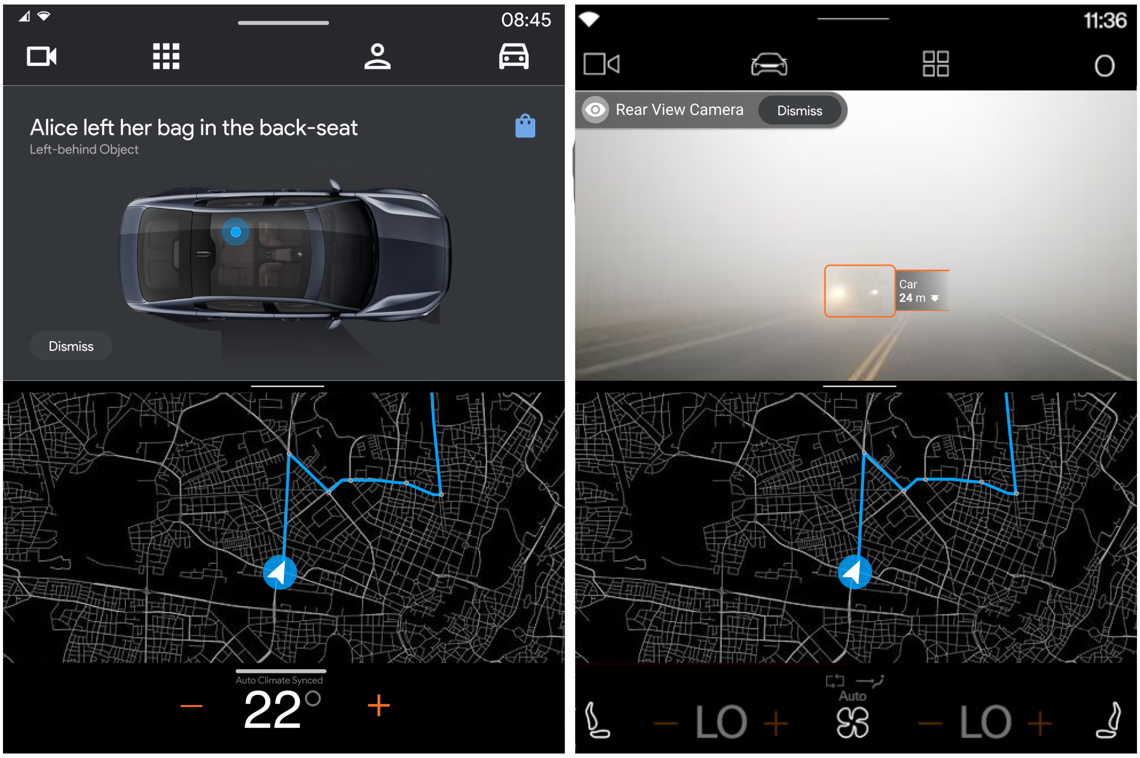

Figure 6.11: A screenshot of the designs *Left-behind object* (left) and *Car behind* (right).

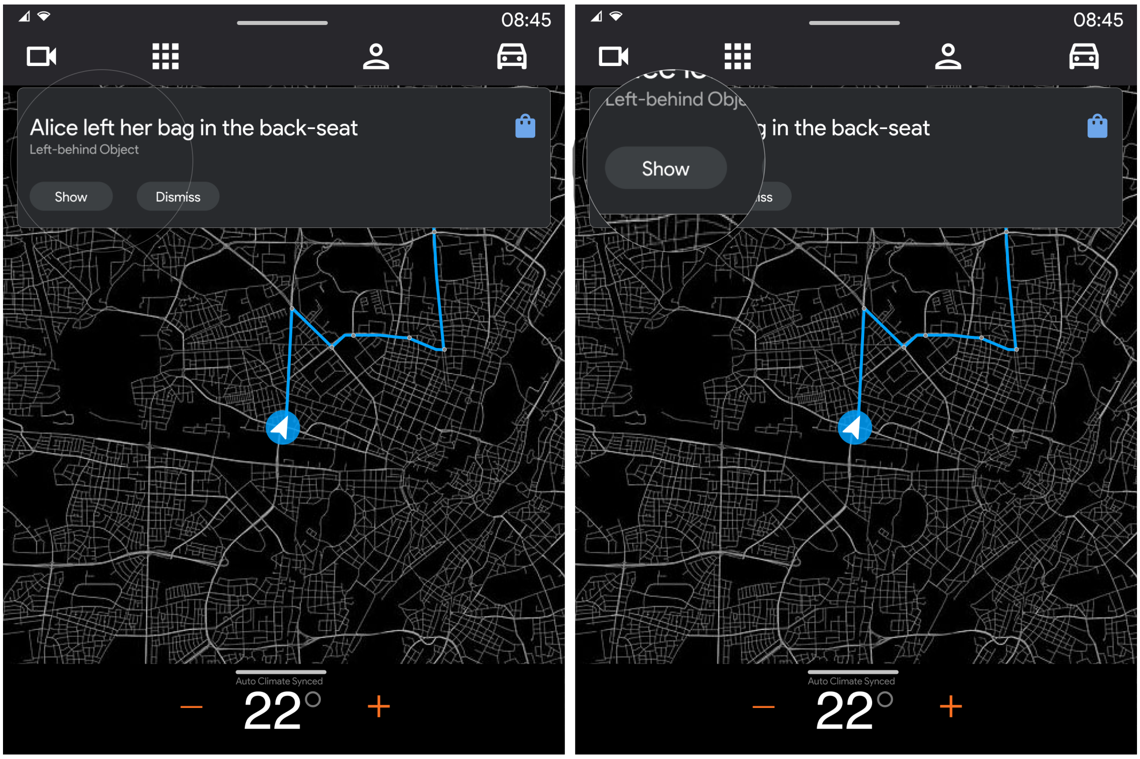

Figure 6.12: A screenshot of the design *Magnification*, showing what happens when a finger nears the screen, transitioning from the left state to the right.

6.6 Iteration 2: Test

To evaluate the video prototypes created during the prototype phase, as well as to adapt to the distance-based testing preferred by most participants at the time (during COVID-19 restrictions), the decision was made to do a Walk-though based evaluation. Participants were recruited, and over Zoom, shown the video prototypes, after which they were asked to describe what they observed. After clarifying any misconceptions, the facilitators discussed and collected responses to the statements in a SUS formula. Notes were taken of any comments made during the evaluations.

Planning of phase

As suggested by Mackay and Fayard, the video prototypes should be evaluated using the method *Walkthrough-based evaluation*, this would involve showing the video prototypes along with a text description to the participants whilst making sure they understood what the prototypes were and the ideas they represented. After having opened for general comments, the ten statements from the SUS method would be asked as opening questions. The interviewer would take note of to what level they replied(strongly disagree - strongly agree), if applicable, as well as any comments made. This strategy was based on the fact that the SUS statements used were constructed to cover a holistic viewpoint of usability, each statement distinguishable from each other, and rephrasing them as questions would thus hopefully uncover as many different usability opinions/thoughts as possible.

During the interviews, the interviewer would also record numerical responses of the SUS statements, so that a second SUS score could be measured, this time for a later iteration and by a different stakeholder group. The SUS scores, as well as the comments made during the interviews, would be analysed together with any comments made by the company stakeholder to better plan for the next iteration, including the selection of ideas for continuation as well as suitable alterations to them.

Execution of phase
For the walkthrough-based evaluation, structured interviews were conducted, involving a set of six, mixed occupation individuals in the ages 24-29. The interviews were structured in the following way:

1. The interviewer shared an invite link to a *Zoom* meeting.
2. Permission to record age and summaries of comments made was asked for by the interviewer.
3. The interviewer shared, through the in-meeting chat interface, a URL link to a *YouTube* page for one of the video prototypes.
4. The participant was asked to explain what they think the design was, as a walkthrough of the design.
5. The interviewer explained any misconception in the participants understanding, as well as took note of the misconceptions themselves.
6. The participant was asked to express any immediate thoughts and comments.
7. The interviewer explains that the coming phase covers ten standardized statements, usually answered with: strongly disagree to strongly agree and that in the interview, they will be used to uncover more thoughts and opinions from the participant.
8. Each statement was rephrased as a question, and the participants' comments documented.
9. The interviewer repeated the process for each video prototype.

Interview notes were taken in a *Google Form*, as it would automatically generate a *Google Sheet* for easy analysis, similar to how the Expert Interview: Questionnaires and the earlier SUS evaluations had been done. SUS scores for the four video prototypes were calculated as the following values, see Figure 6.13 for where the scores placed on the SUS diagram in relation to the first iteration:

- *Intercom*: 96.25 [prev. 95 in Iteration 1]
- *Magnification*: 87.5 [prev. 93.75 in Iteration 1]
- *Left-behind object*: 86.25 [prev. 86.25 in Iteration 1]
- *Car behind*: 72.5 [prev. 85 in Iteration 1]

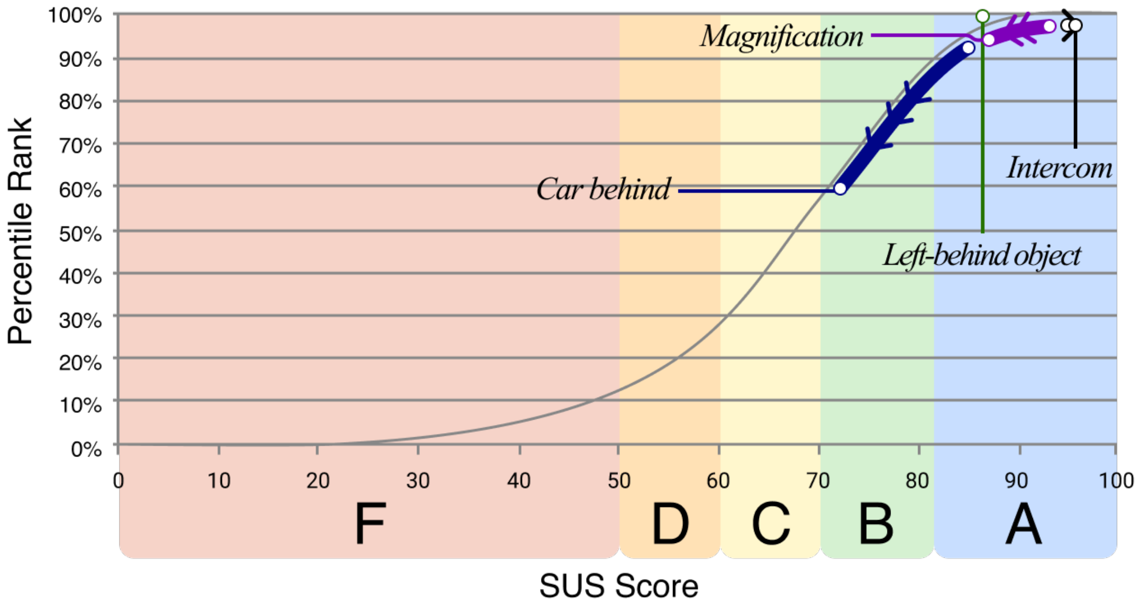

Figure 6.13: How the second iteration's prototypes placed in the SUS diagram (explained in Section 4.8.2) after the Test phase, in relation to the first iteration.

From the collected notes, both from the conducted interviews and oral & written feedback from the company stakeholder, a summary of author suggestions for each idea was written:

MAGNIFICATION
Both parties involved were unsure how much the obscuring of screen space would affect interaction with the centre console. Suggestions were made to introduce a parallel copy of the aimed-at area, showing the magnification - keeping the surface under the finger's aim unchanged. From the companies side, a feasibility concern was raised - implementing/realizing the design as shown in the video would involve, possibly, alterations in the software at a level lower than the operating system the in-vehicle infotainment is run on, effectively putting far out of the scope of a thesis as well as outside the authors' expertise.

LEFT-BEHIND OBJECT
Several participants expressed concerns regarding how consistent the feature's accuracy would be, such as when smaller objects in harder to find/see places are involved. Two participants as well as the company suggested a simple way to also share the notification with the person who left the object behind, such as sending it to their phone.

CAR BEHIND
Requiring manual activation of the feature should be replaced with an option to have it automatically trigger if needed. As some users are not confident and/or prefer other input modalities, alternatives to voice command should be available. For users who prefer to have their eyes on the road, a simplified dashboard or HUD version of the feature could better suit their need. Other low-vision settings, such

as night vision, could be included in the idea - especially since driving in darkness is more common than fog or mist.

INTERCOM
In hindsight, it is not a surprise that this idea scored high on both SUS evaluations, as both participants and the company pointed out preexisting implementations of this idea - and since it has already been done, the thesis should not further this idea.

The SUS scores given by the participants was much in line with the ones calculated for the storyboards in iteration 1, with some key differences. The ideas *Magnification* and *Car behind* both received far more concrete suggestions for improvement by the participants and the company than the other ideas - meaning that if the implementation shown in the videos had included those suggestions, higher SUS scores would probably have been given.

Following the evaluations, the choices on what ideas to cull, or continue with, were made. Due to the idea already existing in the industry, therefore outside the defined design space, development of *Intercom* would not continue. Due to the lack of necessary expertise, but mainly due to being far outside the scope of the thesis, *Magnification* was also judged to be left as is - this left the ideas *Left-behind object* and *Car behind* to be furthered in the next design iteration. As two relatively stronger reasons for the culling choice existed, the SUS scores were not considered in this selective culling process; however, it was thought to be useful in later iterations for comparative reasons.

Due to the change in nature from the initial ideas as well as having too long or ambiguous names, the names of the finalised ideas were reconsidered. *Left-behind object* was simply abbreviated as *LeBO* (Left Behind Object) for simplicity in referencing. *Car behind* was renamed into *ADASight* (ADAS Sight), to indicate the idea's heavy reliance on the *ADAS* technology and the word *"Sight"* was juxtaposed as a reference to the multiple camera feeds utilized in the idea concept.

6.7 Iteration 3: Prototype

The third iteration's prototyping phase was planned in detail, mainly caused by expertise from previous projects in software development, including what tools should be acquired and used and what steps probably were needed to deliver a viable artefact. The same expertise resulted in a somewhat typical development process: following the initial plan, facing and conquering hurdles, and adjusting the plan when facing hurdles too difficult to power through within the scope of the iteration. In the executing part, a development environment was set up, the new tools learned and finally two artefacts were developed.

Planning of phase
To further the fidelity of designed artefacts, the car UI development tool Kanzi Studio (not only suitable for the situation but also recommended by the company

stakeholder) should be applied to create Android applications. To do this, license keys for academic projects must be acquired and the software tools installed. When set up, Kanzi Studio tutorials relevant to the thesis should be followed - to prepare the designers for the development of the applications.

After acquiring the tools, as well as the expertise needed to use them, Android applications based on iteration 2's mock-ups should be created using Kanzi Studio and an Android Automotive emulator. While working with Kanzi, any design improvements identified would be incorporated into the new prototypes.

The prototypes should be developed to such an extent, that they could be evaluated using specific scenarios in the coming test phase. Following completion of the prototypes, setups for evaluating them should be designed. Depending on available facilities, hardware, time budget, as well as access to the participant, the designed setup should cater to not only physical sessions but digital ones as well.

Execution of phase
After contacting *Rightware*, the company behind Kanzi Studio, both software license keys as well as access to documentation and tutorials were acquired. Following the successful setup of the development environment, key, relevant tutorials were completed. While developing the applications, inexperience with Visual Studio, as well as with the Kanzi C++ API, caused several issues that hampered development. With time and the help from Kanzi support, they were resolved.

To aid in development, an emulator running a version of Android Automotive was set up using the built-in AVD (Android Virtual Device) Manager in Android Studio, with the following specifications(excluding default options):

1. Category: Tablet -> Polestar 2 11.13" [1152 x 1536px]
2. Operating System: Android 9.0 x86_64 [Pie]
3. Layout: Portrait
4. RAM: 4GB

The general workflow during development was as follows:

1. Set up a git repository, specifically on GitHub, for collaborative development as well as version control.
2. In Kanzi Studio; Developed basic android application using hard-coded values from Figma mock-ups
3. In Visual Studio; created Kanzi plugin responsible for connecting UI changes to external data feeds, i.e. ADAS sensor data modalities.
4. In Kanzi Studio; installed developed plugin.
5. In Kanzi Studio and Android Studio; replaced hard-coded values, when needed, with dynamic ones, i.e. that could change depending on external data feeds.
6. In Android Studio; added audio-out modality to the prototypes as Kanzi studio did not, to the designers' knowledge, support audio.

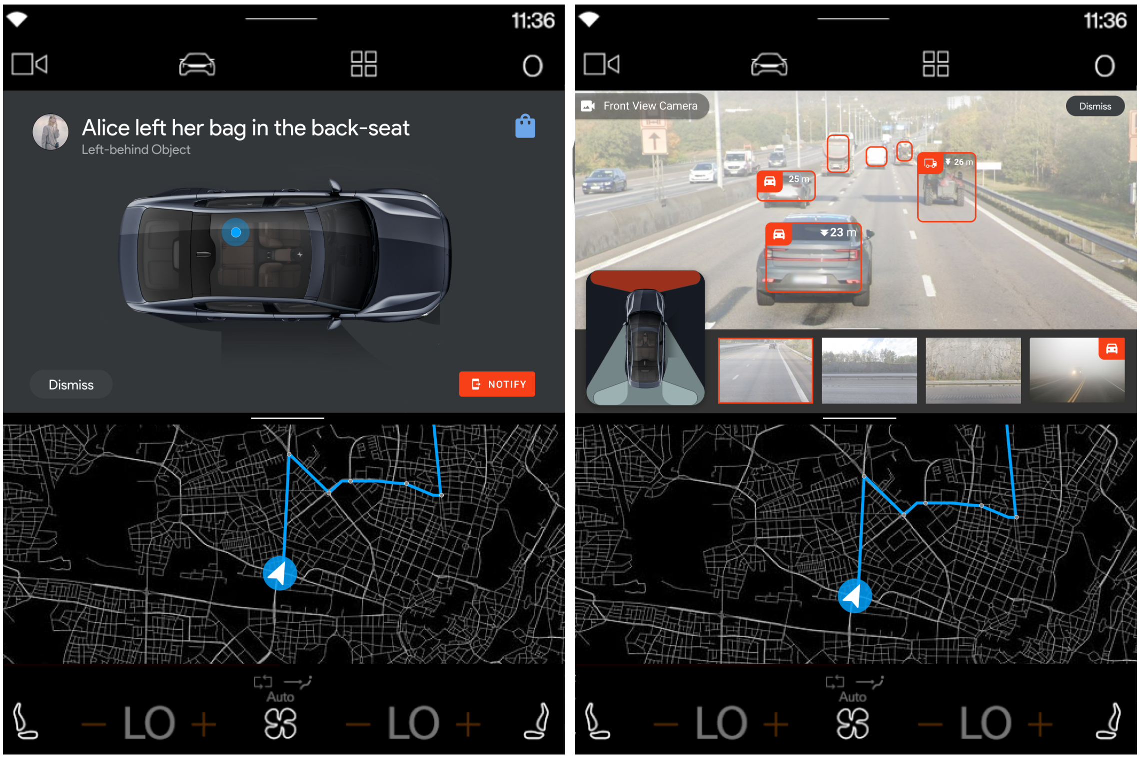

Figure 6.14: A screenshot of the designs *LeBO* (left) and *ADASight* (right).

As a rule, the designers built and deployed the applications onto the emulator between big changes to the Kanzi project and the Android studio code. This was done to minimize the introduction of bugs and application breaking operations. When making design alterations, relative to the UI present at the end of iteration 2, these were first done using Figma, due to the relatively faster sub-design-iterations using a familiar tool, and then through Kanzi.

Due to the unrefined nature of the available real-life sensor log files, proof-of-concept demonstrations were developed using hard-coded data changes in the two applications. For example, in ADASight: the sensed distance to a car - and in LeBO: by who, what object, and where an object was left behind in the car. In Figure 6.14 one can see screenshots taken from the emulator running the end version of the apps. For more screenshots, see Appendix A.3.

6.8 Iteration 3: Test

The test phase was initiated by planning around non-ideal circumstances, laying out the preparatory work for evaluating the final artefacts of the design process. This involved setting requirements on artefacts such as test setups, recruitment material (consent forms, etc.), data collection methodology and recruitment criteria. In the executive part of the phase, all these materials, documents and preparations were made, followed by recruitment, performing the evaluations with participants, and lastly, the outputted data was sorted, parsed and collated.

Planning of phase
Due to unforeseen difficulties in acquiring required hardware in time for the evaluations of the prototypes, iteration 2's interactive mock-ups would be updated and prepared for use in the evaluations, thus looking and interacting as iterations 3's Android applications. Thus, instead of deploying apps directly on a tablet, the evaluation would use the app Figma Mirror [105] creating a Wizard-of-Oz setup with remote-controlled data changes during the session as well as any functionality allowed by the tool Figma. While not ideal, the decision had the added benefit that digital evaluations could be conducted in a somewhat similar way to physical ones, with the main difference being that the participant would have to verbally express their intent and any interactions made with the interface, whereupon the 'Wizard' would have to manually trigger the relevant UI change.

Data recording of the sessions would mainly be through digital forms: a SUS form as was done in previous iterations, a DALI form (to be created using Google Forms as was done with the SUS ones) as participants would perform tasks using the prototypes, as well as a basic demography form. Demography data would be collected not only for posterity but also to act as an aid during data analysis and discussion. For legal and safety reasons, a consent form would have to be written and signed before each session by the participants, primarily to cover the data access laws (GDPR being one) and the hazards of physical proximity & interaction during Covid-19 pandemic restrictions. Session agendas, i.e. what happens when would also need to be written for both physical and digital sessions.

Following the preparations, participants with at least some experience driving, preferably available for physical evaluations, should be recruited. Following recruitment, sessions would be held and documented, both through the forms as well as thorough notes and photos taken during the session. Collected data would be summarized and composed for chapter 7.

The SUS form for the evaluations should be based on the ones previously prepared for iteration 1 and 2.

As no template form was found, the DALI form that would be used was to be constructed based on the original definition of the factors involved (Pauzié, 2008), as well as template forms for NASA-TLX and other HCI work-load indexes.

Execution of phase
Before the task-based evaluations could begin, the necessary forms were prepared and written. The SUS form used was mostly reused parts of the SUS forms used in iteration 2. For the constructed DALI form, the result is the following six questions, answered through a Likert scale from 1[*Very Low*], up to 7[*Very High*]. The leading question has its answer scale flipped, to act as a control questions:

1. To what extent did the system demand your attention while performing task X?
2. To what extent was the visual aspects of the task uncomfortable/irritating when performing task X?
3. To what extent was the auditory aspects of the task uncomfortable/irritating when performing task X?
4. To what extent did you feel hurried or rushed when performing the task?
5. To what extent did you feel distracted from driving when performing task X?
6. To what extent did you feel stressed while performing the task?

