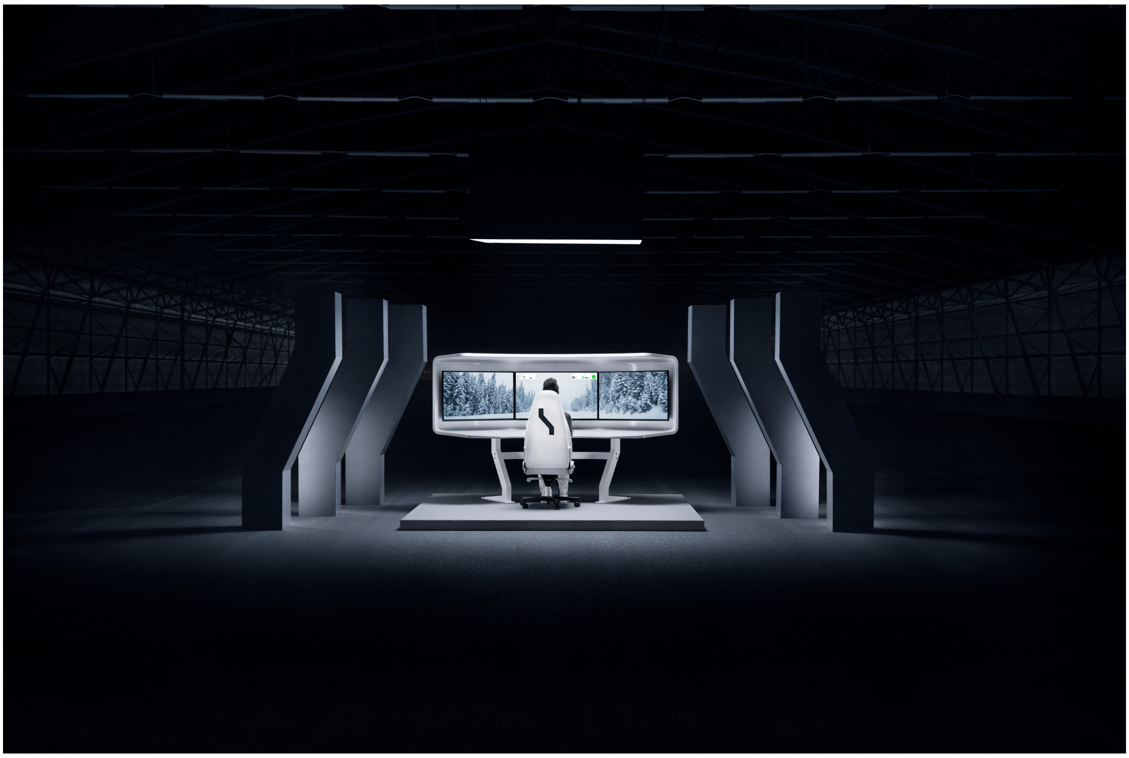


Visualisation and Interaction for Remote Driving

Promoting Situation Awareness and Mitigating Cognitive Load for Remote Operators

Master's Thesis in Interaction Design and Technologies

Julia Flising

Ellen Widerstrand

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2022

Master's thesis 2022

Visualisation and Interaction for Remote Driving

Promoting Situation Awareness and Mitigating Cognitive Load for Remote Operators

Julia Flising
Ellen Widerstrand

UNIVERSITY OF
GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Division of Interaction Design
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2022

Visualisation and Interaction for Remote Driving
Promoting Situation Awareness and Mitigating Cognitive Load for Remote Operators
Julia Flising
Ellen Widerstrand

Supervisor: Josef Wideström, Computer Science and Engineering
Advisor: Richard Brodie, Einride
Examiner: Staffan Björk, Computer Science and Engineering

Master's Thesis 2022
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Remote operator using the remote Pod operation station (Einride, 2021c).

Typeset in LaTeX
Gothenburg, Sweden 2022

Visualisation and Interaction for Remote Driving
Promoting Situation Awareness and Mitigating Cognitive Load for Remote Operators

Julia Flising
Ellen Widerstrand

Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

As the world scrutinises its transportation network in the glaring light of climate change, remotely operated vehicles are emerging as an attractive solution to the problem. Boasting better safety, reliability and sustainability for the future, standardised freight trucks may be the pioneering project of its time. However, the operator of a remotely operated truck faces a unique set of challenges that needs addressing not only for their well-being but that of each stakeholder, from pedestrians to industry actors. In this Interaction Design Master Thesis the aim has been twofold: to investigate what key interface factors may be used to enhance the operator's situation awareness when in demanding driving situations, and to examine what interface elements serve to mitigate cognitive load for the operator.

This Research Through Design project has employed frameworks from Human-Centred Design (HCD) and IDEO's 3i Model of Design Thinking to tackle the challenge through a user-focused process of two separate prototype usability tests. Each prototype was created from a foundation of formative research featuring methods ranging from literature reviews to expert interviews. The final product of this thesis work is a set of guidelines set to inform academia and industry actors alike of best-practice recommendations for the design of a remote operating interface. An iterative design process was carried out to ensure their good quality. These guidelines present ways in which to ensure the interface remains operator-supportive by tackling identified issues within automation, situation awareness and interaction design.

Keywords: Interaction design, remote operations, cognitive load, situation awareness, research through design, graphical user interface.

Acknowledgements

A special thanks to our academic supervisor Josef Wideström for his guidance and support through the entire process of bringing this Master Thesis from proposal to finished product.

Many thanks to our advisor Richard Brodie at Einride for providing invaluable insight and advice into the practical workings of not just the remote operating station, but the company as a whole.

We would also like to extend a sincere thank you to Team Remote Operate at Einride for their continuous support and interest in our work.

To all other Einride employees who volunteered their time and experience by partaking in interviews and recurring user tests, we are deeply grateful for your assistance.

A similar word of thanks goes to the volunteers from the Interaction Design and Technologies Master program at Chalmers University of Technology for lending their perspective and insight to the evaluation of our guidelines.

We are also grateful to Reto Weber at the Division of Maritime Studies at Chalmers University of Technology for taking the time to give us an inside look at the Full Mission Bridge Simulator.

Lastly, Ellen dedicates this thesis to her forever beloved grandmother Inga who never got to see the finished result.

Julia Flising, Gothenburg, May 2022
Ellen Widerstrand, Gothenburg, May 2022

Contents

List of Figures

1

Introduction

According to data from the International Institute for Sustainable Development, emissions in transportation account for around 64% of global oil consumption and 23% of the world's energy-related carbon monoxide emissions (Mead, 2021). Among the recommendations for continued electrification of the mobility sector, governments and organisations are urged to consider the implementation of structural changes to the transit system to reduce the negative environmental impact (Mead, 2021). This vision and intent are shared with industry actors such as Einride, a Swedish transport company specialised in the digitisation, electrification and automation of standardised freight fleets (Einride, n.d.-b).

While there has been a significant movement within the industry towards the development of fully autonomous vehicles run via AI, other projects have put their focus on unmanned vehicles operated through remote driving, also referred to as remote operation (Liu et al., 2017a). These solutions are argued to be more reliable and effective than a fully autonomous system when dealing with complex situations (Ackerman, 2015). Einride's invention, the Einride Pod, is the company's response to the sustainability challenge, employing a remote operation capability on top of implemented self-driving technology to achieve safer and cleaner transportation options (Einride, n.d.-a).

The remote driving concept may host significant potential in terms of efficiency, but it also comes with its own set of challenges. The human driver uses primarily four senses when driving: visual, aural, haptic and vestibular systems (Gnatzig et al., 2013). Remote operation places significant limitations on these kinds of inputs by physically displacing the driver from the vehicle, instantiating a need for conscious, human-centred design to ensure the safety and well-being of relevant stakeholders (Gnatzig et al., 2013). Due to the feedback modality of the human-computer interaction being largely limited to the visual modality, efforts to keep the vehicle operator in the loop instantiate a great need for the application of Interaction Design methods to ensure safe and effective cooperation between system and user (Endsley and Jones, 2014). One of the ways in which Interaction Design may influence the problem space is through the application of graphical user interface (GUI) elements. This factor may, on its own, prevent significant cognitive fatigue via excessive visual excise on the operator.

Situation awareness is another important factor in this work. Already heavily used in fields such as medicine, air traffic control and other automated systems, the concept provides crucial principles and insights into designing for enhanced decision making and performance (Endsley and Jones, 2014).

1.1 Context

Hosseini et al. mention that the main challenge of remotely operating a vehicle in urban areas is the lateral control of that vehicle (Hosseini et al., 2016). Lateral control is, according to Hosseini et al., manoeuvres such as overtaking where the operator requires more information than just the camera feed to be able to operate the vehicle robustly and precisely. The aural and visual systems provide the driver with their primary source of time- and position-based prediction, making these systems especially important for navigating traffic (Gnatzig et al., 2013). Furthermore, Gnatzig et al. have pointed to a lack of situation awareness producing the effect of a non-realistic driving feeling, causing the operator to behave differently compared to when operating the vehicle from directly inside of it (Gnatzig et al., 2013). This creates a potentially hazardous situation where the operator may behave unpredictably, too late, or in error due to their misunderstanding of the information presented to them on the GUI. This may also have consequences for secondary stakeholders such as pedestrians or motorists.

1.2 Design Problem

This Research through Design project will tackle the challenge of investigating and curating best-practice guidelines for developing GUIs within the remote operation sector. By incorporating principles of graphical interface design, cognitive science and human-centred design, this work seeks to address primarily visual challenges faced by remote operators in their interaction with the system.

Formative research will direct the development of several iterations of prototypes and design guidelines. The first iteration of guidelines will be put to the test by providing a theoretical foundation for comparative usability testing of two different prototype solutions. The following iteration will draw upon the learnings and results of the previous round to synthesise the second iteration of guidelines and prototypes. Evaluation of these prototypes and subsequent analysis of the results will help shape best-practice design guidelines. These guidelines will subsequently be put through a final evaluation with the purpose of ensuring their good quality and applicability. Due to the combination of theoretical and practical work involved, this thesis will seek to produce and test iterated design solutions with characteristics of low visual excise to counter a challenging and busy working environment. The goal will be to produce guidelines that support the creation of a comprehensive, continuous and safe interface solution that benefits all stakeholders.

1.3 Research Problem

Trucks with autonomous capabilities in combination with remote driving provide a more office-like work experience, but its physically-removed nature is not without its range of issues. Assessing and addressing a given traffic complication requires considerable cognitive reorientation on behalf of the operator, potentially instantiating heavy cognitive load due to the amount of visual information displayed (Sweller et al., 1998). The visual excise on its own runs the risk of fatiguing the operator and preventing them from reacting to critical traffic situations in time (Cooper et al., 2014). The visualisation of the operator's cabin features overlapping screens and views of considerable size and thus serves to exasperate existing situation awareness issues inherent to the remote driving setup.

Two research questions were identified for closer study, the first of which sought to investigate what principles within the field of Interaction Design may be applied to best support the operator's formation and maintenance of their mental model of situation awareness. The other looked to point out ways in which one might design with the specific purpose of reducing the cognitive load placed on the same remote operator.

- RQ1: What are key interface factors for enhancing the operator's situation awareness in demanding driving situations?
- RQ2: What factors of given interface elements serve to reduce the cognitive load as a part of the interaction between operator and truck in a remote driving setup?

Both research questions look to employ design paradigms with the goal of providing a more human-centred and effective operator experience. Though each question presents a slightly different focus, this outlook remains a common factor between the two. RQ1 questions how proper interaction design may enhance the operator's awareness of situational elements, such as the surrounding traffic, and reduce the risk of user errors, such as the misinterpretation of system messages. An example of a demanding driving situation, as posited in RQ1, would include contexts where one or several actions are required of the operator for the Pod to be able to safely and effectively continue working toward its established objective. RQ2 primarily investigates design measures which prevent information overload and other potential consequences of overbearing cognitive load.

1.4 Delimitations

The scope of this work mainly concerns the perception and interpretation of output from a remote operation interface to the operator and the implications this has for interaction. This work will not consider the very relevant options of aural and haptic feedback as a feedback modality, primarily due to technical constraints. Neither the chair nor physical instruments used in the evaluations could add haptic feedback, making it hard to evaluate its use. The feedback of the remote operating station and how the GUI conveys information will be considered based on how well it strikes a balance between sustaining a high amount of situation awareness with minimal adverse effects on the operator and their continued performance. This work will not investigate AR or VR as a system solution.

