CHALMERS
UNIVERSITY OF TECHNOLOGY

Seating and Driver Interactions in Automated Vehicles

Development of a User Study Investigating Seating Positions and HMIs

Master's thesis in Biomedical Engineering

Melina Makris

Division of Design & Human Factors, Department of Industrial Material and Science
Chalmers University of Technology
Gothenburg, Sweden 2020

Master's thesis 2020

Seating and Driver Interactions in Automated Vehicles

Development of a User Study Investigating Seating Positions and HMIs

Melina Makris

Department of Industrial Material and Science
Division of Design & Human Factors
Chalmers University of Technology
Gothenburg, Sweden 2020

Seating and Driver Interactions in Automated Vehicles
Development of a User Study Investigating Seating Positions and HMIs
Melina Makris

Supervisors: Anna-Lisa Osvalder, Department of Industrial Design
Arun Muthumani, Autoliv
Examiner: Anna-Lisa Osvalder, Department of Industrial Design

Master's Thesis 2020
Division of Design & Human Factors
Department of Industrial Science
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2020

Seating and Driver Interactions in Automated Vehicles
MELINA MAKRIS
Division of Design & Human Factors
Department of Industrial and Materials Science
Chalmers University of Technology

Abstract

The technology for automated vehicles is developing rapidly, turning drivers into passengers, having end users demand a more relaxed seating position. The reclined seating position is one of the most popular expectation among end users, nevertheless, there is a lack of investigations of reclined seating positions in highly automated vehicles in dynamic environments. To fill this gap, the thesis proposes a valid user study methodology to enable assessing the end user preferences and requirements on the reclined seating position, as well as its related HMIs between active and automated driving mode, and between upright and reclined seating position. Moreover, it evaluates usability problems of the reclined seating position and its related HMIs, based on the seating and HMIs of a prototype vehicle provided by Autoliv.

The research approach encompasses methods such as Failure Mode and Effect Analysis, subjective data collection methods and applied meta-methods to develop a valid user study method. Not only did the approach yield useful insights for the development of a user study method of high degree of validity, but it also helped identify ethical aspects to consider when conducting user studies with human subjects. For the evaluation of the reclined seating position and its related HMIs in Autoliv's prototype vehicle, methods such as heuristic evaluation and semi-structured interviews were utilised. This approach generated findings of usability problems of the reclined seating position and HMIs.

The project found that in order to conduct a valid user study assessing end user preferences and requirements on the reclined seating position and its related HMIs, six aspects must be considered. These aspects include; recruiting relevant participants, providing them with consistent tasks, providing adequate time constraints, avoiding social influences, utilising appropriate data collections methods and carrying out a pilot study. Moreover the project recommended a set of brief guidelines on how to refine the reclined seating and its HMIs in Autoliv's prototype vehicle to increase usability. These findings set the basis for future investigation of the end user requirements and preferences of a reclined seating position and related HMIs in highly automated vehicles.

Preface

This project was carried out by Melina Makris, a Master of Science student in Biomedical Engineering with a background in Mechanical Engineering, from Chalmers University of Technology. The scope of the thesis covers a cross-disciplinary field, targeting both engineering topics and social research aspects. The project was conducted for Autoliv in Vårgårda, Sweden, during the spring semester of 2020.

I wish to thank my supervisor at the research department at Autoliv, Arun Muthumani, for his support, advice and help during this project. I would also like to acknowledge Mikael Enanger for his assistance, as well as Martin Östling for his input and support. Moreover, I would like to express great appreciation to my examiner and supervisor Professor Anna-Lisa Osvalder from Chalmers University of Technology, for her encouragement, feedback and patient guidance throughout the project.

Furthermore, I wish to thank the eight experts whom participated in the expert validation sessions. Their willingness to give their time so generously has been greatly appreciated. I would also like to thank the three Autoliv employees who took part in the online study session, providing me with valuable evaluations.

Lastly, I wish to thank the project group of Fordonsstrategisk Forskning och Innovation (FFI), for their support and valuable input. My special thanks are extended to Lotta Jakobsson who enabled this exchange of knowledge by arranging and leading the meetings.

Gothenburg, June 2020
Melina Makris

Contents

Abbreviations

ADS Automated Driving System

AVs Automated Vehicles

FAV Fully Automated Vehicle

FFI Fordonsstrategisk Forskning och Innovation

FMEA Failure Mode and Effect Analysis

HMI Human Machine Interface

RPN Risk Priority Number

SAE L2 Society of Automotive Engineers Level 2

SAE L3 Society of Automotive Engineers Level 3

SAE L4 Society of Automotive Engineers Level 4

SUS System Usability Scale

1 Introduction

The technology for Automated Vehicles (AVs) is developing rapidly, increasing the importance of safety, as well as useful protection principles customised for a wide range of sitting positions. Today, passengers have larger variation of sitting postures compared to drivers, which somehow are delimited due to the driving task. However, in the future, drivers are expected to become passengers as a result of higher levels of automation, and are thus expected to choose their sitting posture with more freedom than today, similarly to passengers. One of the desired seating postures defined is the reclined seating posture [1]. This position does not only pose challenges regarding safety, but also the challenge of avoiding the experience of motion sickness, while at the same time feeling comfortable and safe [2].

In higher levels of automation the driver will not be considered to be a fallback option, implying that the AV will not require the driver to take over control with short notice. These levels of automation open up opportunities for a more relaxed seating posture, as well as they are expected to influence the design of the Human Machine Interface (HMI) between upright and reclined seating. Due to the fact that there will be no rush to take over control in these levels of automation, the HMI between active and automated driving mode may also be developed accordingly. This poses the challenge of designing the optimal HMI from the user's perspective, which feels practical and intuitive when travelling in vehicles of higher levels of automation.

Up until today occupants' experiences of AVs in dynamic environments have not been studied in a broader extent. This raises the interest of investigating how end users will experience the reclined seating position in AVs, and which preferences and requirements they will set of the reclined seating position. These findings are crucial, as defining the preferred seating position would set the basis for development of restraint systems adapted to the position of question, as well as the related HMIs.

This thesis attempts to address the gap in the research by proposing a valid user study methodology to investigate the end user requirements and preferences of the reclined seating position and its related HMIs. It does not present a full investigation, but rather aims to increase the understanding of which user study approach to use when conducting a successful user study of high degree of validity. In addition, a brief evaluation of usability problems related to the reclined seating position and its HMIs is proposed, based on the seating and HMIs of a prototype vehicle provided by Autoliv.

1.1 Background

The introduction of automation in vehicles is expected to change the traditional driving task from active driving to passive monitoring. Due to significant advancement in technology, vehicles are equipped with driving assisted features such as lane keep assistance and collision assistance, in order to enhance the safety quotient of drivers. Today, vehicles may handle the lateral and longitudinal control, while the human driver must monitor the road and immediately respond when the system limits are reached. This type of automation corresponds to Society of Automotive Engineers Level 2 (SAE

L2), *partial driving automation* [3]. At Society of Automotive Engineers Level 3 (SAE L3), *conditional driving automation*, the driver has the possibility to engage in non-driving related tasks, but remains as a fallback driver whom takes over the control upon the take over request, initiated by the automated driving system in case of system failure.

When it comes to higher levels of automation, corresponding to Society of Automotive Engineers Level 4 (SAE L4) and above, the driver is *not* considered to be a fallback option. At these levels, the vehicle would not request for driver intervention, however driver initiated transitions are possible if the driver desires to take over control. Hence, the driver will still be able to actively drive, despite the fact that the traditional driving task will significantly change. As a consequence, smooth human-machine cooperation is essential at these levels, enabling seamless communication between driver and Automated Driving System (ADS) [4].

Furthermore, it has been reported that relaxing, sleeping with reclined seats and living room position (front seats rotated 180°) are the most popular positions preferred by end users in automated cars [5]. Besides automation level, motion sickness is also considered to be one factor affecting preferred seat postures and driver activities. When a vehicle is of automation SAE L3 and above, the driver is not required to monitor the traffic environment, nor attend to the vehicle control, acting similarly to a passenger. As passengers more frequently experience motion sickness compared to drivers, due to the absence of vehicle control, it is expected that drivers of AVs will tend to experience motion sickness more frequently than before [6]. In order to manage the risk of motion sickness, the underlying causes need to be understood, consequently enabling designing the vehicles and the HMI appropriately [7].

1.2 Aim

The aim of this master thesis is to successfully assess the end user requirements on:

1. The reclined seating position and its HMIs used for seat transitions, from "upright" to "reclined" and vice versa, in an automated SAE L4 vehicle.
2. The HMIs for automated transitions, from manual to automated driving mode and vice versa, in an automated SAE L4 vehicle.

1.3 Objectives

The main objective is to develop a valid user study method for successfully assessing the end user requirements and preferences of; the reclined seating position and its HMI for seat transition between "upright" and "reclined" position, as well as the HMI used for automated transitions, between manual and automated mode, in an SAE L4 vehicle prototype from Autoliv. The secondary objective is to provide a first set of guidelines on the end user requirements of the reclined seating position and HMIs, in the SAE L4 vehicle prototype from Autoliv.

1.4 Research Questions

The thesis project aims to answer two research questions, which are phrased in this below.

RQ1. How should a valid user study be conducted in order to successfully assess the end user requirements and preferences on; the reclined seating position and its HMI for transitions between upright and reclined mode; and also the HMI for automated vehicle transitions between manual and automated mode in an SAE Level 4 vehicle?

RQ2. What are the initial guidelines of end user requirements on the SAE L4 prototype vehicle from Autoliv, derived by remote user evaluation?

1.5 Delimitations

Due to the pandemic COVID-19 which deployed during the spring of 2020, the thesis is delimited to exclusively carry out studies by remote. The online study is further limited to only include Autoliv employees as participants, as the study contains confidential material such as the prototype vehicle and its features. Furthermore, the thesis will focus on automation of SAE L4, and not consider any other levels of automation. The online study will not include evaluation of the seat itself, but rather the seating position and HMI between active and automated driving mode, and between upright and reclined seating position. Moreover, the report is delimited when it comes to showing Autoliv's prototype vehicle and its HMIs, hence the results and guidelines on these will be restricted.

The developed user study method which is to be used in a dynamic environment in the future, will focus on the seating position and HMI in short distance drives and short duration time. In addition, it is limited to only study lower speeds, in a controlled, yet dynamic, driving environment. Lastly, the planned user study will exclusively involve Autoliv employees, as the prototype vehicle which is to be used in the user study contains confidential material.

1.6 Ethical Aspects

Increasing automation in vehicles opens up the opportunity of a more relaxed seating position while travelling. This opportunity is expected to yield end user desires of a different seating position than the traditional upright seating position. One reason why it is important to investigate end user preferences and requirements in such vehicles, is to set the basis for development of restraint systems adapted to the end users' preferred seating position. Investigating the end user preferences increases the possibility of developing products which the end user want to use, which further increases the probability that the developed seating position will be used as intended.

When collecting personal data in user studies, the importance of privacy and confidentiality is crucial. Alan Bryman, Professor of Social Research at Loughborough University, states the importance of holding consent from the test persons in order to use the collected data [8, p. 511]. Recorded data from expert interviews and online study has to be deleted after indirect observations. Furthermore, the findings and feedback must thus be documented in such a way that individual test users cannot be identified, according to Jakob Nielsen, Ph.D. in human–computer interaction from Technical University of Denmark in Copenhagen [9, p. 189].

The ethical aspects of performing tests with human subjects also include making the participants feel as comfortable as possible throughout the sessions, as participants have a tendency to feel pressure to perform in user studies, especially when recorded or observed [9, p. 182]. In order to decrease the risk of the users feeling of inadequate or unpleasant during the user study, the purpose of the user study must be stated clearly in beforehand, which is to test the system in question and not the user. It should also be clear to the test user that no information of the performance of any individual users will be revealed.

2 Literature Study

The literature study contains the theoretical framework and includes end users' preferred sitting positions, HMIs and motion sickness in automated vehicles as well as user study methods. This method is used in order to gain a wider comprehension of today's state of the art, and provide the theoretical framework for the master thesis.

Apart from performing literature studies, information is collected from workshops and seminaries where researchers in the industry discuss the topic of seating in future AVs. Participating in such events is considered to be a great opportunity to grasp the different perspectives of experts within the field, and thus develop a broader understanding of obstacles as well as opportunities when it comes to seating in future vehicles. Moreover, participation is taken in Fordonsstrategisk Forskning och Innovation (FFI), a project focused on research, development and innovation regarding safety and environment. FFI is based on a collaboration between the state and the vehicle industry [10]. Taking part in this project enables a more holistic overview of innovations within the field, and opens up the possibility to share ideas and receive valuable input on the thesis project, such as recommendations of relevant literature.

2.1 Seating Posture and Activities While Travelling

Apart from user studies carried out in static environment, a study of train passengers' valuations and activities has been performed, stating that the most common activities as a train passenger is reading, staring/sleeping, talking and working on laptop. Depending on the activity, the passengers had different preferences regarding the posture. As a consequence, the majority of passengers desired adaptability options to fit the seat to the performed activity [11]. When for instance reading in a reclined seat without an adjustable headrest, the headrest cannot be used unless the arms are raised to bring the reading material in a higher position. As a result, the possibility to adjust the headrest was mentioned to improve the comfort when it came to reading in a reclined seat [11]. The study also discusses that window gazing is possible when in the reclined seating position, but that a more upright posture might be feasible for having a view out of the window.

Another study investigated which activities test persons desired to perform in an autonomous car, both when travelling shorter distances alone, and longer journeys with family [12]. The study was performed in a static environment, where the test persons were asked about their preferences. It was shown that when driving shorter distances alone, the "drivers" wanted to window gaze, surf the internet, sleep and rest. When they travelled together with others however, they wanted to perform activities such as playing board and video games, watch movies and socialise together with other occupants. A survey of the public's opinion of autonomous and self driving vehicles in the U.S, the U.K and Australia also shows that the respondents' most frequent desire when riding such a vehicle is to watch the road even when not driving [13].

A user test conducted in the U.S. investigated driver and passenger activities and position in an autonomous cars by asking participants to imagine being in an autonomous car [14]. The study reported the most common activity to be phone usage, often for reading. Reading from the phone was by the participants reported not to cause motion sickness into the same extent as when reading a book. Regarding the seating posture, three primary categories of positions were identified, namely; upright, sprawled outward and fetal. The upright seating position was by the participants associated with a state of alertness. The sprawled outward seating position was used in combination with a reclined backrest and the seat moved backward in the tracks to obtain more space. Conversely, in the fetal seating position the seat was moved forward on the tracks and the knees were lifted to press the dashboard or brought up onto the seat, taking up a smaller space. These preferences for different positions were in the paper associated with various concerns for as well safety, as social norms and physical comfort.

2.1.1 Attitudes Towards Alternative Restraints

Regarding safety in unconventional seating positions, a study of seating configuration and position preferences in fully automated vehicles Fully Automated Vehicle (FAV) showed that the the majority of participants, in particular 80,7%, were willing to wear a different seat belt in a FAV while seated in a reclined seat [1]. The finding was considered positive in the study, as a changed seating position would require development of new restraint systems. Moreover, the finding is consistent with several additional studies, which have reported positive attitudes towards alternative restraint systems. For instance, a study exploring attitudes towards extra restraint systems in highly automated cars reported positive attitude towards extra restraints if the restraint system allowed more freedom when choosing the seating position [12]. In addition, a study exploring passengers' comfort experiences of extra seat belts during on-road driving, also states acceptance towards different restraint configurations amongst both children and adults [15].

2.2 HMI in Automated Vehicles

As automation increases in vehicles, a smooth communication between the driver and vehicle is stated to be of high importance [16]. The driver must understand both the possibilities and limitations of the vehicle, as well as the vehicle must understand which tasks the driver can and cannot perform. This communication between the driver and vehicle is enabled through the HMI, including visual displays and auditory messages as well as haptic control elements. All of these elements may be utilised to provide instructions and feedback when guiding and assisting the driver. In order to achieve a smooth communication and collaboration between human and machine, requirements are set on the HMI, which must be easy to understand, learn, use and remember.

A study states the importance of communicating relevant system states in autonomous vehicles, such as if the automated driving system works reliably, via the human machine interface [17]. The study summarises design recommendations for both visual-auditory and visual-vibrotactile HMIs, derived from empirical research. The guidelines for instance state that usage of colours in the interface used to communicate system states should be intuitive and in accordance with common conversions. Another guideline suggests locating visual interfaces which are used to communicate system states at a suitable position and distance, preferably close to the driver's line of sight. A set of these guidelines were then validated with an empirical study [18]. In this study, it was concluded that researchers in the area of automated driving may gain important insights into usability by utilising a heuristic evaluation approach using the HMI checklist.

When it comes to HMI in reclined seating positions, a study of activities in trains found that visual demands of tasks and reach distance are two parameters causing the driver to lean forward from a reclined seating position [11]. The study assumed the requirement of adjustabilities to increase visualisation or the ability to reach controls with optimal posture. A guideline for HMIs in automated vehicles compiles with this statement, suggesting to position control elements where they are easily reached and seen [17].

2.3 Motion Sickness in Automated Vehicles

Motion sickness is a condition which occurs as a consequence of three contributing aspects, namely; a sensory conflict when the sensed motion differs from the expected motion, inability to anticipate the direction of motion and lack of control over the direction of motion [19]. In autonomous vehicles where the driver will not actively drive but rather act more similarly to a passenger, these aspects are expected to be experienced more frequently, increasing the risk of motion sickness.