The consent form used can be found in its entirety in Appendix B.

The demography form contained questions regarding the following fields:

- Age
- Main occupation
- Driving experience (a Likert-scale 1-5)
- Have used advanced safety features before, e.g. lane-keeping, adaptive cruise control, automatic braking, collision avoidance, auto-dimming headlights, etc.
- Have used a system similar to Android Automotive, Android Auto or Apple Carplay.
- Participates physically or digitally?

The evaluation sessions were structured as follows:

SESSION AGENDA: PHYSICAL
The following agenda is directed at evaluations conducted in-person with the participants with props in a physical setup.

1. Prepare masks and don masks.
2. Sanitize facility and equipment, prepare equipment and refreshments.
3. (Optional) Meet up with the Participants and guide them to the facility.
4. Have participants read and sign the waiver.
5. Offer a mask to the participant.
6. Have participants fill out a demographic form.
7. Introduction of the system and the thesis. Maybe show examples of contemporary Infotainment systems to provide context.
8. Perform a task
9. Have participants fill out a DALI form.
10. Have Participants fill out a SUS form.
11. Offer short breaks and refreshments.
12. Reiterate 4-6 with other apps.
13. Express gratitude for participation and offer refreshments.

SESSION AGENDA: REMOTE
The following agenda is directed at evaluations conducted remotely with the help

of a video conferencing tool with the participants using their own personal device to participate in the evaluation.

1. Welcome and Meet up with the Participant on Video conference.
2. Have the participant read and sign the waiver.
3. Obtain GDPR consent and request the participants turn on their camera.
4. Have participant fill out a demographic form.
5. Introduction of the system and the thesis. Explain contemporary Infotainment systems to provide context.
6. Ask the participant to set their computer off to the side and set the scenario of driving a car. Begin playing the 'Traffic sounds' at this point.
7. Share prototype on a screen share and ask them to think out loud as they perform the task.
8. Have participants fill out a DALI form.
9. Have Participants fill out a SUS form.
10. Offer short break.
11. Reiterate 6-9 with the other app.
12. Express gratitude for participation.

6.8.1 Setup and tasks

The physical setup was to have a laptop in front of the Participants playing a pre-recorded driving seat POV (point-of-view) video with traffic sounds, asking the participant to follow traffic as if they were driving - as a way to simulate real driving in terms of visual and additive stimuli. A tablet with the application running is set up slightly off to the side of the participant - the setup can be seen in Figure 6.15. This was done to roughly simulate the position of the infotainment displays in contemporary vehicles. A mock steering wheel, made out of cardboard is placed in between the laptop and the participant seat. The participants were asked to interact and attempt to perform application-specific tasks on the tablet whilst devoting a sufficient amount of their attention to the road ahead. As for the digital setup, a laptop/display was placed off-centre in relation to the Participants with video conference application on full-screen. This is done to roughly simulate the position of the infotainment displays in contemporary vehicles.

In cases where a participant does not follow the intended interaction pattern, to expose them to more of the app, the instruction to interact in a different way than previously should be given. It is important to instruct the participant that there is no wrong or right, way to perform the task and that this instruction to 're-do it' is only given to allow them to see more of the design before evaluating it. If this information is not given, a false sense of forced interaction could emerge - affecting the evaluation output.

LeBO

The following setup and timeline were used when conducting evaluations of the LeBO design. The driving seat POV video shows a parked car, a facilitator plays

Figure 6.15: Physical setup for the third iteration's evaluation sessions, as seen from the side (left) and from the driving seat point of view (right).

the role of a passenger and sits next to the participant who acts as the driver - both being occupants in this setting.

1. Vehicle pulls into a stop/car is parked.
2. The passenger indicates they are leaving the car.
3. Door-slamming sound is played with a speaker in the direction of the leaving passenger.
4. LeBO-notification is triggered on the device.
5. Let the participant interact how they wish with the device.
6. (In case the participants choose to not notify the departed passenger) Ask them to play out the scene again with the instruction to notify the‘passenger.
7. (If notification is sent) The passenger picks up their phone, pretending to receive a notification.
8. (If the participant indicates in some other way to the passenger) The passenger reacts suitably to it.
9. The passenger returns to the car and opens the door.
10. Play door opening-sound from the relevant direction.
11. The passenger retrieves object(s) and expresses gratitude to the participant for notifying them.

ADASight

The following setup and timeline were used when conducting evaluations of the LeBO design. Driving seat POV feed shows driving in foggy conditions.

1. After 10 seconds of driving, ADASight ‘poor vision’-notification is triggered
2. The participant opens ADASight
3. The participant is allowed to interact with ADASight for >30 seconds
4. The participant is asked to play out the scene again.
5. After 10 seconds of driving, the ADASight ‘car approaching’-notification is triggered.

6. The participant opens ADASight
7. The participant is allowed to interact with ADASight for >30 seconds
8. The participant is introduced to the manual activation modes, verbal and touch.
9. The participant is asked to play out the scene again, this time using the manual activation of the feature.

6.8.2 Evaluation output

This subsection covers the raw data outputted by the last Test phase, such as where the new SUS scores placed the ideas (see Figure 6.16), further analysis and comparisons are covered in the Chapter 7: **Results**.

SUS
- *LeBO*: 84.82 (out of 100 - high is good) [prev. 86.25, 86.25 in Iteration 2, 1]
- *ADASight*: 70.89 (out of 100 - high is good) [prev. 72.5, 85 in Iteration 2, 1]

DALI
- *LeBO*: 1.96 (out of 7 - low is good)
- *ADASight*: 3.89 (out of 7 - low is good)

demographics

- *Number of participants*: 14
- *Age*: AVERAGE = 30.9 | 12 under, 2 above
- *Occupation*: 12 engineering related, 2 other
- *Perceived driving experience (1-5)*: AVERAGE = 2.9 | 6 below, 8 above
- *Have experience using an ADAS feature*: 7 yes, 7 no
- *Have experience using a modern automotive IVI OS*: 3 yes, 11 no
- *Participation location*: 13 physical, 1 digital
- *Session length*: AVERAGE = under 30 minutes

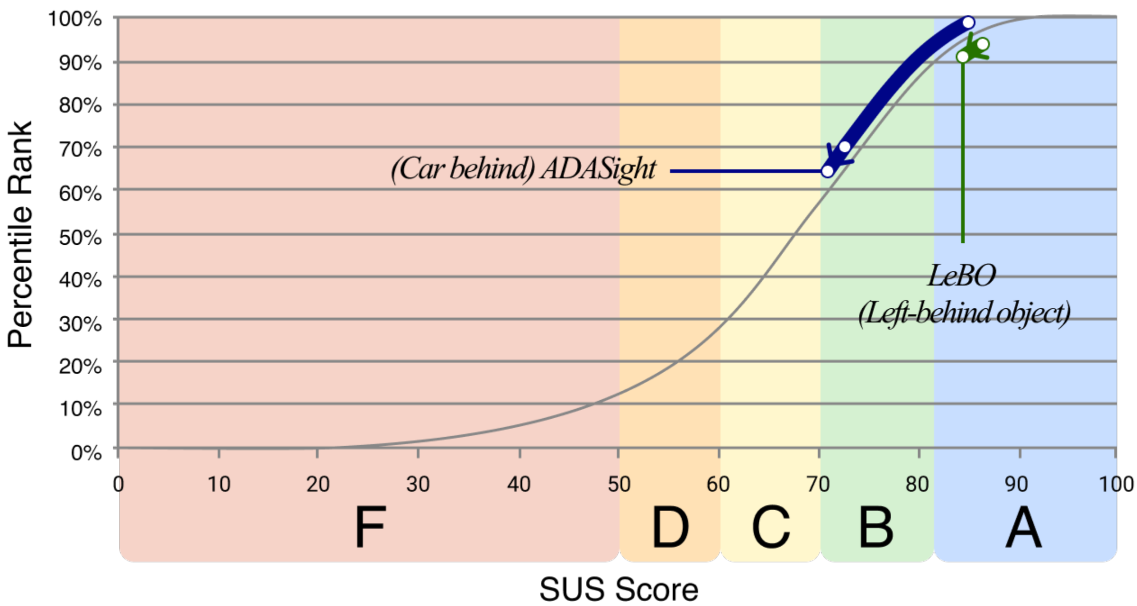

Figure 6.16: How the third iteration's prototypes placed in the SUS diagram (explained in Section 4.8.2) after the Test phase, in relation to the first two iterations.

7
Results

This chapter covers the different results produced by this thesis, starting with the artefacts produces as a result of the validation process of the two hypotheses, then ending with the results of the validations themselves. Briefly, sub-results will be covered for the artefacts, as a way to introduce the final results in context. For the validations, the hypotheses covered and motivated in detail in previous chapters, will be re-introduced to put their results into context.

When interpreting the numerical data measured and analysed in the thesis, which is covered by this chapter, it is important to note the assumptions made. The statistical method used was t-tests, which requires normally distributed data, random sampling of the population (user group) as well as an adequately sized sample group. As a result, this chapter makes those assumptions when mentioning t-test, p-values and statistical patterns and significance. Addressing the assumptions, and what effect they have on the results, is covered in Chapter 8, Section 8.1.

7.1 Final artefacts

To measure the validity of the hypotheses, the thesis designed, developed and evaluated two artefacts for the company Aptiv. The following section covers in short the steps leading up to, and in detail, the final iteration's prototypes: two Android Applications, see Figure 6.14. Each of these subsections will end on the numerical output, and analysis made on it, of the different evaluation methods. The final subsection covers a statistically significant topic discovered in the analysis of the data.

Earlier iterations
Starting from a defined design space, ideas in form of one or two sentence descriptions were developed. From these sentences, use case scenarios were written to better define the imagined interaction. Using the scenarios as a guide, storyboard comic strips were drawn to illustrate them. The GUI components were sketched out in greater detail in a partially interactive digital mock-up using Figma. Based on the illustrations and the GUI mock-up, live-action video prototypes were shot and edited - using an actual car along with the mock-up running on an attached tablet. Working on the mock-up used in the shooting, Android application GUI's were developed through the software tool Kanzi - these applications being the final artefacts. For a step-by-step illustration of this fidelity progression, from idea to app, see Figure 7.1.

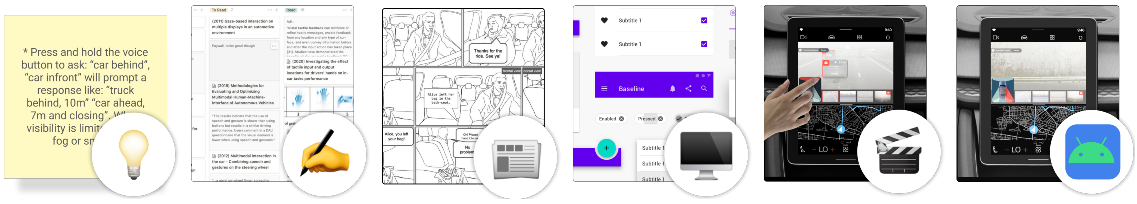

Figure 7.1: How the artefacts produced throughout the process developed from ideas into applications, gradually increasing the level of fidelity.

7.1.1 ADASight

The application ADASight, working name *Car behind*, leverages and reveal to the user what modern ADAS platforms allows vehicles to 'see', or understanding the area surrounding it. When opened, either manually or after taping 'show' in the proactive notification, users are exposed to live camera feeds from different angles around the car - in the artefacts case frontal, rear and side-view camera angles. Layered over the selected camera feed, in the main window, are bounding boxes placed on top of vehicles and objects the ADAS platform is detecting. In relation to these boxes are vital information such as distance to the user as well as if the object/vehicle is approaching or distancing itself fast or slow. Clickable miniatures of the camera feed, acting as feed selectors, can have warning icons for vital information, such as a fast-approaching bike, prompting the user to swap to that feed. To better convey what feed is currently selected, as well as to provide another way of selecting a feed, is a car and feed mini-map, see Figure 7.2.

Figure 7.2: The feed selector in ADASight, the selected camera feed is highlighted in the burnt orange color, not selected ones in grey.

Evaluation results

Looking at the SUS scores derived during the process, ADASight started off with a grade of A and dropped during iteration 2 down to a B, where it remained in iteration 3 - see Figure 6.16. The new, lower score, placed ADASight on the cusp of C - a sign of, if not poor, then average usability. However, looking into the demographic groups within the participants (based on their individual responses to the demographic form), interesting patterns were observed - see Figure 7.3.

To detect any statistically significant patterns, one and two-tailed t-tests were made, both on the total score, as well as the individual aspects for both DALI and SUS on all demographic groups (not 'participation location' due to small 'digital' sample size).

For ADASight, there is enough evidence to suggest that there is a strong statis-

ADASight

one tailed | two-tailed					High value=good
SUS [0 - 100]					
SUS score	p = 0.05 | 0.11	**ADAS experience** Overall = 70.89	Group A: no	63.57	Group B: yes 78.21
Would frequently use	p = 0.00 | 0.00	**ADAS experience** AVERAGE = 2.07	Group A: no	1.14	Group B: yes 3.00
Cumbersome / awkward	p = 0.01 | 0.02	**ADAS experience** AVERAGE = 2.50	Group A: no	1.86	Group B: yes 3.14
Would frequently use	p = 0.03 | 0.06	**Modern OS experience** AVERAGE = 2.07	Group A: no	1.73	Group B: yes 3.33
Unnecessarily complex	p = 0.04 | 0.09	**Occupation** AVERAGE = 2.93	Group A: engineer	3.08	Group B: other 2.00
Need support	p = 0.00 | 0.01	**Occupation** AVERAGE = 3.29	Group A: engineer	3.58	Group B: other 1.5
Pain points	p = 0.01 | 0.02	**ADAS experience** AVERAGE = 62.50	Group a: no	50.00	Group B: yes 75.00

Figure 7.3: The observed patterns from the different demographic groups within the participants, correlations with p-values lower or equal to .05 indicating strong statistical significance.

tical significance for the following statements, i.e. their null hypotheses *'the two population means are equal'* were rejected:

- Participants who had already experienced an ADAS feature perceived a higher SUS score than participants who had not.

- Participants who had already experienced an ADAS feature desired to use the feature more, than participants who had not.

- Participants who had already experienced an ADAS feature perceived the feature as less cumbersome/awkward to use, than participants who had not.

- Non-engineer participants perceived the feature as more unnecessarily complex than participants with engineering-related occupations.

- Non-engineer participants felt a greater need for support than participants with engineering-related occupations.

- Participants who had already experienced a modern automotive IVI OS desired to use the feature more, than participants who had not.

- Participants who had already experienced an ADAS feature perceived a higher collective value for the pain points than participants who had not.

7.1.2 LeBO

The application LeBO, working name *Left-behind Object*, leverages and reveal to the user what modern ADAS platforms allows vehicles to 'see', or understanding the in-cabin events, specifically who the occupants are, what they bring into the car and where they place it. When a left-behind object is detected by the vehicle, after the object's owner has left the vehicle, a notification is triggered which highlights as much relevant information about the event as possible to the occupants. The information includes *who*, *what* and *where*, if the information exists - unknowns are replaced by *someone*, *something* and *in the car*.

Evaluation results
Looking at the SUS scores derived during the process, LeBO remained, throughout the iterations, within a grade of A - see Figure 6.16. Despite its somewhat stellar scores, both in terms of SUS and DALI, interesting patterns were observed within the participant data - see Figure 7.4.

LeBO

one tailed | two-tailed High value=good

<table>
<tr><th colspan="5">DALI [1 - 7]</th></tr>
<tr><td>Attention demand</td><td>Age</td><td>Group A: val < 30.9</td><td>Group B: val > 30.9</td></tr>
<tr><td>p = 0.03 | 0.07</td><td>AVERAGE = 3.79</td><td>3.27</td><td>5.67</td></tr>
<tr><td>Attention demand</td><td>Driving experience</td><td>Group A: val < 3</td><td>Group B: val > 3</td></tr>
<tr><td>p = 0.00 | 0.01</td><td>AVERAGE = 3.79</td><td>2.17</td><td>4.86</td></tr>
<tr><th colspan="5">SUS [0 - 100]</th></tr>
<tr><td>Would frequently use</td><td>Age</td><td>Group A: val < 30.9</td><td>Group B: val > 30.9</td></tr>
<tr><td>p = 0.05 | 0.10</td><td>AVERAGE = 3.57</td><td>2.73</td><td>4.00</td></tr>
<tr><td>Would frequently use</td><td>ADAS experience</td><td>Group A: no</td><td>Group B: yes</td></tr>
<tr><td>p = 0.03 | 0.07</td><td>AVERAGE = 3.00</td><td>2.43</td><td>3.57</td></tr>
<tr><td>Need support</td><td>Modern OS exp</td><td>Group a: no</td><td>Group b: yes</td></tr>
<tr><td>p = 0.01 | 0.02</td><td>AVERAGE = 3.57</td><td>3.91</td><td>2.33</td></tr>
<tr><td>Inconsistent features</td><td>Driving experience</td><td>Group A: val < 3</td><td>Group B: val > 3</td></tr>
<tr><td>p = 0.04 | 0.08</td><td>AVERAGE = 3.57</td><td>4.00</td><td>3.14</td></tr>
<tr><td>Pain points</td><td>ADAS experience</td><td>Group A: no</td><td>Group B: yes</td></tr>
<tr><td>p = 0.03 | 0.07</td><td>AVERAGE = 78.93</td><td>70.00</td><td>87.86</td></tr>
</table>

Figure 7.4: The observed patterns from the different demographic groups within the participants, correlations with p-values lower or equal to .05 indicating strong statistical significance.

To detect any statistically significant patterns, one and two-tailed t-tests were made, both on the total score, as well as the individual aspects for both DALI and SUS on all demographic groups (not 'participation location' due to small 'digital' sample size).

For LeBO, there is enough evidence to suggest that there is a strong statistical significance for the following statements, i.e. their null hypotheses *'the two population means are equal'* were rejected:

- Older and/or driving experienced participants perceived a lower attention demand than younger and/or less experienced ones.

- Older and/or participants who had already experienced an ADAS feature perceived higher desirability than younger and/or participants who had not.

- Participants who had already experienced a modern automotive IVI OS perceived a higher need for support than participants who had not.

- Participants with more driving experience perceived a higher inconsistency in the feature than participants with less experience.

- Participants who had already experienced an ADAS feature perceived a higher collective value for the pain points than participants who had not.

7.1.3 Shared attributes (Pain points)

In the analysis of the designs, the per-SUS attribute development of the designs was analysed - see Figure 7.5 - 7.6. To aid the analysis a similar score-to-grade conversion was made, where the original value was scaled first up to an out-of-a-hundred score, then into an A-F grade using the same limits as for the total SUS score. Using conditional formatting, the same grade colours were applied as in Figure 4.1, making it easier to detect positive and negative changes to the designs. The weakest attributes of both were highlighted, and noticing a partial overlap, merged, into a set of attributes together referred to as *Pain points*:

- Would frequently use
- Easy to use
- Well integrated
- (inverted) Cumbersome / Awkward to use
- Feel very confident using

ADASight

Iteration:	Would frequently use	Unnecessarily complex	Easy to use	Need support	Well integrated
1	2.50	0.50	4.00	1.00	4.00
2	2.33	1.83	2.67	0.33	3.17
3	2.07	1.07	2.64	0.71	3.14
Out of 100:					
1	62.50	87.50	100.00	75.00	100.00
2	58.33	54.17	66.67	91.67	79.17
3	51.79	73.21	66.07	82.14	78.57

Iteration:	Inconsistent features	Learn to use quickly	Cumbersome / awkward	Feel very confident	Need to learn a lot
1	0.00	3.00	0.50	2.50	0.50
2	2.00	3.33	1.17	3.17	0.33
3	0.57	3.21	1.50	2.14	1.00
Out of 100:					
1	100.00	75.00	87.50	62.50	87.50
2	50.00	83.33	70.83	79.17	91.67
3	85.71	80.36	62.50	53.57	75.00
	A	B	C	D	F

Figure 7.5: The cross iteration development of the individual attributes measure by the SUS forms..

LeBO

Iteration:	Would frequently use	Unnecessarily complex	Easy to use	Need support	Well integrated
1	2.50	0.00	2.50	0.50	3.50
2	2.67	1.00	4.00	0.00	3.50
3	3.00	0.43	3.43	0.43	3.00
Out of 100:					
1	62.50	100.00	62.50	87.50	87.50
2	66.67	75.00	100.00	100.00	87.50
3	75.00	89.29	85.71	89.29	75.00

Iteration:	Inconsistent features	Learn to use quickly	Cumbersome / awkward	Feel very confident	Need to learn a lot
1	0.00	3.50	0.50	3.50	0.00
2	0.50	4.00	0.83	3.33	0.67
3	0.43	3.64	0.93	3.29	0.21
Out of 100:					
1	100.00	87.50	87.50	87.50	100.00
2	87.50	100.00	79.17	83.33	83.33
3	89.29	91.07	76.79	82.14	94.64
	A	B	C	D	F

Figure 7.6: The cross iteration development of the individual attributes measure by the SUS forms.

7.2 Validity of hypotheses

The thesis presents two answers (**H1** and **H2**) to the research questions (**RQ1** and **RQ2**) first introduced as a two-part research problem in Section 1.2 - 1.3, and rationalized in Section 3.4, and lastly defined in detail in Section 5.1. The answer to the first research question (**RQ1**) being a suggested development/design process (**H1**) consisting of an ordered method selection, see Section 5.2.1 and Figure 5.1. The answer to the second (**RQ2**), an evaluation strategy (**H2**), is suggested to act as a companion process to that design process, see Section 5.2.2 and Figure 5.2.

To measure the validity of the hypotheses (**H1** and **H2**) two Android applications were developed in collaboration with a company - the first question's answer employed to generate the different iterations' artefacts, the answer to the second employed to steer the development as well as measuring participants' perception of novel artefacts, to simplify; what they thought was good and what they thought was bad.

The following two subsections cover more in detail what the hypotheses (**H1** and **H2**) were, how they were found, as well as what the produced results indicate for their validity. A discussion of to what level the hypotheses were valid is covered in the following chapter 8: **Discussion**.

7.2.1 Hypothesis 1 (H1)

> **RQ1:** *How can designers make better use of vehicle sensor data to create improved vehicular user experiences?*

H1: The suggested design process (in short)
The suggested process to investigate the prospects of leveraging data from interior and exterior sensors in the project is by following the design process structure knows as *Five Stage Process* by Interaction Design Foundation in (2021). This allows a structured progression from idea to prototypes and then on to a well-evaluated and refined artefact. For a more thorough description and motivation, see the planned execution of the H1 in Section 5.2.2 and Chapter 6.

As mentioned in Chapter 5 the *double diamond* process structure is also a suggested method in this use-case for its similarity in approach. First introduced by Design Council (2004), it consists of: *Discover*, *Define*, *Develop* and *Deliver*. While deploying these methods, it is vital to acquire holistic knowledge of the domain first, and then follow it up with the use of an iterative process loop, to produce and evaluate the artefact.