While also relevant to the mission of developing a more human-centred remote operator experience, this thesis will not make recommendations regarding hardware options or operator input devices. This is because designing for the visual modality falls more within the thesis authors' expertise. Furthermore, while the design focus remains within the said visual realm, efforts will not be made to develop unique iconography as this will not be sufficient for a thesis work in interaction design. Any icons created and added to the interface will strictly fulfil the purpose of demonstrating a functionality or evaluating a standing design principle. They will not have gone through any sort of evaluation regarding their suitability in terms of communication or industry standards.

It also bears mentioning that this work specifically will be investigating the remote operation of freight trucks and their unique situational context featuring tasks such as cargo loading, driving and vehicle maintenance. While inspiration will be drawn from multiple sources, including simulators and other remote operation systems, recommendations made as a part of this work will be mainly suited to remotely operated trucks. This work will thus not be looking into developing design advice for maritime operations, manufacturing or mining, even though the result might also be useful for such situations.

2

Background

This chapter covers relevant background regarding the thesis. The first section expounds upon the different stakeholders for this thesis, ranging from industry actors to academia. The last section discusses state of the art solutions within remote operations or related fields.

2.1 Stakeholders

As the most directly affected party of the remote operation issue, one of the primary stakeholders within this project are the operators themselves. They are at the centre of the remote operation shift and will ultimately be held responsible for any accidents that occur.

The practical definition of a remote operator is still in a certain state of flux as technological progress is made in the field, making it rather difficult to clearly outline the profession as a user group. At this current time, a class ‘C’ driver’s licence is required by Einride for operating a Pod on public roads, but some exceptions do apply (Gröndahl, 2022). The Swedish Transport Agency requires drivers to have a class ‘C’ licence to be able to drive heavy trucks Transportstyrelsen, 2021a. Currently, there is not a licence specifically for remote operations and this licence is, therefore, considered to be the best suiting due to the type of vehicle. As this still is a new concept it is likely that the regulations will change, which might allow people without this licence to operate a Pod or similar vehicles.

2.1.1 Einride

The thesis is done in collaboration with the company Einride and they, therefore, constitute a significant stakeholder. Einride helps their customers in scaling their operations and offers benefits from digital, electric and autonomous technology (Einride, n.d.-b). As part of their vision of working toward sustainable and intelligent solutions within standardised freight, the company has created an electric and autonomous truck, called Pod, that operates on public roads (see Figure 2.1) (Einride, n.d.-a). It currently operates on SAE level 4 with its monitored autonomous drive (MAD) mode and wide array of automated systems, though the implications of this will be explained further in section 3.5 (Einride, n.d.-a). The Pod currently requires limited regulatory oversight and can be controlled from the Einride Remote Operating Station. The station could eventually allow operators to simultaneously oversee and control several Pods, which would create a new type of work environment for the operators compared to the environment of truckers on the roads (Einride, 2021b). The goal is that this new type of environment should make it possible for few operators to operate as many Pods as possible.

Figure 2.1: The Einride Pod (Einride, 2021a).

Einride is aware of the fact that the interface of the remote operating station can be perceived as cluttered and overwhelming. As they do not have the resources to improve the interface themselves, they hope that this thesis can give them guidelines on how to best improve it. Einride is, therefore, an important stakeholder in their capacity as the owner of the original interface and product in question, as well as through their role as our client.

2.1.2 Secondary Stakeholders

There are several groups of people, including pedestrians and other traffic participants, that occupy the role of secondary stakeholders. Their presence will require the operator to be responsive, such as in the case of a pedestrian crossing the street, and they will also be directly affected if the system would fail. A system failure could potentially be very dangerous for these people as they are situated close to the Pod. If many accidents would occur with the Pod or other types of autonomous vehicles, they would likely be less well-received by the public and require stricter regulations - potentially reducing their value as a transport option.

2.1.3 Academia

This work seeks to formulate, develop and evaluate principles from a range of academic fields in the effort to provide an answer to the given research questions. As automation and other sectors progress in the development of remotely operated solutions, it will be in the interest of academia to keep abreast with the situation, paying attention to how this phenomenon may differ between application areas. The findings of this work will serve to inform researchers and students about how teachings from established fields measure up in the cutting-edge context of remote operations. Stakeholders from academia naturally include Chalmers University of Technology as they are the academic partner of this thesis work. However, stakeholders from academia also consist of any academic (institution or individual) looking to learn more about these given topics.

2.1.4 Interaction Designers

Beyond the theoretical, this work also provides practical learnings relevant to the field of interaction design and its practitioners. Similarly to the interest of academia, keeping an eye on design solutions applied to the ever-expanding context of human-centred design and human-automation interaction presents a teachable moment for any interaction designer. The solutions and principles part of this work will inform how well current practice fits with the given use case and may serve as relevant insight for designers looking to tackle similar design problems or use contexts moving forward.

2.2 State of the Art

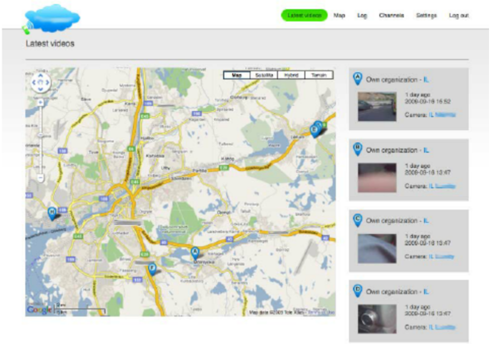
There are different setups for operating services and vehicles by way of remote operation. This section will describe current state of the art solutions by presenting three setups developed within different fields of emergency response, maritime simulation and remote trucking.

2.2.1 Rescue Service Emergency Response Centre

Organization et al. describes that a physical location where resources and information are coordinated during incidents is called an emergency operation centre (EOC) Organization et al., 2015. The purpose of an EOC is, according to Organization et al., to be a central part of incident management by having highly trained experts coordinate the response to an emergency. Organization et al. further describes that each coordinator has a workstation with a computer, displays and a telephone. The coordination done at an EOC is similar to the operations performed by a remote operator. Both have the responsibility of controlling what happens at a different location and they need to consider many different types of information from the interface viewed on a number of screens in front of them.

Bergstrand and Landgren have studied the field of emergency response to explore how live videos broadcast from mobile phones can be used to improve situation awareness between actors in command centre settings and accident sites (Bergstrand and Landgren, 2009). An application, called Liveresponse, was designed by Bergstrand and Landgren to explore the potential benefits of using video to better communicate when working remote. Response actors on site of an incident use cameras on mobile phones to capture clips from the incident. The device is activated by a hardware switch and the video is started and later stopped by pressing a button (Bergstrand and Landgren, 2011). The video clips are available to watch in the command centre just a few seconds later. The web application, used in the command centre, shows not only the videos shot at the incident (see Figure 2.2) but also dynamic map services (see Figure 2.3), along with traffic congestion data and live videos from traffic surveillance cameras (Bergstrand and Landgren, 2009).

Figure 2.2: Screenshot from a video taken at the scene of an accident (Bergstrand and Landgren, 2011).



Latest videos

Own organization - IL
1 day ago
2009-09-16 16:52
Camera: IL

Own organization - IL
1 day ago
2009-09-16 13:47
Camera: IL

Own organization - IL
1 day ago
2009-09-16 13:47
Camera: IL

Own organization - IL
1 day ago
2009-09-16 13:47
Camera: IL

Figure 2.3: The interface of the web application (Bergstrand and Landgren, 2011).

Bergstrand and Landgren let professional responders use the application in over 200 incidents over ten months and collected data on its usage (Bergstrand and Landgren, 2011). The data was later analysed to find out how the videos were used in the different locations, i.e. on-site and in the command centre, and how this affected the communication between the different actors. It was found that the broadcasted videos typically were quite short with a mean time of just under a minute. A reason for the short videos was, according to an incident commander interviewed by Bergstrand and Landgren, that broadcasting takes time away from doing other important activities. Some responders gave verbal comments while filming and the importance of those comments was emphasised by the command centre operators. It was also pointed out that shorter clips along with verbal comments are beneficial since it better highlights important aspects of the situation at hand.

According to Bergstrand and Landgren, video broadcasting adds visual understanding to a commander's report which can be difficult to express verbally (Bergstrand and Landgren, 2011). The video is both broadcasted and stored, making it possible to view the clips when there is time and also rewatch them several times. Bergstrand and Landgren mentioned that video conferencing, which is a common type of workplace video communication, is centred around people. Visual reporting, on the other hand, is object-focused and the possibility of repeatedly viewing videos encourages reflection.

It was concluded by Bergstrand and Landgren that video content is valuable for sensemaking (Bergstrand and Landgren, 2011). People who are not geographically present benefit, according to Bergstrand and Landgren, from viewing short clips of the situation as it gives them better insights into the situation which can help them in their ongoing work.

Another suggested way of incorporating video to improve emergency response is, as proposed by Al-Khafajiy et al., to use an application called Smart Hospital Emergency System (SHES) (Al-Khafajiy et al., 2019). Al-Khafajiy et al. describe that SHES aims to provide faster response time by allowing patients to attach media and their location through their smartphone when requesting an ambulance. The information entered by the patient can then be sent directly to the emergency responders which can help them form an early understanding of the situation and better prepare for treatment of the patient. Al-Khafajiy et al. further describe that the app also allows for video communication between the patient and the emergency operator which adds visual communication to the conventional method of phone calls.

2.2.2 Full Mission Bridge Simulator

Chalmers has nine maritime simulators, used for education and research in shipping, located in a simulator centre at Lindholmen in Gothenburg (Chalmers University of Technology, 2021a). One of the simulators is a full mission bridge simulator, seen in Figure 2.4, that allows complete simulations of operations that would be performed on real ships (Chalmers University of Technology, 2021b). Operations can be carried out in different environments and weather conditions. Ice can for example be simulated to practice operating in cold conditions where ice is present (Aylward et al., 2021).

Figure 2.4: The Full Mission Bridge Simulator at Chalmers.

The Full Mission Bridge Simulator is very similar to the bridge on a real ship, both considering its aesthetics and functionality. The inside of the bridge is almost identical to a real ship, but the major difference is what is seen outside the windows. A representation of the surroundings is used for the simulator. The representation is projected onto a curved projector screen located a few meters from where the windows are located on the bridge. Some of the other maritime simulators used a set of flat screens put together to form one big screen instead of a projector screen. A simulator with several smaller screens put together can be seen in Figure 2.5 and the screen used for the Full Mission Bridge Simulator can be seen in 2.6. One significant difference between the two screen setups is that the former has big black lines between the smaller screens which divides the image. The projector, on the other hand, shows one smooth image, which gives it a more realistic look.

Figure 2.5: A simulator where several smaller screens are put together to form one big.

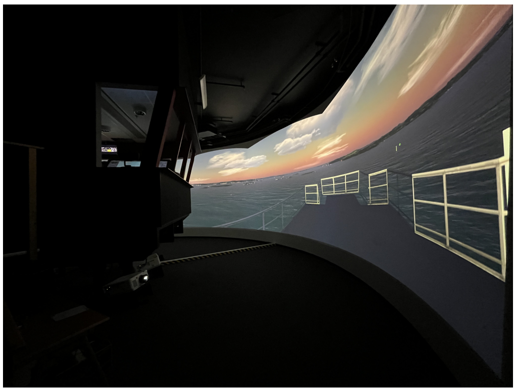


Figure 2.6: The projector screen for the full bridge simulator.

Just as for a real ship, the simulator requires at least two operators where one is the captain, and has the main responsibility of operating the ship, and one is the helmsman, who mainly is responsible for communication and the radar (Olsson and Jansson, 2006). There are many screens and different sets of steering equipment, which can make it hard to operate by yourself. There are many aspects to consider when operating a ship and different functionality is, therefore, used to aid the operators. Sound is, for example, used to notify the operators that actions need to be taken in different situations. One such situation is, for example, that the ship has entered shallow water.