Two of the mentioned factors could be improved for occupants in autonomous vehicles, namely; the sensory conflict when the sensed motion differs from the expected motion, as well as the inability to anticipate the direction of motion [19]. These factors are stated to be influenced by the extent of the visual field, direction of gaze and posture, which all can be designed optimally to decrease the risk of motion sickness. Suggestions made in the study includes large, transparent windows to extent the field of view and displays placed in such a way that the gaze focus is kept nearly straight to optimise the gaze of direction. Moreover, another approach to reduce the sensory conflict between the difference of sensed and expected motion is providing visual stimuli to mimic the perceived motion on a screen.

When it comes to the reclined seating position, researchers state that it is possible to view outside the window, but that the upright posture might be more useful for window-gazing [11]. The reclined position may thus be assumed to increase motion sickness. On the contrary, it is stated that horizontal seating posture contributes with a positive effect improving motion sickness, resulting in the suggestion of optimising the posture with fully reclined seats, having the passenger laying down flat facing upwards [20].

Higher degree of automation in vehicles also opens up the opportunity for different seating orientations. However, a study of motion sickness in automated vehicles showed that reward travelling in automated vehicles compromises the experience of the passenger. The reward seating is stated to lead to increased motion sickness, particularly within low speed urban environments [20]. In addition, research suggest that posture in autonomous vehicles can be optimised by not having swivel seats, as side or rear facing posture increases motion sickness [20].

3 Research Approach

This chapter describes the methods used in order to answer the research questions, along with theory behind the methods. Methods for validating the first draft of the user study are presented as well as methods for identifying usability problems in the existing reclined seating position and its belonging HMIs. Moreover, user study methods aimed to be used in the first draft of the user study are presented.

3.1 Failure Mode and Effect Analysis

A Failure Mode and Effect Analysis (FMEA) is a method where possible hazards are predicted, and a plan is elaborated for how to avoid or decrease severe risks which the hazards may lead to. The predicted risks are given a risk score, a so called Risk Priority Number (RPN), corresponding to the product of the severity of the consequence and the probability of its occurrence [21]. The higher the value of the RPN, the greater the risk.

When carrying out the FMEA, each hazard and its related risk is given a current control measure, aiming to decrease and control the risk [21]. The severity and probability of the risk is thereafter set with the control measure in mind. The smallest severity score used is 1, and corresponds to "Minor injury", whilst the highest severity rate is 5, and is defined by "Catastrophic injury (fatality)". The smallest likelihood score is 1 and corresponds to "Extremely unlikely", whilst the highest likelihood is 5, and is defined as "Almost certain". An RPN between 1-4 is defined as acceptable, whilst an RPN between 5-12 should be moderated and an RPN between 15-25 must to be moderated. Moderation could be achieved by for instance changing the control method or adjusting the conduction of the study. Thereafter, the control measures of all risks are revised in order to ensure an acceptable RPN between 1-4 for all possible hazards and risks. The definitions are derived from a template provided from Autoliv.

Furthermore, the persons affected by the hazards are defined in the FMEA in order to obtain a broader overview of the amount of persons affected by each risk. Lastly, the FMEA matrix may also serve as a tool for documentation of when the revision of the control measures are completed and by whom.

3.2 Storyboard

A storyboard is a visualisation tool which helps designers understand the user interaction with the product in its context [22]. The method consists of developing a frame-by-frame series of the interactions which are to be performed in order to complete the task of question. Moreover, utilisation of this method facilitates understanding the interaction in terms of time by mapping how much time is required for each interaction.

3.3 User Study Methods

There are numerous types of methods for carrying out user studies, developed to achieve different purposes. Consequently, which method to use depends on the issues the study aims to address, as well as the purpose and goal of the study [23]. It is also possible to combine several methods in order to adapt the method to suit the study. Using established methods helps ensuring that the yielded results may be relied on to a greater extent. I addition, following a specific procedure or method may enforce structured work and documentation, facilitating others' interpretation and understanding of the results.

3.3.1 Subjective Data Methods

Subjective data methods collect data by asking the users for their subjective opinion [9, p. 34]. The data is thus collected directly from the participants in the studies, whom are able to express what they feel, think and believe of the situation of question. This approach yields the benefit of assessing the overall user experience of what is to be evaluated [23].

3.3.2 Empirical Study

Empirical studies are characterised by the origin of their collected data, which comes from investigation of users carrying out tasks and handling products in their actual contexts [23]. When the method is used properly, it may lead to a high degree of ecological validity, due to the fact that the collected information comes from actual use or actual situations [8, p. 29]. Commonly used methods within empirical studies are questionnaire scales, interviews, observations and focus groups, as well as usability testing [23].

Questionnaire Scales
There are several types of scales which may be used when assessing the subjective satisfaction of using a system. In these types of questionnaires, users are asked to rate the system of question in rating scales. The semantic differential scales is one type of scale, which list two opposite terms and asks the respondent to place the system on the most appropriate rating along the direction [9, p. 37]. Typically, the rating system consists of 1-5or 1-7 rating scales [9, p. 35]. The results of the rating for the subjective satisfaction are often calculated as the mean of the ratings for the individual answers [9, p. 36]. Using rating scales such as the semantic differential scale for several systems, enables considering the ratings in relation to each other and thus assess which system is more pleasant to use. A common weakness of such scales is however that they may yield validity problems, due to users' tendency of overrating systems [9, p. 37] [24].

An efficient method to assess the perceived usability of an HMI is the System Usability Scale (SUS) [25]. The method consists of a survey scale, and includes 10 statements which are scored on a 5-point scale of strength of agreement. The 1st, 3rd 5th and 7th items are positively-phrased, where the score contribution consists of the scale rating minus 1. The 2nd, 4th 6th, 8th and 10th items are negatively-phrased, where the score contribution is 5 minus the scale rating. After rating, the overall SUS score is achieved by multiplying the sum of the item score contributions by 2,5. The final scores thus lay between the interval of 0 to 100, with intervals of 2.5 where the higher score implies better usability [26]. The method is stated to be flexible when assessing interfaces of technologies. As the SUS provides a single score on a scale, it is considered to be easily understood by a wide range of persons, as well as it is a cost effective method [25]. Moreover, the scale balances positively-phrased statements with negative-phrased ones, which avoids the acquiescence bias, which is the bias of more people are likely to agree rather than disagree with a statement [24].

Interviews
Interviews and questionnaires are quite similar methods, as both imply asking users a set of questions and collecting their answers. In contrast to questionnaires which only require the respondent present however, interviews involve more resources as they require having both an interviewer asking the questions, and a respondent [9, p. 211]. Interviews are however flexible in other aspects, as the interviewer has the opportunity to rephrase questions if they are misinterpreted.

Interviews can be structured in different ways, depending on the available resources such as time and amount of respondents, but also depending on the aim of the interview. Semi-structured interviews are characterised by having a predetermined order of the questions, similarly to structured interviews [8, p. 321]. Nevertheless, they are also characterised by allowing the interviewer to ask follow up questions, similarly to unstructured interview [8, p. 322][9, p. 210]. This approach enables interviewees themselves to raise additional issues. Hence, they are often used in studies with exploratory approach, where one does not yet know what one is looking for [9, p. 211].

3.4 Meta-Methods

Meta-Methods are methods which apply to methods with the aim to ensure a successful user study approach [9, p. 111]. These methods involves extra work when conducting the user study method of question, nonetheless, the utilisation of meta-methods save work in the long term, as they ensure that the planned approach is on the right track.

3.4.1 Validation of User Study Method

Presenting a first draft of a user study method to independent persons who can review and critique it from a new perspective is a method which aims to ensure a user study approach of high quality [9, p. 111]. These persons should preferably hold experience within usability engineering or similar fields. In order to conduct a study of quality and obtain accurate results, the used method must aim to attain a high degree of validity, implying that assessments made by using the specific method corresponds to the true value [8, p. 73] [9, p. 169].

Utilising meta-methods aim to avoid systematic and random errors leading to results which significantly differ from reality, and thereby achieve high degree of validity. Typical validity problems which may occur involve using the wrong participants, providing the participants with the wrong tasks and not including time constraints and social influences [9, p. 169]. Vague formulations of questionnaires and interviews is another factor which may affect the validity, if they are misinterpreted by the respondents [9, p. 36]. By having the method revised by other persons with experience within the field, factors affecting the validity of the study such as these, may be found and avoided.

3.5 Heuristic Evaluation

A heuristic evaluation is performed by letting evaluators inspect the interface of the system of question, and evaluate its advantages and disadvantages. Ideally, such a method is conducted by having the evaluators try out with the system which is to be evaluated, and it is further stated that results which are generated without letting users try out the interface elements should be interpreted carefully [9, p. 209]. Commonly, heuristic evaluations are thus carried out on the basis of the evaluators' own intuition and common sense [9, p. 155].

Advantages with this method that it enables finding usability problems early in the concept phase. Moreover, it is shown that a small amount of evaluators is sufficient to find the majority of the usability problems [9, p. 156]. The literature states the payoff from using more than one evaluator, as different evaluators tend to find different problems, and recommends at least three evaluators, stating that this amount is shown to find approximately 65% of the usability problems. The method is considered to be an efficient usability engineering method, as it is fairly easy to generate guidelines for a revised design, based on the finding from the evaluations. However, it does not provide a systematic way to generate solutions for the usability problem, but rather refers to the usability design principles which are violated [9, p. 159].

Altering between heuristic evaluation and user testing is also considered to be an efficient approach when it comes to refining the design in an iterative design process [9, p. 226]. By utilising the heuristic evaluation, usability problems are identified with a small amount of evaluators. Thereafter, the design should be refined based on the found usability problems. The redesigned interface should then be subjected to user testing, which aims to find remaining usability problems on the iterated design, which could not be find in the heuristic evaluation. This approach avoids "wasting" users and time. The combination of these two methods have been shown to supplement each other, rather than leading to repetitive findings.

3.6 Applied Research Approach

The research is divided into four phases. The first phase defines the *Experimental Setup*. The experimental setup sets the requirements and basis for carrying out the dynamic user study, such as detecting risks using an FMEA, as well as defining the environment, speed, test route, participants, prototype vehicle, operators and timing of the user study with the aid of a storyboard.

The second phase consists of developing the *First Draft of User Study Method*. This phase describes the approach for conducting a user study to assess the end user requirements of the reclined seating position and HMIs. This includes further developing the test route in terms of laps and timing of tasks, by utilising the outcome from the storyboard, as well as defining data collection methods and further developing the questionnaire scales and interview questions.

The third phase consists of the *Validation of First Draft of User Study Method*. In order to validate the first draft of the user study, meta-methods are applied to the developed user study method. More specifically the first draft of the user study is revised by experts through a semi-structured interview.

The fourth phase includes an *Online Study* evaluating the reclined seating position and the desired HMIs. The online study approach consists of heuristic evaluation as well as a semi-structured interview session with the evaluators. Parts of the expert feedback is utilised when conducting the semi-structured interview for the online study. The total research approach with all four phases is illustrated in Figure 1.

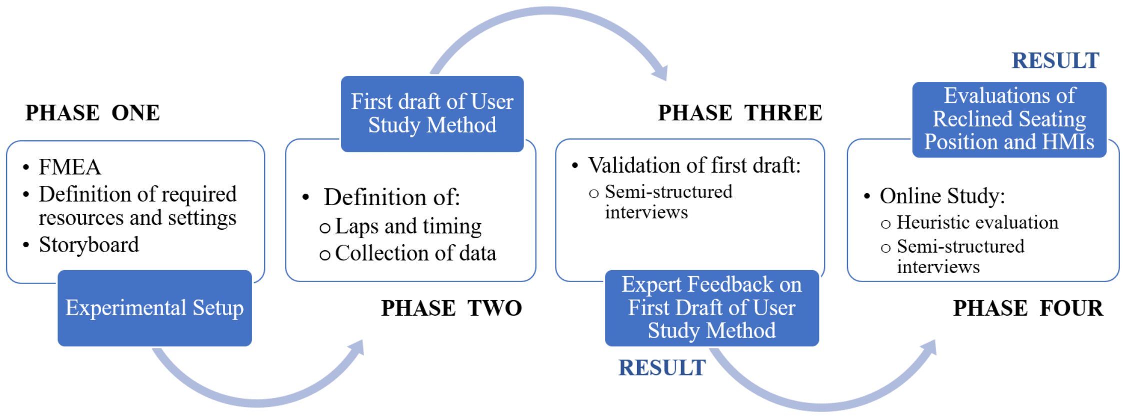

Figure 1: The approach of the thesis project. The white area shows the research approach in each phase, and the blue area shows the output and input for the various phases.

4 Experimental Setup

The first phase of the research includes defining the experimental setup. The experimental setup aims to provide a proper basis to the first draft of the user study in order to safely and efficiently assess the end user requirements of the reclined seating position and its belonging HMIs. The experimental setup therefore includes defining; the environment where the study is carried out, the speed of the vehicle during the user study, the test route, recruitment of participants, the experimental prototype vehicle used in the user study, operators required for executing the user study, the risks of the execution of the user study, how to avoid or decrease the risks and the timing of the user study. These definitions further determine the resource and settings required for the development of the user study.

4.1 Environment

As mentioned in the introduction, the user study is planned to be carried out in a dynamic environment in order to mimic the automated driving situation to the extent possible. Yet, the dynamic environment is required to be safe, in order to avoid exposing the participants and operators to risks in real traffic. With these requirements, the airfield in Vårgårda was selected as it was considered both large and sufficient to provide a safe environment, in the terms of it being free from other road users. Moreover, the airfield is located at approximately a five minute driving distance from the Autoliv's office in Vårgårda, which also is convenient as the participants are most likely to consist of Autoliv employees in Vårgårda. These parameters are considered crucial when designing the experimental setup of the route, since the user studies are required to be executed in a safe and efficient manner.

4.2 Speed

As mentioned in the introduction, the user study is delimited to only study lower speeds. The speed which is planned to be kept during the user studies is 25 km/h, in order to be comparable to studies of motion sickness in automated vehicles were carried out in the same speed [20]. In addition, it is considered suitable to perform the user studies in a lower speed, due to safety reasons and comfort of the test persons who are to experience an entirely new technology which they will not be familiar with.

4.3 Test route

The test route is defined to a path of approximately 1850 m, with two straight paths of approximately 900 m and two turns of approximately 25 m length, see Figure 2. In order to in some extent avoid excessive the influence of motion sickness, the route is designed to test the end users' experience in a straight path with turns at each end, with no complex slalom paths.

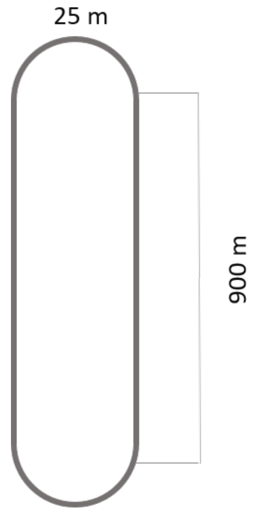

Figure 2: The path of the planned test route.

4.4 Experimental Prototype Vehicle

The experimental prototype vehicle consists of a Volvo S60 which is modified with a flexible driver seat which enables reclined seating position. The vehicle has vehicle initiated transitions between active and automated drive, and between upright and reclined seating, with feedback to driver via auditory and visual modalities. Two cameras are setup inside the vehicle collecting video data, in order to enable indirect observations. One camera aims at the driver, and another camera aims at the steering wheel and dashboard.

4.5 Participants

The aim is to recruit participants representative of the end users, as this is considered to be the main rule when recruiting participants [9, p. 226]. Nevertheless, the participants will be recruited from a pool of engineers within the organisation of Autoliv in Vårgårda, as the user studies contain confidential material such as the prototype vehicle, its features and the developed HMIs. All participants will thus have experience from the automotive industry. They will be in the ages 25-65, as these are the ages available in the company, and the user study will include both men and women, as the end users are expected to be of both genders. The participants should preferably be of varying anthropometric measures in terms of length, as the end users of different lengths are expected to experience the reclined seating position differently.

The participants are required to hold a driver's licence of category B in order to be allowed to participate, since the user study includes driving the vehicle. As the participants are to experience an entirely new driving situation, driving habit is considered important in order to avoid influences from people whom are not entirely confident or comfortable even in everyday driving situations. In total, approximately 16 participants are planned for the user study, as this amount is considered sufficient to cluster the participants having 8 male and 8 female participants of different lengths. If the lengths then are divided into two groups, there are four groups of four participants each.

According to studies, 3-5 participants are recommended for usability testing, which is why this amount is suggested [9, p. 156]. Nonetheless, this is one of all parts in the experimental setup which needs to be reviewed in the validation.

4.6 Failure Mode and Effect Analysis of User Study

An FMEA was carried out in order to predict possible hazards which may occur during the empirical user study and plan how to avoid or decrease severe risks which they may lead to. The total FMEA may be seen in Appendix A. Overall, the FMEA of the user study received moderate RPNs. Nevertheless, 4 out of 11 risks received the RPN of 6, which was the highest RPN reached in the FMEA. These RPNs were therefore moderated, according to the defined research approach where an RPN of 5-12 should be moderated.

4.6.1 Risk of Experimenter Forgetting to Brake in case of Emergency

At first, the experimenter was assigned to be in charge of the emergency brakes. The first RPN which had to be mitigated was thus the risk of the experimenter getting distracted and forgetting to brake in case of an emergency.The first the control measure to this risk was that the experimenter should prioritise safety before talking or taking direct observations during the drive. However, this control measure was insufficient as it would put the experimenter under huge pressure, having to concentrate on both leading the user study while always being ready to use the emergency brake. In the refined control measure it was instead decided that the experimenter should sit in the back seat to observe and lead the study, and not on the passenger side where the extra brake is placed. A safety operator should instead be included in the user studies, sitting in the front passenger seat, with the only task to concentrate on and handle the emergency brake. This mitigated the RPN from 6 to 3.