The *Five Stage Process* structure is to be carried out in the following stages : *Empathize*, *Define*, *Ideate*, *Prototype* and *Test* [89]. The following paragraphs cover the suggested adaptation of the *Five Stage Process* with regards to the project.

In the initial stages of following of the *Five Stage Process* (2021) process, it is recommended to develop an empathetic understanding of the problem areas (*Empathize* phase) and to clearly define the design space (*Define* phase) framed by the research questions. Literature studies are an effective means to gather domain knowledge while expert interviews are a core enabler in developing empathy. The outcome of these literature studies and expert interview will contain some valuable insights and data, therefore it is recommended to organise the resultant information using tables allowing for rapid sorting, management and aggregation of the sources. The authors utilised the embed tables on the online tool Notion [95], to furnish an overview of the domain knowledge. Furthermore, interviews when conducted with domain experts (company personnel in this Thesis' case) yields a cogent view of the design space. These suggested methods put together provide an understanding of prior work, empathy for the users and motivations of the stakeholders in the automotive industry.

To ensure the preservation of the multi-modal aspects of ideas produced, it is recommended that in the *Ideate* phase, tables be drawn based on all the probable combinations of the previously identified modalities. These tables can then used to cross-pollinate modalities which then lead up to a plethora of ideas and solutions. This also has the merit of all probable ideas being made available for filtering out the less-successful ones towards the end. Organised ideation processes such as the aforementioned method of sorting and filtering modalities lead the authors to extract the final set of multi-modal ideas. These ideas can be further filtered by using questionnaire surveys answered by domain experts and also the method, *Theory of Change* [70].

In the *Prototype* phase it is suggested that the ideas be put forth into scenarios as described by Nardi to investigate the feasibility of the few selected multi-modal solutions ideated in the previous step. Once these scenarios are evaluated for conceptually feasibility the ideas can be prototyped into more tangible / higher-fidelity mediums. The ideas' feasibility from a conceptual and practical standpoint is to be examined. The main goal of the *Prototype* phase is to develop solutions, that best addresses the design space defined in the previous three phases.

The goal of the evaluations in the *Test* phase is to determine the best variations of the prototypes. When prototypes fail in evaluations they are to be, set aside for improvement and re-evaluation. These evaluations are to be run with both domain experts and test participants. It is recommended to evaluate the prototypes' both from a subjective and from a usability standpoint. To that effect, a combination of structured interviews, System Usability Scale [82] and a task load index like the NASA TLX [80] or more specifically DALI [81] could be used. Analysis of the progressive trends in these evaluation results illustrates enhancements, and *'Pain Points'*, which in turn indicates faults and compromises in the actual artefact. For more details regarding this analysis see Figure 7.5 and 7.6.

Suggestion: Iterative process

As stated in Chapter 5, an iterative process loop is recommended wherein the results of the *Test* phase is analysed and then its findings are to be re-injected back into the *Prototype* phase resulting in the prototypes being improved or tweaked accordingly. This iterative process enabled the authors to effectively gauge the prospects of leveraging sensor data for enhancing the ride UX.

Results
Based on the last set of evaluation results, it can be observed that both the final artefacts, *LeBO* and *ADASight*, as covered in Section 7.1, placed high on the usability scale, and recorded relatively low scores for task loads in the DALI evaluations.

While it is, to a degree complicated to ascertain if a *improved vehicular user experiences* was attained through means of either one of the final artefacts, but it can be argued that the sensors being leveraged in both of the final artefacts were previously under-utilised by OEMs for user-facing features. The aforementioned evaluation results coupled with the generally positive responses in the surveys are adduced to corroborate for the *better use of occupant, interior and exterior sensor data* being achieved.

Therefore, taking into account prior work in this design space and the results of the artefacts' evaluations, it could be reasoned that H1 was valid.

7.2.2 Hypothesis 2 (H2)

> **RQ2:** *How can designers estimate the value, resulting from UX designs relying on vehicle sensor data, added to the vehicular user experience?*

H2: The suggested evaluation strategy (in short)
For a more thorough descriptions and motivation, see the planned the execution of **H2** in Section 5.2.2 and Chapter 6.

It is suggested to, throughout the design process, apply different evaluations methods, as it enables staying cost-effective as well as to grade ideas using the most suitable tool depending on the fidelity, amount of, form, etc., of the ideas. This is vital as the number of ideas, and their realized level of detail has an inverse relationship. Many vague ideas are best handled using qualitative methods, while few semi-finished artefacts are best handled with qualitative methods.

To cull the large amounts of ideas generated early in the suggested design process, a rapid and effective evaluation of ideas is recommended, as it enables culling ideas based on hard and soft limitations such as desirability, feasibility and viability. For example, in the case of this thesis' process, that involved removing ideas that were already in the works, ideas that were impossible to do, etc. using the method *Expert Interview.* The method utilised field experts know-how bundled with understanding collected from academia by the authors, in the evaluation.

After having produced the first iterations of prototypes, a rapid evaluation method that generates quantifiable values should be applied. This not only allows quick comparisons of ideas but also to follow the changes to furthered ideas' aspects. In the case of this thesis, the rapid but powerful SUS evaluation was used, which allowed the designers to see how, for example, the perceived desirability of the ideas changed across iterations, fidelities and mediums. From the values, positive trends, indicating successful improvements, and negative trends, indicating weaknesses, were observed - see Figure 7.5 and 7.6.

In the last iteration(s) of a design process, more non-rapid qualitative methods should be applied, such as task-based evaluations, as it allows observing how real users interact with the designs as well as allows measuring actual, or semi-actual, task load values. These load values are vital for features in a vehicle environment, as too high loads, such as attentive load, can, directly and indirectly, lead to severe accidents. Just measuring usability, which if too high might actually increase such loads, is not enough. Despite the relevance of load-related evaluations, the time-cost of performing task-based evaluations necessitates performing them, at the earliest, late in the design process when there are fewer artefacts. In this thesis, task-based evaluations were held, having participants fill out DALI forms after performing the tasks.

Results

As intended the amount of ideas/designs handled were progressively lowered after each iteration. Going from 37 to 6, to 4, to 2 based firstly on desirability, feasibility and viability. As the number of ideas decreased, the average quality increased and the differences between them shrunk, the more precise method was applied to differentiate them. Numerical scores allowed, as planned, further prioritisation of the 'best' ideas in terms of being successful results of following H1, the hypothesis to R1.

The last set of evaluations, coupled with earlier SUS evaluations, resulted in easy to understand measurements of the perceived usability as well as suitability (task load) for the design in a vehicle environment. The two final artefacts, *LeBO* and *ADASight*, as covered in Section 7.1, placed high on the usability scale (graded A and B respectively), and low to medium (respectively) on the driving activity load index. Further statistical patterns were, as covered previously, also detectable from the evaluation method outputs - increasing the level of detail designers can measure the value-added, if any, of designs leveraging occupant data. Furthermore, the evaluation methodology suggested is not limited to such designs, and could in fact also be applied to measure added value in other design areas; however the suitability of those, untested applications cannot be answered by the results produced in this thesis.

With the quality of the final artefacts in mind, as well as the level of detail the evaluation method choice gave the designers when measuring added value, one could argue that H2, the second hypothesis, was valid.

7.3 Design area primer

Excluding the research specific results produced as a result of thesis, the content covered by Chapter 1: **Introduction**, Chapter 2: **Background**, and Chapter 3: **Related work**, could also be seen as a result of this thesis. The afforementioned chapters feature design methodology, guidelines and recommendations for similar work. The chapters put together provides UX designers with a primer potentially useful for further work in the fields covered by this thesis, such as Automotive Interaction Design (AID) and Multimodal Interaction Design (MID), etc.

8
Discussion

This chapter covers the authors' thoughts and arguments for and against, the results of trying to answer the research problems, the process as it was conceived and executed, how generalisable the produced results are, what ethical considerations it could have as well as what future work the authors' argue could and should be done for all things produced by the thesis.

8.1 Results

The main results, indicating the validity of the hypotheses, can be surmised in the following ways:

Hypothesis 1 (H1):
Two applications were developed using the suggested process, both scoring above average on three consecutive formative usability evaluations (SUS), and one scoring low, the other medium, on a formative task load evaluation (DALI). Both of these result types are arguably indicative for the designed artefacts being a suitable and effective use of, an ADAS platform's sensor data, that has the potential to improve vehicular ride experiences.

Due to the nature of formative evaluations, the results should not be read as hard proof, nor disproof, of the actual validity of **H1** - instead read as indicators for it. To further the accuracy of these indications, a more summative study should be made as a continuation of the design process performed in this thesis - either as a customer survey after the designs have made it into real vehicles on the market or on prototypes as done in this thesis. More detailed suggestions for future academic work will be covered under Section 8.5.2, and suggestions about continuing with the artefacts in Section 8.5.1.

Hypothesis 2 (H2):
Without issue, the suggested evaluation strategy (**H2**) was executed alongside the design process this thesis' design process. The aim of evaluating, first separating the 'better' ideas from the 'worse', followed by identifying weaknesses to improve designs, was achieved. In the final phase, evaluatory data was produced that allowed conventional statistical tools to be applied - from which statistical patterns such as demographically specific weaknesses and strengths were identified. Disregarding assumptions made for the participant selection, mentioned at the start of Chapter 7,

that it was possible to measure the value added (negative and positive) to such an extent is arguably not only indicative but also, perhaps, proof that the suggested strategy allows *"...designers* to *estimate the value, resulting from UX designs relying on vehicle sensor data, added to the vehicular user experience"* .

Of course, as the term 'estimate' is open to interpretation, false assumptions would lead to poor estimates, and true assumptions would increase the accuracy of the estimates. This thesis did succeed in providing a way to estimate the added value but still has not provided a way to measure, nor increase the accuracy of those estimates. To do this, one would need to test the assumptions made that effects those estimates. In the case of the evaluation methods employed alongside the design process in this thesis, those assumptions were:

1. Normally distributed data.
2. Random sampling of population.
3. Adequately sized sample group.

For the first assumption, depending on the statistical analysis used, different aspects of the data is assumed to be normally distributed. In the case presented in this thesis, t-tests were employed and as such, both compared populations should be tested for normality. As for the second assumption, it is hard, if not impossible to tell if sampling is truly random or not - as each non-random sampling could also be the result of true random sampling. One could add a bias to only use a random sample that represents the global population's demographical data, for example, that would however not be truly random.

As for the third, and probably the only controllable aspect of an evaluation, one can perform what is called *Power Analysis* on the data to measure to what extent the assumption is reliable for a specific amount of participants. If the reliability, or power, is too low for the existing number of participants, designers can roughly calculate how many more participants they need to recruit. See the following subsection for an example of power analysis performed on the data produced in the design process of this thesis.

8.1.1 Power Analysis

As stated previously in section 4.8.3, a common tool for measuring how valid the assumptions of sample size are is that of *Power Analysis*. As this thesis is not a strictly summative study, nor one in the field of statistics, but rather a more formative, *Research through Design* one - the authors selected to utilise a guide for power analysis [86]. (2020). The guide explains the tool, how one might use it as well as gives a ready-to-use formula for calculating the power.

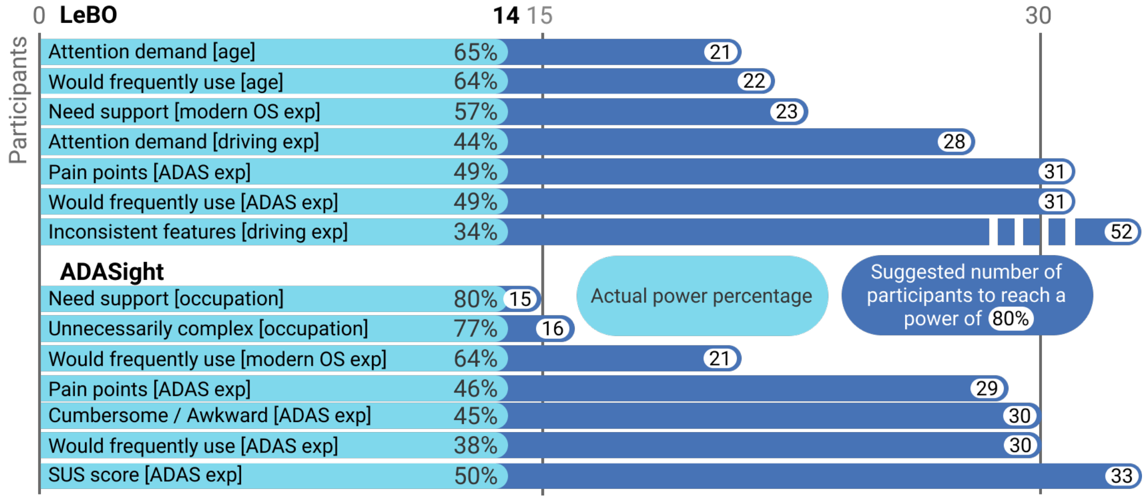

Figure 8.1: Power analysis results for LeBO and ADASight, x-axis showing number of participants.

Looking at the population data related to the p-values presented in Chapter 7, power values and suitable subject numbers for a power of 80% (or 0.8) were calculated - Figure 8.1 showing primarily the scale difference between the actual and the suitable number of participants, Figure 8.2 instead showing the scale of the actual and recommended power for the top (in terms of power) patterns.

Looking at results from the power analysis, it is quite clear that none of the statistical trends identified using t-tests have a high enough level of power, most under 50% and only two being close to the recommended 80% - see Figure 8.2. However, looking at the suggested number of participants, it is clear that with only a rough doubling of the participant numbers, all but one pattern would possibly achieve a power of at least 80% - see Figure 8.1.

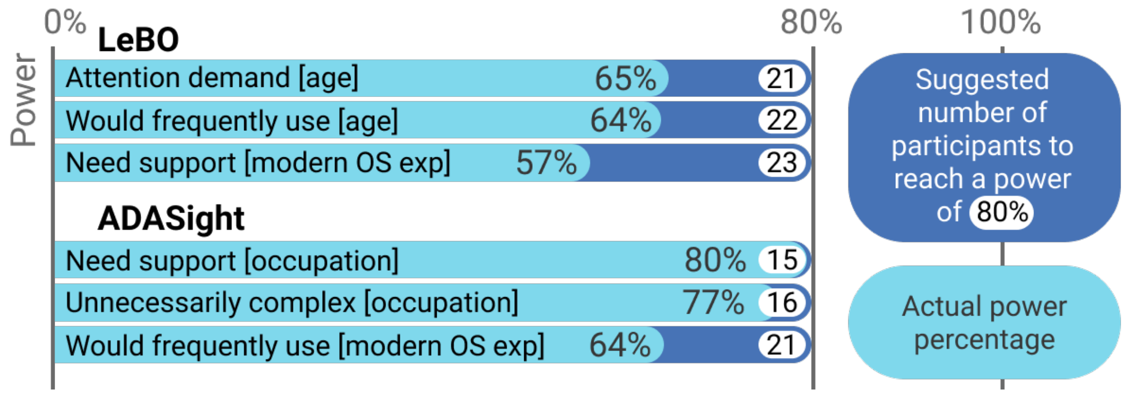

Figure 8.2: Power analysis results for LeBO and ADASight, x-axis showing power percentage.

8.2 Process

To ensure the efficient usage of time and other resources, the decision to closely adhere to a tried and tested design process method was made early on in the thesis. The *Five Stage Process* [89] provided a structured progression in the form of clearly defined phases which the authors followed, without much deviation. On that note, it was also deemed unnecessary to utilise time and resources for learning techniques or tools which do not immediately contribute to the end results - in-detail planning was made between each iteration, making adaptable what was needed, for example: only after ideas had been selected could fitting prototyping mediums be selected. This choice had the benefit of keeping the process cost-effective; however, it came with the detriment that **H1** had to be kept vague, without specific method suggestions - also, preparation/learning time with each method was shorter than if it had been chosen earlier than just before the relevant process phase.

In comparison, the evaluation strategy was largely decided upon in the Ideate phase, and as a result, cross iteration comparisons were possible. As such, that method selection could possibly have been made even before the design process started - allowing for more time to prepare for using them, such as background research into how previous projects have used those methods.

In the post-processing of the evaluation data, post-design processing if you will, power analysis revealed that the participant numbers were too low to provide a definite validation of the design process' success. Actively avoiding to look at the collected data until after all evaluations were made, to stay to conventions as well as avoid inflicting bias upon the evaluation process, hurt the results produced in the way that, had power analysis been applied earlier than it was, more evaluations could have taken place - increasing the reliability of the validations.

Something revealed by reading related work was the fact that commonly A/B testing is commonly applied in automotive interaction design, as a way to measure how to-vehicle-added designs affect different driving tasks as compared to driving without them. The limited scope of the thesis and the late inception of utilizing driving assistive modalities limited the thesis in such a way that using a real vehicle would take too much time and the fidelity of the final designs would have been too low for such tests. This leaves A/B testing to be covered in Section 8.5.2: **Future work - Academia**.

8.3 Generalisability

While the aim of the thesis was to contribute with design knowledge covering a specific design space, the produced result, being inspired by related work, could be applied in other design projects outside the targeted design space. The suitability of doing so, however, is not known by the authors of this theses, as the work performed did not take into account the specifics of such, other projects.

However, depending on the interpretation of the term ’generalisability’, the results of this thesis could be seen as highly generalisable. The defined design space, which was done before generating ideas, was far from fully utilised when looking at the final two ideas. A such, one could argue that by being applicable to all parts of the design space left unexplored, the results are indeed generalisable. A different take would be to ask the question, can the results be useful in a project covering non-design research? Perhaps not, and if that is one’s perspective, then no, the results produced are not generalisable; however, the authors deem such a perspective trivial, and such summarize by saying that the results are highly generalisable within the design space and fairly generalisable in similar design projects.

8.4 Ethical considerations and consequences

Any system which collects and records data in any form from a personal environment like the interior cabin of a vehicle has to carefully tread the line between convenience and privacy. What users share from their experiences, and who users share it with, should be up to their discretion. This is especially an issue when a financially driven organisation, e.g. a company, is situated to gain from unethical choices.

Privacy and Data gathering
The technology powering the in-vehicle user experiences, like the infrared cameras and in-cabin radars, also happens to be at the core of the safety sensing system, therefore establishing complexities regarding transferring complete control to users over these sensors. However, giving users control over their information, ensuring strict data privacy while ensuring a frictionless user experience is not always an easy and efficient option, but it is a necessary and desirable one.

The artefact or academic contribution may affect the field in such a way that the adoption of the technology leveraged is altered. This may be both positive or negative, with better UX, demand might increase for the technologies, high enough and companies might disregard caution in favour of profits - in which case society and consumers might be affected as a result. The complete opposite might also happen however, the results produced by this thesis might act as a proof of concept that good UX can be achieved through a careful study and as a result, automotive companies will go with the proven route for more satisfied customers.

Doing research in collaboration with a company
Writing a thesis for a company may put authors in ethical conundrums, e.g. if the company instructs the designer to hide unethical features such as constant surveillance from users, rendering that feature invisible for them. For students writing a thesis at a company, it may be hard to know what to do, either go against the company’s wishes for the principle of it or go along with it as ’they are the experts’, ‘if they say that is how you do it. . . ’, etc.

The solution, sometimes, is to do the ‘right’ thing and go against the company if needed, what is more, common and perhaps the most agreeable choice, however, is to merely report on what was done in the project as well as discuss what, if any, ethical considerations the reader should take note of in regards to the thesis.

In the case of this thesis' collaboration, no ethical issues, as far as the authors know, of such natures constrained the thesis. The company did not ask to hide nor alter the presented results. The company collaborated with, particularly the specific department, specializes in experimental development. The department and supervisors at the company desired a research through design thesis, specifically exploring future use possibilities of their products. This fact helped keep the thesis focused on the academic contributions, rather than on the artefact contribution.

8.4.1 Trust in automated systems

With the rapid advances and growth in ADAS features, it can be hypothesized that modern vehicles are safer and more connected than ever before, however, the acceptance and normalization of these systems are heavily reliant on the users' willingness to use these systems in their daily lives. This is also inferred by Trübswetter and Bengler as she states *"drivers' intention to use is crucial factor for system acceptance"* and by Adell as *"Acceptance is the degree to which an individual intends to use a system and, when available, to incorporate the system in his/her driving"*.

While acceptance and trust earned by a system on account of its user experience and feature effectiveness are generally positive; however, the over-reliance on these features by users on account of an abundance of trust in the same system is not. Overtrust is the poor calibration in which users' trust exceeds system capabilities; as a result, may potentially cause major safety issues when a road scenario falls short of the system's capabilities. Over-trust can be the compounded result of several factors, but the primary reason, in this case, appears to be false marketing and exaggerated branding.

Compared with traditional manually controlled vehicles, self-driving cars can reduce human error-induced crashes, which account for 93 per cent of those in the U.S. [107]. Furthermore, with better route planning and more efficient vehicle operation, self-driving cars can reduce road congestion and fuel emissions [108]. As Clark et al. state, riding in a self-driving car frees the driver from tasks and enables him to engage in his choice of leisure or productive non-driving activities.

As Bradsher notes, vast amounts of government funds, incentives and programs have accelerated the development and manufacture of electric vehicles and more notably, vehicles with next-generation technologies such as ADAS. China, evolving to become the world's biggest proponent of electric vehicles [111], is forcing automakers globally, to pick up the pace of development, production and market capture [110]. To counter the growing competition, car OEMs like Tesla [112], GM [113], etc are actively leveraging ADAS as differentiating factor, branding their respective systems as *"AutoPilot"* [114] and *"SuperCruise"* [115] respectively.

While the ADAS solutions in contemporary cars are advancing technologically to become more capable with each iteration, these systems are still primarily assistive in nature. However, OEMs continue marketing and branding these assistive technologies as fully autonomous or capable of *"hands-free driving"* [115]. This results in uninformed or non-technically savvy consumers misplacing a large amount of their trust in these systems and treating these systems are replacements for human attention, situation awareness and caution while on the road. This phenomenon of misplaced or over-trust has unfortunately resulted in a number of mishaps and road accidents resulting in loss of life and property [116], [117].

This has resulted in numerous lawsuits and several regulating bodies including the NTSB (National Transportation Safety Board) taking a closer look at not only the ADAS technology but the branding and marketing that is feeding consumers a false sense of trust and security. A 2020 report on a 2018 accident involving a Tesla running on *'AutoPilot'* by the NTSB [118] concluded that the system was not designed to handle certain scenarios and therefore should have been automatically disengaged when the scenario deemed to be beyond the capabilities of the system.

Furthermore, the regional court of Munich had ruled in 2020 [119], stating *"The use of the term 'Autopilot' and other formulations suggest that the vehicle is technically capable of fully autonomous driving..."* Autopilot is *"...a driver assistance system..."* the court said. *"While underway, driving without human intervention is not possible"*.