2.2.3 Remote Trucking Operations

Phantom Auto is a company that offers teleoperation safety technology that enables remote driving of different types of logistics vehicles such as pallet jacks and trucks (Phantom Auto, n.d.). Phantom Auto describe that their platform can be used to assist autonomous vehicles by remotely handling edge cases that are difficult to handle autonomously. They further describe that it can also be used for distanced driver training which means that it enables real-time telepresence for supervisors when new drivers are training.

One company that has partnered with Phantom Auto to create remotely operated trucks is ITS ConGlobal (ITS ConGlobal, 2020). ITS ConGlobal is an integrated intermodal service provider that operates at over 120 locations in North America

(ITS ConGlobal, n.d.). Moving freight by more than one type of transportation is called intermodal shipping and it is often more cost-effective and sustainable compared to long-haul trucking (Yeager, 2020). The cargo to be shipped is loaded into intermodal containers, which eases the move between different types of transportation.

ITS ConGlobal has several types of vehicles involved in their operations, but ITS ConGlobal describe that it is specifically yard trucks they currently have integrated with Phantom Auto's technology (ITS ConGlobal, 2021). ITS ConGlobal describe that they started remotely operating trucks within the AllianceTexas Mobility Innovation Zone (MIZ) in November 2021. The MIZ is, according to ITS ConGlobal, a unique landscape that covers 27,000 acres and includes a full-ecosystem environment that can be used to test the latest technologies. ITS ConGlobal's container depot is also located in AllianceTexas, making it convenient for them to use. According to them, they mainly use the MIZ as a proving ground where they test remote operations and train using distance drivers. They further state that distanced driver training has been especially beneficial during the COVID-19 pandemic as it limits the need for physical interaction.

3

Theory

In the following sections, the theoretical background of the thesis is presented. These sections are organised according to the main fields of study as identified by formative research, ordered not according to relevance but rather to facilitate a linear exploration of the problem space.

3.1 Graphical User Interface

Sharp et al. describe that the first type of graphical user interface (GUI) was called WIMP as it consisted of windows, icons, menus and a pointing device (Sharp et al., 2019). Even though the design of the first generations of WIMP interfaces was boxy, the core of WIMP interfaces is quite similar to today's GUIs, according to Sharp et al. One major difference is, however, how the user interacts with the interface. The user interacted with WIMP interfaces using mouse and keyboard, but users today can interact with interfaces in many different ways, including touch. Similar principles and studies within the field of GUIs are thus directly relevant to the scope of this thesis. Interaction Design makes ample use of its learnings when working with problem spaces such as flight control towers or control rooms for nuclear facilities, looking to find the right setup to suit user needs.

There are more ways in which graphical user interfaces have evolved since the WIMP interfaces. Sharp et al. mention that there now is a much bigger selection of different types of icons and menus (Sharp et al., 2019). Audio can be incorporated for both icons and menus, icons can be animated in 3D and menus can be icon-based to be able to fit on a smaller screen.

3.1.1 Flow

When users interact with a product, the goal is for them to be productive and stay in the right mindset (Cooper et al., 2014). When someone is very focused on an assignment, they can lose track of their surroundings. The state one reaches when wholeheartedly focusing on an activity is called flow according to Cooper et al. This state is something that should be promoted as it often makes users more productive and less aware of distractions and may be achieved through the attentive addition of elements such as the call-to-action. A call-to-action is an interactive element typically aimed at inducing people into taking certain actions (Tubik, 2018). Things that disturb the users' flow should, however, be avoided as much as possible according to Cooper et al. Modeless feedback is, as described by Cooper et al., one way of avoiding the use of dialog windows, i.e. a regular disruptor of user flow. Modeless feedback avoids this by displaying in-depth visual information in the main view of an interface (Cooper et al., 2014).

3.1.2 Excise

Cooper et al. describe that a user who interacts with an interface typically does so because they have a goal that they want to achieve and to be able to achieve it, the user has to do some work (Cooper et al., 2014). That tasks involved in said work can, according to Cooper et al., be categorised as either goal-directed or excise. A goal-directed task is, as described by Cooper et al., something that directly brings the user closer to the goal while an excise task does not directly contribute to reaching the user's goal. Excise tasks are something the user has to do that is unnecessary or inefficient and should, therefore, be eliminated wherever possible. If excise is eliminated, users become more productive and will most likely have a better user experience as every step takes them closer to the goal.

Cooper et al. describe that excise can be divided into four types: navigational, skeuomorphic, modal and stylistic excise (Cooper et al., 2014). Actions the user has to do to get to a different part of the interface or locate functionality are rarely aligned with their goals and can be classified as navigational excise according to Cooper et al. Furthermore, Cooper et al. also describe that skeuomorphic excise is when mechanical representations are used in the interface. Mechanical systems are intended to be operated by hand and it is often harder to interact with a digital representation of a mechanical system. Elements in the interface should, therefore, not mimic something physical according to Cooper et al.

The third type of excise that Cooper et al. mentions is modal excise and it refers to interruptions that break a user's flow such as a modal error message or confirmation dialogue (Cooper et al., 2014). The user's flow should, according to Cooper et al., never be interrupted without a good reason. The authors also mention that users should be able to make the changes they want to without having to ask for permission. Another thing to consider regarding modal excise, as mentioned by the authors, is that input should be allowed wherever output is shown. This means that a user should be able to modify options, if they are changeable, where they are displayed.

Visual style should make information and interface behaviour clear to the user and if the user instead has to decode the information shown on the screen, it is classified as stylistic excise (Cooper et al., 2014). There are many different ways to eliminate excise and one way that Cooper et al. mention is to properly map controls to functions. It should be clear to the user what and how a control affects an object. Poor mapping both stops the user's flow and increases the user's cognitive load.

3.2 Information Visualisation

Spence describes information visualisation as the activity of appropriately representing data to help the user form a mental model of said data (Spence, 2014). Fekete et al. mentions many benefits of using information visualisation (Fekete et al., 2008). These benefits comprise of mainly cognitive aspects and the core of them is that visual representations can aid the processing and analysis of data and also aid memory. Other benefits that Fekete et al. mentions include reducing the user's search for different types of information and making it easier to identify and recognise patterns.

Fekete et al. describes that vision is the sense that has the largest bandwidth according to Information Theory and the visual modality is, therefore, the most suited for transferring information to the brain (Fekete et al., 2008). How a user perceives features and shapes can, according to Fekete et al., be divided into two main psychological theories called preattentive processing theory and the Gestalt theory. The first one, preattentive processing theory, focuses on pattern recognition and describes what visual features a user perceives rapidly. An important aspect of information visualisation is to choose a visual encoding that can be done preattentive for the most interesting parts of the data. Visual features that are possible to preattentive process include colour, orientation, width and length of lines, curvature and many more.

The second theory, Gestalt theory, focuses more on the whole visualisation rather than its details (Fekete et al., 2008). A user's understanding of visualisations is based on a set of principles. One of these principles is proximity, which describes that users perceptually group things that are located close to each other. Users often tend to also group similar items which is stated by the principle similarity. Another principle, closure, describes that a closed contour typically is perceived as an object. The list of principles described by Fekete et al. continues and the principles are also an important aspect to consider in information visualisation.

3.2.1 Norman's Action Cycle

How a user interacts with something can, in information visualisation, be described using Norman's Action cycle (Spence, 2014). Norman's Action Cycle, seen in Figure 3.1, consists of two gulfs: the Gulf of Execution and the Gulf of Evaluation. A user has a goal that they want to achieve and to be able to achieve that goal, they have to change the world. The world in this case can be anything from a door to a computer.

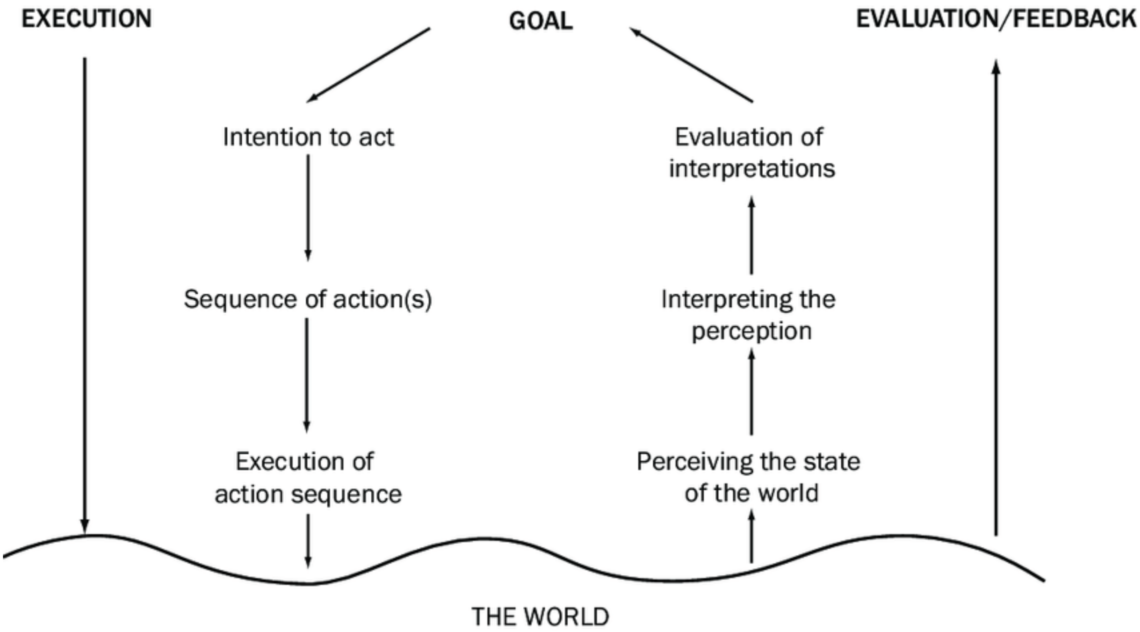


Figure 3.1: Norman's Action Cycle (Hooijdonk, 2022).

The Gulf of Execution comes first in the cycle and it reflects the gap between what a user perceives is possible to do and what interaction possibilities the world actually has (Spence, 2014). The Gulf of Execution consists of the three steps *formulating an intention*, *formulating an action plan* and *execution* according to Spence. Once the action plan has been executed, a change will often (but not always) occur in the world and the cycle continues to the Gulf of Evaluation. The Gulf of Evaluation starts with *perception*, moves on to *interpretation* and ends with *evaluation*.

3.3 Cognition

Cognition refers to a wide range of mental processes involved in the acquisition, storage, manipulation and retrieval of information - all divided across different functions that underpin many daily activities (Riesberg, 2016). Examples of these include processes such as memory, perception, reasoning and executive control (Riesberg, 2016). These processes are highly relevant for the development, maintenance and evaluations of an operator's understanding of a given remote operated problem space, such as a flight control room (Endsley and Jones, 2014). The simplicity of labels such as reasoning belies the complexity of each process, as they are all engaged in higher cognitive function (Riesberg, 2016). Furthermore, there are a plethora of proposed models attempting to best describe the workings of even a single process such as memory, further hinting at the complicatedness of defining mental processes (Riesberg, 2016). Starting with memory, this section will cover components and processes relevant for designing an interface that supports an operator's cognitive state and workings.

3.3.1 Memory

Memory as a function wears many hats in human cognition, supporting not only the short-term workspace of attention and reasoning but also the long-term storage of information and past experiences (Baddeley et al., 2015). The main difference between the two comes down to properties regarding the acquisition, storage and retrieval of information; working memory is limited in size whereas long-term memory is enormous. The acquisition of information into working memory involves a much simpler process than incorporating it for the long term, and working memory storage is much more fragile and prone to decay than that of long-term memory (Riesberg, 2016). As our understanding of memory evolved so did the models depicting its various functions, even to the point of which researchers argue whether short-term memory and working memory are equivalent concepts (Riesberg, 2016). This work takes the stance that the two are indeed interchangeable and will use working memory to refer to the short-term memory system.

Working memory is the workspace where a limited amount of information is stored temporarily, requiring active effort to maintain and is used to create and update mental models and projections for the future (Endsley and Jones, 2014). Originally described without the episodic buffer, this updated model of working memory was proposed by Baddeley in 2000. Baddeley et al. describes working memory as comprised of three parts: the phonological loop, the visuospatial sketchpad and the central executive (Baddeley et al., 2015). The model containing the three parts can be seen in Figure 3.2.