4.6.2 Risk of Collision

In case of a collision, there is no airbag on the driver side, as the steering wheel in the prototype vehicle does not hold one. The first control measure for this risk was the seat belt, which is considered to be sufficient as the highest speed which the test requires is 25 kph. However, in order to increase the safety, the refined control measure once again consisted of giving the safety operator control of the extra brake in case of an unexpected event. The aim with this control measure was to mitigate the risk of not having an airbag in the steering wheel on the driver side and lead to decreasing the RPN from 6 to 3.

4.6.3 Risk of Malfunctioning Vehicle

Another predicted risk which received an RPN of 6, and thus needed to be moderated, was the risk of malfunctioning vehicle function. At first, the control measure for this risk was to have the vehicle served and maintained according to manufacturer's instructions. However, due to several modifications in both software and hardware, it was decided that the technicians that had modified the vehicle should inspect it prior to the usage in the user studies as well. The technician will serve as a so called system operator, who is situated in the nearby during the empirical user study being ready to handle technical issues or system crashes. This refined control measure decreased the RPN from 6 to 3, which lays within the interval of 1-4, an is thus considered acceptable according to the defined research approach.

4.6.4 Risk of Vehicle disconnection from GPS

The last hazard causing a risk of RPN 6 was the risk of the prototype vehicle disconnection from GPS and thus autonomous driving mode, which would lead to the vehicle switching to driving mode without a warning. The first control measure to this risk was the fact that there is an extra brake pedal on the passenger side, which the safety operator may reach if needed. Moreover, if the test person is in the upright seating position, he or she will be able to reach both the pedals and steering wheel, and thus be able to steer away from a possibly dangerous situation. However, when refining the control and discussing with the technician responsible for the GPS, it was stated that the likelihood of disconnection only is high in the beginning of the test when starting the prototype vehicle, rather than during the drive. A technical inspection of the vehicle prior to the test, performed by a system operator, would therefore decrease this RPN as well, in this case from 6 to 2.

4.7 Experimenter and Operators

A number of different types of operators will be required in order to carry out the empirical user study, with different tasks and responsibilities before, during and after the user study.

4.7.1 Experimenter

An experimenter will be needed, which is the person in charge of running the user studies. Preferably the experimenter should have previous experience of the used methods, which in this case means experience of running similar empirical studies and leading interviews [9, p. 179]. The experimenter will be seated in the back seat to the right, being able to observe the participant who sits in the driver seat. However, the experimenter should not interfere with the participants during the user study [9, p. 183].

Before the test, the experimenter should hold an introduction of the purpose of the test and give the instructions of the test to each participant [9, p. 188]. The participants should be informed of the test procedure, and be reminded of that the participation is voluntary and may be stopped by the participant at any time [8, p. 511] [9, p. 188]. Moreover, the experimenter should also reassure the participant that the results of the test will be kept confidential [9, p. 183].

During the user study, it is also the experimenter's responsibility to make the participants feel as comfortable as possible [9, p. 182]. The experimenter should take the time needed for small talk with participants before the user study as well as during breaks, with the aim to calm down the participants and achieve a relaxed user study atmosphere.

After the user study a session of debriefing may be utilised, where the experimenter is able to answer additional questions which could not be answered during the user study due to fear of biasing the results [9, p. 184]. Such an opportunity is stated to help the participants repair their confidence if they felt that they have made errors during the user study.

4.7.2 Safety Operator

A person serving as a safety operator will be present in the vehicle during the test. The safety operator will be seated in the front passenger seat, and will be in charge of the extra brake pedal which is placed there as mentioned in the FMEA of the user study. The safety operator will not be allowed to interact with the participant during the user study.

4.7.3 System Operator

In addition to these operators, a third person, serving as a system operator has to be present during the user studies. However, this person remains outside the prototype vehicle. The system operator is in charge of the prototype vehicle, making sure that hardware and software systems work as intended [9, p. 182]. In case of a system crash, it is the system operator whom is expected to have the necessary skills to handle the system. It is thus of great importance that the system operator has a high degree of system knowledge. Prior to each test, the system operator inspects the vehicle, and is throughout the user study sessions responsible for the maintenance of the vehicle.

4.8 Storyboard

In order to obtain a clearer overview of the interactions which aim to be studied, all interactions were broken down into smaller tasks. This was accomplished by studying a video of a technician performing the interactions in the prototype vehicle in a static environment. A storyboard in the form of a presentation was thereafter made, summarising each individual task and interaction, along with the approximate time it takes for the task to be carried out and transitions to be executed according to the video. The storyboard also included the length and timing of the various audio voice messages, as well as the various animations appearing on the screen instructing the driver. The detailed storyboard provides the prerequisites in terms of timing for the experimental setup of the study. The timing of the tasks in the storyboard may be seen in Table 1, however the complete storyboard including illustrations and detailed description of the interaction is not shown in the thesis due to confidentiality reasons.

Table 1: Storyboard notes including tasks and timing of each interaction.

1. Participant opens the door (2 s)
2. Participant gets in the vehicle (4 s)
3. Participant closes the door (2 s)
4. Participant buckles seat belt on and upright (driving) position is activated (5 s)
5. Participant grabs steering wheel and drives actively (1 s)
6. Audio voice informs about availability of automated drive (6 s)
7. Participant performs interaction to activate automated drive - Automated drive activated (4 s)
8. Audio voice informs about availability of reclined seating position (8 s)
9. Participant performs interaction to activate reclined seating position (2 s)
10. Seat reclines (8 s)
11. Audio voice informs about deactivating automated drive (9 s)
12. Participant performs interaction to go back to upright position (2 s)
13. Seat goes back to upright position (8 s)
14. Audio voice instructs about interaction to take control (3 s)
15. Participant takes over control– Active drive activated (1 s)

5 First Draft of User Study Method

The second phase of the research approach consists of developing the first draft of the user study, utilising the definitions from the experimental setup. For the first draft of the user study, an empirical approach was chosen in order to study the users while performing the tasks in the actual system and a mimicked environment [23]. This type of method is used in order to achieve a high degree of validity, as the information gathered comes from actual use of the HMIs and seating positions in a dynamic environment, similar to the real environment of the vehicle [9, p. 185]. The developed first draft of the user study method consists of four laps, aiming to collect data of the user experiences of the reclined seating positions in comparison to the upright seating position and the belonging HMIs separately in each lap.

5.1 Lap 1

The first lap aims to collect data on user experience of the HMI between active and automated driving. Ten transitions, consisting of the interaction of letting go of and taking over control, are planned throughout the lap, see Figure 3.

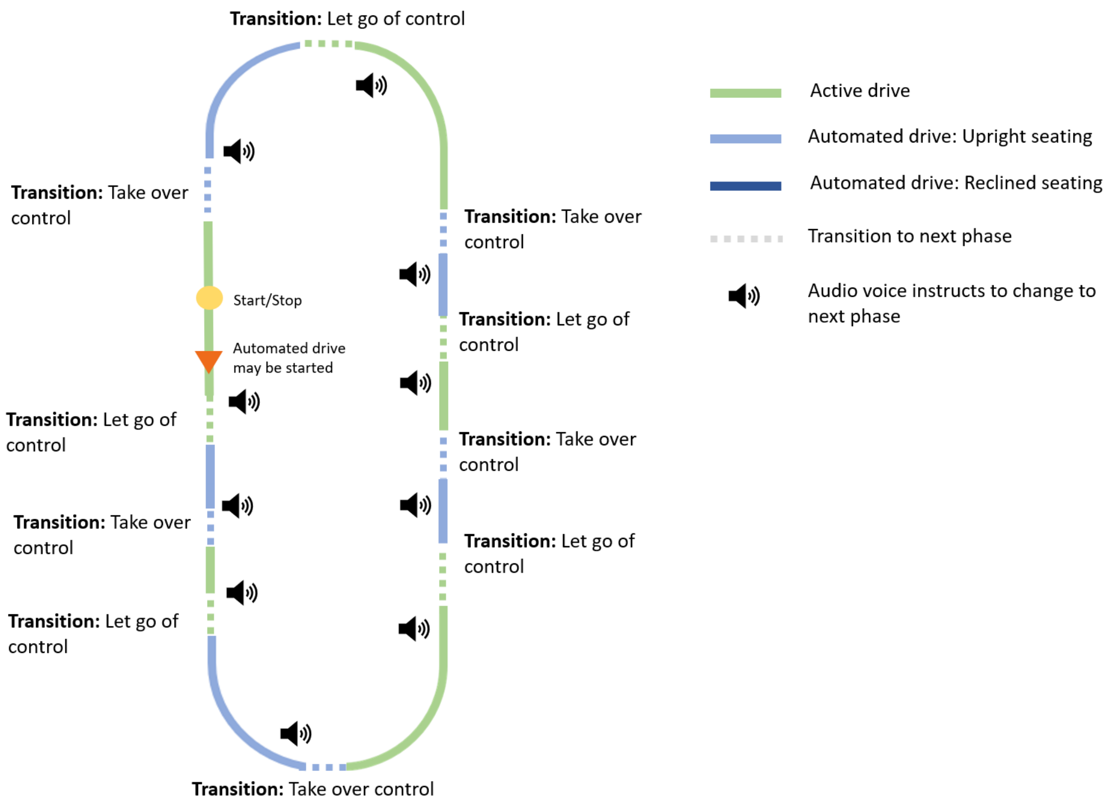

Figure 3: Lap 1, investigating the HMI between active driving mode and automated driving mode Mote that the figure is not proportional to time, but rather aims to show the amount of transitions, driving modes and seating positions

The transition of starting the automated drive, letting go of control, takes approximately 4 seconds. The transition of stopping the automated dive and starting the active drive, taking over control, takes approximately 1 second. Including the time to start and stop the lap, which is estimated to take 10 seconds, the interactions are calculated to take approximately 26 seconds in total throughout the whole lap. The total lap takes approximately 4 minutes and 20 seconds, which in theory means that the remaining 3 minutes and 44 seconds are divided into manual and automated drive. The remaining time is intended to be divided equally, leaving approximately 24 seconds driving time between each transition. The drive starts and ends with active drive, and both the curves are scheduled to consist of both automated upright drive and active drive to have the participants experience both driving modes in the turns.

5.2 Lap 2

The second lap aims to collect data on upright seating position in automated mode, in order to enable comparison with the reclined seating position in automated mode. In this lap, the vehicle will initiate the transition to go from active driving to automated drive in upright seating position in the very beginning of the lap. The automated drive in upright seating mode is then kept until right before the end of the lap, where the vehicle initiates transition to active driving, see Figure 4.

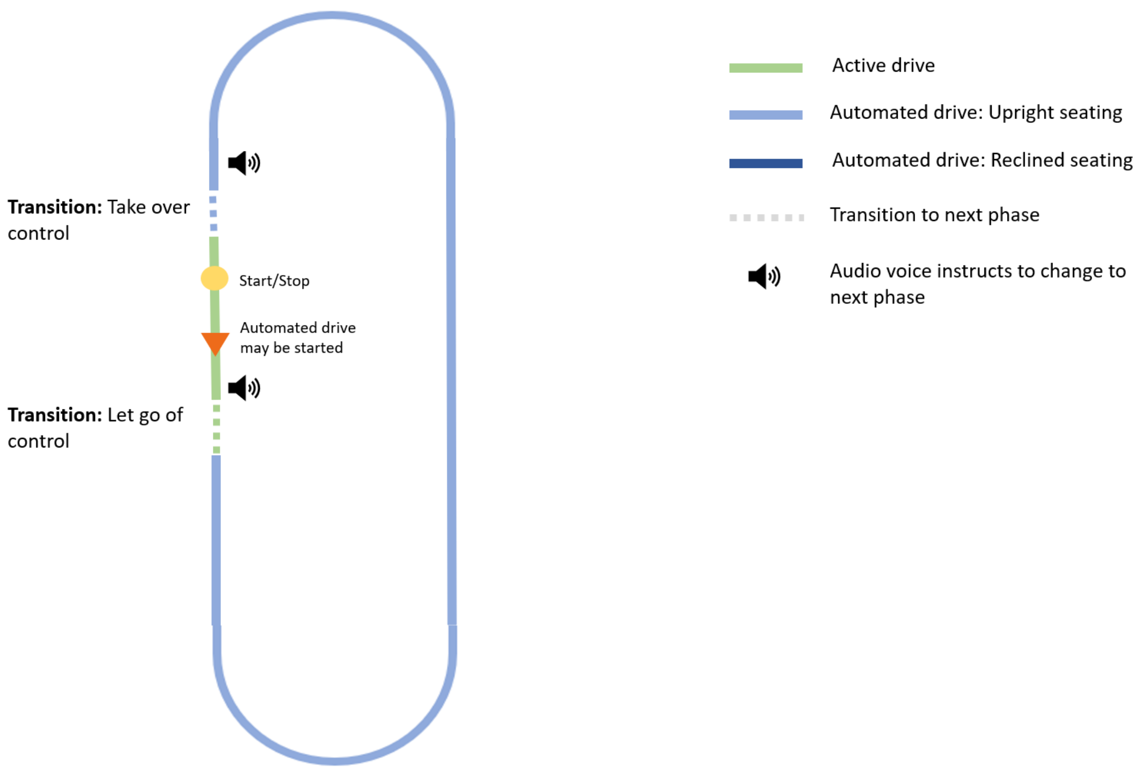

Figure 4: Lap 2, investigating the upright seating position in automated driving mode. Note that the figure is not proportional to time, but rather aims to show the amount of transitions, driving modes and seating positions.

The second lap contains only two transitions; the transition for letting go of control taking approximately 4 seconds, and the transition for taking over control taking approximately 1 second. Including the time to start and stop the lap, which is estimated to take 10 seconds, the interactions will take approximately 15 seconds. This leaves approximately 4 minutes and 5 seconds to the experience of automated upright seating, including two turns and one straight distance. This setup aims to allow the test person to experience the upright seating in automated drive for as long as possible within the lap, in order for the test person to become used to and comfortable with the automated driving in an upright seating position, before testing the reclined position which is considered to be an entirely new experience. Moreover, the aim of having the test person experiencing both the upright and reclined seating position is to enable comparison of the experiences of the different seating positions.

5.3 Lap 3

The third lap aims to collect data on reclined seating position in automated driving mode, in order to assess the user experience of sitting in a reclined seat in an automated driving vehicle. Similarly to the previous lap, the vehicle will initiate the transition to go from active driving to automated driving in upright seating in the very beginning of the lap, and thereafter from automated driving in upright position to automated driving in reclined seating. The automated driving in reclined seating mode is then kept until right before the end of the lap, where the vehicle initiates transition to upright seating position and thereafter active driving, see Figure 5.

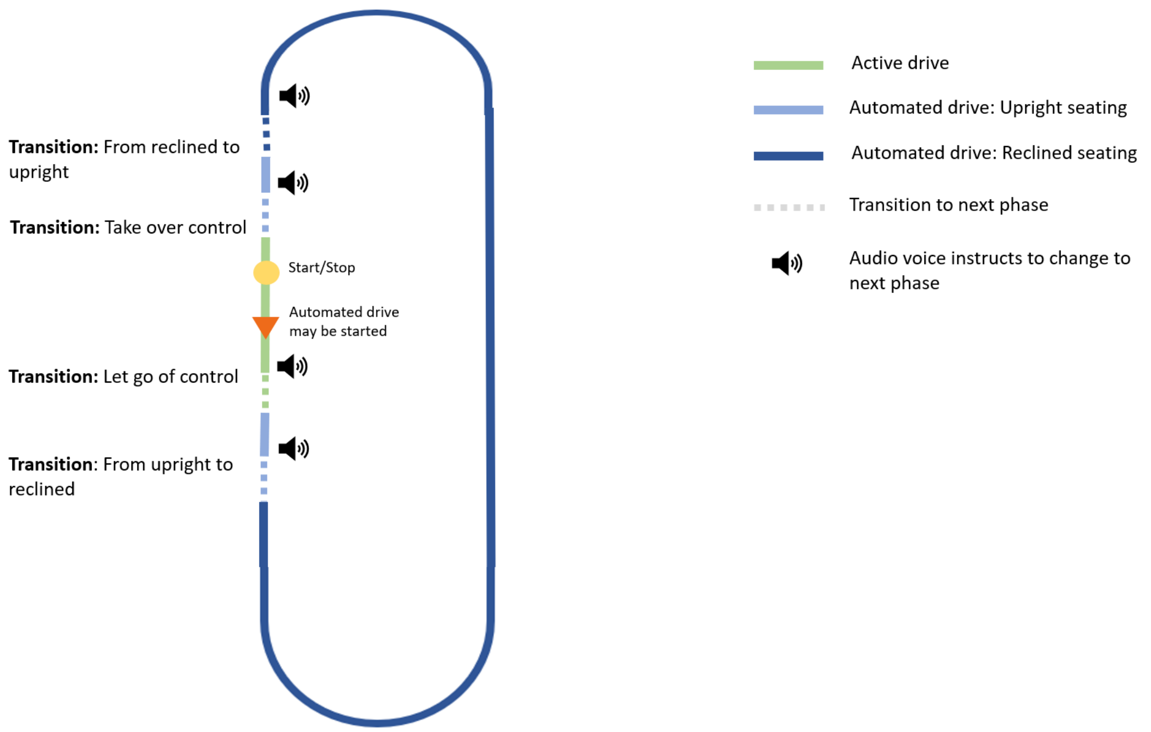

Figure 5: Lap 3, investigating the reclined seating position in automated driving mode. Note that the figure is not proportional to time, but rather aims to show the amount of transitions, driving modes and seating positions.

The third lap contains four transitions, taking approximately 25 seconds in total. The transitions are made in the beginning and in the end of the lap. These transitions in addition to the estimated 10 seconds it takes to start and stop the lap, leaves 3 minutes and 35 seconds for the experience of reclined autonomous seating position, including the two turns and one straight distance. This setup aims to allow the test person to experience the reclined autonomous seating for as long as possible within the lap, in order for the test person to have the longest possible time to concentrate on the experience.