8.5 Future work

This section covers the authors' perspective on what future work could and should be done in regards to the artefacts produced, followed by the academic areas and contributions covered in this thesis. The first subsection touches upon immediate as well as long term suggestions. The second on how future academic and non-academic design work could further the work generated in this thesis, but also how and what, works using or learning from this thesis' work, should take place.

8.5.1 Artefacts

As with any artefact or product, there are means and avenues available for the improvement and developmental continuation. As a first step towards getting the artefacts produced in this thesis into the vehicles of everyday people, the faked and hard-coded parts would need to be replaced with real data, then later by live real data from ADAS platforms. To best reach usable, desirable, viable and feasible products, these last stages of development should be interlinked with task-based evaluations, such the one employed in iteration 3, inside real vehicles, with real people - as a low driving task load is vital for any addition to, and alteration of, in-vehicle interactions. To reach product status the applications would need to, at least, go through the following stages:

Figure 8.3: Additional visual spaces where the artefacts could be realized in the future. Clockwise from left, ADASight as a Heads-up Display, ADASight on the rear-view mirror, LeBO phone notification and ADASight on the digital instrument cluster.

1. Replace mock data with actual data logs from an ADAS platform.
2. Deploy applications on actual hardware, i.e. in a vehicle running Android Automotive.
3. Perform task-based evaluations using real-life driving tasks, A/B testing in applicable situations, to compare the difference the applications make.
4. Analyse and react to findings.
5. Replace data logs with live-connection to an ADAS platform.
6. Reevaluate applications again in the test track, and/or in traffic to further the accuracy of the evaluations.
7. At some point includes or provision other modalities of input and output, to allow for further preferences of interaction.

Another interesting avenue further development might take, is into the use of other display surfaces in vehicles, such as the HUD, rear-view mirrors, personal device or dashboard, see Figure 8.3 - some having the benefit of diverting the driver's gaze from the road less than the centre console does. In fact, feedback from the last evaluations, as well as from the company indicated that such avenues are not only possible but also suitable to pursue. Of course, other devices such as ones using augmented or virtual reality, might also prove to be wise choices for future explored platforms.

8.5.2 Academia

As mentioned at the start of this document, the areas explored in this thesis were purposely selected to be as unexplored as possible, while still relevant to both the company, the needs of the industry as well as interaction designers in general. By exploring, to the authors' knowledge, unknown areas, many aspects were not fully, or even partially addressed nor attempted. This is in line with what Gaver (2012) defined as the essence of *Research through Design*, that by cumulative effort design knowledge is created [56]. The hypotheses made and validated in this thesis proposed a suitable way of doing something, not the *most suitable* way; however, future academic endeavours can combine the experiences generated by the thesis with their own works, to encroach further towards it.

As for where future studies, and other academic works, should head, the authors suggests avenues that further validates but also alters the hypotheses made by this thesis - as by the essence of research through design, by cumulative work can new understanding be discovered. Interesting alterations could be changing the scope of the design process, either in time, design space, fiscal budget, etc. Designers could compare the suggested methodologies with a completely different way of achieving the same thing. Reworking discarded or kept ideas, exploring other modalities or perhaps technologies not yet developed as of the time of writing this thesis - the list goes on.

To further validate the outcomes of this thesis, more participants could be gathered and the final test setup recreated, to justify the assumptions of adequately sized participant groups - at least 30 would be needed, as stated in Section 8.1.1. Further demographical data could be gathered as well, to enable an analysis of more population groups.

To increase the accuracy of the data, the testing environment could be made more accurate by using a real vehicle, a real IVI system, real sensor data as well as performing real driving tasks on a real road. Having participants performing the same task without the designs would, over time, enable A/B testing of how beneficial the addition of the designs actually is. These types of test would, however, take far too much time for a similarly sized project, as long time users would be needed before testing, to mitigate both any wow factor new features have, as well as letting the participant get used to the new features. It is not hard to see, that this type of long term testing, is more suitable for on the market-artefacts, that have accustomed users as well as highest-fidelity realisations.

9
Conclusion

Based on the output of the final phase of development and evaluation of the process in the thesis, the following results were produced: two artefacts were designed and evaluated, the evaluations allowing the fine grained measurement of added strengths and weaknesses of the designs.

Power Analysis revealed that the participant group was inadequately sized, making the validation of the artefacts as 'good' only formative, and not strong enough to warrant the first hypotheses fully validated. The second set of results showed that the hypothesised evaluation strategy (**H2**), did allow measuring both positive and negative value added to vehicle riding experiences - coupled with the power analysis, it also allowed checking the assumptions of sample size, as well as addressing it. With such possibilities, proven by the use of t-tests and power analysis onto the resultant data, the second hypothesis (**H2**), is believed with confidence to be valid, i.e. that it allowed measuring positive and negative value added.

So, does the results produced diverge from current thinking? No, the underlying mentality the thesis was based on was research through design, and as mentioned several times throughout the thesis, only through several design projects is new understanding discovered. This thesis did not produce any controversial findings, primarily as it was based on best practises, guidelines and previous research. The thesis did not aim to explore alternate ways of doing established methodologies, instead it aimed to apply a methodology in a, to the authors' knowledge, unexplored design space. The methodology in this case being inspired by established thinking on more explored design spaces.

As was done in the early parts of the project, it is recommended to rely on both academic and industry knowledge when initiating a design process. As was the case in this thesis, the different parties come with not only a unique perspective, but also unique knowledge. When performing similar work to the work covered within this thesis, it is a good idea to utilise tools such as *Power Analysis* to make sure the assumptions made are true - for example, in thesis the participant group size was revealed to be too small. If statistically sound understanding of the value added is desired, the evaluation should aim to be summative rather than formative. To that end, a participant group size of at least 30, preferably higher, would provide the research a better chance at providing more reliable measurements.

As revealed by the power analysis, the sample size of the final evaluation was not of adequate size - making the validation of the suggested design process only partial; however, as stated at the start of the thesis, the process aimed to fill a void in design knowledge, and as such a partially validated methodology could perhaps be better than none when planning future design work. The truth remains though, that it should be known limitation of this thesis' knowledge contributions. The scope of the thesis did only allow acquisition of background knowledge up to a point, as a result the primer this thesis gives on the covered fields should not be interpreted as either final or all-encompassing. The same goes for any conclusions drawn by the results of the design work performed in the thesis. The final artefacts of the thesis were never tested in a fully accurate environment nor sufficiently compared to the alternative of not having the designs at all - these limitations should be known for any who aims to plan their work based on this thesis contributions.

To reconnect back the introduction of this thesis, and the research problem it aimed to solve; two things are of focus in this thesis - finding a good way to use fused sensor data going forward with automotive interaction design, and how to measure what is good and bad about designs leveraging it. While the artefacts and the methods choices behind them were not proven to be a definitive answer, the knowledge collected in preparation, and the learnings from the process and the formative evaluations is sure to be useful for any future work or related studies. The thesis was successful in providing a fine-grained way of measuring positive and negative impacts of such designs from early to late stage artefacts. The thesis also aimed to, and succeeded with, contributing generated ideas and artefacts to the company this thesis was done in collaboration with.

References

[1] L. Li et al. “Cognitive cars: A new frontier for ADAS research”. In: *IEEE Transactions on Intelligent Transportation Systems* 13.1 (2011), pp. 395–407.

[2] B. C. G. Aptiv LLC. *Path to Autonomous Driving Begins with Consumer and ADAS Adoption*. 2018. URL: https://www.aptiv.com/en/insights/article/path-to-autonomous-driving-begins-with-consumer-and-adas-adoption.

[3] A. Inc. *Apple CarPlay*. 2021. URL: https://www.apple.com/ios/carplay/.

[4] G. LLC. *Android Auto*. 2021. URL: https://www.android.com/auto/.

[5] J. Schneider and T. Nett. “Safety Issues of Integrating IVI and ADAS functionality via running Linux and AUTOSAR in parallel on a Dual-Core-System”. In: *Automotive-Safety & Security 2014* (2015).

[6] B. Latour. “Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts”. In: (1992).

[7] M. Hassenzahl et al. “It’s More Fun to Commute” An Example of Using Automotive Interaction Design to Promote Well-Being in Cars”. In: *Automotive User Interfaces*. Springer, 2017, pp. 95–120.

[8] A. Heide and K. Henning. “The “cognitive car”: A roadmap for research issues in the automotive sector”. In: *Annual reviews in control* 30.2 (2006), pp. 197–203.

[9] C. M. Wickens, M. E. Toplak, and D. L. Wiesenthal. “Cognitive failures as predictors of driving errors, lapses, and violations”. In: *Accident Analysis & Prevention* 40.3 (2008), pp. 1223–1233.

[10] R. Okuda, Y. Kajiwara, and K. Terashima. “A survey of technical trend of ADAS and autonomous driving”. In: *Technical Papers of 2014 International Symposium on VLSI Design, Automation and Test*. 2014, pp. 1–4. DOI: 10.1109/VLSI-DAT.2014.6834940.

[11] A. LLC. 2021. URL: https://www.aptiv.com/en/about.

[12] K. P. Clark. “Building Trust in Driverless Vehicles”. In: *Aptiv LLC.* (2020). URL: https://www.aptiv.com/en/insights/article/building-trust-in-driverless-vehicles.

[13] M. Arpe and C. Ensley. *The Future of Automotive Data Connectivity*. Tech. rep. 2021. URL: https://www.aptiv.com/en/insights/article/the-future-of-automotive-data-connectivity.

[14] R. A. Bolt. ““Put-That-There”: Voice and Gesture at the Graphics Interface”. In: *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH ’80. Seattle, Washington, USA: Association for Computing Machinery, 1980, pp. 262–270. ISBN: 0897910214.

DOI: 10.1145/800250.807503. URL: https://doi.org/10.1145/800250.807503.

[15] ACM. *Results for: All: "multimodal interaction"*. 2021. URL: https://dl.acm.org/action/doSearch?AllField=%5C%22Multimodal+interaction%5C%22&startPage=0&sortBy=Ppub_asc.

[16] ACM. *Results for: All: "multimodal interface"*. 2021. URL: https://dl.acm.org/action/doSearch?AllField=%5C%22Multimodal+Interface%5C%22.

[17] E. Freeman et al. "Multimodal Feedback in HCI: Haptics, Non-Speech Audio, and Their Applications". In: *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1*. Association for Computing Machinery and Morgan & Claypool, 2017, pp. 277–317. ISBN: 9781970001679. URL: https://doi.org/10.1145/3015783.3015792.

[18] J. R. Williams. "Guidelines for the use of multimedia in instruction". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 42. 20. SAGE Publications Sage CA: Los Angeles, CA. 1998, pp. 1447–1451.

[19] R. Sharma, V. I. Pavlović, and T. S. Huang. "Toward multimodal human–computer interface". In: *Advances in image processing and understanding: a festschrift for Thomas S Huang*. World Scientific, 2002, pp. 349–365.

[20] Y. Cao and M. Theune. "The use of modality in in-vehicle information presentation: A brief overview". In: *sage* 21 (2010), p. 37.

[21] W. H. Organization. *World health statistics 2019: monitoring health for the SDGs sustainable development goals*. World Health Organization, 2019.

[22] P. George. *The Infotainment Systems Are Still Bad*. 2020. URL: https://www.thedrive.com/tech/34350/the-infotainment-systems-are-still-bad.

[23] T. Winograd. "From Computing Machinery to Interaction Design, i: Denning, P & Metcalfe, R (Eds.): Beyond Calculation: The Next Fifty Years of Computing". In: (1997).

[24] K. Salminen et al. "Unimodal and Multimodal Signals to Support Control Transitions in Semiautonomous Vehicles". In: *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '19. Utrecht, Netherlands: Association for Computing Machinery, 2019, pp. 308–318. ISBN: 9781450368841. DOI: 10.1145/3342197.3344522. URL: https://doi.org/10.1145/3342197.3344522.

[25] A. Schmidt et al. "Automotive User Interfaces: Human Computer Interaction in the Car". In: CHI EA '10. Atlanta, Georgia, USA: Association for Computing Machinery, 2010, pp. 3177–3180. ISBN: 9781605589305. DOI: 10.1145/1753846.1753949. URL: https://doi.org/10.1145/1753846.1753949.

[26] J. Löchner et al. "Validating advanced driver assistance systems (ADAS) using comprehensive, loss-free in-vehicle measurements". In: *16. Internationales Stuttgarter Symposium*. Springer. 2016, pp. 1143–1154.

[27] G. De Vos. *Safety and Convenience Innovations Powered by Interior Sensing*. 2020. URL: https://www.aptiv.com/insights/article/safety-and-convenience-innovations-powered-by-interior-sensing.

[28] D. Kern and A. Schmidt. "Design Space for Driver-Based Automotive User Interfaces". In: *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '09. Essen, Germany: Association for Computing Machinery, 2009, pp. 3–10. ISBN: 9781605585710. DOI: 10.1145/1620509.1620511. URL: https://doi.org/10.1145/1620509.1620511.

[29] B. Pfleging, S. Schneegass, and A. Schmidt. "Multimodal interaction in the car: combining speech and gestures on the steering wheel". In: *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2012, pp. 155–162.

[30] L.-B. Bilius and R.-D. Vatavu. "A synopsis of input modalities for in-vehicle infotainment and consumption of interactive media". In: *ACM International Conference on Interactive Media Experiences*. 2020, pp. 195–199.

[31] J. Campbell et al. "Human factors design guidance for driver-vehicle interfaces". In: *National Highway Traffic Safety Administration, Washington, DC, DOT HS* 812 (2016), p. 360.

[32] N. Trübswetter and K. Bengler. "Why should I use ADAS? Advanced driver assistance systems and the elderly: Knowledge, experience and usage barriers". In: (2013).

[33] G. Prabhakar et al. "Interactive gaze and finger controlled HUD for cars". In: *Journal on Multimodal User Interfaces* 14.1 (2020), pp. 101–121.

[34] F. Roider et al. "The effects of situational demands on gaze, speech and gesture input in the vehicle". In: *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2017, pp. 94–102.

[35] M. Knobel et al. "Clique trip: Feeling related in different cars". In: *Proceedings of the designing interactive systems conference*. 2012, pp. 29–37.

[36] S. Krome et al. "Autojam: Exploring interactive music experiences in stop-and-go traffic". In: *Proceedings of the 2017 conference on designing interactive systems*. 2017, pp. 441–450.

[37] D.-B. Vo and S. Brewster. "Investigating the effect of tactile input and output locations for drivers' hands on in-car tasks performance". In: *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2020, pp. 1–8.

[38] A. Farooq, G. Evreinov, and R. Raisamo. "Reducing driver distraction by improving secondary task performance through multimodal touchscreen interaction". In: *SN Applied Sciences* 1.8 (2019), pp. 1–14.

[39] L. M. Reeves et al. "Guidelines for multimodal user interface design". In: *Communications of the ACM* 47.1 (2004), pp. 57–59.

[40] Google, Inc. *Material Design*. 2021. URL: https://developers.google.com/cars/design.

[41] A. Ratzka and C. Wolff. "A pattern-based methodology for multimodal interaction design". In: *International Conference on Text, Speech and Dialogue*. Springer. 2006, pp. 677–686.

[42] N. B. Sarter. "Multimodal information presentation: Design guidance and research challenges". In: *International journal of industrial ergonomics* 36.5 (2006), pp. 439–445.

[43] T. Tourwé et al. "Ontology-driven elicitation of multimodal user interface design recommendations". In: *23rd Benelux conference on Artificial Intelligence (BNAIC 2011)*. 2011.

[44] I. Pettersson and W. Ju. "Design techniques for exploring automotive interaction in the drive towards automation". In: *Proceedings of the 2017 conference on designing interactive systems*. 2017, pp. 147–160.

[45] F. Flemisch et al. "Towards a dynamic balance between humans and automation: authority, ability, responsibility and control in shared and cooperative control situations". In: *Cognition, Technology & Work* 14.1 (2012), pp. 3–18.

[46] S. Diefenbach, N. Kolb, and M. Hassenzahl. "The'hedonic'in human-computer interaction: history, contributions, and future research directions". In: *Proceedings of the 2014 conference on Designing interactive systems*. 2014, pp. 305–314.

[47] K. R. May, T. M. Gable, and B. N. Walker. "A multimodal air gesture interface for in vehicle menu navigation". In: *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2014, pp. 1–6.

[48] C. Müller and G. Weinberg. "Multimodal input in the car, today and tomorrow". In: *IEEE MultiMedia* 18.1 (2011), pp. 98–103.

[49] AUDI. *MMI Touch Response System - AUDI*. 2020. URL: `https://www.audiusa.com/us/web/en/innovation/audi-mmi.html`.

[50] BMW. *Introducing BMW iDrive - BMW*. 2021. URL: `https://www.bmw.com/en/events/ces/2cars.html`.

[51] V. Fang and L. Ainsworth. "Designing & rapid prototyping a gesture-enabled steering wheel". In: *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2012, pp. 17–19.

[52] C. D. Wickens. "The structure of attentional resources". In: *Attention and performance VIII* 8 (1980), pp. 239–257.

[53] A. LLC. *Advanced Safety - CREATING A SAFER TODAY AND TOMORROW*. 2021. URL: `https://www.aptiv.com/en/solutions/advanced-safety`.

[54] "Free for All". In: *Management Science* 14.4 (1967), B-141-B–146. DOI: `10.1287/mnsc.14.4.B141`. eprint: `https://doi.org/10.1287/mnsc.14.4.B141`. URL: `https://doi.org/10.1287/mnsc.14.4.B141`.

[55] H. W. Rittel and M. M. Webber. "Dilemmas in a general theory of planning". In: *Policy sciences* 4.2 (1973), pp. 155–169.

[56] W. Gaver. "What Should We Expect from Research through Design?" In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. Austin, Texas, USA: Association for Computing Machinery, 2012, pp. 937–946. ISBN: 9781450310154. DOI: `10.1145/2207676.2208538`. URL: `https://doi.org/10.1145/2207676.2208538`.

[57] IDEO. *IDEO HCD Toolkit Methods*. 2021. URL: https://www.designkit.org/methods/rapid-prototyping.

[58] K. Popper. *Conjectures and refutations: The growth of scientific knowledge*. routledge, 1963.

[59] I. Lakatos. "Science and pseudoscience". In: *Philosophical papers* 1 (1978), pp. 1–7.

[60] R. Collins. "Why the social sciences won't become high-consensus, rapid-discovery science". In: *Sociological forum*. Vol. 9. 2. Springer. 1994, pp. 155–177.

[61] E. B.-N. Sanders and P. J. Stappers. "Co-creation and the new landscapes of design". In: *Co-design* 4.1 (2008), pp. 5–18.

[62] M. Muller and A. Druin. "Participatory Design: The Third Space in HCI". In: *Handbook of HCI* (Jan. 2002).

[63] Y. Yamauchi. "Power of peripheral designers: how users learn to design". In: *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*. 2009, pp. 1–12.

[64] A. Light and R. Luckin. "Designing for social justice: people, technology, learning". In: (2008).

[65] Chalmers Tekniska Högskola. *Chalmers Library*. 2021. URL: http://www.lib.chalmers.se/en/.

[66] Scopus. *Scopus - Expertly curated abstract & citation database*. 2021. URL: https://www.elsevier.com/solutions/scopus.

[67] IDEO. *Expert Interview*. 2021. URL: https://www.designkit.org/methods/expert-interview.

[68] S. T. Y. Friis D Rikke. *Learn How to Use the Best Ideation Methods: SCAMPER*. 2020. URL: https://www.interaction-design.org/literature/article/learn-how-to-use-the-best-ideation-methods-scamper.

[69] A. F. Osborn. "Applied Imagination: Principles and Procedures of Creative Thinking". In: (1953).

[70] IDEO. *Theory of Change*. 2021. URL: https://www.designkit.org/methods/explore-your-theory-of-change.

[71] IDEO. *Rapid Prototyping*. 2021. URL: https://www.designkit.org/methods/rapid-prototyping.

[72] usability.gov. *Prototyping*. 2021. URL: https://www.usability.gov/how-to-and-tools/methods/prototyping.html.

[73] B. A. Nardi. "The use of scenarios in design". In: *ACM SIGCHI Bulletin* 24.4 (1992), pp. 13–14.

[74] H. L. McQuaid, A. Goel, and M. McManus. "When You Can't Talk to Customers: Using Storyboards and Narratives to Elicit Empathy for Users". In: *Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces*. DPPI '03. Pittsburgh, PA, USA: Association for Computing Machinery, 2003, pp. 120–125. ISBN: 1581136528. DOI: 10.1145/782896.782926. URL: https://doi.org/10.1145/782896.782926.

[75] K. N. Truong, G. R. Hayes, and G. D. Abowd. "Storyboarding: an empirical determination of best practices and effective guidelines". In: *Proceedings of the 6th conference on Designing Interactive systems*. 2006, pp. 12–21.

[76] J. F. Kelley. “An Empirical Methodology for Writing User-Friendly Natural Language Computer Applications”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’83. Boston, Massachusetts, USA: Association for Computing Machinery, 1983, pp. 193–196. ISBN: 0897911210. DOI: 10.1145/800045.801609. URL: https://doi.org/10.1145/800045.801609.

[77] J. F. Kelley. *CAL - a Natural Language Program Developed with the OZ Paradigm: Implications for Supercomputing Systems*. Research Report. IBM Thomas J. Watson Research Center, 1985. URL: https://books.google.se/books?id=gNgaHQAACAAJ.

[78] O. Karras et al. “Video as a by-product of digital prototyping: capturing the dynamic aspect of interaction”. In: *2017 IEEE 25th International Requirements Engineering Conference Workshops (REW)*. IEEE. 2017, pp. 118–124.

[79] W. E. Mackay and A. L. Fayard. “Video brainstorming and prototyping: techniques for participatory design”. In: *CHI’99 extended abstracts on Human factors in computing systems*. 1999, pp. 118–119.

[80] S. G. Hart. “NASA-task load index (NASA-TLX); 20 years later”. In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 50. 9. Sage publications Sage CA: Los Angeles, CA. 2006, pp. 904–908.

[81] A. Pauzié. “A method to assess the driver mental workload: The driving activity load index (DALI)”. In: *IET Intelligent Transport Systems* 2.4 (2008), pp. 315–322.

[82] usability.gov. *System Usability Scale : SUS*. 2021. URL: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html.

[83] R. Likert. *A technique for the measurement of attitudes.* Vol. 22. 140. 1932, pp. 11–55. URL: https://legacy.voteview.com/pdf/Likert_1932.pdf.

[84] J. Sauro. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. CreateSpace Independent Publishing Platform, 2011. ISBN: 9781461062707. URL: https://books.google.se/books?id=BL0kKQEACAAJ.

[85] Z. Bobbitt. *How to Perform t-Tests in Google Sheets*. 2020. URL: https://www.statology.org/t-test-google-sheets/.

[86] A. Coppock. *10 Things to Know About Statistical Power*. 2020. URL: https://egap.org/resource/10-things-to-know-about-statistical-power/.