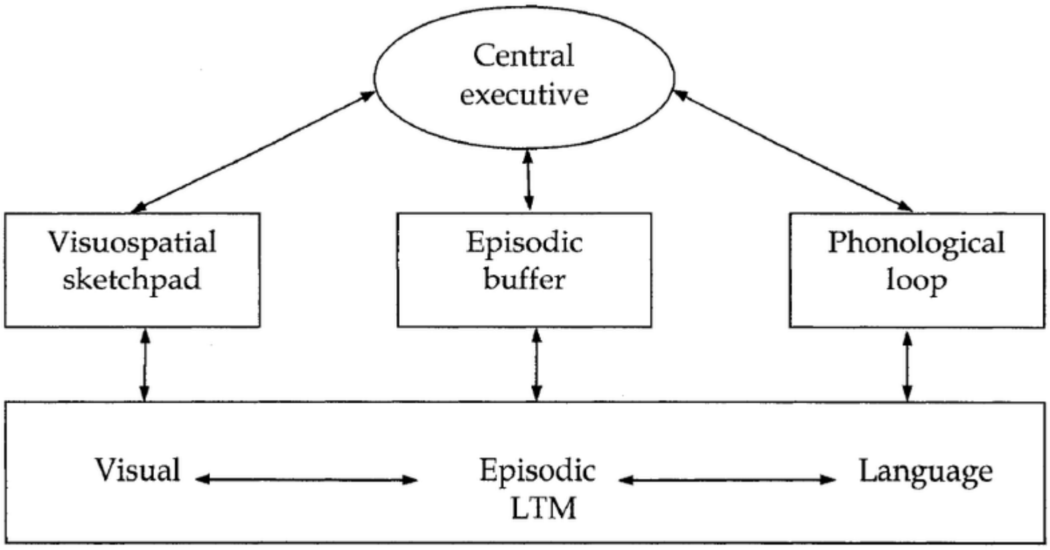


Figure 3.2: Baddeley multicomponent model of working memory (Baddeley et al., 2015).

In short, the phonological loop is a model of verbal short-term memory, the visuospatial sketchpad represents visual and spatial information while the episodic buffer allows for the various subcomponents to interact with long-term memory (Baddeley et al., 2015). The central executive of working memory has two separate modes of attentional control: the automatic and the supervisory attentional system (SAS) (Baddeley et al., 2015). The automatic mode maintains the responsibility of tasks that are based on existing habits and skills, requiring little conscious awareness whereas the SAS is triggered by novel situations requiring the formation or alteration of strategies (Baddeley et al., 2015).

Long-term memory is, according to Baddeley et al., the system with the capacity to store information over long periods of time (Baddeley et al., 2015). Baddeley et al. further describe that it typically is divided into explicit (declarative) or implicit (nondeclarative) memory, as seen in Figure 3.3. Explicit memory refers to memories open to intentional retrieval of either personal events (called episodic memory) or words and knowledge (known as semantic memory) (Squire et al., 1993). Implicit memory instead refers to information in long-term memory retrieved via the performance of actions rather than conscious recall or recognition, an example of which would be learned skills such as biking or driving (Squire et al., 1993).

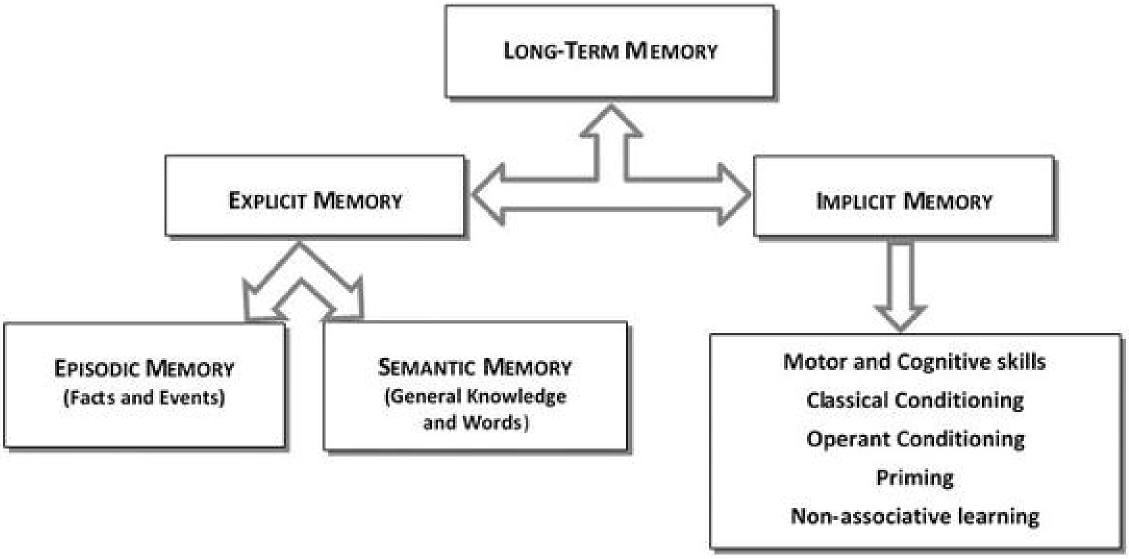


Figure 3.3: Components of long-term memory (Squire et al., 1993).

3.3.2 Mental Models and Schema

Long-term memory structures, known as schema and mental models, play a significant role in improving situation awareness (Endsley and Jones, 2014). Schemas are well-integrated chunks of knowledge about the world that are stored in semantic memory in the form of either scripts (the knowledge of events and their consequences) or frames (knowledge structures containing fixed structural information) (Baddeley et al., 2015). They can be said to be the prototypical stage of a mental model, allowing people to rapidly classify and understand information perceived similarly to shortcuts (Endsley and Jones, 2014). Mental models are defined as mechanisms by which humans can generate descriptions of a system's purpose and form, explanations of its functioning and observed system states as well as predictions of future states (Rouse and Morris, 1986). They are complex structures used to model the behaviour of systems, be they related to aviation, automation or medicine (Endsley and Jones, 2014). These models help people understand what information should be attended to at a given time and help fill in the blanks of a procedure should any information be missing (Endsley and Jones, 2014).

One significant advantage of mental models and schemas is, according to Endsley and Jones, that the situation for a given event does not need to be exactly like the previous one for it to be recognised as a match (Endsley and Jones, 2014). Endsley and Jones state that it instead may be formed outside of direct personal experience, such as through reading, and each factor establishes them as relevant tools for enhancing situation awareness. There is a mismatching issue within the domain of this use case where the implementation models of technological systems do not match the mental model of its users, potentially causing usability issues and reduced performance (Cooper et al., 2014).

3.3.3 Attention and Cognitive Load

Despite every effort, there are limits to human cognition, making the maintenance of ideal awareness and understanding of the environment an impossible task (Riesberg, 2016). Counter-intuitive cognitive phenomenons such as change blindness (the observer's inability to detect changes in scenes they are actively observing) makes evident that perception is, in fact, an active process (Riesberg, 2016). Apart from mental models and schemas, human cognition is equipped with tools such as selective attention to help us parse our environment and direct our attention more purposefully (Riesberg, 2016).

Selective attention is the taking possession of one of what appears to be several possible objects or trains of thought, implying the withdrawal from some matter to attend to another (James, 1890). In short, it is a tool for maintaining a reasonable amount of cognitive load (Riesberg, 2016). Sweller et al. describe that cognitive load refers to the number of resources used in working memory and may come in three different forms: (1) intrinsic cognitive load, (2) extraneous cognitive load and (3) germane cognitive load (Sweller et al., 1998). Intrinsic cognitive load is, according to Sweller et al., the effort associated with a specific topic and its inherent level of difficulty while extraneous cognitive load refers to the way information is presented to a learner. Finally, germane cognitive load refers to the effort put into creating a schema, a permanent store of knowledge (Sweller et al., 1998). Since these types of loads share a pool of resources, a task that may be considered difficult (high intrinsic and germane cognitive load) should present information in a manner that reduces extraneous cognitive load according to Sweller et al.

3.4 Situation Awareness

Situation awareness (SA) is a user's awareness of information around them and the understanding of what this information means for the present and future (Endsley and Jones, 2014). However, due to natural limits of human cognition, said awareness is often defined and delineated in terms of what information is important for a goal at a given time (Endsley and Jones, 2014). Strongly associated with principles in user-centred design, situation awareness and its themes make frequent appearances in human-centred design literature (HCD) and works within the field of human-computer interaction (HCI) (Endsley and Jones, 2014). SA has been subject to many attempted definitions ranging from the "*continuous extraction of environmental information, and integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing future perception and anticipating future event*" (Dominguez, 1994) to "*externally directed consciousness*" (Smith and Hancock, 1995). The most prominent of these, however, is of SA as "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*" (Endsley, 1995). This is the working definition for the scope of this thesis.

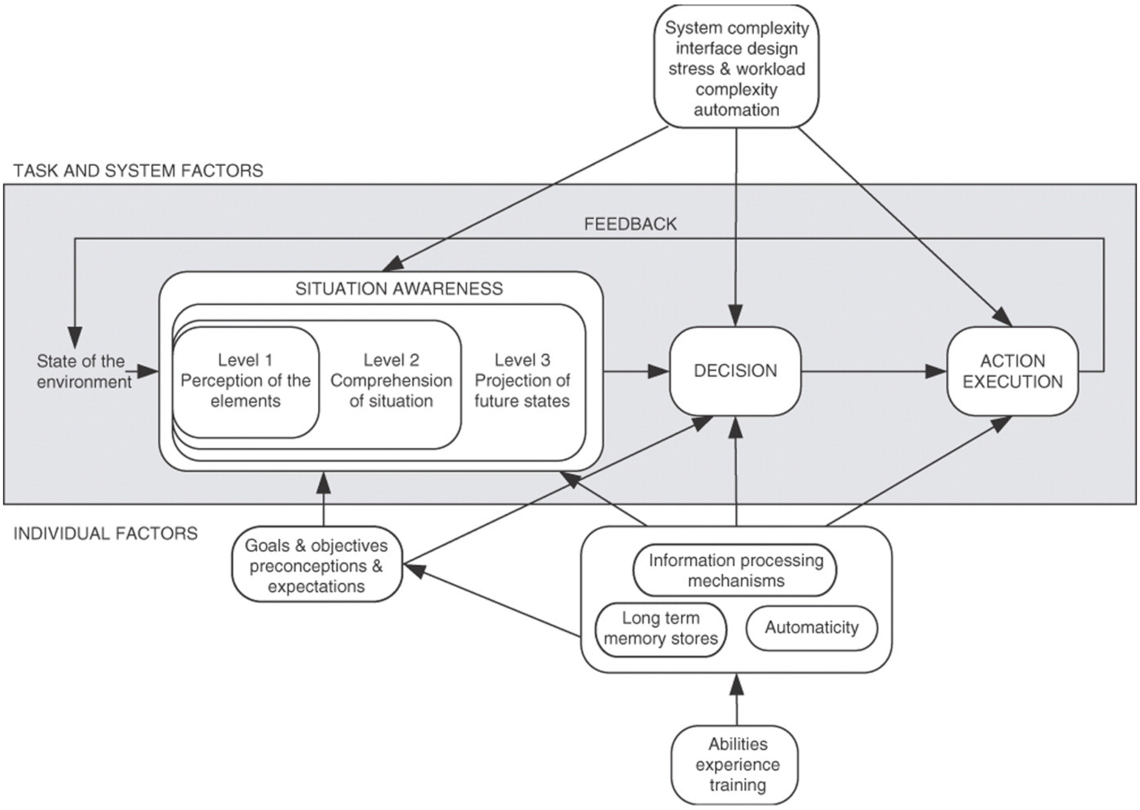


Figure 3.4: The three-level model of situation awareness (Endsley and Jones, 2014).

Complementary to this working definition is the three-level model of situation awareness (see Figure 3.4). The model posits the idea of SA acquisition and maintenance as being influenced by three elements; individual, task and systemic factors (Endsley, 1995). Examples of individual factors include practical experience, training and workload, while task and systemic factors may be represented by complexity and interface design respectively, as described by Endsley. The SA itself is, as the name of the model suggests, divided into three levels described by Endsley (Endsley, 1995). Endsley describes that Level 1 SA involves perceiving the status and dynamics of task-related elements in the surrounding environment. Level 2 SA, according to Endsley, has the decision-maker form a holistic picture of the environment as they comprehend the significance of objects and events around them. Endsley further describes that Level 3 SA involves the creation of mental models forecasting the future states of the system and elements in the environment by combining knowledge from both preceding levels.