5.4 Lap 4

The fourth lap aims to collect data on the HMI between upright and reclined seating position, in order to assess the user experience of the developed interface. Similarly to the previous lap, the vehicle will initiate the transition to go from active driving to automated driving in upright seating in the very beginning of the lap, and thereafter from automated driving in upright position to automated driving in reclined seating position. Thereafter the vehicle will continue initiating transitions between these two modes, before ending up with active drive right before the lap is finished, see Figure 5.

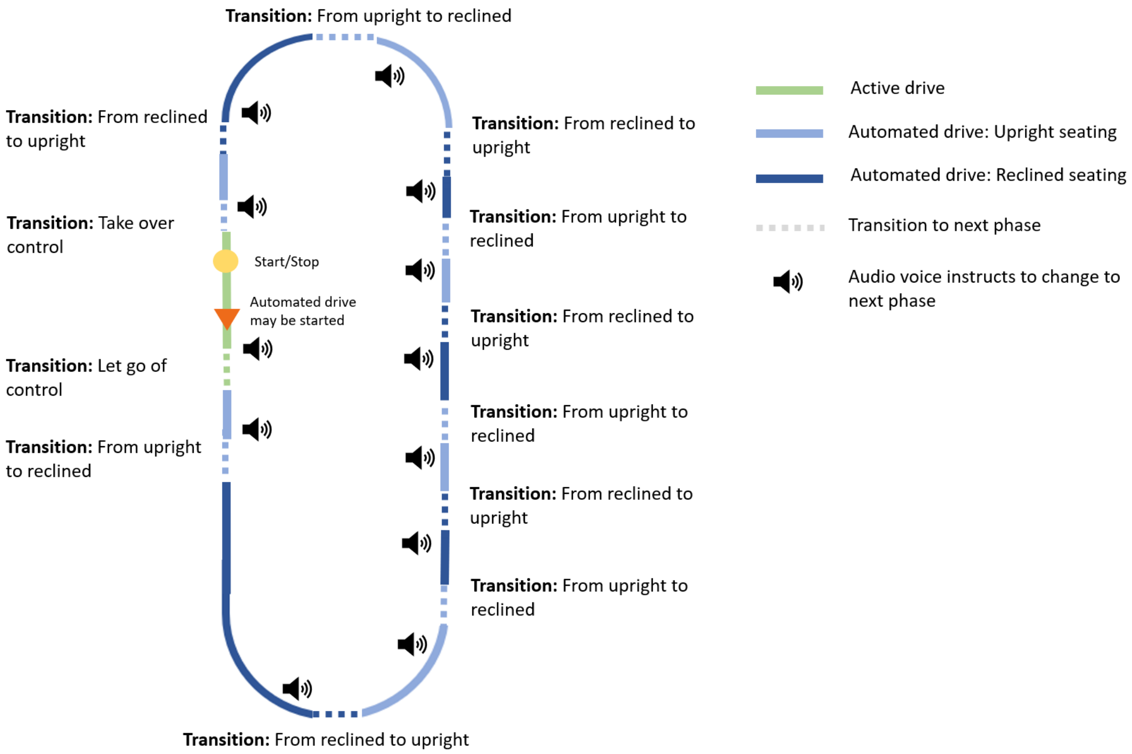

Figure 6: Lap 4, investigating the HMI between upright and reclined seating. Note that the figure is not proportional to time, but rather aims to show the amount of transitions, driving modes and seating positions.

The fourth lap contains 12 transitions, and 10 out of these transitions consist of the transition between upright and reclined seating. Altogether, the transitions take 105 seconds. When adding the estimated 10 seconds it takes to start and stop the lap, the total time of the interactions take 115 seconds. This means that the remaining 2 minutes and 25 seconds are divided into upright and reclined autonomous seating position. The 2 minutes and 35 seconds are intended to be divided equally, leaving approximately 13 seconds for each position between each transition. This setup aims to allow the test person to experience the transition between upright and reclined seating position for as many times as possible within the lap, in order for the test person to have time to become used to the interaction and concentrate on the experience.

5.5 Collection of Data

The data collection was planned to consist of observations, questionnaires and interviews. The observations were planned to be made both during the user study, in so called direct observations, by the experimenter. However, as audio and visual recordings were planned to take place inside the vehicle during the user study, most of the observations were planned to be made indirectly. After each lap, data was planned be collected by the aid of questionnaire scales, which would be filled in by the participants. Thereafter the participants would be interviewed. As the setup consists of four laps, the collection of data was planned to consist of four questionnaires and four interviews, one after each lap.

5.5.1 Observations

Observing the users provides knowledge and insights into how the seating and HMIs are handled. This method enables finding out how the tasks are performed and identifying problems which may arise [9, p. 224]. By identifying problems occurring in the empirical user study, guidelines on how to improve the evaluated position or HMI may be be elaborated. Both direct observations from the experimenter sitting in the back seat were planned, as well as indirect observations by the aid of video and audio recordings from cameras setup inside the vehicle.

5.5.2 Questionnaire Scales

The data collection was also planned to include questionnaires with differential scales, which the user would answer after each lap. The scales were to used in order to gather semi-qualitative data. This data may be used to compare the different seating positions with each other, as the same semantic differential scale was planned to be used. The developed semantic differential scale consisted pairs of opposite terms which were to be scored on a 6-point scale. A 6-point scale was used instead of a 5- or 7-scale, in order to avoid neutral answers. Two separate scales were developed, one which aimed to evaluate the subjective satisfaction of the HMIs in lap 1 and 4, see Table 2, and another one which aimed to evaluate the subjective satisfaction of the seating positions in lap 2 and 3 , see Table 3. Moreover, if the design of the HMIs is to be modified after interpreting the results and refining the design, the same scales can be used in order to assess the experience of the new design compared to the previous one.

Table 2: Differential Scale used to collect semi-qualitative data of the general user experience of the HMIs, by asking *"How did you experience the interface for altering between active and automated drive/upright and reclined seating position?"*

Quick feedback							Slow feedback
Easy/Simple							Difficult
Safe							Unsafe
Practical							Impractical
Intuitive							Not intuitive

Table 3: Differential Scale used to collect semi-qualitative data of the general user experience of the seating positions, by asking *"How did you experience the upright/reclined seating position?"*

Comfortable							Uncomfortable
In control							Out of control
Safe							Unsafe
Practical							Impractical
Natural							Unnatural
Good vision							Limited vision (of the road)
Good vision							Limited vision (for window gazing)
Necessary							Unnecessary

In addition, the SUS was also planned to be used when assessing the user experience of the usability of the HMIs. It was thus planned to be used in lap 1 and lap 4. By utilising the SUS, a quantifiable measure of the usability is generated. This would enable comparison with other HMIs which have used the same scale, such as HMIs outside the particular study. In order to make the questionnaire more easily understood by participants and maintain a high degree of validity, the word "awkward" has replaced the word "cumbersome" in statement 8, as this replacement was showed to be understood more widely in cases where participants do not consist of native English speakers [27]. The SUS can be seen in Table 4.

Table 4: The SUS which is used to collect data of the user experience of the usability of the HMIs after lap 1 and 4.

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

5.5.3 Semi-structured Interviews

After the questionnaire, data is collected with semi-structured interviews which the user will answer after each lap. The semi-structured interview aims to collect qualitative data, providing a more in-depth description of the users' experiences. The interview starts up with general questions regarding the participant's experience of the HMIs or seating positions, and end up with more specific questions regarding for instance preferences of feedback and which parameters the users consider important for the seating position. The general questions which are to be asked after each lap can be seen in Table 5. This approach was chosen in order to avoid influencing the participants' opinions, allowing them to answer open and general questions prior to going into details [8, p. 326]. This is also the reason why the interview is the final step in the data collection, to avoid influencing the responses in the questionnaires.

Since the interview will be semi-structured, the experimenter has the possibility to ask follow-up questions, facilitating the collection of the end users requirements of the seating position and HMIs in an automated SAE L4 vehicle. The specific questions for lap 1 and 4 which assess the user experience of HMIs can be seen in Table 6, and questions regarding the seating positions asked after lap 2 and 3 can be seen in Table 7.

The interviews are audio-recorded with consent from the participants, in order to enable indirect observations. This approach also allows for the experimenter to carefully pay attention to what is being said, enabling following up interesting points made and further allowing a vast examination of what participants say. The drawback of this method is that it is time consuming and requires equipment such as a tape recorder and microphone of high quality in order not to let valuable data go to waste [8, p. 330]. Moreover, recordings of these types can easily increase the pressure which the participants may feel during a user study [9, p. 181]. However, the approach minimises the risk of distraction of the experimenter, who must stay alert and concentrated on the interviewee's answers [8, p. 329].

Table 5: General open-ended interview questions asked after each lap.

General questions after every lap
1. How did you experience the HMI/seating position?
2. What did you like, and not like about the HMI/seating position?
3. What advantages and disadvantages do you see with this HMI/seating position?
4. Why do you think people would, or would not want to use this HMI/seating position?

Table 6: Specific questions, asked after lap 1 and 4, assessing more detailed aspects of the HMIs.

Specific questions after lap 1 and 4
1. How did you experience the feedback of the HMI?
2. In what way would you prefer to receive feedback?
3. In which modality would you prefer to receive feedback?
4. If you could choose freely, how would you like to interact with the vehicle to take over and let go of control/ to recline and upright the seating position?
5. To what extent do you think this HMI is suitable in the context of an automated SAE L4 vehicle?

Table 7: Specific questions, lap 2 and 3, assessing more detailed aspects of the seating positions.

Specific questions after lap 2 and 3
1. If you could choose freely, how would you like to be seated when in automated SAE L4 drive, and why?
2.Which seating position would you prefer (upright or reclined) and why?
3. Which parameters of the seating position in an automated SAE L4 vehicle, are of highest importance for you to feel comfortable?

6 Validation of First Draft of User Study Method

In the third phase of the research approach, a meta-method was applied to the first draft of the user study. Semi-structured interviews with experts within the field were carried out, where the first draft of the user study was revised, aiming to achieve a user study of higher degree of validity and quality.

6.1 Expert Validation Approach

The first draft of the user study method along with its experimental setup, was presented to experts. The presentation was given online, and consisted of a Power Point presentation which was sent to the each expert before the validation session. Each expert was encouraged to look through the Power Point prior to the session. The Power Point was also shared during the validation, and included background and aim of the study, as well as a description of the experimental setup, laps and data collection methods.

6.2 User Study Experts

Eight experts with great experience of usability engineering and of conducting and executing user studies participated in the expert validation. The experts consisted of academic persons such as researchers and Doctors of Philosophy within design and human factors, as well as technical experts within the automotive industry with expertise in user study methods. The aim of collecting data from experts of various backgrounds was to receive feedback of different perspectives, and presumably on various aspects of the user study method.

6.3 Collection of Expert Validation Data

During the validation sessions, each expert was interviewed with a semi-structured approach, in order to enable follow up questions and a deeper understanding of their suggestions. The experts provided feedback regarding both the setup including the amount of participants and the design of the test route, as well as the scales for the questionnaire and questions for the interviews. Each expert was interviewed separately, in order to enable collection of independent feedback from each expert. The duration of each sessions varied between approximately 50 and 80 minutes. The interviews were recorded with the experts' consent, enabling indirect observations after each session. The collected feedback was thereafter compiled anonymously in an Excel sheet, see Appendix B. After the indirect observations, the recordings from the sessions were deleted, as agreed with the experts.

6.4 Feedback from Expert Interviews

The expert interviews provided insights on how to conduct and execute user studies of high degree of validity. This section presents a summary of feedback and findings from the interviews with the experts, which are considered relevant in the modification of the user study method, not only when it comes to validating the study, but also regarding ethical aspects.

6.4.1 Feedback on the Experimental Setup

Firstly, the experimental setup was revised, including aspects such as the order of the laps, social influences and the setup of the reclined seating position in the prototype vehicle.

Introductory Test and Approach

All of the experts stated the importance of trying out the system and its different HMIs and seating positions, prior to starting the actual user test. Most of the experts suggested having a so called *lap zero*, where the user tries out the system in the user study environment. This approach makes it possible for the user to become familiar with the system before the collection of data, in order to avoid collecting data influenced by learning effects. In lap zero, the participant will be able to ask questions regarding any uncertainties while in the user study environment.

Another suggestion was to have a static introduction of the seating positions and HMIs in the prototype vehicle, where the participant could try out the interactions and positions in the given context before the actual test starts. A third suggestion was to have a so called "think aloud" lap before the actual test starts, where the participant is asked to think aloud while using the system. However, this method was also stated to be demanding, as it puts the participant to extra workload.

Lastly, a suggestion was made on having a more exploratory approach for lap 0-3, followed up by open questions in interviews after each lap. Thereafter, a final lap could be carried out to test and assess the experience as a whole. This lap could be followed up with both interview and questionnaire scales, providing both qualitative data in the form of interviews, and semi-qualitative data in the form of the questionnaire scales. Thus the participants have time to get used to the system in the 0-3 laps, before evaluating the system as a whole in the final lap.

Balancing Laps

In the first draft of the user study method, the lap testing upright seating positions was planned to be performed before the lap testing reclined seating position. However, three of the experts stated the importance of balancing these two laps, as they both aim to assess the experience of the seating position, meaning that 50 % of the user studies should test upright seating position before reclined, and 50 % of should test reclined seating position before upright. By balancing the laps, the influence of the experience over time is decreased in the results. Furthermore, as lap 1 and 4 both evaluate HMIs, it was also suggested to divide the user study into two sections: first focusing on HMIs, thereafter focusing on seating position.

Social Influences

Another factor in the setup which was mentioned to affect the participants was the two operators present in the experimental prototype vehicle during the user study. Three experts suggested to remove the test leader from the vehicle, as indirect observations can be made later with the aid of the recordings inside the vehicle.

Optimal Reclined Position for each Participant
Lastly, one expert stated the importance of using the optimal reclined position for each participant, when assessing the experience of the reclined seating. If the tested reclined position is the same for everyone, the results regarding if the end users want to be seated in a reclined position or not will not be valid. Having an exploratory phase of the study could help finding out which parameters affect the users preferences. The exploratory phase could be carried out statically, where the participants adjust the reclined seating positions according to their preferences, and explain why they chose the specific settings.

6.4.2 Feedback on the Recruitment of Participants

The number of participants and which type of participants to recruitment was a factor which received much feedback, which is presented below.

Number of Participants
Three out of eight experts suggested increasing the number of participants, and clustering them into different groups. However, it was discussed that the clustering was mostly relevant if the study aimed to yield statistics. The mentioned parameters; age, gender and length, should in that case all be evaluated to see how they affected the results. Another expert discussed the choice between actively dividing the participants into groups before recruiting, or having a more demographic approach documenting the background of the participants without targeting specific groups.

Two experts suggested performing fewer user studies at a time, with around 2-5 participants, as the study is in an exploratory phase. Fewer participants could thus be used to verify if the HMI is usable, before redesigning and performing another 2-5 studies on the modified design, and repeating this procedure until achieving the desired result. However, it was also mentioned that if a quantifiable result is wanted, for instance when assessing the user experience of the seating position, more participants at a time would be useful.

Background of Participants
Regarding the anthropometric measures, one expert suggested to measure the proportions of the body in addition to measuring the length, as two persons of the same length may have different proportions, leading to different experiences. Furthermore, measuring the participants' attitudes towards technology with the aid of a technology adaption scale was also suggested, to assess their attitude towards new technology. Lastly, as motion sickness also is a factor which will be evaluated in the user study, data should also be collected on how commonly the participants feel motion sick when travelling.

Furthermore, experts commented on the fact that the participants will consist of Autoliv employees, an thus be experienced within the automotive industry and presumably have a great interest in vehicles. This factor will bias the results, as the participants will not correspond to the actual end users, which will include people outside the automotive industry. This was however presented as a limitation of the study, and experts then suggested recruiting participants from different departments to assess a broader perspective.

6.4.3 Feedback on the Questionnaire scales

The questionnaires which were planned for the first draft were revised. Both general and detailed feedback was received on the questionnaires, including the differential scale as well as the SUS. The approach itself was revised, such as the chosen scales, along with more detailed feedback on the words used in the scales.

Scale Steps

The general feedback regarded how many steps to divide the differential scale into. In the first draft the differential scale was set to a scale of six steps. One expert recommended using a 5- or 7-scaled Likert-scale instead of the developed 6-scaled differential scale, as the Likert-scale is a commonly used method. The expert reasoned that by using the Likert-scale, the results could be compared to results outside the particular test, which have been achieved with the same scale. On the contrary, another expert stated that if the aim is to make comparisons between the answers within the specific user study, which is the case when evaluating the seating positions, a self made scale could be accurate and thus adequate. Another expert stated to avoid using scales of 5 or 7, in order to not collect neutral answers, whereas another expert reasoned that the two scale steps closest to the centre in the 6-scaled questionnaire can be interpreted as quite a natural answers as well.

Selection of Scales

Another topic was if the questionnaire should include both the differential scale and the SUS, in the laps evaluating the HMIs. One argued that the SUS could be quite time consuming, and recommended turning the relevant statements from it into interview questions. Two experts suggested to use only a few statements from the SUS, which is a method showed to yield a high degree of validity, in order to collect a quantifiable measurement since the interview covers qualitative data [28]. These items are statement 3 and 8, and are shown to predict the SUS with 96 % accuracy. On the contrary, another expert preferred using the whole SUS rather than only a few statements, since using the whole SUS will make the results comparable to results outside the particular test as well, arguing that it would still be possible to only consider a few statements when the data is collected for all the 10 statements. Another expert stated that the differential scale seemed to overlap with the SUS, and suggested to instead include value words, collecting more qualitative data which the SUS does not collect.