[87] J. Neyman. *Mathematical Statistics: A Historical Sketch with Particular Reference to the United States, in: On the History of Statistics and Probability*. Ed. by D. B. Owens. Marcel Dekker Inc, NY, 1976.

[88] M. McCloskey. *Turn User Goals into Task Scenarios for Usability Testing.* 2014. URL: https://www.nngroup.com/articles/task-scenarios-usability-testing/.

[89] Interaction Design Foundation. *Design Thinking*. 2021. URL: https://www.interaction-design.org/literature/topics/design-thinking.

[90] Design Council. *Double Diamond Framework - Design Council.* 2004. URL: https://www.designcouncil.org.uk/news-opinion/what-framework-innovation-design-councils-evolved-double-diamond.

[91] Skype Communications S.a r.l. *Skype*. 2021. URL: https://www.skype.com/en/.

[92] Microsoft. *Microsoft Teams*. 2021. URL: https://www.microsoft.com/en/microsoft-teams/group-chat-software.

[93] Zoom Video Communications, Inc. *Zoom*. 2021. URL: https://zoom.us.

[94] Instructure, Inc. *Canvas LMS*. 2021. URL: https://www.instructure.com/product/canvas/higher-education/lms.

[95] Notion. *Notion.so*. 2021. URL: https://www.notion.so/product.

[96] Overleaf. *Overleaf*. 2021. URL: https://www.overleaf.com/.

[97] D. Gray-Grant. *Why you should consider keeping a research diary*. The Publication Coach. 2018. URL: https://www.publicationcoach.com/research-diary/.

[98] D. Gray. *David Gray Discusses Keeping a Research Diary*. SAGE Publishing. 2014. URL: https://www.youtube.com/watch?v=5fLLe6m0ZhM.

[99] Miro. *Miro.com*. 2021. URL: https://www.miro.com.

[100] Procreate. *Procreate*. 2021. URL: https://procreate.art/ipad.

[101] Figma, Inc. *Figma*. 2021. URL: https://www.figma.com.

[102] Google, Inc. *Material Design*. 2021. URL: https://material.io/design/introduction.

[103] Aptiv LLC. *Aptiv Brand Guidelines*. 2021. URL: https://www.aptiv.com/aptiv-brand-guidelines.

[104] Google, Inc. *Material Theming*. 2021. URL: https://material.io/design/material-theming/overview.html#material-theming.

[105] Figma Inc. *Figma Mirror*. 2021. URL: https://apps.apple.com/us/app/figma-mirror/id1152747299.

[106] E. Adell. *Driver experience and acceptance of driver support systems-a case of speed adaptation*. Lund University, 2009.

[107] P. Liu and Z. Xu. “Public attitude toward self-driving vehicles on public roads: Direct experience changed ambivalent people to be more positive”. In: *Technological Forecasting and Social Change* 151 (2020), p. 119827.

[108] D. J. Fagnant and K. Kockelman. “Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations”. In: *Transportation Research Part A: Policy and Practice* 77 (2015), pp. 167–181.

[109] H. E. Clark et al. “Limited Self-Driving Vehicle Automation: Age Differences in the Takeover of Vehicle Control, Engagement in Non-driving-related Activities and Opinions of the Technology.” In: (2016).

[110] K. Bradsher. “China hastens the world toward an Electric-Car future”. In: *New York Times* 7 (2017).

[111] L. J. Hui He. *How China put nearly 5 million new energy vehicles on the road in one decade*. 2021. URL: https://theicct.org/blog/staff/china-new-energy-vehicles-jan2021.

[112] I. Tesla. *Tesla Official Website*. 2021. URL: https://www.tesla.com/.

[113] G. M. Company. *General Motors Company Official Website*. 2021. URL: https://www.gm.com/.

[114] I. Tesla. *Tesla Autopilot*. 2021. URL: https://www.tesla.com/autopilot.

[115] G. M. Company. *GM SuperCruise*. 2021. URL: https://www.cadillac.com/world-of-cadillac/innovation/super-cruise.

[116] B. REUTERS. “Tesla Autopilot crash driver ’was playing video game’”. In: (2020). URL: https://www.bbc.com/news/technology-51645566.

[117] K. Lyons. “Two people killed in fiery Tesla crash with no one driving”. In: (2021). URL: https://www.theverge.com/2021/4/18/22390612/two-people-killed-fiery-tesla-crash-no-driver.

[118] N. T. S. Board. *Report : Collision Between a Sport Utility Vehicle Operating With Partial Driving Automation and a Crash Attenuator*. 2020. URL: https://www.ntsb.gov/news/events/Documents/2020-HWY18FH011-BMG-abstract.pdf.

[119] J. Ewing. “German Court Says Tesla Self-Driving Claims Are Misleading”. In: (2020). URL: https://www.nytimes.com/2020/07/14/business/tesla-autopilot-germany.html.

A

Prototypes

A.1 Iteration 1: Storyboards

A.1.1 Car behind

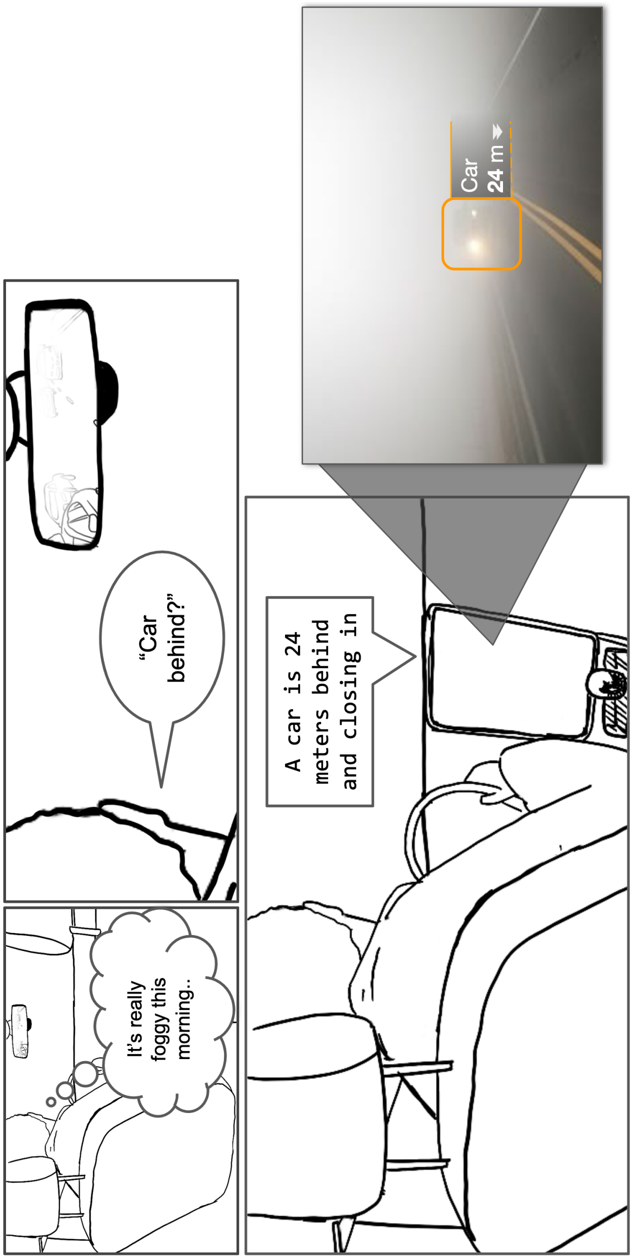

Figure A.1: Storyboard for the idea *Car behind*.

A.1.2 Add to route

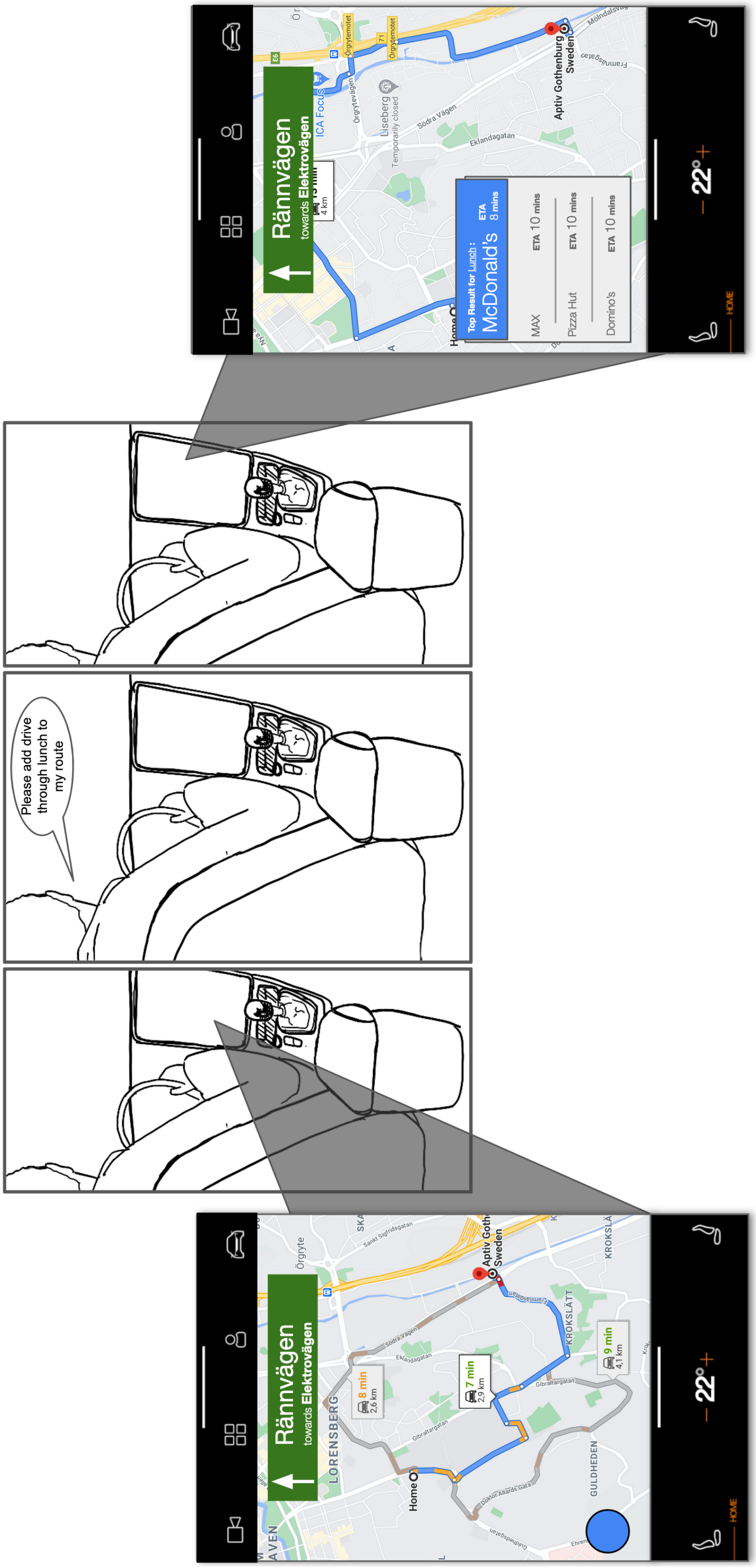

Figure A.2: Storyboard for the idea *Add to route*.

A.1.3 Intercom

Figure A.3: Storyboard for the idea *Intercom*.

A.1.4 Left-behind object

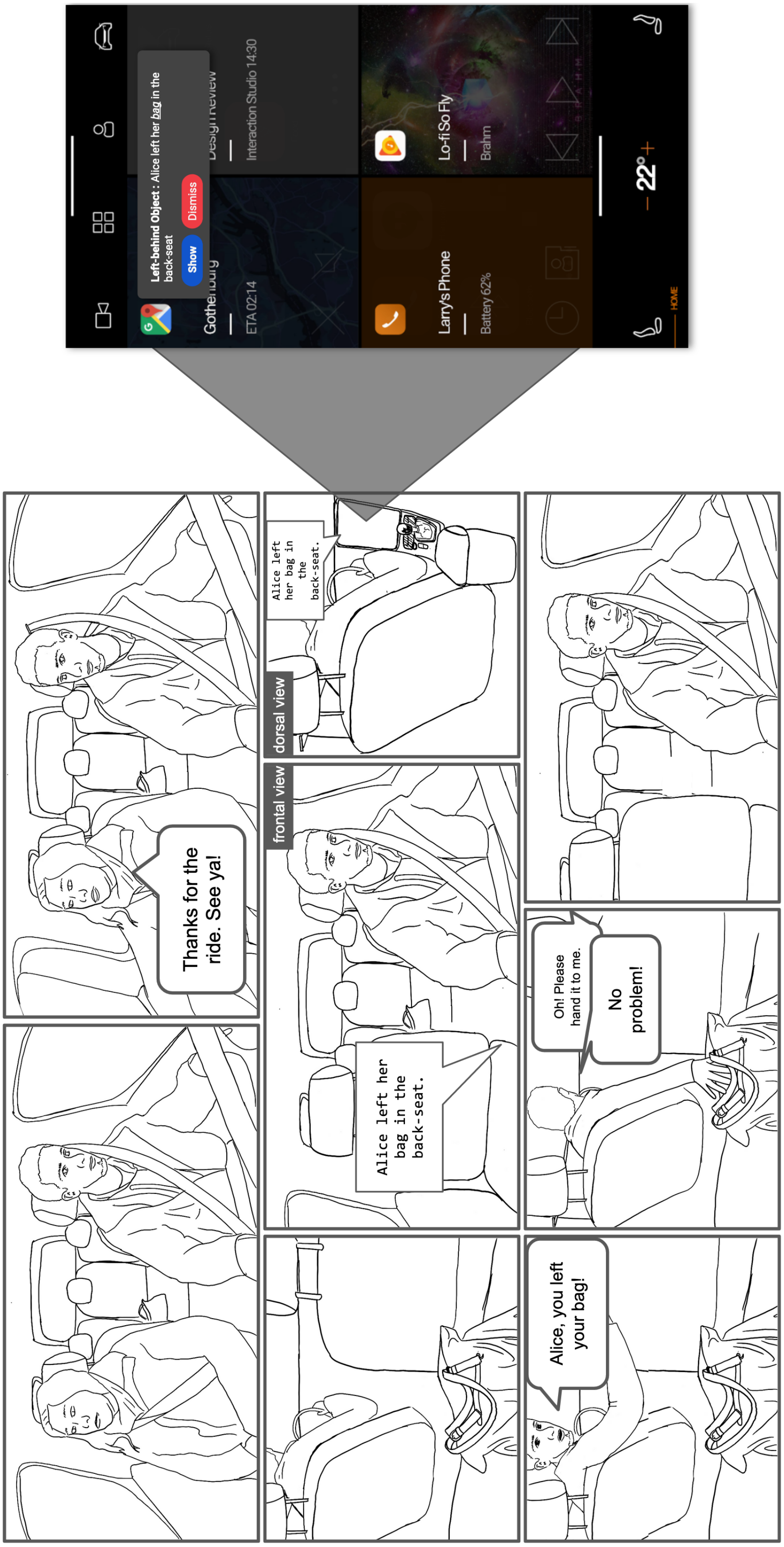

Figure A.4: Storyboard for the idea *Left-behind object*.

A.1.5 Magnification

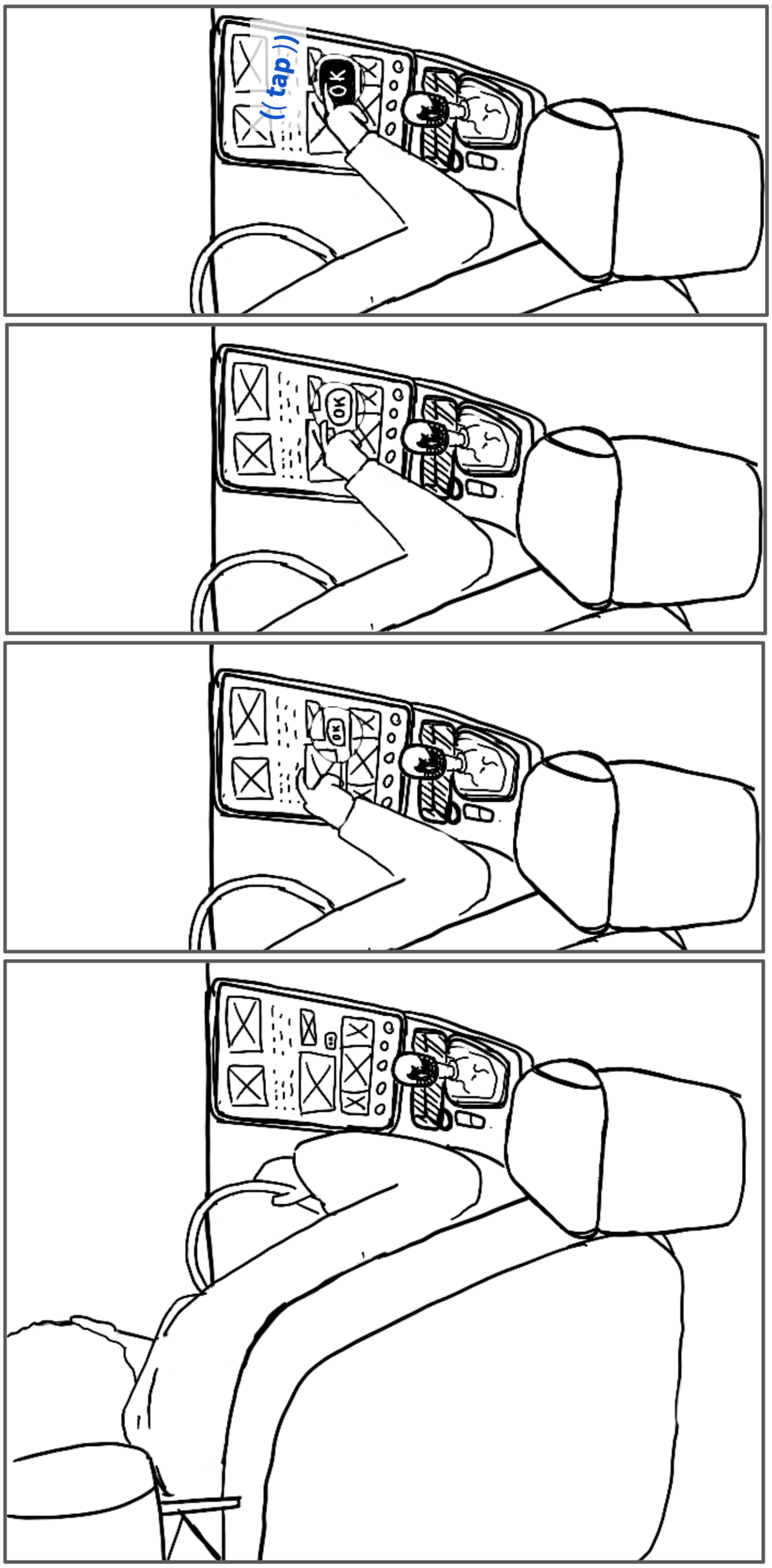

Figure A.5: Storyboard for the idea *Magnification*.

A.1.6 Tooltip

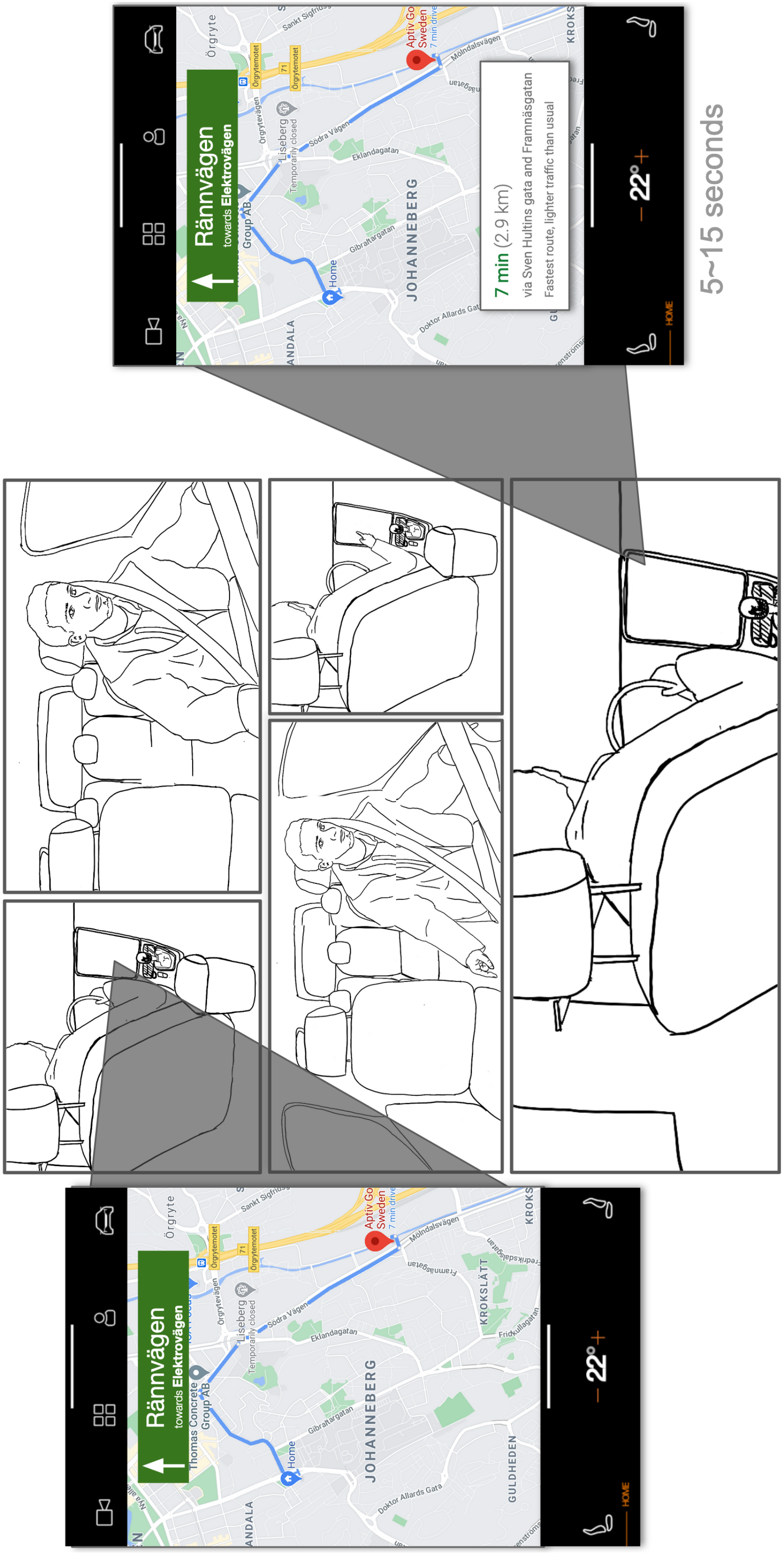

Figure A.6: Storyboard for the idea *Tooltip*.