Mental models facilitate the achievement of each level of SA in different ways; through directing attention to critical elements in the environment (Level 1), integrating said elements to grasp their meaning (Level 2) and generating future states (Level 3) (Salmon et al., 2008). Furthermore, time plays a significant role in how operators in various fields filter information in a given task environment (Endsley and Jones, 2014). Not only spatial factors make the time a significant contributor to the SA process, but also temporal ones, such as how soon a given element will come to affect their goals (Endsley and Jones, 2014).

3.4.1 Human Error

Reason mention that the human error problem may be approached in one out of two commonly cited ways: the person approach and the system approach (Reason, 2000). Reason state that the person approach focuses on the errors of individuals and blames them for inattention or moral weakness. The system approach instead concentrates on the conditions under which individuals work and attempts to build defences against errors and their adverse effects according to Reason.

Within the field of HCD, errors in the use of products or services are referred to as use errors as opposed to human errors - the principle stating that errors arise from a mismatch between human capabilities and the designed system (ISO, 2019). This perspective argues that proper design is the best way of mitigating errors instead of blaming humans for making mistakes during interaction and use (Reason, 2000). The Swiss cheese model of system accidents is commonly used to illustrate how an accident trajectory may bypass defences and safeguards used to prevent such an outcome (Reason, 2000). Furthermore, Reason describes that this model illustrates how the defensive layers inherent to high technology systems, such as those at Einride, are not intact and may eventually lead to a loss (see Figure 3.5). In reality, such systems feature many holes that individually will not instantiate a bad outcome, but that together may temporarily align to admit the trajectory of an accident opportunity to pass through (Reason, 2000).

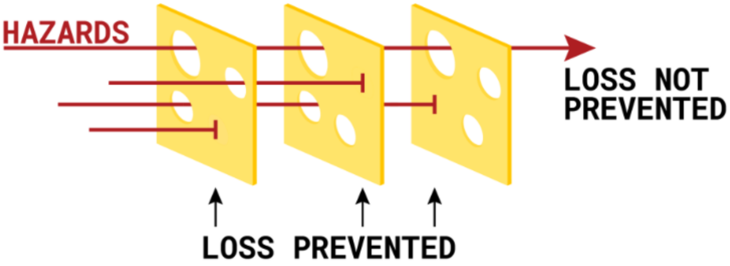


Figure 3.5: Swiss Cheese Model of System Accidents (Baerkahn, 2020).

According to Reason, the holes in the defences arise either through active failures (unsafe acts committed by people in direct contact with the system) or latent conditions (arising from decisions made by management, designers and procedure writers) (Reason, 2000). Reason also describes that active failures may come in the form of slips, mistakes, lapses and procedural violations and are difficult to predict in their specific form. Latent failures may, according to Reason, take the form of error provoking conditions such as time pressure, understaffing, fatigue and other factors that eventually tear long-lasting holes in the defences. These weak points may include things such as unworkable procedures, design deficiencies and untrustworthy indicators (Reason, 2000). Reason further describes that elements may be inactive for years before combining with active failures to generate an opportunity for accidents (Reason, 2000).

3.4.2 Use Errors and Situation Awareness

Good situation awareness is difficult to design for as its success is based on features of the human information processing system and complex domains (Reason, 2000). A complex domain is described by Reason as an environment full of opportunities for latent conditions to arise, conditions that may in turn combine with operator lapses to create a bad outcome (Reason, 2000). A range of eight factors have been identified as the most common enemies of SA: (1) attentional tunnelling, (2) requisite memory traps, (3) workload, anxiety, fatigue and other stressors (WAFOS), (4) data overload, (5) misplaced salience, (6) complexity creep, (7) errant mental models and (8) out-of-the-loop syndrome (Endsley and Jones, 2014). Proper understanding of these issues will help identify potential holes in the defence of the remote operations use case and give insight into ways of providing support using the system approach to the human error problem.

To summarise each effect described by Endsley and Jones; attentional tunnelling is a phenomenon where the user fixates on one set of information to the exclusion of others (Endsley and Jones, 2014). Endsley and Jones state that this presents an obvious issue where other information sources in the environment remain unattended for too long and hinder the maintenance of an accurate mental model of the use situation. Requisite memory traps occur when users rely too heavily on limited short-term memory, presenting a complication in the form of memory decay or bottlenecking of the memory cache according to Endsley and Jones. They further describe that WAFOS illustrates the effect of mental and physical fatigue on SA by co-opting parts of working memory and thus limiting the user's capacity for processing and holding information in memory. Similarly, they describe that data overload occurs when the intake of necessary information outpaces the ability to meet that need by the user's sensory and cognitive system. Distractions can, according to Endsley and Jones, be caused when the user's attention is drawn to less important or irrelevant information. These effects are all problematic since there is a need for many features of the situation to occupy memory and be combined with new insights to update the user's model of SA (Endsley and Jones, 2014). Such similar tasks will draw on similar cognitive resources and thus increase the chance of them running out completely (Riesberg, 2016).

The complexity creep effect has, according to Endsley and Jones, more to do with mental models than the limitations to cognitive resources, but remains related to it since systems with too many features make the formation of mental models difficult (Endsley and Jones, 2014). They state that the errant mental models' effect implies that the user deploys the wrong mental model for interpreting incoming information, leading to misunderstandings, mode errors and representational errors that are especially insidious in automated systems. Mode errors could be the erroneous perception that the vehicle is in fully-automatic mode when it is in remote driving mode. Related to this, representational errors occur, according to Endsley and Jones, as the user fails to realise they are using the wrong mental model. They state that the user has a tendency to explain away conflicting information rather than reevaluate their mental model. Out-of-the-loop syndrome is another SA issue with close ties to automation. This effect occurs when automaticity actively undermines SA by taking the operator out of the loop, keeping them ignorant of system performance as well as the state of the elements the system is supposed to be controlling (Endsley and Jones, 2014). Thus this constitutes an important issue in the present case study.

3.5 Automation

Effectiveness in sociotechnical systems, such as remote operations, depends on the cooperation between autonomous systems and humans, necessitating the design of automated solutions as collaborative systems (Behymer and Flach, 2016). It has been observed in most work domains that human capacities are a limitation and a crucial factor for the system's ultimate success since all automated systems eventually will run into problems they have not prepared solutions or procedures for (Behymer and Flach, 2016). As part of this growing field, vehicle automation is emerging as an innovative technology with significant potential of increasing traffic safety and effectiveness within the freight traffic sector (Wang et al., 2021). The benefits of automated vehicles (AVs) are numerous, ranging from enabling more efficient demand-responsive transit and accessibility to said transit by also increasing service frequencies and service hours at lower operating costs (Wu et al., 2021). Furthermore, secondary stakeholders such as pedestrians and bicyclists are provided with a safer traffic environment by AVs eliminating potentially careless and reckless human driving behaviours (Wu et al., 2021).

A common way of classifying the performance of motor vehicle driving automation systems is to interpret them using the SAE Levels of Driving Automation. Published by SAE International, each of these levels are defined according to the properties and degree of automation in a given system, ranging from no driving automation (Level 0) to full driving automation (Level 5)(SAE International, 2021). As previously mentioned, the system being designed for, the Pod, is of SAE level 4 (Einride, n.d.-a). This implies that the Pod fulfils the following criteria (SAE International, 2021):

1. The human is not driving when the automatic driving features are engaged.
2. Said automated features do not require the human to take over driving.
3. The features can only drive under specific conditions and only when all of them are met.
4. May include features such as a steering wheel or pedals.

3.5.1 Approaches to Automation

Human-machine collaborations already hold the potential to outperform solutions by only humans or computers in domains such as meteorology and medicine, but simply pairing the best computer with the best human operator will not necessarily achieve the best performance (Behymer and Flach, 2016). The unfortunate fact is that system developers often focus on increasing the abilities and capacities of the autonomous agent as opposed to developing solutions for an operator-supportive interface with human agents (Behymer and Flach, 2016). This line of thinking goes well in hand with the prosthesis approach to automation, where the goal is to design to compensate for, or overcome, human limitations (Behymer and Flach, 2016). This approach, however, is flawed in that autonomous systems have their own limitations that the human element actively mitigates. In high-reliability organisations, such as the military, human variability and ability for operational adaptability is considered to be one of the system's most important safeguards (Reason, 2000). High-reliability organisations are typically defined as operations that experience fewer adverse events than expected, especially considering the hazardous conditions they deal with (Reason, 2000).

On the other side of the proverbial coin is the collaborative systems approach, one that strives to bridge the Gulfs of Execution and Evaluation by empathising the salience of clarity of constraints, domain states and displays to the human-machine team (Behymer and Flach, 2016). The size of the Gulf of Execution depends on how effective the human-machine team's actions are in terms of them achieving their goal, while the Gulf of Evaluation is sized depending on the team's ability to observe, perceive and understand the state of the world in terms of their intentions (Behymer and Flach, 2016).

3.5.2 Human-Automation Interaction

Human-automation interaction (HAI) is, according to Sheridan and Parasuraman, a novel kind of relationship between humans and autonomous systems which places the human agent in a supervisory role relative to the machine (Sheridan and Parasuraman, 2005). The typical process of an automated large-scale system involves four stages, described by Sheridan and Parasuraman, that may be designed to operate autonomously to a varying degree. They state that the four stages are: (1) the acquisition of information, (2) the analysis of that information, (3) the decision about what actions to take based on said information and (4) the implementation of that action. Much of the existing understanding of HAI stems from accident research, according to Sheridan and Parasuraman. Three prevalent causes of bad outcomes have been identified from this: insufficient feedback on system states, a misunderstanding or lack of understanding of automation (often attributed to mismatching mental models of operation) and an overreliance on automation (Sheridan and Parasuraman, 2005).

Other key behavioural issues that have remained unattended within HAI are the concerns surrounding complacency and satisficing the achievements of task goals (Kaber, 2018). Complacency is the lack of suspicion of system states in the presence of limited awareness of operational modes and knowledge of more efficient methods of operation (Kaber, 2018). The user tendency of satisficing leans on the “good enough” heuristic where operators accept the most accessible solution at the expense of efficiency (Kaber, 2018).

One way of facing these issues is to pay attention to user interface design and investigate ways of developing representations that enhance the human and automated system’s ability to respond to unanticipated variability in a creative and complexity-robust manner (Behymer and Flach, 2016). The Skills, Rules, Knowledge (SRK) framework defines three ways in which people represent symbols and signs and the subsequent ways in which this serves to distinguish three levels of human performance (Behymer and Flach, 2016). Skill-based behaviour is characterised by being highly practised with virtually no conscious monitoring and is typically initiated by some specific event (Embrey, 2005). Rules-based behaviour features intermediate levels of conscious control as it mainly pertains to applying learned solutions when specific conditions are at play (Embrey, 2005). Finally, knowledge-based behaviour has high conscious control as it typically occurs in novel situations where no procedure exists and significant mental effort is needed to assess the situation at hand (Embrey, 2005). Insight into these levels of human performance provides a fundamental understanding of the mental modes a remote operator deals with at different levels of experience and under varying circumstances.

3.5.3 Situation Awareness in Autonomous Solutions

The need for situation awareness often increases with the addition of automation to a system, as the operator needs to maintain SA over basic system information as well as information from the new automated one (Endsley and Jones, 2014). Operators need to maintain high levels of SA regarding the functioning of both the automation and underlying system parameters necessary for answering those more fundamental questions (Endsley and Jones, 2014). The loss of SA in human operators has been observed to happen through three main mechanisms (Endsley and Kiris, 1995):

1. Changes in vigilance and complacency associated with monitoring.
2. Assumption of a passive role instead of an active role in processing information for controlling the system.
3. Changes in the quality or form of feedback provided to the human operator.

Complacency in monitoring has been connected to operators' subjective confidence in the system's reliability (Wiegmann et al., 2010) as well as their sense of self-efficacy in performing the tasks without autonomous support (Prinzel and Pope, 2000). However, the tendency to attribute unwarranted levels of trust seems to differ from operator to operator in accordance with their predisposition, the perceived characteristics of the system and their understanding of the system's actual functioning (Endsley and Jones, 2014). Furthermore, the complacency problem has been found to be less of a visual problem and more of a matter of attention, as the superimposition of task-relevant information has yet to resolve the issue (Metzger and Parasuraman, 2001a).