Words in Scales

The detailed feedback regarded specific words and opposites in the semantic differential scale, such as clarifying that the opposite to "easy" is "difficult", and the opposite to "simple" is "complex", and that the latter of these two scales refers to a system of several variables and system states. A few of the experts stated the importance of having these differential scales clearly defined, as a few of them are quite similar, such as the mentioned scales.

6.4.4 Feedback on the Interviews

The interview sessions were revised by the experts. General comments were made, as well as suggestions on additional questions, and rephrasing of existing questions. Moreover, it was discussed if a few of the interviews from the different laps could be combined.

General Comments
The open, general questions were found suitable as the study to some extent has an exploratory approach. The fact that the general interview questions were similar in each interview session was also considered to be a suitable, as this approach would make the participant recognise the questions, making them feel prepared.

Additional Questions
One suggested to start the interview with a question regarding their experience of automated driving system, in order to separate if they did not like the automated drive itself, or if they did not like the interaction or seating position. Another question which could be added when assessing the experience of the HMIs is if the participants would like to add or remove something from the interface. Another recommendation was to ask for additional comments, or if there was a question which the participant expected to be asked which they did not get a chance to answer. This question could be asked in the end of each interview and in the end of the whole user study session, in order not to miss valuable input from the users. Regarding the influence from the operators present in the vehicle during the test, it would also be relevant to ask how the participants would feel about trying out the system alone, in order to target the social influence from the participants' perspective.

Rephrasing Questions
Regarding the seating position and which parameters of it the participants consider important to feel comfortable, open questions could be asked. Firstly, the participants could be asked about preferred parameters in general. Thereafter, the questions could be focusing on specific parameters of interest, such as being able to window gaze, asking to what extent the parameter of question is important for the user to feel comfortable when travelling in an automated SAE L4 vehicle.

Avoiding Leading Questions
Regarding motion sickness, one of the experts suggested asking for signs of motion sickness such as tiredness, instead of asking a leading question using the words "motion sickness". Another expert similarly suggested including a question regarding the degree of unpleasantness or discomfort, and stated the importance of asking the question in the end of the interview, in order not to affect the whole interview. Nevertheless it was discussed that participants themselves might bring up if motion sickness was felt. On the contrary, one expert stated that even if mentioning motion sickness will make the participants start thinking about it, the results will be comparable to each other, as all participants will receive the same information.

Combining Interviews from Different Laps
It was discussed that the laps assessing the experience of the seating position could have one combined interview after both of the laps are tested instead of one interview per lap, since the seating positions are meant to be compared. On the contrary, the general questions regarding how they experienced the lap, and the experience of the individual positions should be asked right after each lap, in order to avoid risking that the participants forget or mix up the experiences.

6.4.5 Feedback on the Method as a Whole

Lastly, the method as a whole was revised. This feedback included suggestions of performing a pilot study and factors such as time constraints and practical aspects regarding where to be situated during the questionnaires and interviews. Feedback was also received upon ethical aspects to consider when performing user studies involving humans.

Pilot Study
When having decided upon an experimental setup, a pilot study should be carried out. The pilot study should aim to assess practicalities which need to be taken into consideration, as well as test safety and further verify the user study method. If participants for instance ignore instructions from the vehicle, the operators involved in the study should be aware of the consequences, in order to be prepared to act in such a situation, but also in order to be able to answer participants' questions before the session. In addition, the pilot study should also test the questionnaire scales and interview questions, to make sure that they are interpreted correctly by the users and thus ensure that they are valid, as well as time constraints.

Time Constraints
One crucial factor was by the experts considered to be the timing during the user study, both in the individual laps, but also the timing of each user study session. The amount of transitions carried out in the laps assessing the experience of the HMIs is therefore one of the factors which was recommended to be tested in the pilot study. Enough time must be scheduled between each interaction, in order for the user to be able to focus on the experience, without constantly thinking of the next transition. Furthermore, enough time must be scheduled for each session, in order to avoid a stress and to provide an environment where the participants have enough time to evaluate and express their experience.

Practical Aspects
After each user study lap, one expert suggested that the participants would fill in the questionnaire and be interviewed while in the vehicle. Keeping the participants in the same environment as the user study was by the expert stated to facilitate the process of remembering the experience, and would thus avoid wasting time between the laps as well as valuable input from he participants.

Ethical Considerations

From a research ethical perspective, one of the experts mentioned the importance of informing the participants of all possible risks prior to the user study. In particular, the risk of experiencing motion sickness was brought up. The same expert highlighted the importance of making sure that every participant receives the same information, and that they know of their possibility to stop the user study at any time.

7 Online Study

The fourth and final phase of the research approach consisted of the Online study, including heuristic evaluations and semi-structured interviews regarding the reclined position and HMIs. The aim of the online study was to collect initial guidelines on requirements on the reclined position and HMIs. The online study was developed after receiving feedback from the experts on how to conduct and execute the empirical user study. To the extend possible, revised interview questions from the validation with the experts were utilised in the online study, in the form of rephrased questions. The online study collected data on usability problems and end user requirements on the reclined seating position and HMIs in Autoliv's prototype vehicle.

7.1 Online Study Approach

The online study was carried out online with heuristic evaluations with the aid of a video of the experimental prototype vehicle, as well as semi-structured interviews. The video showed the developed interaction for taking over and letting go of control, and the interaction for changing between upright and reclined seating position, as well as it showed the upright and reclined seating positions. As the aim of the online study was to find usability problems, it was carried out with an exploitative approach, collecting qualitative data.

In each online study session, the participants were instructed not to communicate their findings with each other, and were also informed of that their evaluations were to be made anonymous. The session thereafter included a power point presentation of the background and aim of the study, along with introduction of the HMIs and seating positions which was to be evaluated. The session was thereafter divided into three parts, focusing on evaluation of; the HMI between active and automated drive, the reclined seating position, and the HMI between upright and reclined seating position. For each part, one sequence of the video was shown, focusing on the specific part. Each sequence was shown at least twice, aiming to give the participant a general scope of the system as well as allowing the participant to focus on the specific interface elements.

After inspecting each part, the evaluators were asked to vocalise their thoughts of the specific part, finding advantages and disadvantages and overall general comments. Thereafter, the evaluators were interviewed with a semi-structured approach. The interview questions used in the online study were to the extent possible derived from the feedback from the expert interview in terms of phrasing. Only the questions suitable in the context of an online study were utilised.

7.2 Evaluators

Three evaluators participated in the online study. This amount of evaluators is stated to find the majority of the usability problems, namely approximately 65% of them, and is thus a recommended amount of evaluators from literature [9, p. 156]. All three evaluators were Autoliv employees, due to the fact that the video contained confidential material. Optimally, the heuristic evaluation is performed with usability specialists, or even double specialists, with expertise in both usability an in the specific interface, in order to find as many problems as possible. However, the evaluation may also be performed by people with little or no usability expertise [9, p. 162]. For this study, specialists were not available. Evaluators were instead recruited from the department of engineering and development, as well as from the department of administration, aiming to achieve a broad perspective and finding more usability problems.

7.3 Collection of Online Study Data

The data from the online study was collected through recordings of the heuristic evaluation and interview session, enabling indirect observations. The sessions were recorded with consent from the participants, and were deleted after the evaluations were documented anonymously.

7.4 Evaluations from Online Study

The collected data from the online study includes both the statements from the heuristic evaluation, as well as answers from the semi structured interview. The collected data is divided into three categories, evaluating; the reclined seating position, the HMI between active and automated drive, and the HMI between upright and reclined seating position.

7.4.1 Reclined Seating Position

The evaluation of the reclined seating position consisted of finding advantages and disadvantages with the position, describing preferred seating position in automated SAE L4 vehicles, as well as describing features which would be desired in the reclined seating position.

Advantages and Disadvantages
Two of the evaluators stated the advantage of being able to relax and perform other activities such as sleeping when in the reclined seating position, rather than just watch the road and steer. Yet, one of these evaluators stated not to like the feeling of letting go of control, which would be required for the reclined seating position. Two evaluators mentioned the importance of feeling in control, suggesting more signals of the system state to be given to the driver.

Preferred Seating Positions
When it came to preferred seating positions in automated SAE L4 vehicles, the evaluators had differing opinions. One would prefer not to be reclined, stating the desire to be able to look around and see the road, having an interest in what happens outside the vehicle and liking the feeling of control. However, the same evaluator still thought that the reclined seating position looked comfortable, and mentioned the importance of the comfort of the seats. Another evaluator wanted more freedom and wanted more options when choosing seating positions, such as the possibility of lying on the side. The third evaluator would like to turn 180°to socialise and play board games. When solely comparing the upright and reclined seating positions, one of the participants preferred

upright to feel in control, another one preferred reclined to be able to relax, whilst the third one stated to prefer both, depending on activity.

Desired Features in Reclined Seating Position
Features such as massage and lower leg support were two parameters which one of the participants considered important to feel comfortable in the reclined seating position. Another evaluator instead mentioned having an emergency stop function as an important parameter to feel comfortable and safe, enabling stopping the vehicle in case of for instance a broken tire. whereas the third evaluator stated the importance of knowing that the seat belt forces, or other restraint systems, are adapted to the position of question, to feel comfortable in the reclined seating position.

One of the evaluators acknowledged the desire to be able to see the surroundings and the woad while reclined, without having to change to upright seating position, whereas another one had not thought of the possibility to window gaze while reclined and reasoned that often when you lay back, you fall asleep. Nevertheless, the evaluator was positive towards trying out a system which made it possible to window gaze while reclined.

7.4.2 HMI Between Active and Automated Drive

The evaluations of the HMI between active and automated drive are reduced in the report due to confidentiality reasons, as stated in the chapter of Introduction. The presented evaluations in the report are presented briefly, consisting of evaluations on the feedback of the system.

Feedback of the System
The evaluators liked the visual feedback in form of the colour coding and the auditory feedback. However, one of the evaluators pointed out that colour blindness could be a problem, suggesting to include other visual feedback in the form of symbols. Apart from colour blindness, difficulties of understanding the language in which the audio feedback was given was also one factor to consider when designing the feedback. All of the participants were positive towards receiving feedback in as many modalities as possible; visual, auditory and haptic. Haptic feedback was suggested by all of the participants, either in the steering wheel, or in the seat belt, signalising that the system state has changed. One of the evaluators mentioned the importance of being able to see the system state in the display in the dashboard, reasoning that drivers are used to look there.

7.4.3 HMI Between Upright and Reclined Seating Position

The evaluation of the HMI between upright and reclined seating position are reduced as well, due to confidentiality reasons. The presented evaluations consist of advantages, intuition problems, and the impression of the feedback.

Advantages
One of the evaluators was sceptic towards the interaction for altering between upright and reclined seating, and rhetorically asked "Would this interaction really be the most natural interaction for me?" After a second look of the interface however, the evaluator reasoned that the control element was easily reachable while in the reclined position. The interaction was thus considered suitable after all, and the evaluator considered the risk of performing the interaction when not intended to be low. The other two evaluators also considered the interaction to be convenient. One of them reasoned that it is easily reachable, whilst the other one said it seemed easy and smooth.

Intuition Problems
Two of the evaluators considered the interaction for altering between upright and reclined position to be confusing. They found the interaction intended for reclining the seat to be suitable for uprighting the seat, and the interaction intended for uprighting the seat to be more suitable for reclining the seat. The third evaluator also stated this preference at first, but changed opinion after a second thought after trying out the interaction while being seated in a chair during the interview. After mimicking the interaction, the evaluator reasoned that it felt more natural to perform the interaction intended for uprighting the seat while being reclined and wanting to go to the upright position.

Feedback of the System
Regarding the feedback of the system, two of the evaluators desired receiving feedback in several different modalities in this system as well. These evaluators suggested haptic feedback in the form of vibrations in the seat or a pull in the seat belt, as well as visual feedback in the centre console, communicating that the system is in the process of changing the seating position.

8 Discussion

This chapter presents a discussion of the results from the expert interviews and online studies, as well asa discussion of the methods used to obtain the results. The chapter is divided into two parts, where the first part focuses on the results, whereas the second part focuses on the methods.

8.1 Discussion of Result

This section discusses the results from the expert interviews and online study, including the feedback and suggestions which are received upon the user study method, and the seating positions and HMIs respectively. Interesting findings which raise further aspects which are important to investigate are highlighted, as well as results verifying findings from literature.

8.1.1 Discussion of Results from Expert Interviews

The discussion of the results from the expert interviews covers the feedback on the experimental setup, questionnaire scales, questions in interviews as well as the method as a whole.

Discussion of Results on Experimental Setup
It should be noted that all of the eight experts suggested to allow the participants to try out the system prior to performing the actual user study. Most of the experts suggested a so called lap zero, where the system could be tried out in the user study environment. Another suggestion was to have a static introduction in the prototype vehicle before starting the user study, or even a think aloud lap. These suggestions imply the importance of developing a user study in such a way that influences from learning effects are avoided, and are therefore considered important to regard to when conducting a user study of high degree of validity.

Regarding the setup of the user study, a result which significantly differed from the other ones was the suggestion of having a more exploratory approach for lap 0-3. These laps would then be followed up by interviews after each lap and thereafter a final lap with questionnaires to assess the experience as a whole. Even if the suggestion itself differed from the other ones, the suggestion possibly once again shows the importance of avoiding learning effects. This assumption may be made as the reasoning behind this suggestion was to provide the participants with enough time to get used to the system before finally evaluating it as a whole.

When it comes to the order of the laps, it was suggested that lap 2 and 3 which assessed the experiences of the upright and reclined seating position respectively, should be balanced. This approach was stated to decrease the influence of the experience over time, which may affect the results. This result is thus important to take into consideration when conducting a user study which aims to achieve a high degree of validity.

Discussion of Results on Consistent Tasks
Another important input pointed out aspects such as providing the participants with consistent tasks in the user study. More specifically, the feedback regarded providing each participant with their optimal reclined position when assessing the experience of the reclined position in order for the results to be valid. This result confirms literature stating that providing participants with the wrong tasks yields validity problems [9, p. 169]. Hence, it should be investigated if this suggested approach would be possible to utilise in the user study.

Discussion of Results on Social Influences
Another interesting result was the suggestion of removing the leading operator from the vehicle during the user study, as having two operators present in the vehicle would affect the participant. This statement amplified recommendations from literature, which stated that social influences are typical sources of validity problems [9, p. 169]. This suggestion raises the question regarding if all the operators present in the prototype vehicle are necessary. The safety operator is necessary due to safety reasons, however, the experimenters only task in the vehicle is to take direct observations. Since the observations can be made indirectly as well, due to the cameras setup in the vehicle, the experimenter may be considered to not be required inside the prototype vehicle during the tests. Either way, it is considered to be relevant to include an interview question regarding how the participant would feel about trying out the system alone, in order to target the aspect of social influences during the user study. Moreover, the feedback on the operators present raised the question regarding if the system operator also could act as a safety operator, in order to handle resources more efficiently.

Discussion of Results on Recruitment of Participants
Regarding the recruitment of participants, three out of eight experts suggested increasing the amount of participants, and also clustering them into groups. The reasoning behind these suggestions was that this approach would enable a more statistical analysis. On the contrary, another expert stated that there is a choice between actively clustering participants before recruiting, or having a more demographic approach without targeting specific groups.

These differences in the suggestions could have several different explanations. One possible explanation could be that the experts suggesting clustering assumed that statistical analysis of the results was of priority, whereas the expert stating that a choice has to be made perhaps wanted to raise the question of choosing approach depending on which results are of highest priority. In addition, the available resources may also be a factor leading to different opinions of this question, as more participants requires more time and staff which yields a more expensive study. This finding is considered important, as it brings up the importance of defining the correct participants for the user study of question, not only depending participants' correspondence to the end user, but also depending on what type of results are of priority and the available resources. The finding is also in compliance with literature which states that one common factor yielding validity problem in user studies involves using the wrong participants [9, p. 169].

Another result which is considered interesting involves the suggestion of performing fewer user studies at a time, with 2-5 participants, to thereafter refine the design of the HMI and enable a redesign before performing another 2-5 user studies. This procedure was suggested to continue until the desired result was achieved, as the HMIs are in an early phase of the development. On the contrary, the same expert mentioned that having more participants would be useful if a quantifiable result is required when assessing the experience of the seating position. Another expert also suggested dividing the user studies into two parts, one studying the seating position and the other one studying the HMIs. These result are interesting, as they raise two concerns. Firstly, they imply the importance of further defining the characteristics of the desired results. Secondly, they question if it is reasonable to develop a study which aims to target both the experience of two HMIs, as well as the experience of the reclined seating position. These question should be taken into consideration when further planning the user study.

The last feedback involving participants consisted of comments on the fact that the participants will consist of Autoliv employees, which will be a factor affecting the result and validity of the study. Even though this is a stated limitation of the study, it raises the importance of taking this factor into account when interpreting the results, as this factor may bias the results according to the literature stating that validity problems often are caused by involving the wrong type of participants [9, p. 169].

Discussion of Results on Questionnaire Scales

Regarding the differential questionnaire scale, one suggestion was to use a 5- or 7-scaled Likert-scale instead of the developed 6-scaled semantic scale, in order to enable comparison with results outside the user study. On the contrary, another expert stated that if the aim is to make comparisons between the answers within the specific user study, which is the case when evaluating the seating positions, a self made differential scale could be accurate. Again, this result raises the importance of further defining to what extent the results are to be used in the long run. However, as the user study is in such an early stage, the results are not likely to be compared to systems outside the study yet. It might thus be sufficient to use a self-made scale as this enables targeting the questions which are of interest for the specific study.