A.2 Iteration 2: Mockups

A.2.1 Car behind

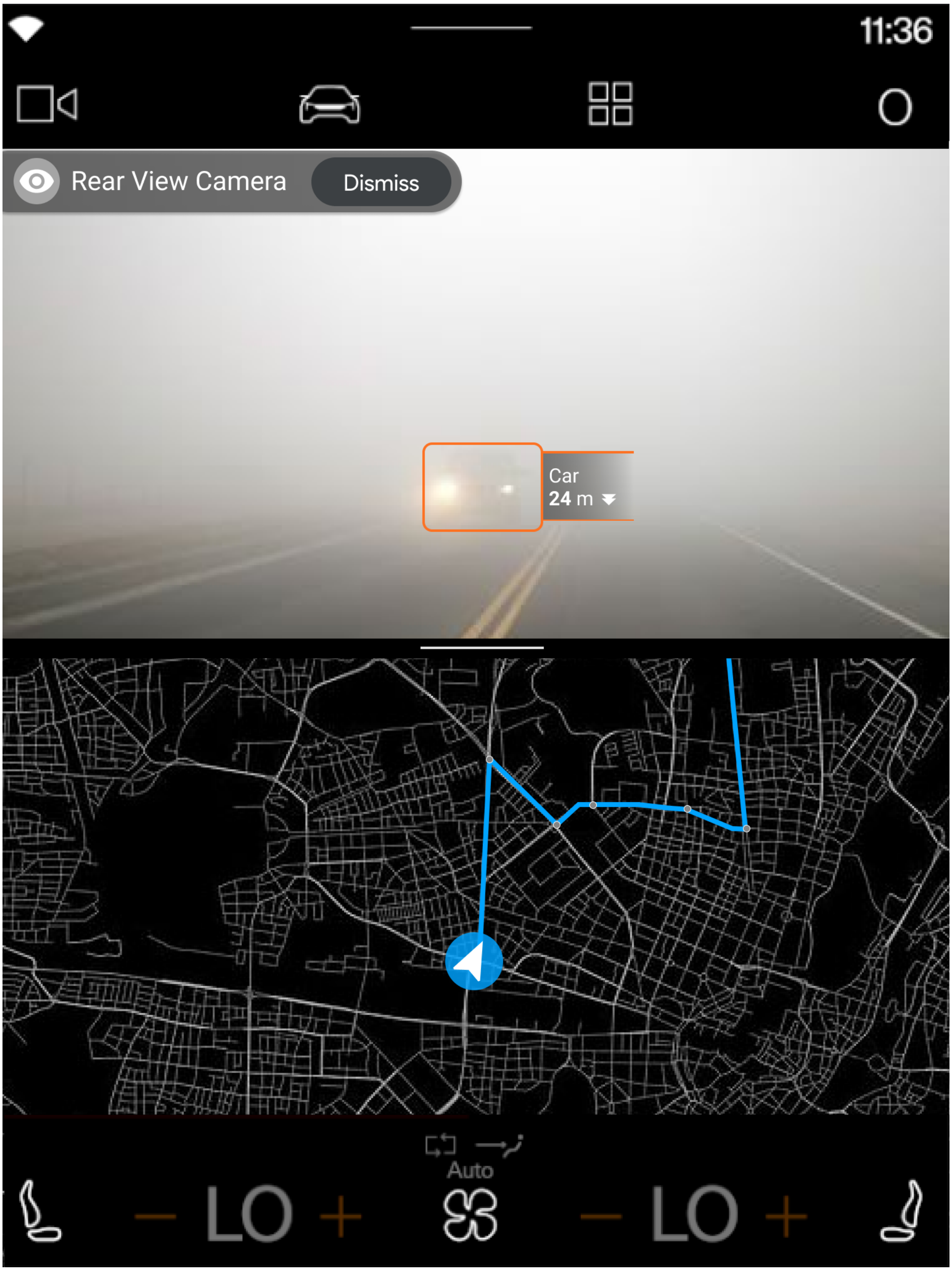

Figure A.7: Screenshot of the mockup for the idea *Car behind*.

A.2.2 Left-behind object

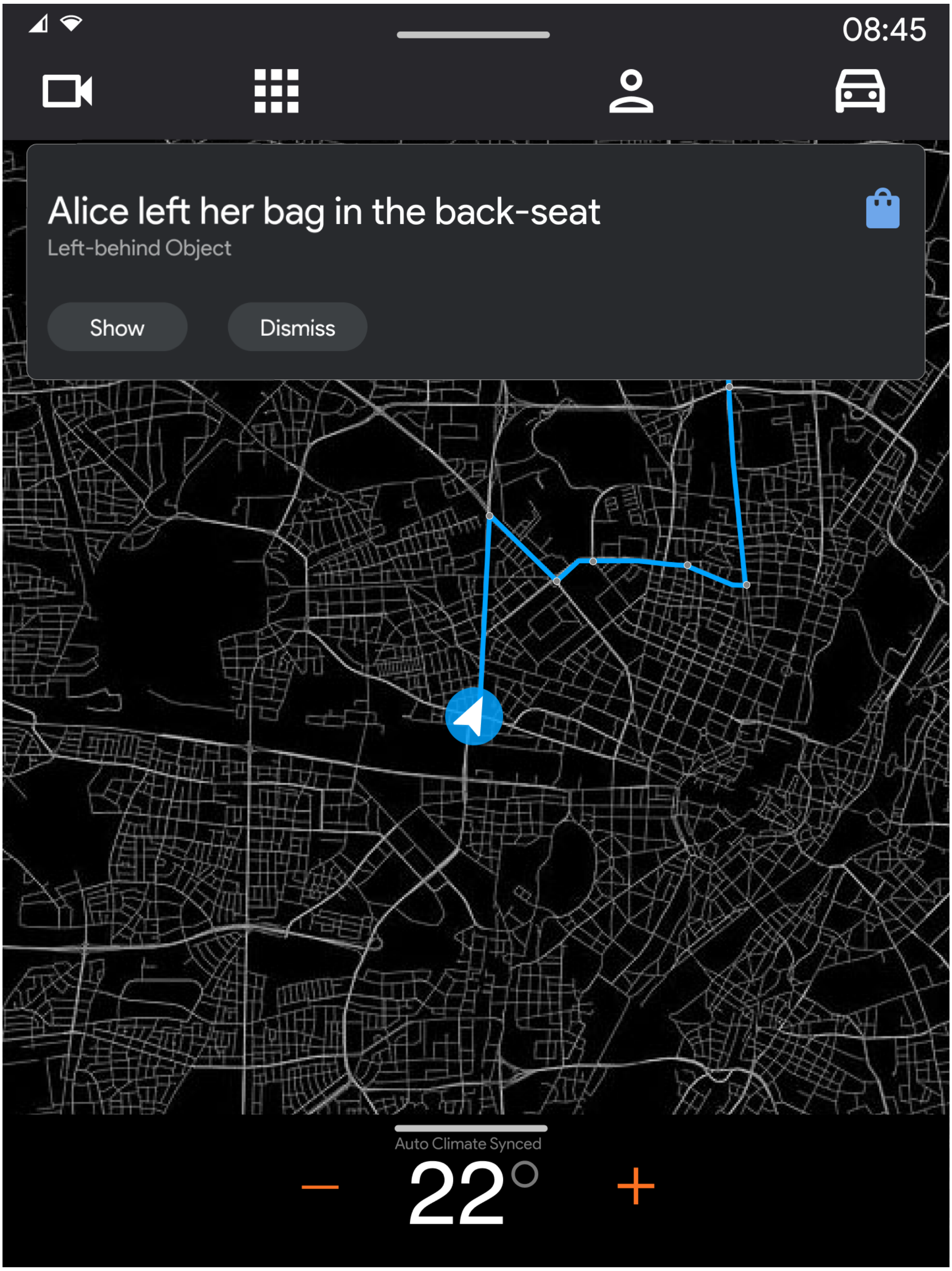

Figure A.8: Screenshot of the mockup, view A, for the idea *Left-behind object*.

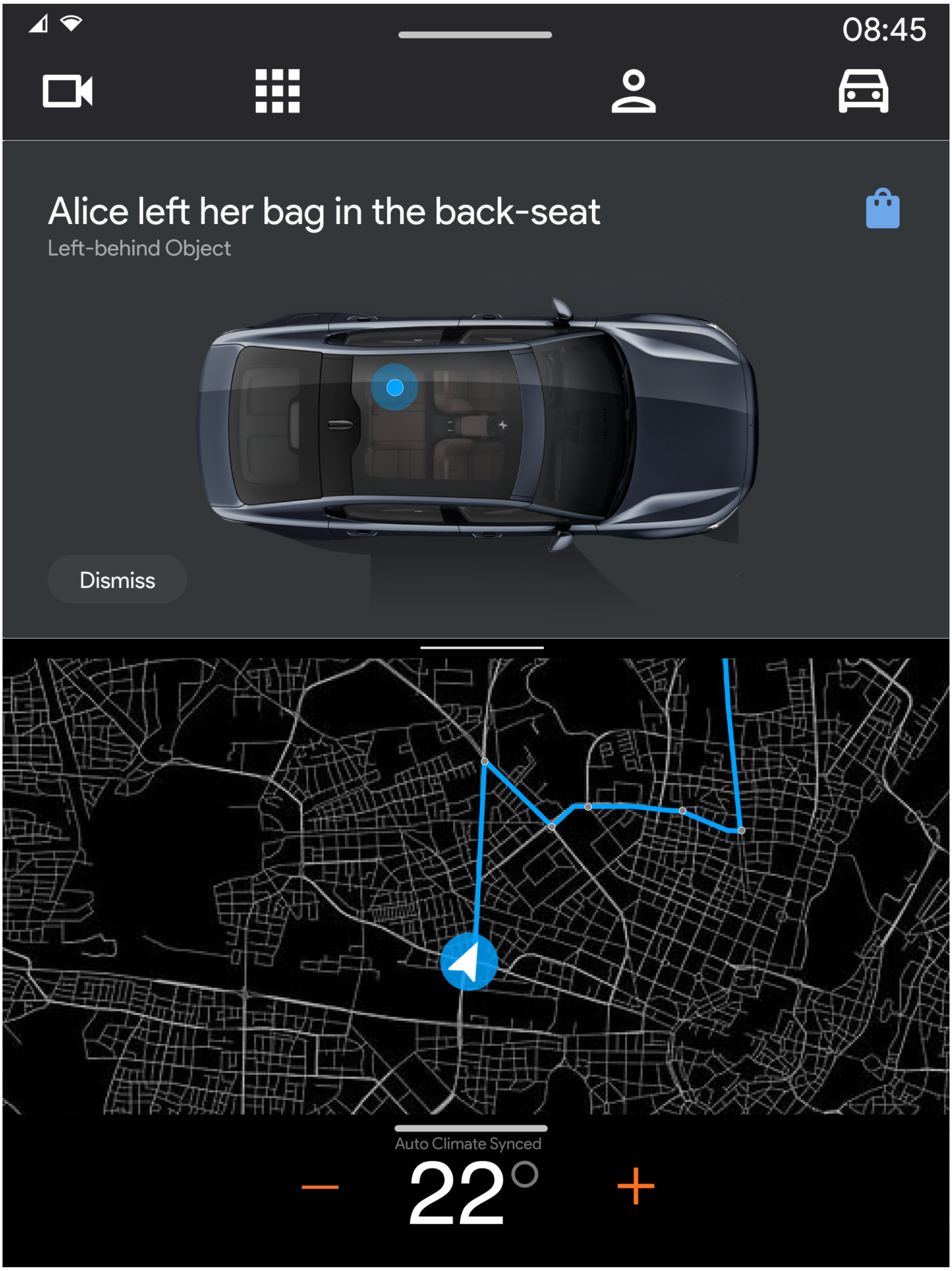

Figure A.9: Screenshot of the mockup, view B, for the idea *Left-behind object*.

A.2.3 Magnification

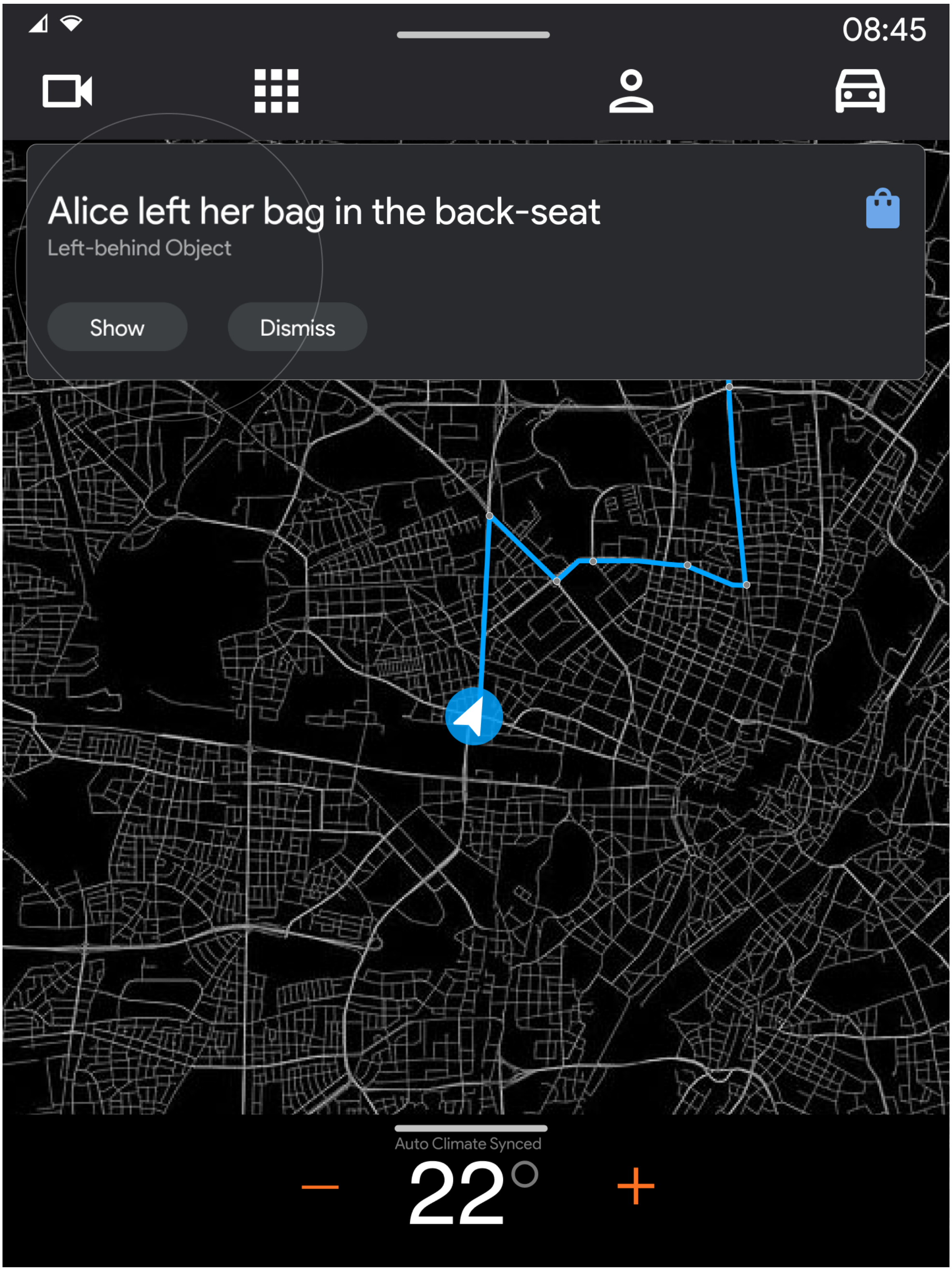

Figure A.10: Screenshot of the mockup, stage A, for the idea *Magnification*.

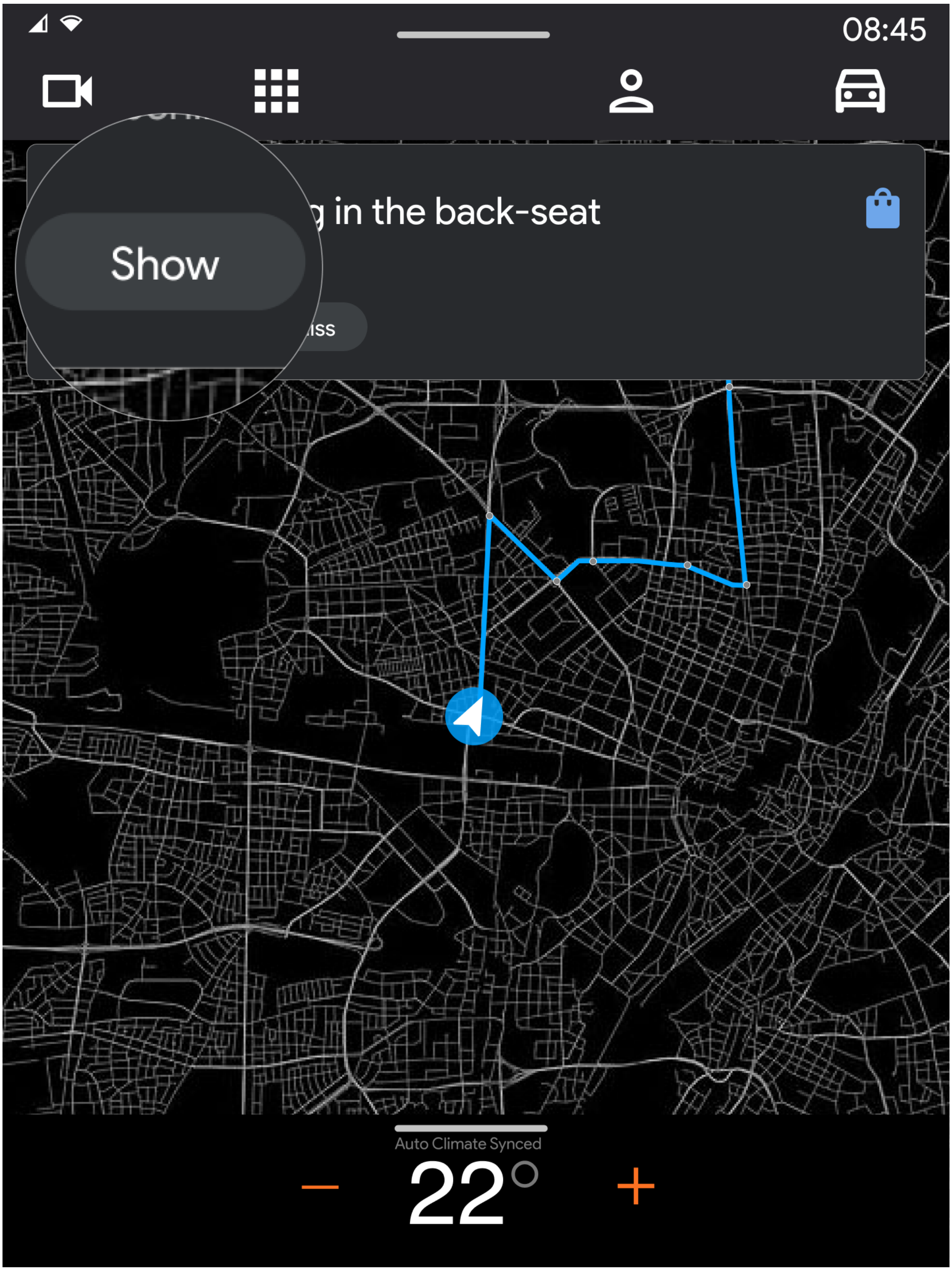

Figure A.11: Screenshot of the mockup, stage B, for the idea *Magnification*.

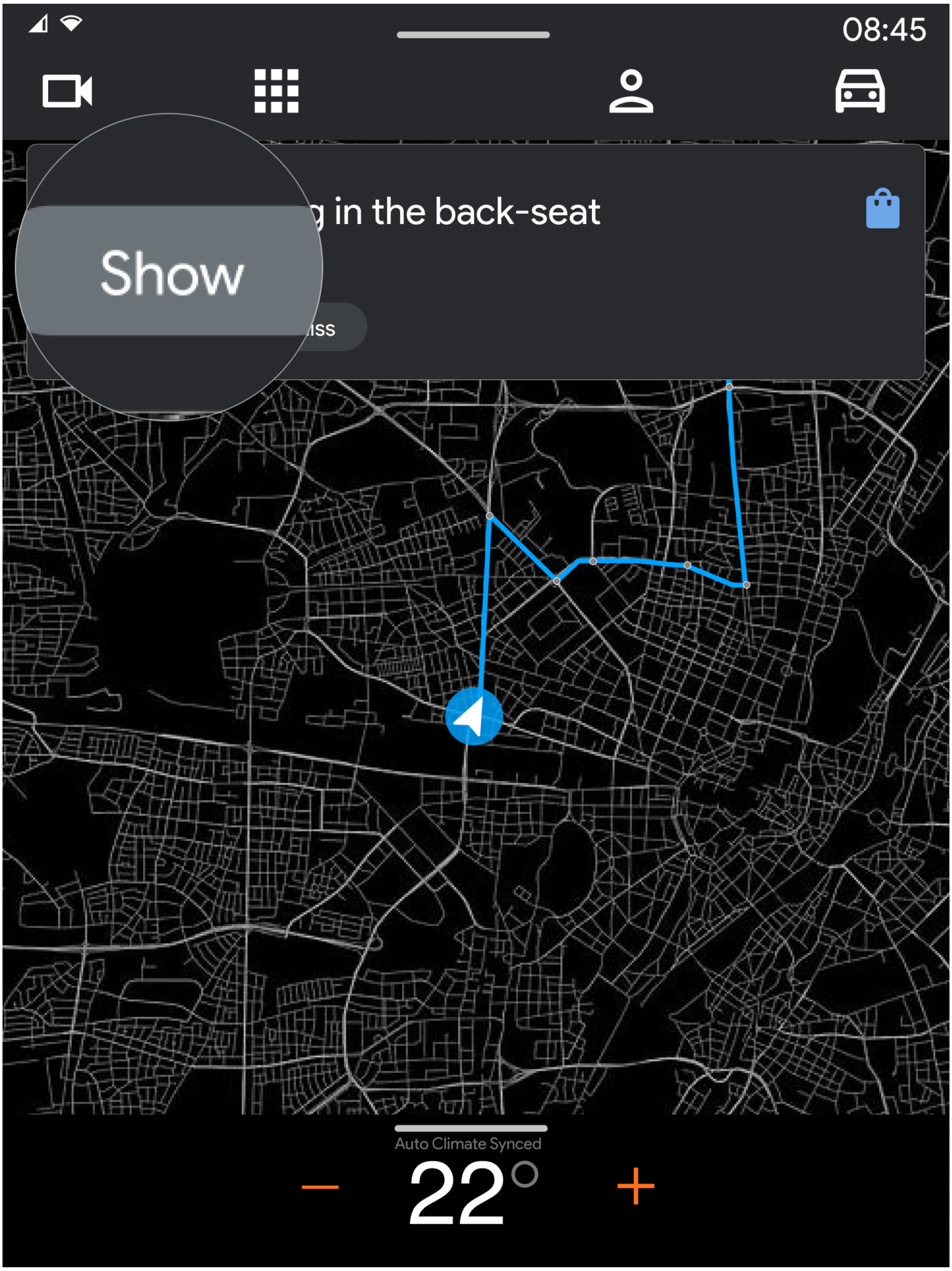

Figure A.12: Screenshot of the mockup, stage C, for the idea *Magnification*.