When it comes to the assumption of passive versus active roles in the HAI-loop, most of the SA loss could be attributed to how information was internalised and stored when monitoring the performance of the automation as opposed to the human operator carrying it out themselves (Endsley and Jones, 2014). A passive role has been observed to cause a significant decline in SA in air traffic controllers with aircraft they simply monitored as compared to those that they actively controlled (Metzger and Parasuraman, 2001b). The feedback mechanism raises concerns regarding the user interface of a given autonomous system. Lacking appropriate feedback, operators have no idea whether their requests have been received or implemented, directly causing them to be out-of-the-loop (Norman, 1990). The use of a multitude of screens of information that may be called up via menus may engender poor feedback quality by obscuring key data from view or misunderstanding its importance within increased complexity (Endsley and Jones, 2014). Some systems may also have issues with salience and visual hierarchy to support the operator's sense of SA (Endsley and Robertson, 2000).

Another approach relevant to the case study at hand is adaptive automation (AA). AA recognises that there are times when human operators need to assume control, and looks to exploit this opportunity by allocating system control to the operator at periodic intervals to combat complacency and improve human monitoring performance (Rouse, 1988). The transfer of control need not be time-based but could also be determined by other factors (Scerbo, 1996):

- The occurrence of critical events.
- Detection of human performance below a certain criterion level.

- Use of psychophysiological monitoring to detect losses of arousal or other cues of poor performance (such as loss of consciousness).
- The use of models of human performance to predict the best times to intervene.

An alternative take on the automation problem is to have another look at the four stages of automated large-scaled systems (Sheridan and Parasuraman, 2005). Automation is significantly helpful in the implementation of tasks as opposed to in the purpose of generating solutions and action plans, where there is instead an observable decline in performance (Kaber and Endsley, 2004). Overall, human agents have less difficulty with autonomous systems purposed for information gathering and action initiation as compared with those systems that involve higher-level cognitive functions (Endsley and Jones, 2014).

4

Methodology

This chapter covers the frameworks and methodologies considered in the thesis work. The first section will comment on the nature of the problem space at hand while the second, Frameworks, describes frameworks considered for use in the design process. The following section, Methodologies, proceeds to mention the methodologies contemplated during each step of the process. The following chapter, Process, will then present how these methodologies and frameworks were applied in this thesis.

4.1 Wicked Problems

Problems that initially can be difficult to describe or understand are, according to Plattner et al., referred to as wicked problems (Plattner et al., 2010). Plattner et al. describe that these problems instead tend to become more understandable and easier to describe as one gets closer to a solution. They also describe that this is a common type of problem that design teams face.

Rittel and Webber list ten criteria that characterise wicked problems (Rittel and Webber, 1973). One such criterion is that a solution never is true or false, but rather good or bad. The judgement of a solution's validity is often heavily impacted by stakeholders differing opinions which can make it hard to define which solution is best (Ritchey, 2013). Another criterion that Rittel and Webber mention is that every wicked problem is unique, making it impossible to achieve a type of solution that fits all problems. Rittel and Webber describe that wicked problems differ from, for example, mathematical problems where many problems are similar and there often is clear which techniques to use for each type of problem. They further describe that even if problems seem similar, there is almost always something that distinguishes them and greatly impacts the solution.

4.2 Frameworks

Overarching frameworks that can provide structure to the design process are the Research through Design (RtD) approach, the Human-Centred Design approach (HCD) and the 3i Model of Design Thinking. The three frameworks are all described in further detail in this section.

4.2.1 Research through Design

Research through Design (RtD) is a framework that conducts research through methods, practices and processes of the design field to generate new knowledge - viewing design inquiry as distinct from scientific inquiry (Zimmerman and Forlizzi, 2014). Drawing on the reflective practice of continuously reinterpreting a given problem space through a process of critiquing artefacts proposed as being solutions, RtD asks researchers to speculate on what could and should be (Zimmerman and Forlizzi, 2014). Four proposed evaluation criteria may, according to Zimmerman et al., be used to determine the quality of RtD work within the field of interaction design and HCI: process, invention, relevance and extensibility (Zimmerman et al., 2007). Zimmerman et al. describe that process refers to the need for interaction designers to maintain methodological rigour and appropriate documentation so that their process can be reproduced. The next criterion, invention, argues that designers must demonstrate that their creation, addressing a specific situation, is an unexplored integration of different subject matters, as described by Zimmerman et al. The third criterion mentioned by Zimmerman et al., relevance, refers to a need for interaction designers to articulate the preferred state for their design attempts and why the community should also consider this state to be desirable (Zimmerman et al., 2007). Finally, the extensibility criterion is defined by Zimmerman et al. as the ability to base new research on the contributions. This indicates that the documentation must be written in a way that allows others to benefit from the knowledge obtained from the research work. Adhering to these criteria ensures a design process that is reflective and produces new and valuable knowledge in forms ranging from novel perspectives to new methods (Zimmerman and Forlizzi, 2014).

4.2.2 Human-Centred Design

Another design framework is human-centred design (HCD) which is described by Giacomin as one out of three major design paradigms of design (Giacomin, 2014). Giacomin characterises HCD as an approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors, ergonomics and usability knowledge to the process. HCD as a field is especially concerned with incorporating user perspectives into the development process of both technical and functional requirements, having identified a set of key principles to be followed (Maguire, 2001):

- The active involvement of users and clear understanding of user and task requirements.
- An appropriate allocation of function between user and system.
- Iteration of design solutions.
- Multi-disciplinary design teams.

HCD refers to the term human-centred quality as a collective phrase for the intended outcomes of interaction with a product or system, defined by its adherence to requirements for usability, accessibility, user experience and avoidance of harm (see Figure 4.1) (ISO, 2019). Usability is the extent to which a product can be used by specific users to achieve their goals with effectiveness, efficiency and satisfaction in a specified context or use (ISO, 2019). Accessibility refers to designing for a wider target group and adhering to prescribed code requirements for use by people with disabilities (Story, 1998a). User experience is a term used in several ways within the field of human-computer interaction (HCI) but may generally be described as a concerning itself with a holistic view of the user's interaction with interactive products (Bargas-Avila and Hornbæk, 2011). Harm of use entails the avoidance of negative consequences for health, safety, finances or the environment that can result from product usage (ISO, 2019).

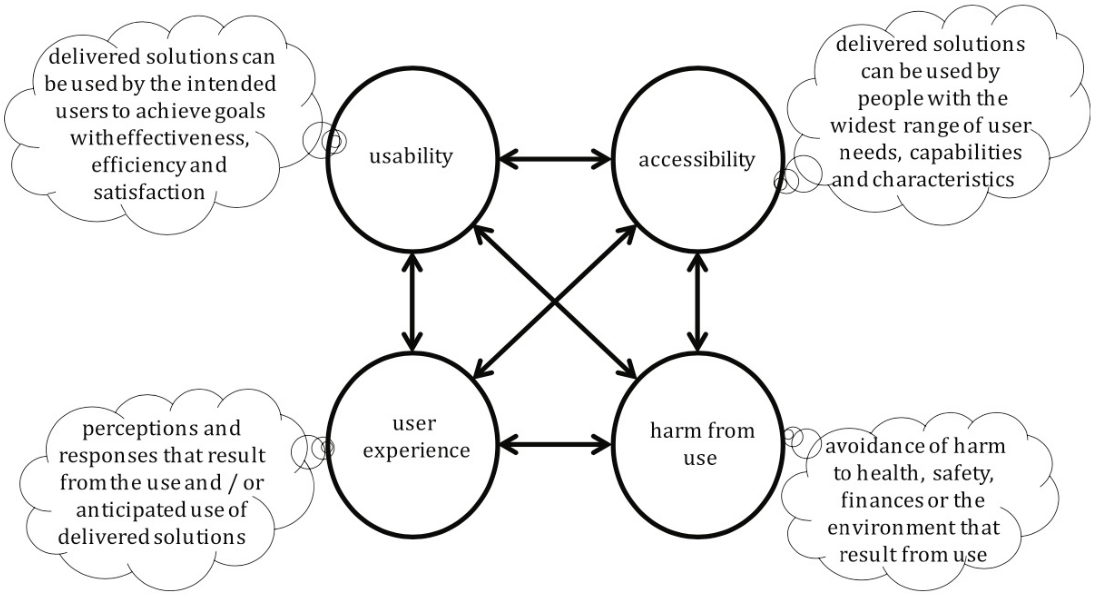


Figure 4.1: The interrelationship between components of HCD's human-centred quality and summarisation of each component's outcome (ISO, 2019).

4.2.3 3i Model of Design Thinking

The final framework described in this section is the 3i Model of Design Thinking. The 3i model involves three core activities: Inspiration, Ideation and Implementation, continuously moving through thought processes in both convergence and divergence stages, as seen in Figure 4.2 (Chung, 2016). Each of these phases features its own suggested methods and procedures. Inspiration is a highly formative part of the design process, focusing on divergent and expansive thinking as the designer inform themselves of the context they are designing for (IDEO, 2015). Ideation follows, featuring both convergent and divergent procedures as the project scope is narrowed and more specific questions are asked as the designer explores the opportunity in front of them and how exactly they plan to tackle it (Chung, 2016). This phase is one of iterative switches between convergent and divergent design methods to ensure a human-centred solution (IDEO, 2015). Finally, Implementation converges the focus on a solution, questioning factors such as performance, sustainability and feasibility (Chung, 2016). The 3i Model of Design Thinking promotes repeated cycles of convergent and divergent design thinking, with the iteration of proposed solutions being heralded as central to the field of HCD itself (IDEO, 2015).

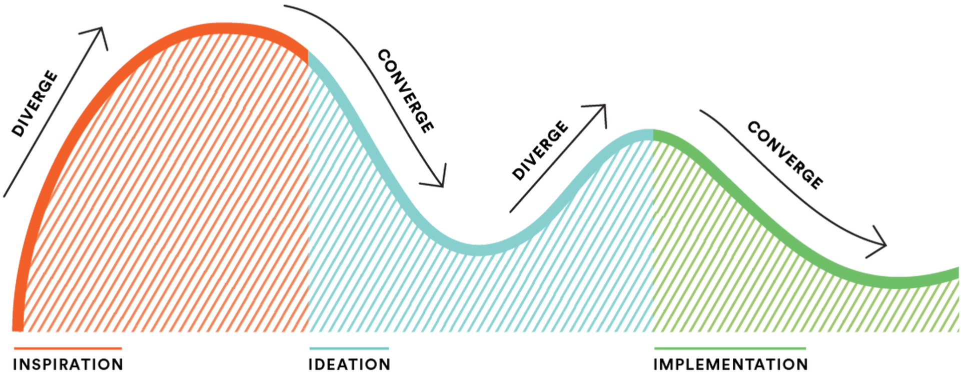


Figure 4.2: The 3i model of design thinking (IDEO.org, n.d.).

As mentioned previously, the field of human-centred design is especially focused on the incorporation and attention to user perspectives in the development process, but the iteration of design solutions is another principle held in the highest regard (Maguire, 2001). These two topics of attending user perspective and iterative design work are not unrelated, as an iterative process places received user feedback critical to the evolution of a given solution (IDEO, 2015). An iterative process ensures the design process is nimble, responsive and allows for the exploration of a wider range of ideas before narrowing in on a finalised concept (IDEO, 2015). The second activity in the 3i model of design thinking, Ideation, is typically an iterative process where the solutions continuously are tested and refined (IDEO, 2015).

4.3 Methodologies

This section describes the methodologies considered during the thesis work and is structured according to the three phases of the 3i Model; Inspiration, Ideation and Implementation. It is structured in such a way due to the fact that the 3i Model offers the most clear structure out of the three considered frameworks. The methods are all mentioned in the phase where they were considered to be applied.