Regarding usage of both the SUS and the differential scale, the results mostly concern if using two scales would be too time consuming, and further to make sure that the scales do not overlap each other or the interview questions. There was no common recommendation which the experts agreed upon regarding how to use the scales. Instead, what is considered interesting is that these results showed that there is a great flexibility of using these type of scales, and that they can easily be adapted to the purpose of the study of question as well as to the available resources in the form of time. As there are differing opinions on how to use the questionnaires, it may be advisable to further test different approaches in a pilot study.

Out of the eight experts, one commented on the specific words and the importance of having a clear definition of them. This could be of importance in order to be able to explain to the participants how to interpret the words, in case of questions. Moreover, this finding was in line with previous findings in literature, stating the importance of revising questionnaires and scales in order to decrease the risk of misinterpretation by the respondents, which further affects the validity of the results [9, p. 36]. Another expert suggested to add values in the differential scale to put more focus on measuring subjective satisfaction. Another important aspect which was mentioned to yield validity problems when using such scales was the users' tendency of overrating systems, which amplifies findings in literature [9, p. 37] [24]. This should be taken into consideration when interpreting results from such scales.

Discussion of Results on Interview Questions

When it came to the interviews, one suggestion was to start with a question regarding the participants' experience of the automated driving system itself, before continuing on the more specific questions regarding either HMIs or seating positions. This was suggested in order to enable separation of these two aspects. This suggestions is considered important, as the experience of the automated driving system is likely to affect the experience of the rest of the system, which raises the question of assessing the participants' attitude towards new technology and automation in general.

Another important recommendation was to ask for additional comments in the end of each interview, in order not to miss valuable feedback from the participants. Moreover, when it came to assessing the experience of the HMIs, one suggestion was to ask an open question regarding what the participants would like to add or remove from the interface. These questions could be considered when further deciding upon which questions to include in the interviews, in order to ensure successfully assessing the end user requirements and preferences.

Discussion of Results on Motion Sickness Assessment

When it comes to motion sickness, the suggestions regarding how to target the feeling were differing. The recommendations consisted of either asking directly for it in the end of the interviews, or asking for signs of motion sickness such as tiredness or unpleasantness to avoid the risk of a leading question. Asking directly for motion sickness was assumed to yield a risk of affecting the responses in the remaining interviews, having the participants thinking about motion sickness in the upcoming laps. The important finding within these results was however that even if the question is formulated to directly ask for motion sickness, at least the results will be comparable as all participants will receive the same information. As there are differing opinions on how to target this parameter, it may be advisable to further test the approaches in a pilot study.

Discussion of Results on the Method as a Whole

A suggestion which is considered important is the timing of each user study session, in order to avoid stress and further provide the participants with the prerequisites needed to evaluate and express their experience without feeling pressured. This suggestion may amplify findings from the literature which state that validity problems often occur due to inadequate time constraints [9, p. 169].

Another reason why the time aspect is crucial is to ensure that participants are given the prerequisites to feel as comfortable as possible during and after the test, which confirms recommendations from literature stating the importance of providing a relaxed atmosphere [9, p. 182]. Another meaningful expert comment involves the ethical aspect of informing the participants of all possible risks prior to the study, such as experiencing motions sickness, and ensuring that every participant receives the same information in the introduction. This comment further highlights the importance of conducting user studies with respect for the participants' well-being and emotion.

8.1.2 Discussion of Results from Online study

The discussion of the results from the online study covers the evaluations of the reclined seating position as well as a brief discussion of the results on the HMIs. Moreover, the affect which the evaluators may have had on the results is discussed.

Discussion of Results on Reclined Seating Position

The results from the online study to some extent confirmed previous findings from the literature studies regarding preferred seating orientations and positions in AVs if the ADS works reliably [11]. The results from the study for instance amplified that the seating position is chosen depending of activity, showing that the reclined seating position is preferred when wanting to relax or sleep, whereas the the upright seating position is preferred when wanting to window gaze and socialise [12]. Another interesting desire was to be able to watch the road and surroundings while reclined, without having to change to the upright seating position. This finding may be important when developing the reclined seating position and features in automated SAE L4 vehicles according to user preferences in the future.

Regarding which parameters that were desired to feel comfortable in the reclined seating, the preferences of the participants diverged. Comfort features such as lower leg support and massage were mentioned, as well as safety features such as an emergency stop enabling to stop the vehicle while reclined. Furthermore, another preference was to receive feedback from the vehicle showing that the restraint systems adapt to the seating position of question when altering between positions. This result verifies findings from literature stating the importance of communicating system states such as if the automated driving system works reliably [17].

Even though the suggestion of having an emergency stop feature was made by only one of the evaluators, it aligns with the concerns mentioned by the rest of the evaluators regarding the importance of feeling in control while in the reclined seating position. On the contrary, such a feature could yield safety risks if used by accident and would require to be developed carefully. These findings may be interpreted to be related to the feeling of trust towards the vehicle and its ADS, rather than the seating position itself. However, the findings still imply the desire of feeling in control and being aware of the surroundings while reclined. These findings are thus considered important for the future development of the reclined seating position.

Discussion of Results on HMIs

For both HMI systems, evaluators wanted feedback in all modalities, and more specifically feedback of the system state was greatly desired. This finding verifies previous findings from studies stating the importance of communicating relevant system states in automated vehicles [17]. Regarding the feedback, possible issues were found regarding visual colour coding feedback, which could cause problems for users suffering from colour blindness. This evaluation is considered important for further development of the HMI of question.

The interaction for altering between upright and reclined seating position was found to be easy to perform. The control element was easily reachable from both the upright and reclined seating position, which is an important factor when designing HMIs according to the literature [17]. However, the interaction was found to be confusing when it came to intuitively know which interaction was intended to upright the seat, and which interaction was intended to recline the seat. This finding is considered crucial when further developing the HMI for altering between upright and reclined position, and is further confirmed in literature stating the requirement of developing HMIs which are easy to remember in order to achieve a smooth communication [16].

8.2 Discussion of Method

This section discusses the method which were used for the expert validation and online study, including the semi-structured interview and heuristic evaluation.

8.2.1 Discussion of Semi-structured Interviews

The semi-structured interviews with the both the experts and evaluators consisted of pre-defined open questions with the possibility to ask follow up questions, and narrowed down into more specific questions towards the end of the interview sessions. The approach was used with the intention to allow the interviewees to raise additional issues in order not to miss out on valuable feedback. However, due to the flexibility of this approach, it may be difficult for people other than the interviewer to interpret why a specific topic was focused upon in one interview rather than another. The flexibility of the semi-structured interview also makes it challenging to generalise findings. On the contrary, the possibility of asking follow up questions was frequently utilised allowing a more in-depth understanding of the feedback.

Another drawback of this method is that qualitative findings to some extent rely on the researcher's view about what is important, whereas another researcher may empathise with other issues. This approach of semi-structured interviews may thus suffer from a compromised objectivity. Due to these reasons, it is difficult to replicate research of qualitative approach, thereby weakening the reliability of the findings.

When it comes to the semi-structured interviews in the validation sessions with the experts, meaningful input was received on many aspects. Not only did the session enable collecting valuable feedback regarding the validation of the study, but also on ethical aspects to consider when conducting user studies. This shows the flexibility of the method, as well as it shows its reliance on the researcher's view about what is important.

Every expert seemed to have their own area of interest when it came to the development of the study, which generated feedback covering a broad scope of the first draft. This shows the benefit of collecting feedback from several experts of different backgrounds and experiences.

8.2.2 Discussion of Heuristic Evaluation with Video

The heuristic evaluation was limited to only consist of investigation of the system through a video, and not through interaction with the system itself. Performing a heuristic evaluation while using the system may possibly yield different results, as the used approach only collected the evaluators' impression of the system from the inspection of a video, without putting evaluators into the real context. One particularly interesting phenomenon from the online study which amplifies this assumption is when one evaluator changes opinion of the HMI after mimicking the interaction while thinking out loud. This could imply that performing the interaction yields an experience which not necessarily corresponds to the first impression obtained by only watching the interaction in a video, which is further stated to be common according to findings from literature [9, p. 209]. The results from the online study should thus be interpreted carefully.

Discussion of Recruited Evaluators

Three evaluators were used in the heuristic evaluation. Even if this amount was stated to find 65% of the usability problems, the proportion of problems found also depends on the evaluators and how the session is conducted. Due to this factor, it is difficult to say how big proportion of the usability problems that were found with this method.

The evaluators in the online were all Autoliv employees, which may have an effect on the results, as the employees are assumed to have a greater interest and knowledge within the subject compared to end users. The evaluators do thus not correspond to the end users. Nevertheless, they were not specialists within the field of the subject, nor within usability, which is ideal when carrying out heuristic evaluations, as they are shown to find many of the usability problems [9, p. 162]. However, even though the evaluators were not specialists in the field of usability engineering, nor within the interface or prototype vehicle itself, it may be assumed that their expertise within their specific automotive area may have been an advantage when pointing out problems and possibilities in the seating positions and HMIs.

9 Recommendations on User Study Method

From the results and discussion of the expert validation, six important parameters are identified when it comes to developing a user study of high degree of validity, assessing the requirements on the reclined seating position and HMIs in automated vehicles. These six parameters are the following; recruit relevant participants, provide the participants with consistent tasks, provide the participants with adequate time constraints, avoid social influences, use applicable data collection methods, and lastly a pilot study is recommended to target aspects which could not be validated with solely expert validation. In addition to recommendations for conducting a user study of high degree of validity, recommendations of ethical considerations derived from the findings from the expert validation are also presented in this chapter.

9.1 Recruiting Relevant Participants

Recruiting participants which correspond to the end users is of great importance as the aim is to investigate end user preferences and requirements. It is thus important to acknowledge that the limitations of exclusively carrying out the study with Autoliv employees decreases the validity since the participants will not correspond to the end users.

Another recommendation derived from the feedback from the experts is to collect the participants' attitude towards technology prior to the user study, as this may affect their answers on their experience of the seating position and HMIs. Moreover, it is recommended to also assess if the participants commonly experience motion sick while travelling, in order to distinguish if the experiences of the user study features specifically triggers motion sickness, or if it is something they experience while travelling in today's vehicles as well.

9.1.1 Further Defining Prioritised Characteristics of Results

Recruiting the correct participants for the user study, requires choosing between clustering participants before recruiting, or having a more demographic approach without targeting specific groups. In order to chose the most suitable approach, the results of highest priority must be defined. On one hand, as the user study is in such an early stage aiming for an exploratory approach, not targeting groups but rather having a more demographic approach collecting background data on the participants available might be sufficient. On the other hand, if more quantitative results are of priority, with the aim of carrying out a statistical analysis, clustering on parameters which are expected to yield differences in the results would be appropriate. If so, the parameters used in the clustering should be defined carefully depending on which parameters are to be statistically analysed. The recommendation is thus to clearly define which results are of priority prior to recruiting participants, in order for the study to yield valid results, as well as to avoid collecting excessive amount of data, wasting valuable time of staff and participants involved in the study.

The chapter of discussion also raises the question of whether or not it is reasonable or not to develop a study which aims to target both the experience of two HMIs, as well as the experience of the reclined seating position. If it concluded that the desired results of the reclined seat and HMIs are of different characteristics, two separate user studies would be recommended. For instance, if it is of priority to statistically analyse parameters such as if the preferences on the reclined seat depend on anthropometric measures, clustering and recruiting more participants would be required for investigation of the reclined seat. Whereas if the investigation of the HMI aims to yield more qualitative data, less participants could be used for a more exploratory approach. If this would be the case, one study could thus aim to focus on the reclined seating position, whereas the another one could prioritise the. This further verifies the need of a more detailed definition of the characteristics of the desired outcome, which is one crucial recommendation when it comes to recruiting relevant participants.

9.2 Consistent Tasks

Providing the participants with consistent tasks is important in order to ensure that the outcome of the user study to the extent possible corresponds to the real situation. It is thus recommended to provide each participant with their optimal reclined seating position, enabling the participants to choose the angle themselves according to their desires, when aiming to assess the end user requirements of the reclined seating position.

Furthermore, the user study should be conducted to avoid influences from learning effects and experience over time. It is thus recommended to provide the participants with the opportunity to try out the system in the so called *lap zero*, where they may become familiar with the system before starting the actual user study where they are to express their experience and thoughts of the system. This approach further provides participants with the consistent tasks, as the aim is not to assess the learnability, nor the first experience of the system. For the same reason, lap 2 and 3 are recommended to be balanced, having half of the user studies carried out with lap 2 before lap 3, and half of the user studies carried out with lap 3 before lap 2.

9.3 Adequate Time Constraints

Providing the participants with enough time to execute each task and to try out the system in general is recommended in order to avoid stressful user study sessions. Enough time must be scheduled ensuring that the participants can take the time they need to express their experience and thoughts. Moreover, this sets the prerequisites for a relaxed and comfortable user study environment. Inadequate time constraints are factors which may influence the user experience negatively, and yield results of lower validity. Hence, the recommendation is to carefully plan the time constraints when further developing the user study.

9.4 Avoid Social Influences

Social influences should be avoided in order to mimic the real environment to the extent possible throughout the user studies. Therefore it is recommended to reduce the amount of operators present in the prototype vehicle during the user study, to the extent possible without compromising the safety. Thus, it is recommended to remove the leading operator from the back seat, whose only reason to be in the vehicle is to observe the participant, as cameras will be setup in the vehicle enabling indirect observations. The only operator required in order to not compromise safety is safety operator sitting in the front passenger seat by the emergency brake pedals, whom thus is recommended to be kept in the vehicle during the user study. Furthermore, in order to handle resources more efficiently, the system operator and safety operator could be the same person. This would reduce the number of persons present during the user study session, and thus decrease the social influence. Lastly, in order to target the social influence, it is recommended to include a question regarding this aspect in the interview.

9.5 Applicable Data collection Methods

The words in the differential scale are altered based on the expert comment in an attempt to reduce the risk of overlap or misinterpretation. Moreover, the intended way of how to interpret the pair of words is explained.

9.5.1 Questionnaires for HMIs

The adjusted differential scale which is recommended for the HMI is shown in Table 8. The scale which rated if the feedback was fast or is removed from the questionnaire scale, and instead phrased in an interview question targeting the speed of the transition, allowing the participants to express how they experienced it. The scale rating between easy and difficult is removed as it is considered to overlap with statements 2 and 3 in the SUS, see Table 9. Moreover, in order to measure subjective satisfaction of the usage, which also was suggested in the feedback from experts, the pair of pleasing and irritating is added.

Table 8: Revised Differential Scale used to collect semi-qualitative data of the general user experience of the HMIs. The scale is to be used along with the question "How did you experience the interface for altering between active and automated drive" and "How did you experience the interface for altering between upright and reclined seating position?" The instruction will be for the participant to mark the box which best reflects their experience of the HMI.

Safe							Unsafe
Practical							Impractical
Intuitive							Not Intuitive
Pleasing							Irritating

A further description of how to interpret the differential pairs is shown below:

- *Safe* or *Unsafe* regards to whether or not interacting with the interface is experienced to put the user at risk.
- *Practical* or *Impractical* regards to whether or not the interaction with the interface is experienced to be appropriate to execute the task of question.
- *Intuitive* or *Not Intuitive* regards to whether or not the intended interaction with the interface is experienced to come natural.
- *Pleasing* or *Irritating* regards to whether or not the interaction with the interface is experienced to be enjoyable.

The SUS is recommended to be used in order to collect a quantifiable measure of the HMIs, which may be used to compare the initial HMIs with revised versions of the HMIs. It is also recommended as it is a well know and commonly used method in usability studies. It is discussed if it will be too time consuming to use both the SUS and differential scales, however it is recommend use the whole SUS in the pilot to start with. If it shows to be inefficient it may be reduced to only use statement 3 and 8 which are shown to yield a high degree of validity, or be converted into interview questions as suggested by the experts. The SUS may be seen in Table 9.

Table 9: The SUS which is recommended to be used to collect data of the user experience of the usability of the HMIs.

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

9.5.2 Questionnaires for Seating Positions

The adjusted differential scale which is recommended for the seating positions is shown in Table 10. The first draft of this questionnaire included the pair "in control and out if control", which was removed as it did not directly answer to the question. The recommended pairs of words are considered interesting as they indicate to what extent the experience of the reclined seating position would be wanted in an automated SAE L4 vehicle. By using the same scale for the reclined and upright position, it is possible to consider ratings in relation to one another.

Table 10: Differential Scale used to collect semi-qualitative data of the general user experience of the seating positions. The scale is to be used along with the question "How did you experience the upright seating position?" and "How did you experience the reclined seating position?" The instruction will be for the participant to mark the box which best reflects their experience of the seating position.

Comfortable							Uncomfortable
Safe							Unsafe
Practical							Impractical
Natural							Unnatural
Suitable for road gazing							Not suitable for road gazing
Suitable for window gazing							Not suitable for window gazing
Necessary							Unnecessary

A further description of how to interpret the differential pairs is shown below:

- *Comfortable* or *Uncomfortable* regards to assessing in what extent usage of the seating position is experienced to be comfortable or not.
- *Safe* or *Unsafe* regards to assessing in what extent usage of the seating position is experienced to put the user at risk.
- *Practical* or *Impractical* regards to assessing in what extent usage of the seating position is experienced to be appropriate to in the context of an automated L4 vehicle.
- *Natural* or *Unnatural* regards to assessing in what extent usage of the seating position is experienced to be natural.
- *Suitable for road gazing* or *Not suitable for road gazing* regards to assessing in what extent usage of the seating position is experienced to enable vision of the road.
- *Suitable for window gazing* or *Not suitable for window gazing* regards to assessing in what extent usage of the seating position is experienced to enable vision out the windows.
- *Necessary* or *Unnecessary* regards to assessing in what extent usage of the seating position is experienced to be a requirement in an automated L4 vehicle.