A.3 Iteration 3: Applications

A.3.1 ADASight

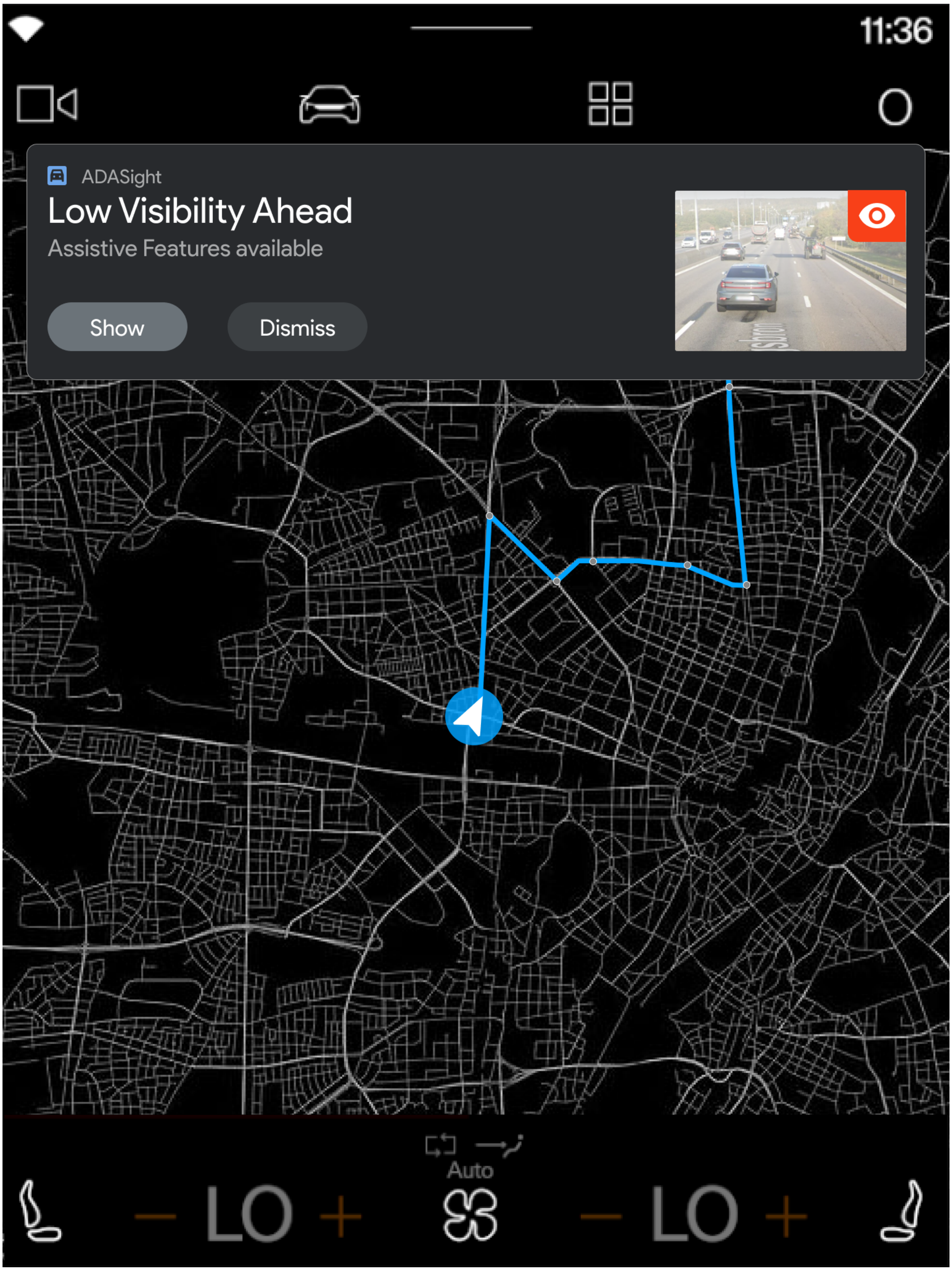

Figure A.13: Screenshot of the application, view A1, for the idea *ADASight*.

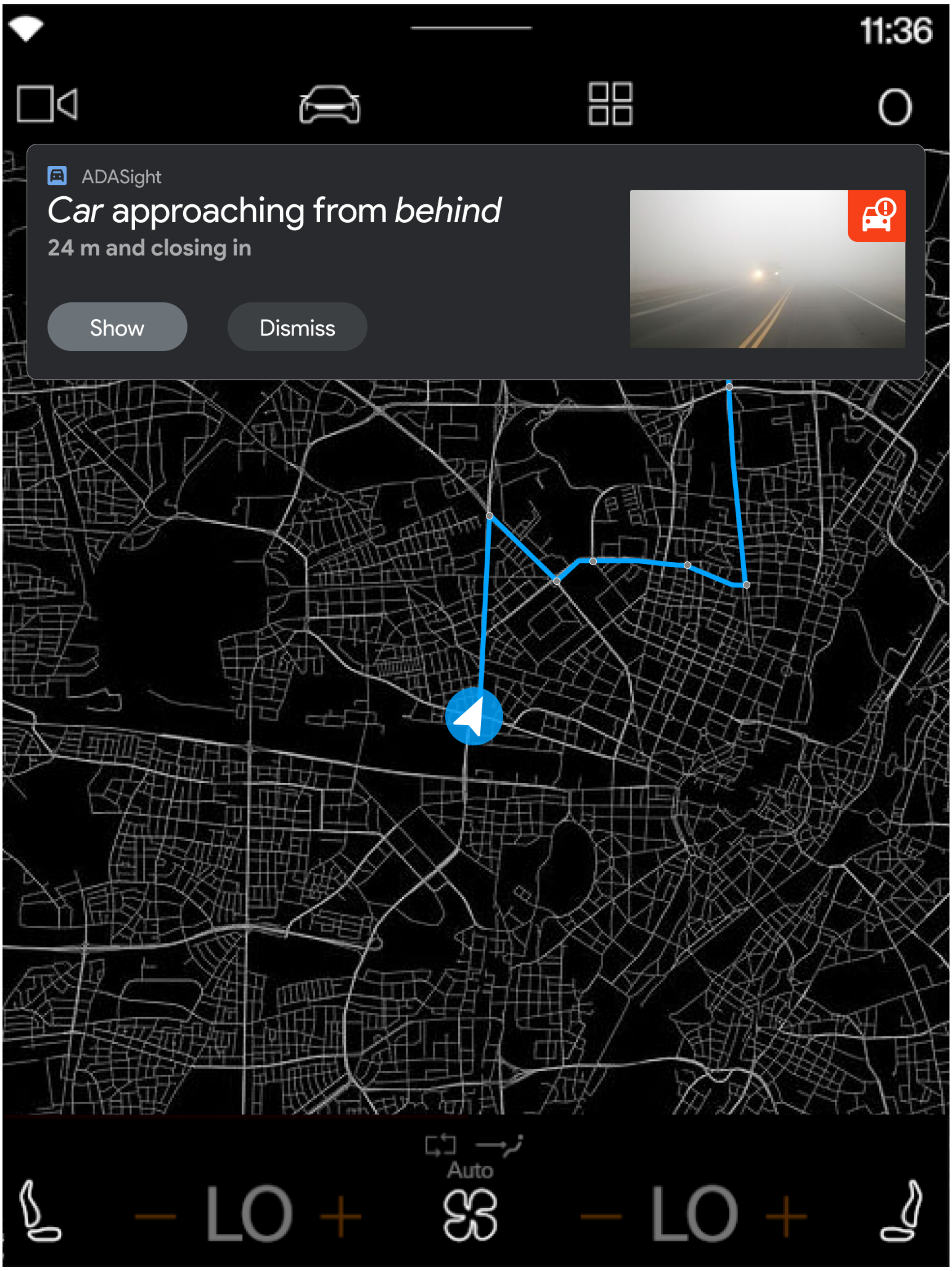

Figure A.14: Screenshot of the application, view A2, for the idea *ADASight*.

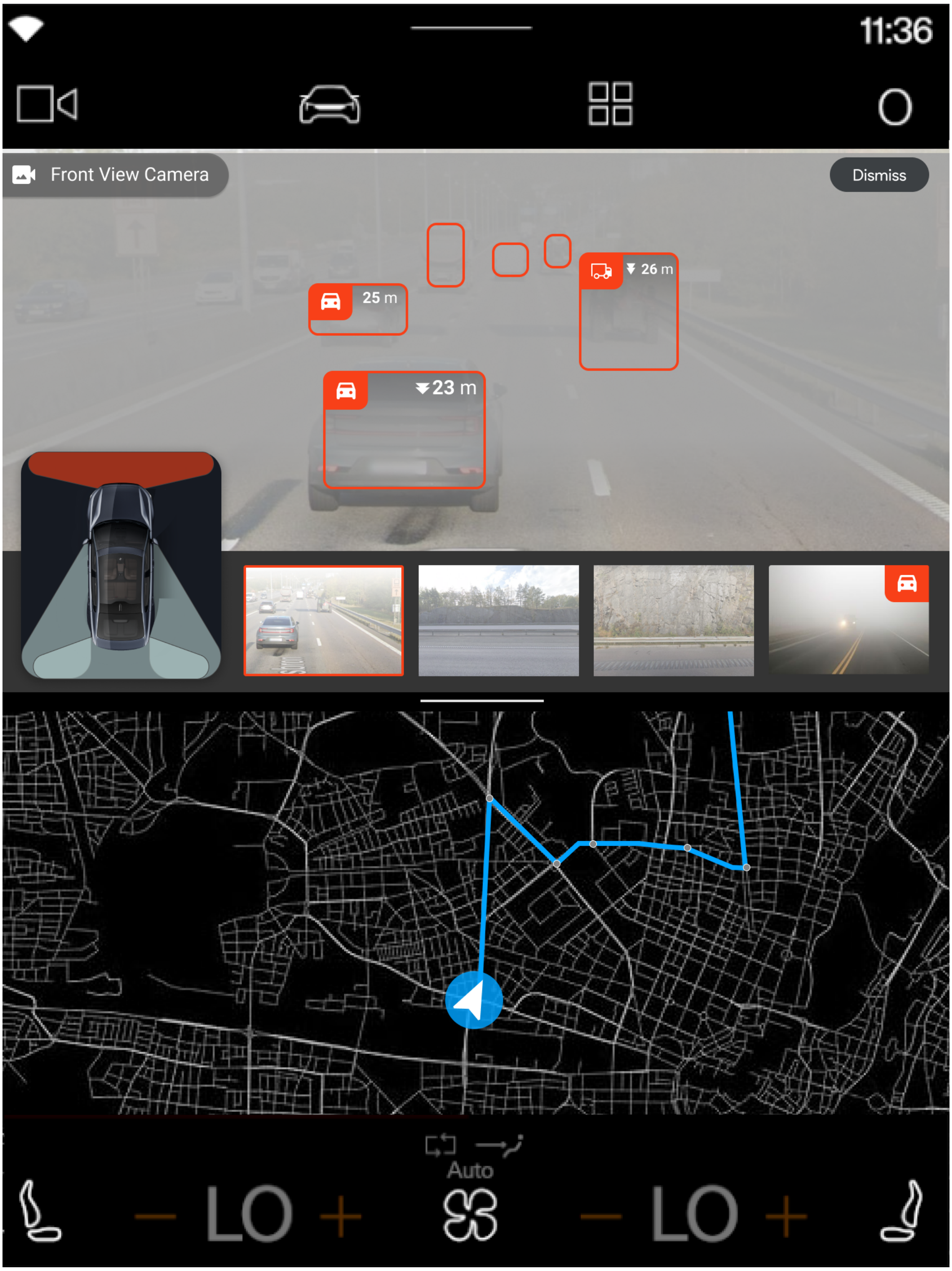

Figure A.15: Screenshot of the application, view B1, for the idea *ADASight*.

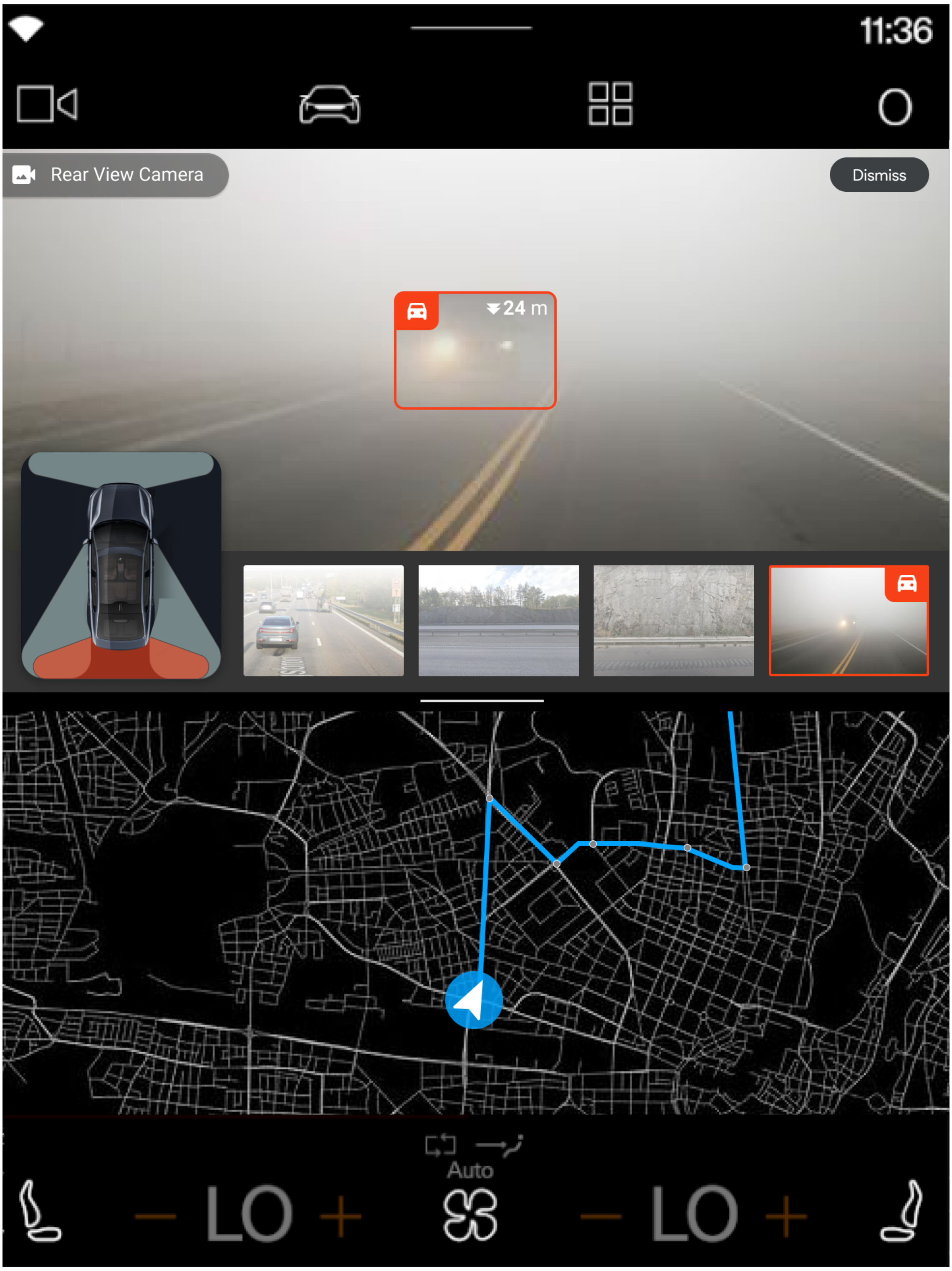

Figure A.16: Screenshot of the application, view B2, for the idea *ADASight*.

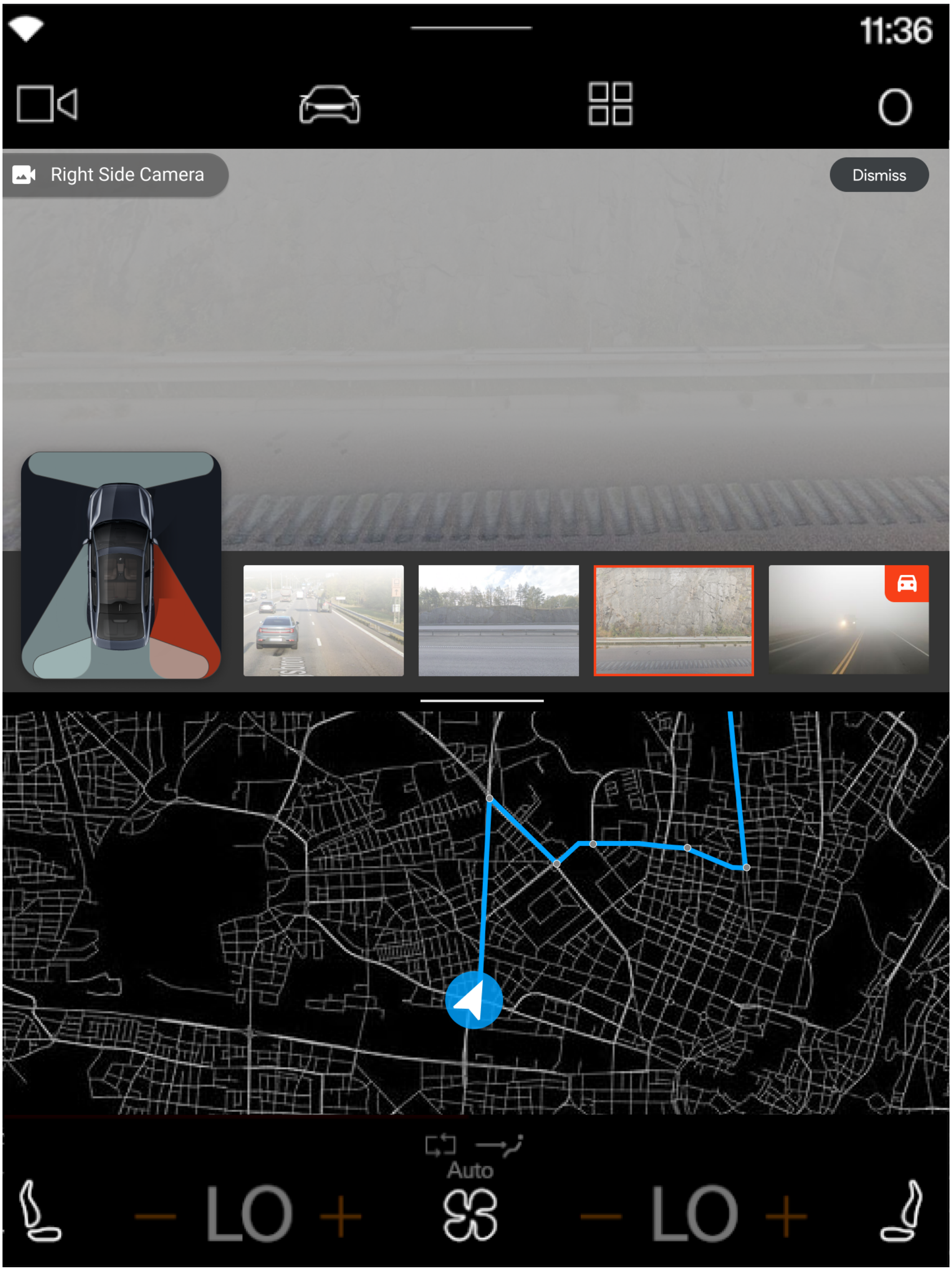

Figure A.17: Screenshot of the application, view B3, for the idea *ADASight*.

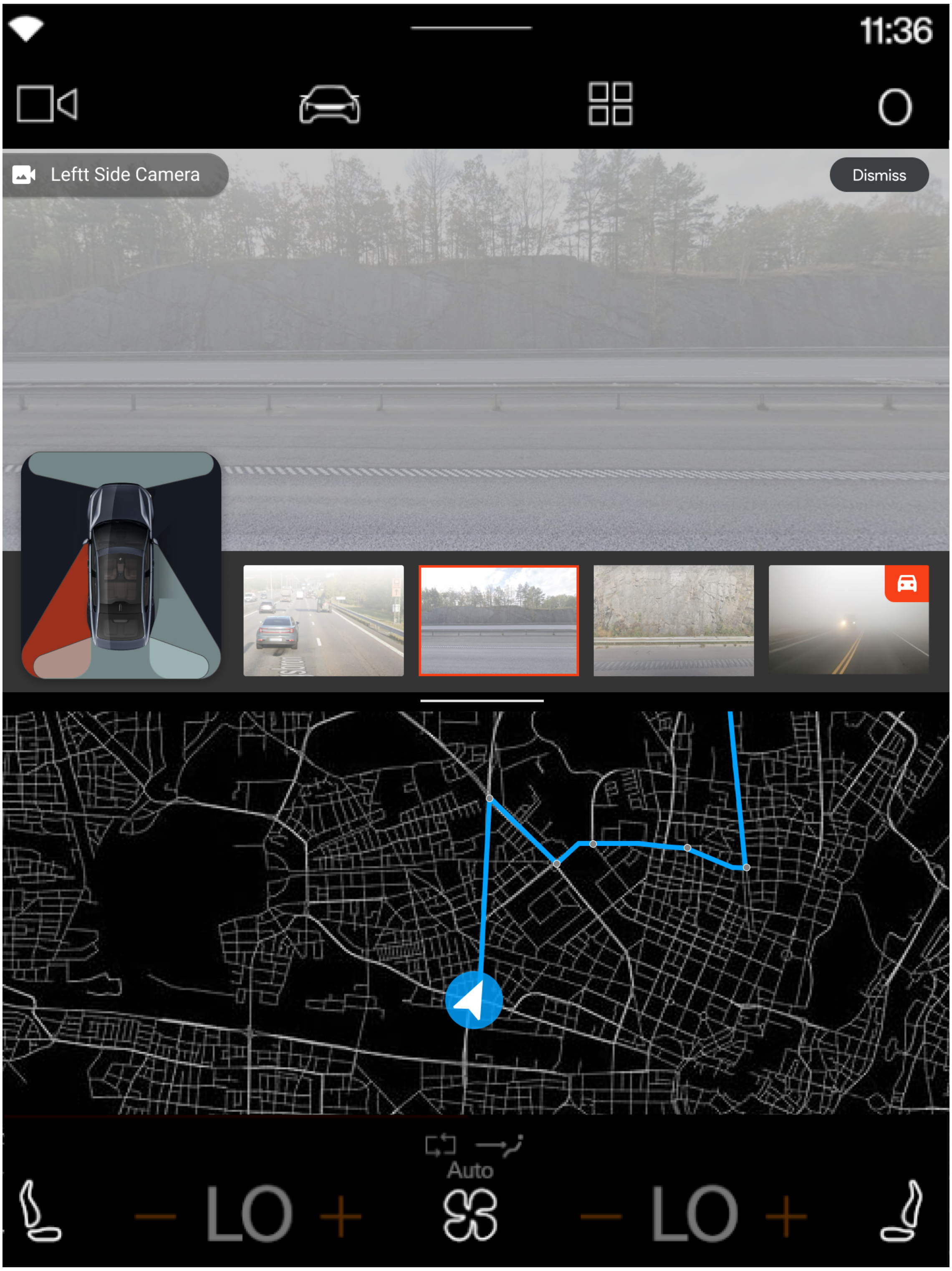

Figure A.18: Screenshot of the application, view B4, for the idea *ADASight*.