4.3.1 Inspiration

As previously mentioned, the Inspiration stage is characterised by divergent thinking and the exploration of the use context (Chung, 2016). As they move through this phase, designers seek a broad view of their chosen field of research and thus draw on convergent methods to give better context to their design problem (IDEO, n.d.). An important component in this phase involves state of the art research, looking into current solutions to not only identify the theoretical foundation for the discussion but also to clarify the work's focus and define the relevance of given research questions. Methods for state of the art research may include observational studies, literature reviews and study visits, to name a few. A literature review can be likened to the secondary research activity as suggested by IDEO in that it involves designers getting themselves immersed and up to speed in modern innovations and findings currently published (IDEO, 2015). A more practical approach to this is the study visit, which can be likened to the more established method of field trips. The field trip is an ethnographic method that involves travelling to a location to speak directly to local stakeholders, gaining the opportunity of getting insight into firsthand experiences within the given environment (Eden et al., 2019). Investigating current paradigms and principles in relevant fields combined with practical solutions is one way of building an understanding of the relevant context.

Another informative method for the Inspiration phase is the semi-structured interview. The interview is a core method within the field of HCD as it allows direct input and knowledge from a given target group, such as experts or future users (IDEO, 2015). The expert interview is useful in that key insights and history of the context and innovations may be provided expediently, but the selection of relevant experts of varying perspectives is crucial (IDEO, 2015). A semi-structured interview features a relatively open framework of open-ended questions or topical trajectories that allow for focused, conversational, two-way communication for good-quality qualitative data (Sharp et al., 2019). This method works especially well when combined with the secondary research of a literature study, providing facts and figures to the more subjective data of the interview (IDEO, 2015).

4.3.2 Ideation

The Ideation phase is, as mentioned previously, one of both convergent and divergent design thinking (Chung, 2016). Generally considered the most creative phase in the design process, this phase looks to interpret the information learned in the more immediate scope of a project. One way of accomplishing this is to analyse findings from formative studies such as interviews, mining the feedback for topics and themes via thematic analysis and affinity diagramming. Thematic analysis is, as described by Braun and Clarke, a widely used method for analysing qualitative data to identify and report patterns within said data (Braun and Clarke, 2006). Braun and Clarke further describe that organising and describing the data set in detail maps out themes that capture important aspects of the data in relation to the research question. One incarnation of thematic analysis is affinity diagramming which, according to Hanington and Martin, is used to externalise and meaningfully cluster observations such as interview insights. They further describe that either a deductive or inductive approach can be used to perform the coding and analysis of data. The inductive approach is described by Hanington and Martin as the creation of categories and code during analysis, resulting in the themes being derived from the data. The deductive method is instead described by Hanington and Martin as a top-down approach where the data is analysed with established themes in mind. Hanington and Martin describe that analysis methods such as these can help keep designers grounded in data as suggested solutions begin development.

Following the convergent section, there is a shift back to divergence as idea generation and the development of suggested solutions begin (IDEO, 2015). One method that can be used to promote conversations between designers and to explore different designs is, according to Hartson and Pyla., sketching (Hartson and Pyla., 2012). Sketching is, as described by Hartson and Pyla., both quick and inexpensive which makes it suited for the early stages of exploration. When a sketch portraying a design has been created, prototypes are typically used in the next step to concretise the designs and use them for testing. Prototypes are especially useful as mediums for communication in different stages of the design process: during evaluation with stakeholders, in ideation work between designers or as a way of exploring solution ideas in implementation-related inquiries (Sharp et al., 2019). Furthermore, prototyping in itself encourages reflection in design, incentivising the iteration reconsideration of different layouts and interface characteristics (Sharp et al., 2019).

Hartson and Pyla. mention that there are many different types of prototypes one can use and that they have different levels of interactivity (Hartson and Pyla., 2012). Most prototypes rely on interaction from the user to show different states, but another way to bring a prototype to life can be to use video animations according to Hartson and Pyla. Hartson and Pyla. further describe that videos typically can show the flow of an application better than static prototypes while still being able to keep a lower fidelity to encourage feedback and discussion.

The Wizard of Oz (WOz) method is another method that can be used to keep a lower fidelity while simulating system responses of a prototype so that it appears to be more functional than it actually is (Hanington and Martin, 2019). Dow et al. describe that the WOz method is used during user evaluations where an operator, often referred to as wizard, simulates parts of the interaction with the interface (Dow et al., 2005). They mention that this allows the users to explore the interface and decide which path they want to take, avoiding the user to follow a predefined flow and missing potentially crucial feedback.

A good way to refine the designs is to use usability testing to get feedback on them (IDEO, 2015). Usability testing is, as mentioned by Sharp et al., an important part of assessing digital products and the controlled settings enable designers to control influences that may impact a user's performance (Sharp et al., 2019). Sharp et al. describe that the goal of a test is to evaluate whether intended users can achieve the tasks for which the design solution is designed. The goal is also to evaluate whether users are satisfied with their experience - often employing data collection in the form of video recordings, logged keystrokes, movements and responses.

One type of user testing is A/B testing which is a method used to compare two versions of the same design to see which one performs statistically better against some form of predetermined goal (Hanington and Martin, 2019). A large user group is typically involved where half of the group gets to use one version while the other half gets to try the other version (Kaufmann et al., 2014). This means that each user is shown only one of the versions and the performance of them can then be compared to determine which version is advantageous.

Another method for user testing that also is focused on comparing different versions of a design is comparative usability testing (Ross, 2017). Anderson describes that comparative usability testing is used to compare different design solutions based on how efficient and effective they are (Anderson, 2021). Two or more designs that allow the users to complete the same task are produced according to Anderson. He describes that these designs then are tested by users who complete the tasks and compare the different design solutions. Based on observations made during the tests and feedback received from the users, it can be determined which design, or parts of the designs, work better for the users.

In conjunction with data being logged, participants may be asked what they are doing and think out loud - a method called the Think Aloud technique which complements the quantitative data of the recordings and action logs with qualitative data (Sharp et al., 2019). The Think Aloud technique serves to get more direct insight into the user's way of thinking, how they plan their actions and how they understand the interface (Sharp et al., 2019).

While user reflection is key for providing feedback in an iterative design process, said reflection may disrupt the flow of events if held immediately in the moment (Chong et al., 2015). As observed in feedback studies, immediate reflection has the benefit of being unaffected by memory decay but nevertheless distracts users from carrying out their main task (Chong et al., 2015). Thus, allowing users to engage in delayed reflection may serve to help them maintain focus on the here and now and provide more thought-through feedback as they are removed from the context of interruption.

4.3.3 Implementation

Implementation is a phase of convergence where solutions are elected and realised. As previously mentioned, this activity poses questions in regard to the performance, sustainability and feasibility of suggested solutions (Chung, 2016). It employs convergent processes to better understand how to bring solutions into the real world, ensuring success by keeping those for whom the solution is intended close at hand and involved (IDEO, 2015). Other activities and principles related to the Implementation phase have, within the scope of the work, already largely been covered by the Ideation section. Prototyping and the iteration of solutions are the key tools by which stakeholders and target users are kept within the feedback loop and ensure the previously mentioned success.

The Implementation phase often includes activities not strictly considered a part of design work but instead a part of project management and goal setting within the design process (IDEO, 2015). This may include the development of action plans for funding and the execution of resource assessments but may also feature more thesis-relevant activities such as the forming of partnerships (IDEO, 2015). The latter can be implemented through the involvement of fellow student designers as part of the evaluation work within this phase. Evaluations at this stage can, for example, feature different interviewing methods such as unstructured and semi-structured interviews. As mentioned previously, semi-structured interviews provide focused two-way communication for good quality qualitative data (Sharp et al., 2019). Unstructured interviews, on the other hand, lack any kind of predetermined categories for either questions or answers as a way of avoiding any sort of limitation to the field of inquiry and seeks to expose the researcher to unanticipated themes, all to better understand the interviewees' perspectives (Zhang and Wildemuth, 2009).

An alternative method that can be used for evaluations at this stage is the AGREE II framework. AGREE II is an instrument intended to provide a framework for the assessment of the quality of guidelines, for the provision of methodological strategy for the development of guidelines and for the guidance on what information should be reported in guidelines and how this is done (AGREE Next Steps Consortium, 2017).

5

Process

This chapter will expound on the practical application of the previously listed methods, ordered according to the phase of the design process in which they were carried out. The chapter also features comments on the benefits and relevancy each chosen method provided to the thesis work itself, giving more specific insight into the contribution of each. The three frameworks Research through Design, Human-Centred Design and 3i Model of Design Thinking were all applied to the design work conducted for this thesis. The section is structured in the order of phases included in the 3i Model of Design Thinking to better illustrate the order in which they were carried out and the types of design thinking involved at a given time. This work employed a systematic approach to the human error problem in accordance with the intention to develop an operator-supportive design solution.

5.1 Inspiration

This section will document the methods employed as a part of the highly divergent first phase of the 3i Model of Design Thinking, Inspiration. The listed methods provided the thesis work with formative research insights that carried over into, transformed and informed subsequent design process iterations.

5.1.1 Literature Review

To ensure that the project work is based on current and robust academic information, getting familiarised with the relevant topics at hand is a critical part of the Research through Design process. The literature review served to inform the upcoming design work by outlining the current state of the art in relevant fields and outlining opportunities for future work in which this thesis project might delve into. Before the literary search began, a selection of major research fields was outlined as being of special interest to the project. While the specific theoretical outline went through some revision as the Inspiration phase ran its course, these five major topics remained consistent: Graphical User Interface, Information Visualisation, Cognition, Situation Awareness and Automation. This set of academic fields then served as keywords in the search for relevant literature, alongside other filtering conditions such as the age of the article and peer-reviewed status. The principal contribution of the literature review came in the form of it narrowing down the scope of the proposed research questions to ensure their relevancy and specificity within the broader field of chosen design paradigms. This was achieved as the review informed the design team of State of the Art academic findings and ensured that the stated research questions did not seek to investigate already covered topics and solutions.

5.1.2 Study Visit

As previously mentioned, investigating state of the art solutions within fields such as remote operation and automation may serve to immerse designers and researchers in the context they will be investigating. This enhances not only understanding of the physical context but also the implications it has for users' ability to carry out mental tasks within the specific environment. Topics such as these are crucial when designing for situation awareness since many factors outside of the GUI may serve to affect users' performance and satisfaction with a given design solution.

As such, a study visit to the Chalmers Full Mission Bridge Simulator was scheduled and carried out as a part of the Inspiration phase. A guided tour of three maritime simulator bridges were given during which the purposes, current usage and limitations of each were explained by a staff member of the Chalmers Division of Maritime Studies. Follow-up questions were asked and photos were taken for later analysis and discussion.

It was astonishing how similar the bridge simulators were to bridges on real ships. The size of the simulator and the setup of the screen(s) had a big impact on how real it felt. The smallest simulator did not feel as real as the other two and its size and placement in the room might have contributed to this perception. It had five smaller screens put together to form one big screen which also made it feel less real compared to the other two which had one big projector screen.

As for real ships, the simulators had many different complex instruments that looked very similar to those on real ships. These instruments were a clear example of skeuomorphic excise and it was interesting to see that being used. Due to the complexity of the instruments, it was overwhelming to see it all for the first time and it seemed very reasonable that several people are needed to operate the ship. Some instruments were in the form of screens that showed different types of information. Some data was shown on several screens using different visualisations and it was interesting to see how different visualisation strategies can be used to represent the same data.

5.1.3 Interviews

As the project addresses the limitations of human cognition and the need for human-centred design within the field of automation, the strengths and weaknesses of Einride's current solution needed to be investigated. Expert interviews with several specialists within Einride's remote operate team provided rich insight into this. A total of six participants were interviewed, all of whom have a class 'C' licence meaning that they are allowed to drive trucks. Three of the participants have also had an important role in the development of the current interface of the remote operator station.

Each of the participants was asked to read and fill out a consent form delineating the conditions of the interview. All were then subsequently asked the same selection of 19 questions (the first 19 questions seen in Appendix A). These questions were formulated in a semi-structured manner featuring open-ended questions and space for each expert to apply their experience as they saw fit. The three participants that reported involvement in the development of the current solution were also asked an additional four questions (questions 20-23 in Appendix A).

Responsibility was split between the role of interviewer and notetaker, a division that stayed consistent for all six interviews. To augment the real-time note-taking, each interview was recorded using the Voice Memos application (Apple, 2022) on iPhone and later transcribed in its entirety. The keynotes written by the notetaker were used for overview purposes while transcriptions provided the fine details of each response.