9.5.3 Interviews

When it comes to the interviews, a question of the experience of the lap is recommended amongst the general questions aiming to assess the experience of the automated driving system in general, as suggested by experts. Moreover, a question regarding how the user would feel about using the system alone, without the operator present is recommended, see question 1 and 6 in Table 11. Lastly, a question regarding additional comments is recommended to be asked after each lap, see question 7 in Table 11.

Table 11: General interview questions recommended to be asked after each lap.

General questions after every lap
1. How did you experience this lap?
2. How did you experience the HMI/seating position?
3. What did you like, and not like about the HMI/seating position?
4. What advantages and disadvantages do you see with this HMI/seating position?
5. Why do you think people would, or would not want to use this HMI/seating position?
6. How would you feel about using the system alone, without the operator present?
7. What additional comments do you have, which you feel you did not get the chance to express?

Regarding the HMIs, a question regarding what the users would like to add or remove from the interface is recommended, see question 5 in Table 12.

Table 12: Specific questions recommended to be asked after the laps assessing the experience of the HMIs.

Specific questions after HMI laps
1. How did you experience the feedback of the HMI?
2. In what way would you prefer to receive feedback?
3. In which modality would you prefer to receive feedback?
4. If you could choose freely, how would you like to interact with the vehicle to take over and let go of control/ to recline and upright the seating position?
5. What would you like to add to, or remove from the interface?
6. To what extent do you think this HMI is suitable in the context of an automated SAE L4 vehicle?

When it comes to the interviews of seating positions, instead of asking for important parameters of the seating positions, the question is rephrased to ask for additional features, see question 4 in Table 13, as this phrasing was understood more easily in the online interviews. Thereafter, it is recommended to ask of features which are of extra interest in the study see question 5 in Table 13, in order not to influence the users when asking about their preferences.

Table 13: Specific questions recommended to be asked after the laps assessing the experience of the seating positions.

Specific questions after seating position laps
1. If you could choose freely, how would you like to be seated when in automated SAE L4 drive, and why?
2.Which seating position would you prefer (upright or reclined) and why?
3. Which additional features of the seating position in an automated SAE L4 vehicle, are of highest importance for you to feel comfortable?
4. To what extent is *other feature of interest in the study* important for you to feel comfortable when travelling in an automated SAE L4 vehicle?

The recommendation is to keep the general interviews after each lap, in order not to miss any valuable input from the users. However, the specific questions after the seating position laps can be asked exclusively after the last seating position lap, as the laps assessing the seating position experience are recommended to be balanced. This is recommend in order to avoid an overlap.

9.6 Aspects to Test in Pilot Study

Prior to carrying out the user study it is recommended to conduct a pilot study. The aim of the pilot would be to find out if instructions, questionnaires, or interview questions are difficult to interpret, as well as to test if there is enough time for each task and user study session in total. Overall, pilot testing may reveal inconsistencies in the definitions of the user study method which are difficult to find without a trying out the method. One particular question which was discussed to be in need of further investigation in the pilot study is how to target the experience of motion sickness. It is thus recommended to try out different approaches, both asking directly for motion sickness in the end of the interview in order to avoid influencing the direction of the interview too much, as well as asking for signs of motion sickness such as unpleasantness or tiredness.

9.7 Ethical Aspects of Recommended User Study

Apart from recommendations for developing a valid user study, this section focuses on ethical aspects to consider when carrying out the recommended method, which were derived from the semi-structured interviews with experts. When using human subjects in user studies, it is of great importance to conduct studies which show respect for the participants' well being and emotions. This implies that the user study must give the participants the prerequisites to feel as comfortable as possible before, during and after the test.

Before the test, participants must be provided with clear introduction of the user study procedure, as mentioned in the chapter of The Experimental Setup. In addition, the expert interviews highlighted the importance of making sure that every participant receives the same information. This sets the basis for recommending a checklist covering the most important parts, including:

- Reminding the participants that the purpose is to evaluate the reclined seating and HMI, not the participant
- Reassuring participants that results from the user study will be held confidential, and that their individual test will not be identified
- Reminding the participants of their possibility to stop the test at any time, and that the user study is voluntary
- Informing about the potential risk of experiencing motion sickness

where the three first points were recommended already in the experimental setup, whereas the last point is a recommendation derived from the expert interviews. Regarding ethical aspects during and after the user study, the recommendations is to follow the suggestions stated in the chapter of Experimental Setup, in the section of Experimenter and Operators.

10 Guidelines on End User Requirements

This chapter presents guidelines of the end user requirements on the reclined seating position and HMIs in in Autoliv's SAE L4 prototype vehicle. The guidelines may be used to refine the existing design, to thereafter subject it to user testing. Such an iterative approach enables finding further end user requirements and increasing usability.

10.1 Guidelines on the Reclined Seating Position

The guidelines presented below are developed with the aim to increase usability of the reclined seating position.

Window and Road Gazing while Reclined
By enabling window and road gazing while reclined, the driver is provided with the possibility to remain reclined while still being able to watch the road and surroundings, maintaining the feeling of control while reclined.

Feedback on Adapting Restraint System when Altering Between Positions
By providing feedback regarding that the restraint systems have adapted to the seating position of question, the driver receives information of the system state and if it works reliably, increasing the communication between vehicle and driver.

Emergency Stop Function
By enabling an emergency stop function, the driver will have the possibility to take control quickly in case of an emergency, increasing the feeling of safety while reclined. On the contrary, such a feature could yield safety risks if used by accident and would require to be developed carefully.

Leg Support and Massage
By enabling leg support and massage, the relaxing activity in the reclined seating position can be improved.

10.2 Guidelines on the HMIs

As mentioned in the delimitations in the chapter of Introduction, only brief guidelines on the HMI will be provided due to confidentiality reasons. The guidelines are presented below.

Feedback in All Modalities
By providing feedback in all modalities, the risk that the driver misses valuable communication from the vehicle is decreased.

Redundant Visual Feedback
By providing visual feedback which is redundant, the risk that the driver misses out on valuable communication due to conditions such as colour blindness is decreased. This could be achieved by not only providing visual feedback in the form of colour coding, but also with symbols.

Investigate Intuitiveness of HMI for Altering Between Seating Positions
By further investigating which of the interaction is intuitive for reclining and righting up the seat respectively, the HMI can be refined to avoid confusion.

11 Future Work

In the future, studies are recommended to be performed beyond the limitation of the pandemic COVID-19, which deployed during the execution of the thesis and forced the studies to be carried out by remote. There is a great need of studies in the context of the vehicle, not only the recommended user study, but also heuristic evaluations inside the vehicle. Interaction with the reclined seating position and its belonging HMIs is needed to further identify usability problems, and enable continued refinement of the existing system. It is recommended to continue with studies in a dynamic environment after refining the design, in order to utilise the benefits of altering between heuristic evaluation and user testing.

When it comes to the dynamic driving environment, more meta-methods would be beneficial. In particular, a pilot study should be conducted in order to ensure scheduling enough time for the studies as well as to verify that the interview questions and questionnaires are interpreted properly by the participants. The pilot study should be used to address the inconsistencies which could not be validated with solely the expert validation.

Finally, the developed and validated user study method should be carried out to assess the end user requirements and preferences on the reclined seating position, the HMI between active and automated drive, and the HMI between upright and reclined seating. In the future, more excessive studies with participants other than Autoliv employees would also need to be carried out, as well as it would be interesting to conduct dynamic studies of longer duration, as well as dynamic studies on road to increase the ecological validity.

12 Conclusions

This chapter presents the conclusions by answering the research questions phrased in the chapter of Introduction.

RQ1. *How should a valid user study be conducted in order to successfully assess the end user requirements and preferences on; the reclined seating position and its HMI for transitions between upright and reclined mode; and also the HMI for automated vehicle transitions between manual and automated mode in an SAE Level 4 vehicle?*

A successful user study of high degree of validity is achieved by; recruiting relevant participants which implies clearly defining the prioritised characteristics of the results, providing participants with correct tasks, providing participants with adequate time constraints, avoiding social influences, utilising applicable data collection methods including questionnaires and semi-structured interviews, and finally subjecting the study to pilot testing.

RQ2. *What are the initial guidelines of end user requirements on the SAE L4 prototype vehicle from Autoliv, derived by remote user evaluation??*

The initial guidelines of end user requirements on the reclined seating position include enabling window and road gazing while reclined, providing feedback on adapting restraint systems when altering between positions, providing an emergency stop function, and enabling leg support and massage.

The initial guidelines of end user requirements on the HMIs include providing feedback in all modalities, making visual feedback redundant and investigating the intuitiveness of the HMI for altering between positions to avoid confusion.

References

[1] S. Koppel, J. Jiménez Octavio, K. Bohman, D. Logan, W. Raphael, L. Quintana Jimenez, and F. Lopez-Valdes, "Seating configuration and position preferences in fully automated vehicles", *Traffic injury prevention*, pp. 1–7, 2019.

[2] L. Jakobsson, "Assessment of Passenger Safety in Future Cars", *Application within FFI Trafiksäkerhet och automatiserade fordon*, 2017.

[3] V. A. Banks, A. Eriksson, J. O'Donoghue, and N. A. Stanton, "Is partially automated driving a bad idea? Observations from an on-road study", *Applied Ergonomics*, vol. 68, pp. 138–145, 2018.

[4] C. Sentouh, J.-C. Popieul, S. Debernard, and S. Boverie, "Human-machine interaction in automated vehicle: The abv project", *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 6344–6349, 2014.

[5] I. Pettersson and I. M. Karlsson, "Setting the stage for autonomous cars: a pilot study of future autonomous driving experiences", *IET intelligent transport systems*, vol. 9, no. 7, pp. 694–701, 2015.

[6] T. Wada, "Motion sickness in automated vehicles", in *Advanced Vehicle Control: Proceedings of the 13th International Symposium on Advanced Vehicle Control (AVEC'16), September 13-16, 2016, Munich, Germany*, 2016, p. 169.

[7] C. Diels, J. E. Bos, K. Hottelart, and P. Reilhac, "Motion sickness in automated vehicles: the elephant in the room", in *Road Vehicle Automation 3*, Springer, 2016, pp. 121–129.

[8] A. Bryman, *Social research methods*. Oxford university press, 2016.

[9] J. Nilsen, *Usability Engineering*. Morgan Kaufmann, 1993.

[10] Vinnova, *Detta är FFI*, `https://www.vinnova.se/m/fordonsstrategisk-forskning-och-innovation/om-ffi2/`, Online; accessed 22 January 2020, 2019.

[11] L. Groenesteijn, S. Hiemstra-van Mastrigt, C. Gallais, M. Blok, L. Kuijt-Evers, and P. Vink, "Activities, postures and comfort perception of train passengers as input for train seat design", *Ergonomics*, vol. 57, no. 8, pp. 1154–1165, 2014.

[12] S. Jorlöv, K. Bohman, and A. Larsson, "Seating positions and activities in highly automated cars–a qualitative study of future automated driving scenarios", in *International Research Conference on the Biomechanics of Impact*, 2017.

[13] B. Schoettle and M. Sivak, "A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia", University of Michigan, Ann Arbor, Transportation Research Institute, Tech. Rep., 2014.

[14] H. P. Ive, D. Sirkin, D. Miller, J. Li, and W. Ju, "" Don't make me turn this seat around!" driver and passenger activities and positions in autonomous cars", in *Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2015, pp. 50–55.

[15] A.-L. Osvalder, I. Hansson, and K. Bohman, "Children's and adults' comfort experience of extra seat belts when riding in the rear seat of a passenger car", *Traffic injury prevention*, vol. 16, no. sup2, S46–S51, 2015.

[16] O. Carsten and M. H. Martens, "How can humans understand their automated cars? HMI principles, problems and solutions", *Cognition, Technology & Work*, vol. 21, no. 1, pp. 3–20, 2019.

[17] F. Naujoks, K. Wiedemann, N. Schömig, S. Hergeth, and A. Keinath, "Towards guidelines and verification methods for automated vehicle HMIs", *Transportation research part F: traffic psychology and behaviour*, vol. 60, pp. 121–136, 2019.

[18] Y. Forster, S. Hergeth, F. Naujoks, J. F. Krems, and A. Keinath, "Empirical validation of a checklist for heuristic evaluation of automated vehicle HMIs", in *International Conference on Applied Human Factors and Ergonomics*, Springer, 2019, pp. 3–14.

[19] M. Sivak and B. Schoettle, "Motion sickness in self-driving vehicles", University of Michigan, Ann Arbor, Transportation Research Institute, Tech. Rep., 2015.

[20] S. Salter, C. Diels, P. Herriotts, S. Kanarachos, and D. Thake, "Motion sickness in automated vehicles with forward and rearward facing seating orientations", *Applied ergonomics*, vol. 78, pp. 54–61, 2019.

[21] R. K. Sharma, D. Kumar, and P. Kumar, "Systematic failure mode effect analysis (FMEA) using fuzzy linguistic modelling", *International Journal of Quality & Reliability Management*, 2005.

[22] C. Van der Lelie, "The value of storyboards in the product design process", *Personal and ubiquitous computing*, vol. 10, no. 2-3, pp. 159–162, 2006.

[23] M. Bohgard, S. Karlsson, E. Lovén, L.-Å. Mikaelsson, L. Mårtensson, A.-L. Osvalder, L. M. Rose, and P. Ulfvengren, *Online Course: Work and technology on human terms*, 2009.

[24] D. Travis, "Measuring satisfaction: Beyond the usability questionnaire", *Online at http://www. userfocus. co. uk/articles/satisfaction. html*, 2008.

[25] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale", *Intl. Journal of Human–Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[26] J. R. Lewis and J. Sauro, "The factor structure of the system usability scale", in *International conference on human centered design*, Springer, 2009, pp. 94–103.

[27] K. Finstad, "The system usability scale and non-native English speakers", *Journal of usability studies*, vol. 1, no. 4, pp. 185–188, 2006.

[28] J. Sauro, "Measuring U", *Denver, Colorado*, 2018.

Appendix A FMEA Matrix

Likelihood	Severity
Extremely unlikely	Minor injury (first
Unlikely	Moderate injury (lost
Likely	Serious injury
Extremely likely	Major injury (RIDDOR
Almost certain	Catastrophic injury

Likelihood \ Severity	1	2	3	4	5
1	1	2	3	4	5
2	2	4	6	8	10
3	3	6	9	12	15
4	4	8	12	16	20
5	5	10	15	20	25

RISK ASSESSMENT

Plant/Location:	Dept/Area:	Activity/Process being	Assessors:	Date: 2020-

Hazard	Persons	Risk	Current Control Measures	Current Risk L	Current Risk S	Current Risk T	Revised Control Measures	Revised Risk L	Revised Risk S	Revised Risk T	Date Complete
OP gets distracted during the test	OP+ SOP +TP	OP forgets to brake in case of an emergency	OP makes sure to prioritize safety before talking. Questions by the TP should be handled outside the car.	2	3	6	Add a Safety Operator (SOP) in the vehicle, whose only task is to handle the emergency brake. Have the OP seated in the back seat to the right, leading the user study.	1	3	3	
TP drives too fast in the transitions (faster than instructed by OP).	OP+ SOP +TP	Anything that could happen would happen faster with less time to respond.	Extra set of brake pedal available to the SOP.	2	2	4	OP will instruct TP to follow OP's commands.	2	2	4	
Tire gets a puncture	OP+ SOP +TP	Flat tire causes vehicle to sway off road.	Inspect tire pressure before every test session. Keep speed below 25 km/h. Extra set of brake pedal available for SOP.	1	1	1	Extra brake pedal should allow the SOP to help slow down and stop the vehicle.	1	1	1	
No Airbag on driver side (TP) due to new type of steering wheel	TP	In case of crash there is no Airbag on driver (TP) side.	The seatbelt on the driver's side may be enogugh protection in case of a crash, as the speed will be kept below 25 km/h.	2	3	6	The SOP has control of the extra brake pedal and can stop the vehicle if any unexpected event should occur.	1	3	3	
Unexpected Object/Person/Animal on track	OP+ SOP +TP	Unexpected object enters test track during ongoing test.	SOP aborts test by using extra brake pedal.	2	2	4	Get instructions from test site management on how to behave on test track.	2	2	4	
Vehicle drives faster than 25km/h.	OP+ SOP +TP	Driver pushes throttle to increase speed above 25km/h.	SOP can press separate brake pedal to slow down vehicle. The AD function of the car is set to 30-35km/h.	3	1	3	OP informs TP not to drive above 25km/h at any time.	2	1	2	

FMEA of user study in dynamic environment, part 1.

Vehicle standard function failure	+SOP	Vehicle is not fit to drive due to standard vehicle function malfunctioning	Vehicle is serviced and maintained according to manufacturer's instructions.	2	3	6	SysOP carries out vehicle inspection prior to any usage of the vehicle and is situated in the nearby during the study.	1	3	3	
Unexpected/emergency landing on airfield	+SOP	Risk of colliding with the vehicle landing on the airfield.	SOP will need to be able to recieve calls from the emergency landing centre to be able to steer the vehicle away from the airfield.	1	4	4	Book the airfield and make sure that the Airfiled centre knows who to call in case of emergency.	1	4	4	
Vehicle disconnects from GPS	+SOP	Prototype vehicle disconnects from autonomous driving mode and switches to manual driving mode without a warning	SOP has an extra brake. If the TP is able to reach the brake and wheel, he/she may also stop the vehicle or steer away from dagerous situation.	3	2	6	The likelihood of this to occur is higher in the beginning, when starting the preototype vehicle, rather than druing the drive. The risk may be mitigated by a technical inspection prior to the test. Thus, the likelihood decreases.	2	2	4	
Unexpected behavoir/issues with AD functionality of prototype vehicle.	+SOP	Any unforeseen behavior of the AD functionality of the vehicle causes a safety critical event.	SOP has an extra brake. If the TP is able to reach the brake and wheel, he/she may also stop the vehicle or steer away from dagerous situation.	2	1	2	SOP's resposibility to abort test by using extra brake pedal	2	1	2	
TP becomes motion sick during the study.	+SOP	TP vomiting in the vehicle, causing delay of the user studies	Stop the test if TP feels uncomfortable. OP gives clear intstructions to the TP, regarding the TP's possibility to stop the tests whenever the TP wants to. Make it clear that every reason to stop the test is adequate.	2	2	4	Provide motion sickness bag in driver's door.	2	1	2	

Persons affected:
OP - Operator, TP - Test person, SOP - Safety operator, SysOP - System Operator, OE - Other employee, C - contractor, T - trainee, YP - Young person, PW - Pregnant worker

FMEA of user study in dynamic environment, part 2.