A.3.2 LeBO

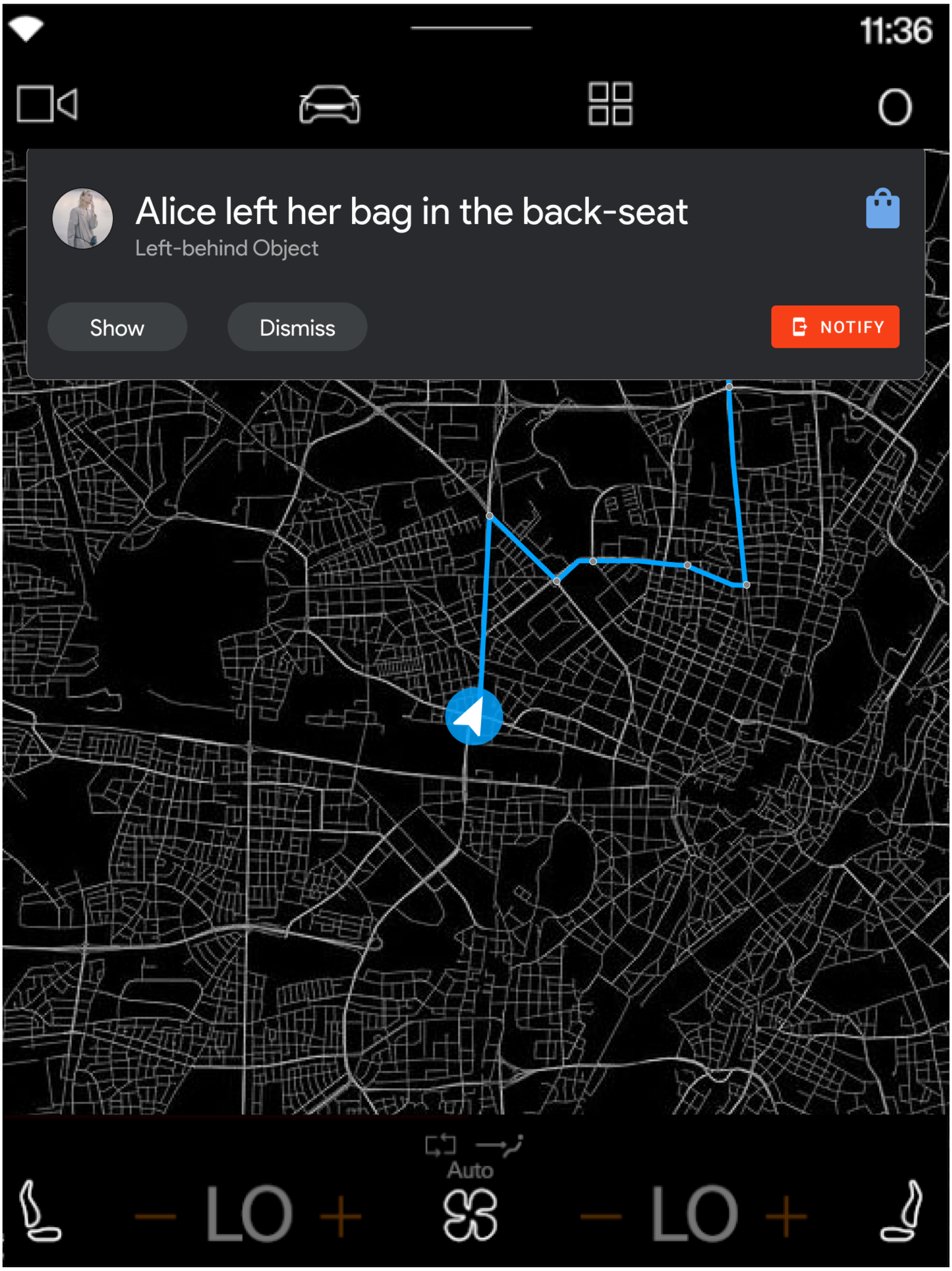

Figure A.19: Screenshot of the application, view A1, for the idea *LeBO*.

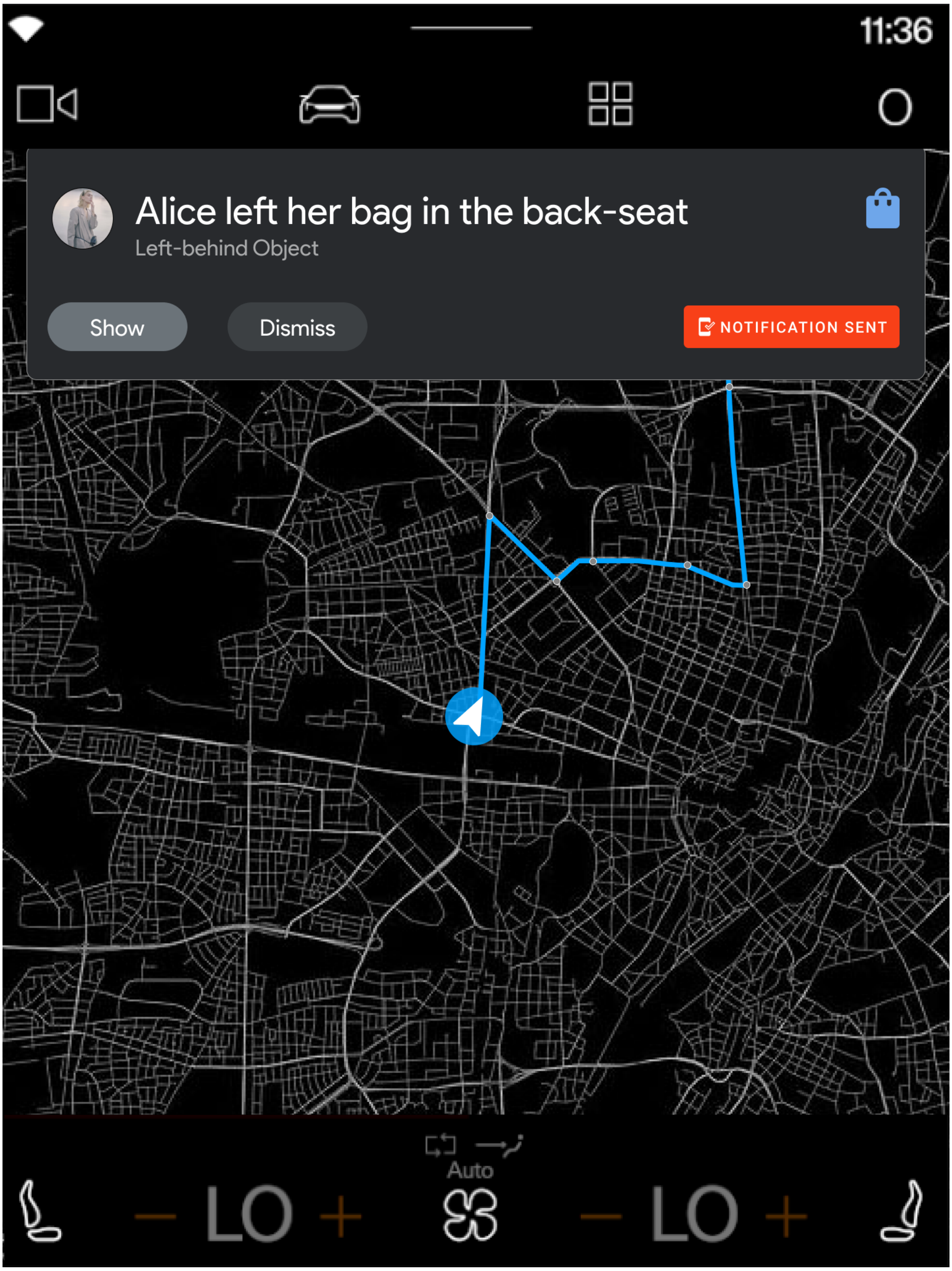

Figure A.20: Screenshot of the application, view A2, for the idea *LeBO*.

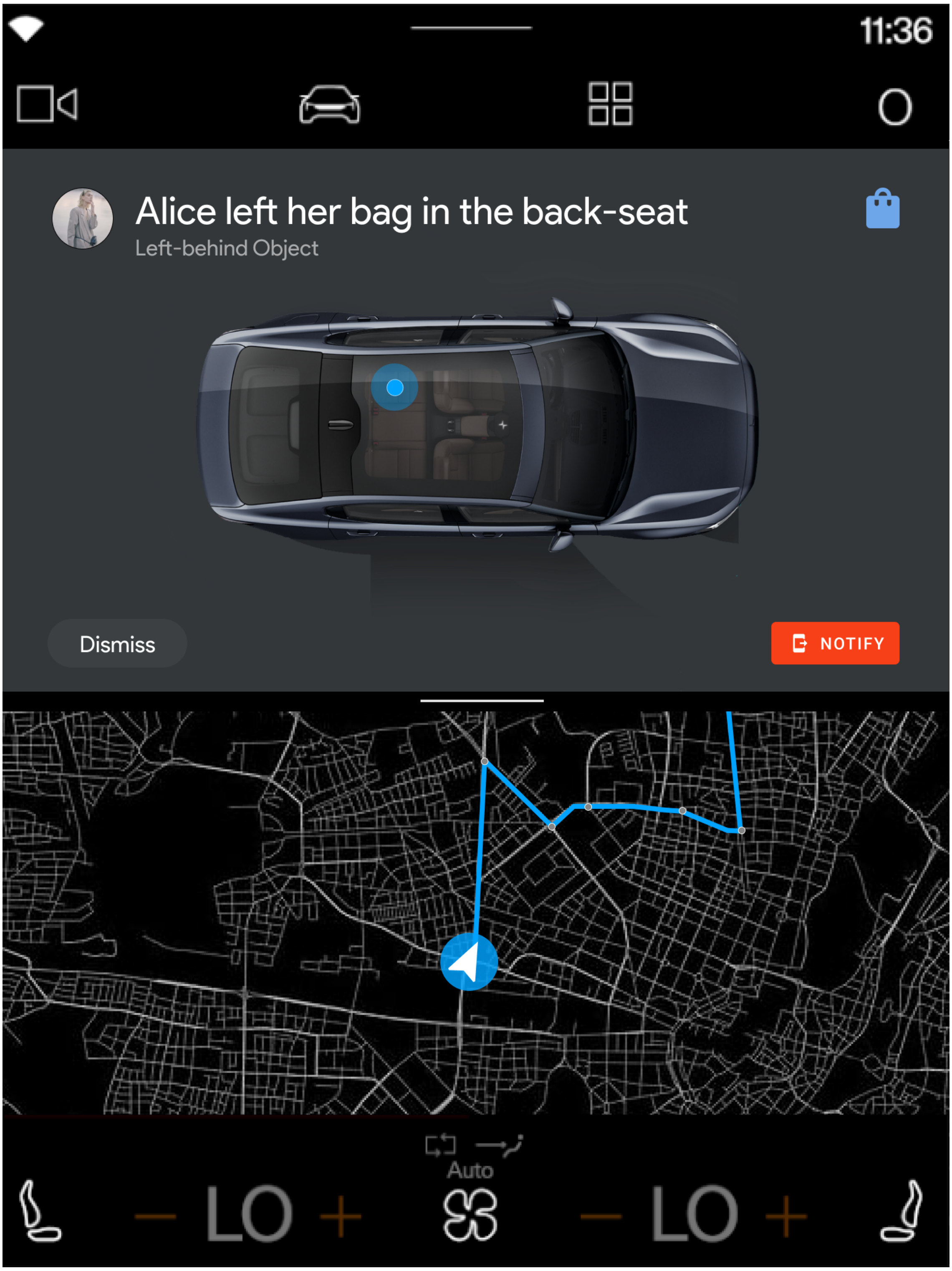

Figure A.21: Screenshot of the application, view B1, for the idea *LeBO*.

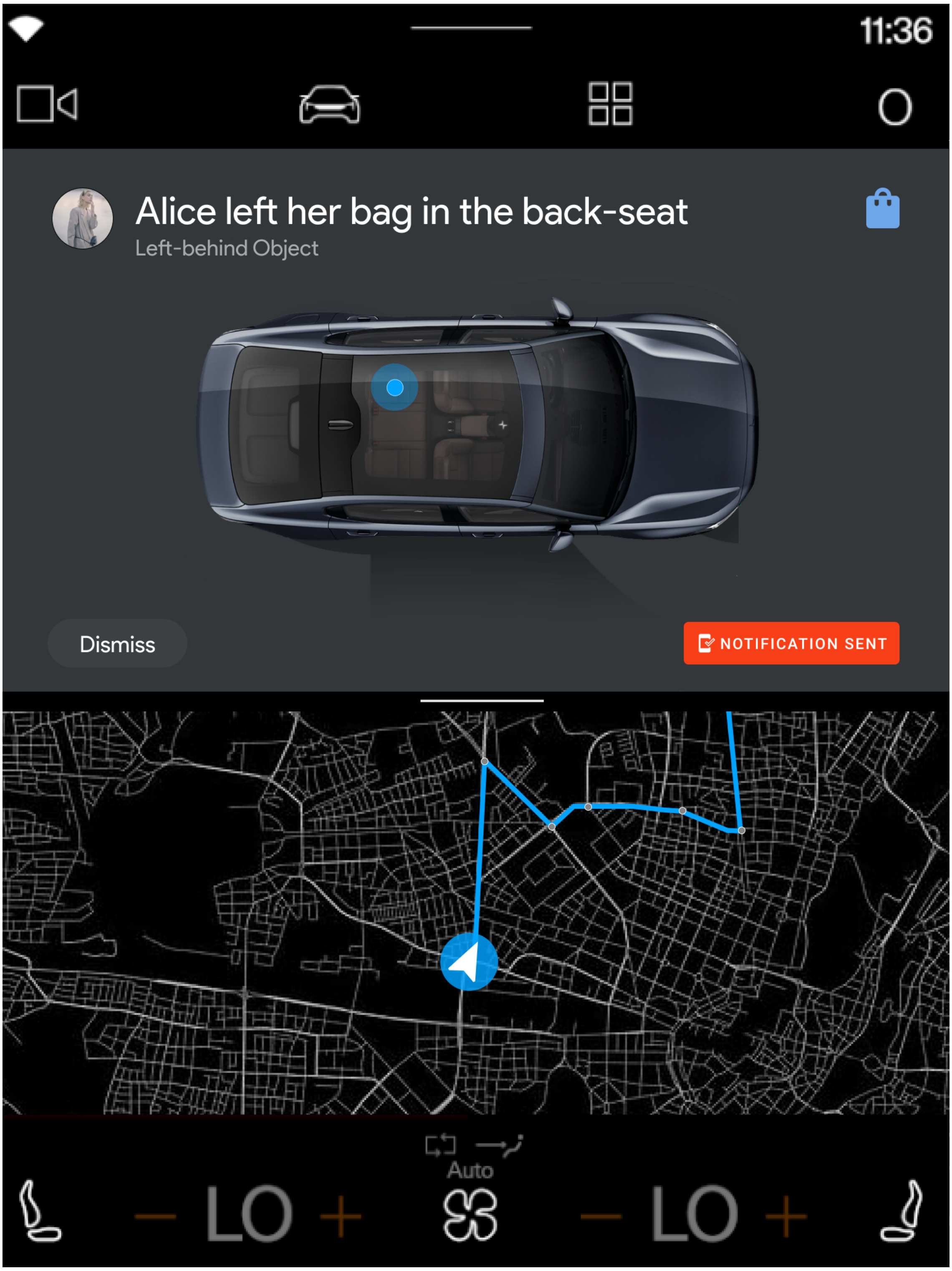

Figure A.22: Screenshot of the application, view B2, for the idea *LeBO*.

B
Consent forms

Human-Computer Interaction Study

Information sheet, risk management policy, and consent form.

You are invited to participate in a research study on two multimodal "**Android Automotive apps**" using advanced safety features. The aim of this study is to investigate how these in-car applications which use interior and exterior sensors would help bolster the UX of the car.

The study will approximately take 60 minutes and your participation in this study is voluntary. You are free to choose whether or not you will take part in the study. You have the right to withdraw from the study at any time, including withdrawal of any information provided.

The following material will be collected during the study for further analysis:
1. A video recording
2. Questionnaires
3. Interview observations/notes

The researchers may record you on video. Videos will be used by the researchers conducting the study only to analyze your interaction with the aforementioned applications. Participants cannot take part without being video recorded. After taking part in the study, you are free to withdraw your data, by emailing one of the study researchers within seven days of the study being completed. If you decide to withdraw, your data will be removed and destroyed.

All collected data will be treated as confidential and stored securely. Recordings, written notes and videos will not contain any identifying information about you. All collected data (including photographs and videos) will be anonymized. Only the researchers of the study will be granted access to the data for legitimate research purposes. We may show video recordings of our study in future project presentations, but only if you agree on this by giving your consent below. Any and all participant data not included in the final thesis report will be removed and destroyed.

Risk management policy

You agree to interact exclusively with the touch-screen and the equipment provided by the test facilitators, placed at the test setup. The touchscreen employed are standard off-the-shelf devices, and warrant standard safety, similar to any commercial screen solution used in the EU. All surfaces that the participants interact with will be thoroughly sanitized after each iteration of the test. The test facilitators will be using face-masks for the entirety of the study. You agree to also wear a face-mask while participating in the study. Respecting these rules, we expect that there is lowered risk of any infection or incident.

Figure B.1: Page one of the in-person consent form used in the last evaluations.

Statements of Understanding and Consent

- I have read and understood the participant form, and I have had the opportunity to ask questions if necessary and have had these answered satisfactorily.
- I have read and understood the risk management policy, and agree to respect it at any time.
- I understand that my participation is voluntary and that I am free to withdraw at any time during the study without giving any reason, and that I can do that within seven days of completion of the study, by emailing one of the study researchers. If I withdraw my data will be removed and destroyed.
- I understand that the data and recordings collected in this study will be used as detailed in the participant information form.
- I agree to be video footage recorded during the study.
- I agree to participate in this study, and I consent to the publication of the results of the study with the understanding that anonymity will be preserved.

The study is being carried by Aditya Giridhar (Master student, Chalmers) and Gustav Svensson (Master student, Chalmers) under the supervision of Morten Fjeld (professor, Chalmers). They can be contacted via email: gaditya@student.chalmers.se , gussvens@student.chalmers.se and fjeld@chalmers.se

Name (IN BLOCK LETTERS): ______________

Signature ______________________

Date (YYYY.MM.DD) :___

Additionally, by signing below I agree that video and data recorded during my participation in the study can be used in scientific papers, conferences and events

Signature ....

Date...

Researcher / Witness

A copy of the signed and dated consent form and the participant information leaflet should be given to the participant and the original should be retained by the researcher to be kept securely on file.

Name of Researcher ....

Date...

Signature ....

Figure B.2: Page two of the in-person consent form used in the last evaluations.

C

Timeplan

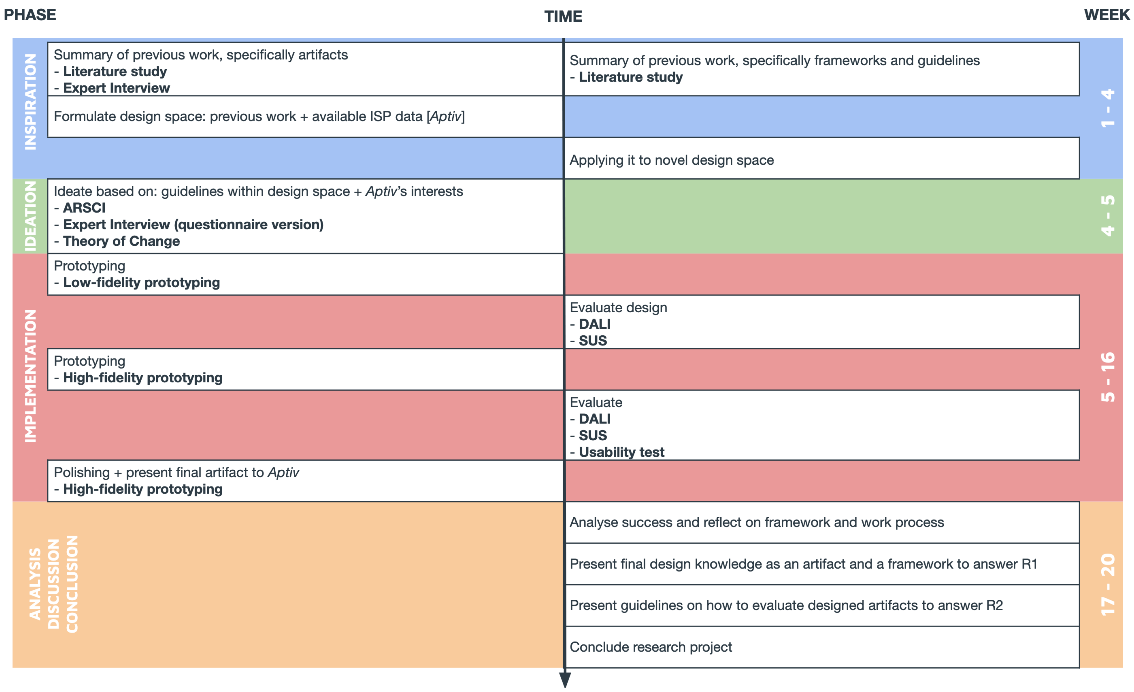

Figure C.1: The proposed timeplan for the thesis project. The figure is split vertically to highlight sub-processes directly contributing to the company or to the academic contribution respectively. On the left and right appropriate phase or week sections can be seen. In **bold** one can observe the methods planned for the different phases.

DEPARTMENT OF COMPUTER SCIENCE AND TECHNOLOGY
CHALMERS UNIVERSITY OF TECHNOLOGY | UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden
www.chalmers.se | www.gu.se