Appendix B Expert Feedback

Subject	Participants		
Expert	**Amount and age**	**Anthropometry and background**	**Driving habit**
1	Suggests a minimum of 28 participants. Age and gender can be divided into 4 groups: Men 25-45 years, men 46-65 years, and the same with female. 7 participants in each group	Divide in length (150-160, 161-170, 171-180, 181-190, 191-200 cm)	Scale of driving habit - use to recruit participants
2	Suggests having 5 studies -> redesign, another 5 studies, redesign, etc. As the study of HMI is in such an exploratory phase, this apporach might be suitable according to Nielsen.	Chose between actively dividing the participants into groups before recruiting, or having a more demographical approach "these backgrounds are what we got"	Apart from driving habit we could also use the techology adaption scale, to assess their attitude towards new technology.
3	No comments.	Only measuring length in the antropometric measures excludes the proportions. Two persons of different length may have different experiences, perhaps due to different proportions.	No comments
4	It the study is exploraory - fewer participants is ok (2-5) to verify if the system is usable. If a more quantifiable result is wanted, more participants could be good, which might be the case when assessing the user experience of the seating position.	No comments.	Combination of driving times per week and km previous year. Add what type of drive? City or urban?
5	No comments.	No comments.	No comments.
6	If the participants are to be clustured into groups of age, gender, length and driving habit, more participants will be required. If gaze is important for the study, the sight of the participants should also be checked. Depends on if the study aims to yield statistics. There should be an aim with all the chosen parameters. The chosen parameters should all be evaluated to see how they affected the results. Ask if they commonly feel motion sickness.	No comments.	No comments.
7	No comments.	No comments.	Days/week or days month could be a definition. Could also be formulated in scales "Had a drivers licence for a)less than a year, b) 1-5 yeears, c) over 10 years d) cannot answer". And "I drive a)every day, b) a few times per week, c) a few times per month, d) a few times per year, e) never, f)cannot answer"
8	A few more than 16 could be needed	No comments.	Important to state what kind of drivers the participants are; professional drivers, urban drivers, etc. Suggest measuring habit in km/week and km/year.

Expert feedback on experimental setup, part 1.

Subject Expert	Test route	Operators	Collection of data
1	No comments.	No comments.	No comments.
2	Hava a lap 0, as behaviour changes over time. This enables trying out the interactions to "get used to/learn" the system before evaluating it.	Having several operators present will affect the participant. Ask how the setup affects the user. How would the user feel about trying out the system alone?	No comments.
3	Hava a lap 0, as behaviour changes over time. Try out the interactions to "get used to/learn" the system. Change order of laps to avoid influeces from learning effects.	Is the test leader required to be in the vehicle during the test? As indirect observations can be made later thanks to the cameras, perhaps the test leader can be skipped in the vehicle.	Questionnaire first, interview after. Some questings might have to be more specific and "leading", which is why the interview should be second. This order also makes it possible to ask questions regarding their answer on the scales.
4	Have a test lap 0 as the user is introduced to a totally new environment. Ask the participants to think aloud in the initial test lap, could answer their questions afterwards before the "real" test. However, if the interactions are considered easy, it might not be neccessary to change the order of the laps.	As both the interface and test person is recorded, the leading operator might not be required in the vehicle.	No comments.
5	Have a lap 0. If a "think aloud"-lap is to be used, it should be after lap 0 and before lap 1.	No comments.	No comments.
6	Have a lap 0. If a "think aloud"-lap is to be used, it should be after lap 0 and before lap 1.	Remove the test leader, as there are recordings in the vehicle.	No comments.
7	Have a static introduction where the participants can try out the interactions and positions in the given context, before the actual test starts.	No comments	No comments
8	Could hava a static introduction and a lap 0. Therafter a brake in order for the participants to take in the information, before starting the actual test session.	No comments	Suggests to use only a few statements from the SUS, which are stated to yield a high degree of validity, in order to collect a quantifiable measurement since the interview covers qualitative data. Will send the article stating this.

Expert feedback on experimental setup, part 2.

Subject	Lap1		
Expert	**Setup**	**Scales**	**Questions**
1	Test the amount of interactions in a pilot to avoid too many transitions, making the situation stressfull.	Scales: 1-5 or 1-7 Likert-scale, instead of current 6-scale.	Start with general questions. "How was this experience in general?" If you could choose freely, how would you chose to have the information given to you? Could narrow down the question regarding feedback by asking about modalities (audio/visual/haptic feedback) as well.
2	Perhaps many transitions, however, probably suitable if the purpose is to test if the users use the system as intended. Will not match the behaviour in a real travel situation. In a real situation one might be seated in reclined for a long time and become confused when suddenly wanting to take over control. With this setup, it is rather an "automated" behaviour as the transitions are very frequent.	Remove the "unnecessarily" complex, in the scale to make it the opposite to "simple". Subjective assessment manikin SAM-scale to assess the feeling (calm/positive/tierd). Regarding scales: if the aim is to make comparisons between the answers in the test, a self made scale is ok. If the aim is to compare with systems outside the particular test, use a common method.	The open questions are suitable if the study is exploratory. Add the question "how did you experience the automated drive", in order to separate if they did not like the AD itself, or if they did not like the interaction.
3	No comments.	No comments.	No comments.
4	Seems like a valid setup regarding the amount of transistions, as the aim is to evaluate the usability of the transistions.	SUS seems suitable in this context as it tests a system. Could exclude some of the questions if necessary.	Consider if everyone understands the word "feedback". Have a definition ready in case of misinterpretation.
5	10 transitions in one lap might be too many. Could be divided into two laps, or decrease the amount of transitions.	The differential scale seems to overlap with the SUS. Could instead include value words, collecting more qualitative data which the SUS does not collect. SUS seems to work good if the system is quite ready. The six-graded scale could be used if it is important to not collect neutral data. However, participants might find it diffucult without a natural option.	Add an open question in the "grey box", such as "Would you like to add something, (or remove something) from the interface?"
6	Might be too many transitions	The opposite to easy is difficult, simple and complex are opposites. Tries to avoid scales of 5 and 7 to not collect neutral answers.	Avoid yes/no-questions. Connect more to the aspect of UX which is of interest, in this case the usability. Ask more questions about what the study aims to evaluate.
7	Might be too many transitions	Suggests to use only a few statements from the SUS, which are stated to yield a high degree of validity, in order to collect a quantifiable measurement since the interview covers qualitative data. Will send the article stating this.	No comments.
8	Might be too many transitions. Participants might become stressed, and might not have adequate time to actually take in the experience as they will alwas be focused on being ready for the next transition.	The scales are not achored to anything, making the difficult to interpret. However, if the HMI is to be redesigned and studied again, the same scale could be used to compare the two HMIs. Could skip SUS as the interview covers the questions.	One can interpret and evaluate the answers from the interview with the aid of the SUS, yielding a quantifible measure.

Expert feedback on lap 1.

Subject Expert	Lap 2		
	Setup	Scales	Questions
1	No comments.	Generally good. Comfort could be misinterpreted, fysical or psycological comfort? Make sure that the participant understands that it is not the seat itself that is evaluated. How they feel in the position is what's important.	Could make most of questions into questionnair with likert-scale. Regarding motion sickness, one may ask: "Did you feel any unpleasentness?"
2	One interview after 2&3 could be suitable , since the comparison between the positions is of interest in the study. However, should pilot this to understand what's more suitable.	"Natural/unnatural" - users will probably prefer different upright and reclined positions, depending on habit, length, preferences. "Reliable/unreliable" - what do we mean? Do we mean the AD function itself and if the user trusts it or not? Or do we mean the HMI, if the user trusts the information given from the vehicle.	"How do you experience being upright when in automated drive?" to point out that the "general" upright position in AD is evaluated. Regarding which parameters of the seating position are of highest importance to feel comfortable, perhaps this could be tested statically instead. The user could set the desired settings, and thereafter be asked which "parameters" made them chose the setting.
3	No comments.	Define natural/unnatural - one might thik that the optimal position is the reclined seating, but it might feel unnatural when actually testing. Define difference between safe/unsafe and reliable/unreliable.	Separate "safe" and "comfortable" in the question regarding preferences of parameters. Parameters needed to feel safe might not align wih parameters needed to feel comfortable.
4	No comments.	No comments.	Combined interview for lap 2 and 3 would be suitable. Perhaps lap 2 should be before lap 3, since the comparison relates to the upright position. However, the experience of upright autonomous positions is also new, which might make it suitable to have a varying order.
5	No comments.	Could add a few more differential scales, as it is quite easy to answer these scales. Perhaps use the attractdiff from Hassensal, ex "ugly/attractive", "human/technical", "simple/complicated".	Add options in the question regarding which parameters of the seating position they prefer, and have the participants rate the options and give the opportunity for them to add another option to include in the rating as well.
6	No comments.	Reliable/unreliable will possibly confuse the user. In control/out of control does not answer the question asked in the questionnaire.	Could have a ranked question regarding important parameters in the questionnaire. Rephrase to ask about experiece rather than impression, to achieve a more natural flow in the interview.
7	No comments.	The words seems adequate for the purpose. However, make sure that no important aspect is missed.	Ask for important parameters without giving options. Thereafter, add the parameters stated by the participants, and have the participants rate their own stated parameters along with provided options. Instead of asking if if they are able to look out the window, rather ask to what extent it is imoportant for them to look out the window.
8	No comments.	No comments.	Could ask about each parameter independently instead of ranking, e.g: To what extent is "parameter" important for you to feel comfortable when sitting in an SAE4 vehicle?

Expert feedback on lap 2.

Subject	Lap 3		
Expert	Setup	Scales	Questions
1	No comments.	Same as previous lap	Ask questions regarding comparison of the seating positions.
2	From a research-ethcal perspective, the participants should be aware of the risk of motion sickness before the test. Even if mentioning motion sickness will make them start thinking about it, the results will be comparable to each other, as all participants will recieve the same information. Important to have a checklist to ensure that everybody gets the same information. Every participant must be aware of the possibility to stop the test in case of motion sickness.	Same as previous lap	Same as previous lap.
3	No comments.	Same as previous	Same as previous lap.
4	No comments.	No comment s.	Ask specificly about the experience of upright and reclined, before starting the comparison. Could specify The question regarding their preferences by having them rate the options, and giving them the possibility to add an additional parameter. Ask about the extremes in their questionnaire - specifically about extra interesting parameters, to find out if their length had an influece of their experience. The final question could consider motion sickness.
5	Good to balance the laps, starting with lap 2 in 50% of the cases and lap 3 with 50% of the cases. Could have one interview for both laps. However, there could also be a point of having one interview/lap, as the last lap will be remebered more clearly than the first lap. For instance, the question regarding how they exerienced the lap, the experence of the individual positions, ams the questions in the grey box should be asked right after each lap. The comparing questions can be asked in a joint interview.	Same as previous lap	No comments.
6	Good to balance the laps, starting with lap 2 in 50% of the cases and lap 3 with 50% of the cases. Could have one interview for both laps.	Same as previous lap	Same as previous lap
7	Good to balance the laps, starting with lap 2 in 50% of the cases and lap 3 with 50% of the cases. Could have one interview for both laps.	Same as previous lap	Same as previous lap.
8	No comments.	No comment s.	Regarding motion sickness, one could ask for the degree of unpleasantness or discomfort in the end of the interview in order not to affect the whole interview. However, the first question regarding how they experienced the lap might enough for participants to mention if motion sickness was felt.

Expert feedback on lap 3.

Subject	Lap 4		
Expert	Setup	Scales	Questions
1	Decrease amount of transitions. Could rather have two identical laps with fewer transitions instead.	If one wants to assess if the participants have understood the interactions correctly, a scale could be used. Divide the tasks and have the participants number in which order they were performed.	Rephrase one of the open question to this: "If you could choose freely, how would you like to interact with the vehicle to recline/uprighten the seat?"
2	Same as lap 1.	Same as lap 1.	Ask regaring the speed of the transitions - slow or fast?
3	The transitions could be a factor affecting motion sickness, but also the speed of the transition itself. Is it neccessary to have 10 transitions? Could there be two laps instead? Perhaps 6-8 transitions are enough.	No comments.	Ask about additional commets, or if there was a question they hav expected which they did not get the cance to answer. Ask about motion sickness as a final question in the interview, in order to avoid leading questions in the beginning of the interview.
4	As the interaction itself is not complex, perhaps fewer transitions would be ok, between 6-8. Changing betweeen the seating positions too often might increase the risk of motion sickness.	Same comments as lap 1.	Suitable that the questions are "re-occuring" in each interview session, making the participant recognize the questions. If it is observed that the participant makes many incorrect interactions/seems confused, it would be good to ask about it in the interview.
5	Decrease amount of transitions. Could rather have two identical laps with fewer transitions instead.	Suggests using the whole SUS rather than only a few statements, since using the whole SUS will make the results comparable to other SUS-results as well. Moreover, one can still choose to only consider a few statements when the data is collected for all the 10 statements.	Same comments as lap 1.
6	Too many interactions.	Same comments as lap 1.	Same comments as lap 1.
7	Suggests to use only a few statements from the SUS, which are stated to yield a high degree of validity, in order to collect a quantifible measurement since the interview covers qualitative data.	Same comments as lap 1.	Same comments as lap 1.
8	Might be too many transitions. Participants might become stressed, and might not have adequate time to actually take in the experience as they will alwas be focused on being ready for the next transition.	Same comments as lap 1.	Same comments as lap 1.

Expert feedback on lap 4.

Subje Expert	The method as a whole	Additional comments	Articles, methods etc.
1	Suggests questionnair after each lap, and interviews after lap 1, 2&3, and 4. Balance lap 2 and 3, having 50% of the participants test upright first, and 50% test reclined first.	Using Autoliv employees will yield a biased result, and will not correspond to the end user, which is what the study aims to assess. This is a source of error. SUS would be suitable to use as it is comparable with other studies. Have an "Additional comments"-question after each lap.	
2	Could have a more explorative setup for lap 0-3, with open questions and feedback on the feeling of the different aspects. Last lap: more ratings, providing quantifible data, in order to assess the experience as a whole. Thus the participants have time to get used to the system.	The reclined position used in the study must be the optimal reclined position for each participant. If the tested reclined position is the same for everyone, the reslults regarding if the end users want to be seated in a reclined position or not will not be valid. Having an exploratory phase of the study could help finding out which parameters affect the users preferences. "Tall/short/men/women prefer this/that".	Subjective assessment manikin SAM-scale to assess the feeling (calm/positive/tierd) .
3	Could be beneficial to have the interview in the vehicle, as the user is still in the correct environment and be not lose time between the test and questions. Could be beneficial to change the order of laps, as it could be a source of error. Remove leading operator inside the vehicle?	Add more "why"-questions in the interviews. Be as minimalistic as possible - ask the questions and use the scales which will give valuable data. Combining questionnaires and interviews is however recommendable. Ask open question regarding their experience of automated drive after lap 0. Combine interview for lap 2 and 3 if the important evaluation is the comparison between the seating positions. The experience could change if the test route would be longer.	
4	No comments.	Make sure to test all possible scenarios: What happens if the participant ignores the instruction from the vehicle? Regarding number of scales, the six-scale is also able to indicate more neautral answers - the ones nearest the centre. Too many scales can be misleading as some people are very careful while some are more capable ofrating in the extremes.	
5	No comments.	Should hava a pilot trying out amount of transitions. The setup yields a quite artificial travel experience, since the airfield is empty. Could discuss this in the report, and suggest further studies which could include more distractions/mimic the real environment more to when studying the seating positions.	Value words which may be used in the differential scales https://www.userfocus.co.uk/articles/satisfaction.html
6	Perhaps skip the SUS and include the intereesting statements from SUS rephrased into questions in the interview instead.	Questions regarding motion sickness could be asked in the end of the interview. There are standardized scales for evaluating motion sickness. Define which variables are of interest, how concepts (of ux, usability) are defined and how are the operationalized. Based on thet the experiement can be developed. The last lap could build in more interactions to evaluate the system as a whole.	
7	Suggests to use only a few statements from the SUS, which are stated to yield a high degree of validity, in order to collect a quantifiable measurement since the interview covers qualitative data. Lap 1 and 4 both evaluate HMIs, thus it might be suitable if lap 4 comes second, dividing the user study into two sections: first focusing on HMIs, thereafter focusing on seating position	There are validated "simulator motion sickness questionnaire" scales which have been used in simulators, could look this up. However, perhaps one should not ask directly for motion sickness, but rather for signs of motion sickness, such as tiredness etc.	How to reduce statements in SUS: https://measuringu.com/single-item-sus/
8	Make sure enuogh time is scheduled for each participant, and having a break between lap 0 and the actual test.	Discuss the ecological validity of the user study, as it is performed in a controlled driving environment in such a short time, compared to real driving situations.	

Expert feedback on the method as a whole, including additional comments and suggested articles and methods.