Taking place toward the end of the Inspiration phase, expert interviews provided the insights and answers most narrowly relevant for the actual development of a user interface for the remote driving station. As was previously mentioned, this included aspects such as current strengths and weaknesses while also covering inherent opportunities within the remote driving setup and reflections on the work and reasoning put into development thus far.

5.2 Ideation

This section will cover the workflow and methods employed as part of the Ideation phase as it pertains to the 3i Model of Design Thinking. The section is initially split according to methods carried out and later divided after the iteration in which they were employed.

5.2.1 Thematic Analysis

As the first step into the Ideation phase of the design process, thematic analysis in the form of affinity diagramming was applied. This initiated the convergent process of scouring through interview transcripts for findings that would serve as the foundation for the first preliminary set of design guidelines and pooling them together into a body of material to be mapped. Each interview transcript was read through and had sentences deemed relevant to the established research questions highlighted. The procedure was conducted by one design team member at a time to ensure thoroughness before each sentence was transferred over to individual post-it notes on a Miro board (Miro, 2022). These notes were subsequently clustered and labelled according to themes emerging from their content into increasingly high-ordered categories (see Figure 5.1).

Figure 5.1: The resulting clusters of notes on the Miro board following the first thematic analysis.

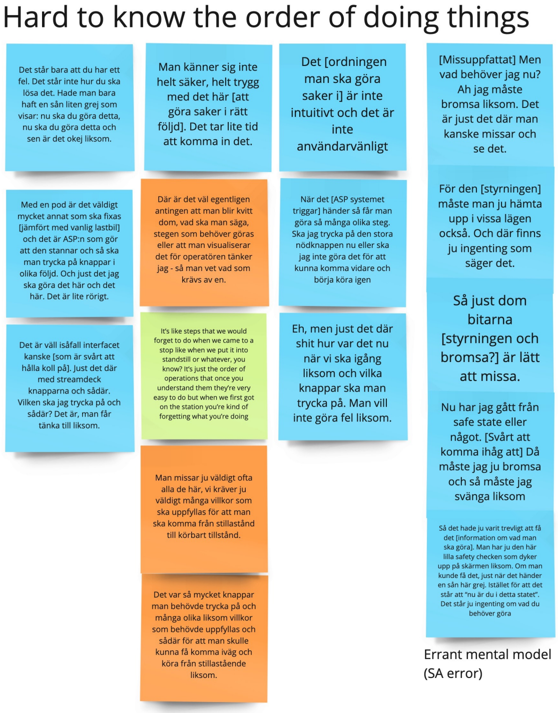
These themes were discussed to find which were explicitly relevant to actionable design input within the scope of the project work. Some themes were considered to be outside the scope and were, therefore, not further deliberated upon. One such theme was *Dangerous blind spots* as it was not possible to change the cameras' field of view when testing. Other themes considered to be outside the scope included themes regarding haptic or aural feedback.

The themes that were considered relevant were combined to form bigger themes. One such combination includes the themes *Show information about cargo/tires*, *Show more information about brakes* and *Physicality issues* that all refer to not being able to know the truck's physical condition. These combined themes were condensed into a total of five preliminary guidelines:

1. Design for dynamic camera behaviour for flexibility in use.
2. Visualise the required steps for the initiation of action plans.
3. Visualise root causes for triggered automated responses.
4. Visualise relevant parameters of the Pod's current status.
5. Design for the limitation of data overload at operator-demanding driving situations.

The first guideline, *Design for dynamic camera behaviour for flexibility in use*, was inspired by themes such as *Cargo view is not used* and *Flexibility in use.* These themes showed that the camera layout the operator wants can differ depending on the situation. It was, for example, mentioned by one interviewee that the cargo view never was used while driving but that it was very helpful when loading the truck. Another interviewee said "I kinda wished that in the reverse, when you put the camera into reverse, that you could see a much bigger picture behind you even though it is on a larger screen" which points out the need for a flexible behaviour of the layout.

The main inspiration for the second guideline, *Visualise the required steps for the* *initiation of action plans* was the theme *Hard to know the order of doing things*. The post-its forming that theme in Miro can be seen in Figure 5.2. Three different interviewees all pointed out that it often is required to perform a set of actions in a specific order and that it can be hard to remember them all. The current interface solution does not give the operator any clues as to what actions they need to do and it can be perceived as frustrating to have to memorise the specific order. The theme *Safety checks* was also considered for this guideline as it points out the need to do a safety check, i.e. perform a set of actions in a predefined order, before being allowed to start driving the Pod.



Hard to know the order of doing things

Det står bara att du har ett fel. Det står inte hur du ska lösa det. Hade man bara haft en sån liten grej som visar: nu ska du göra detta, nu ska du göra detta och sen är det okej liksom.

Med en pod är det väldigt mycket annat som ska fixas [jämfört med vanlig lastbil] och det är ASP:n som gör att den stannar och så ska man trycka på knappar i olika följd. Och just det jag ska göra det här och det här. Det är lite rörigt.

Det är väll isåfall interfacet kanske [som är svårt att hålla koll på]. Just det där med streamdeck knapparna och sådär. Vilken ska jag trycka på och sådär? Det är, man får tänka till liksom.

Man känner sig inte helt säker, helt trygg med det här [att göra saker i rätt följd]. Det tar lite tid att komma in det.

Där är det väl egentligen antingen att man blir kvitt dom, vad ska man säga, stegen som behöver göras eller att man visualiserar det för operatören tänker jag - så man vet vad som krävs av en.

It's like steps that we would forget to do when we came to a stop like when we put it into standstill or whatever, you know? It's just the order of operations that once you understand them they're very easy to do but when we first got on the station you're kind of forgetting what you're doing

Man missar ju väldigt ofta alla de här, vi kräver ju väldigt många villkor som ska uppfyllas för att man ska komma från stillastånd till körbart tillstånd.

Det var så mycket knappar man behövde trycka på och många olika liksom villkor som behövde uppfyllas och sådär för att man skulle kunna få komma iväg och köra från stillastående liksom.

Det [ordningen man ska göra saker i] är inte intuitivt och det är inte användarvänligt

När det [ASP systemet triggar] händer så får man göra så många olika steg. Ska jag trycka på den stora nödknappen nu eller ska jag inte göra det för att kunna komma vidare och börja köra igen

Eh, men just det där shit hur var det nu när vi ska igång liksom och vilka knappar ska man trycka på. Man vill inte göra fel liksom.

[Missuppfattat] Men vad behöver jag nu? Ah jag måste bromsa liksom. Det är just det där man kanske missar och se det.

För den [styrningen] måste man ju hämta upp i vissa lägen också. Och där finns ju ingenting som säger det.

Så just dom bitarna [styrningen och bromsa?] är lätt att missa.

Nu har jag gått från safe state eller något. [Svårt att komma ihåg att] Då måste jag ju bromsa och så måste jag svänga liksom

Så det hade ju varit trevligt att få det [information om vad man ska göra]. Man har ju den här lilla safety checken som dyker upp på skärmen liksom. Om man kunde få det, just när det händer en sån här grej. Istället för att det står att “nu är du i detta statet”. Det står ju ingenting om vad du behöver göra

Errant mental model
(SA error)

Figure 5.2: Post-its grouped to form the theme *Hard to know the order of doing things* in Miro.

The Pod has an automatic safety brake (ASB) system that can cause the Pod to suddenly brake if a sensor notices an obstacle. In the current interface, very little information on why the Pod braked gets displayed which may confuse the operator. The need for better visualisations of this was pointed out by three different people and their comments formed the theme *Operator-supportive information*. This theme was the main inspiration for the third guideline, *Visualise root causes for triggered automated responses*.

The fourth guideline, *Visualise relevant parameters of the Pod's current status*, was based on the themes considering the truck's physical condition. According to one of the interviewees, truck drivers are typically responsible for loading the cargo. This gives them insight into what type of cargo it is and how much it weighs. Being distanced from the truck, the operator does not get this type of information and a visualisation of it is therefore needed. Other types of information the interviewees wanted to have visualisations for included tire pressure and the status of the brakes.

The inspiration for the final guideline, *Design for the limitation of data overload at operator-demanding driving situations*, was taken from several different themes. This resulted in it being broader and less focused on a specific event compared to the other four guidelines. One interviewee mentioned there being too much information in the current setup causing him to get tired much quicker. Another interviewee said that he really has to think when deciding what buttons to press. A third person said that he misinterpreted the surroundings and braked when it was not needed.

To conclude the result of the thematic analysis, while expert interviews provided the project with raw insight into the current solution, the information needed to be narrowed down and aligned with the scope of this work for it to move ahead in addressing the design problem at hand. Thematic analysis filtered expert insight according to established research questions and charted the course for the development of the first iteration of prototypes by way of providing the project with its first set of design guidelines.

5.2.2 Sketching and Wireframing

With the first set of guidelines established, the next step was to start development of sketches for interface design modifications that adhered to said guidance. Screenshots were taken of the most up-to-date remote driving interface and placed into a Figma design file to serve as reference during the sketching and wireframing process (Figma Incorporated, 2022). Following this, a total of eight sketches and subsequent wireframes were created - two proposed design solutions to investigate the effects of each of the four guidelines. The fifth of said guidelines was to be evaluated as a part of each of the four prototypes. The details and specifics of each design solution were developed through team discussion and collaborative sketching before moving on to producing more high-fidelity wireframes in an independent manner. This initial discussion was crucial in that it ensured a shared vision and understanding of the characteristics and behaviours of each solution ahead of individual wireframing.

In the case of guideline number one (from hereon referred to as G1) regarding the design of dynamic camera movement for flexibility in use, the currently used GUI developed by Einride was used as one of the two solutions in order to provide a baseline measurement in future user testing. This is the only proposed solution within the project work not completely original to the thesis authors. The second solution for the fulfilment of G1 was developed by the authors of this thesis and the sketch showing the layout of elements and views for that solution can be seen in Figure 5.3. Activities undertaken during this phase of the design work consisted of icon creation, the rearrangement of camera views and the development of a few new features.

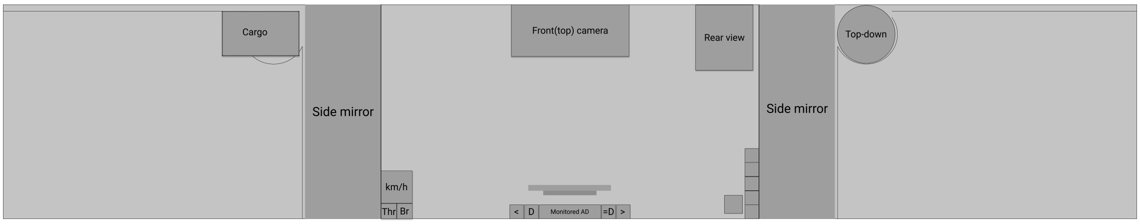


Figure 5.3: Sketch portraying a proposed solution for the placement of GUI elements and camera views.

Sketching brought the project work back into the divergent part of the Ideation phase of the design process, returning focus to the opportunities and potential solutions for designing a more operator-supportive interface within the information refined through thematic analysis.

5.2.3 Iteration 1

When sketches of the GUI had been created, the next step was to create video prototypes based on those sketches. Adobe After Effects was used to create these videos as it is a tool that offers functionality to create motion graphics and visual effects (Fridsma and Gyncild, 2017). Screen recordings of the Pod driving at a test track had been recorded at the remote drive station and these screen recordings were used as a base for the video prototypes. The live recording had been recorded by Einride engineers but the subsequent screen recordings were carried out by the thesis authors. The file format of the live recordings was only possible to view through a specific software in the station which is why a screen recording had to be done. Two different versions had been recorded; one without any GUI elements and one with the current setup. Several layers, all containing different GUI components, were then added on top of the screen recording to create the same setup as in the sketches. The recording that contained the current setup of GUI elements was used for elements such as the speedometer. Mocking these elements would take unnecessary time as detailed animations would have been needed to perceive them as realistic.

A total of eight videos were created, exported and uploaded to YouTube to simplify the viewing process during the upcoming user tests. Two versions were created for each guideline and the goal was to compare those to each other. Both A/B testing and comparative usability testing were considered when deciding on which method to use for the user testing. A/B testing typically require a high number of participants and as the participants were limited to employees at Einride due to secrecy, it was not possible to reach the desired number of people. It was, therefore, decided to use comparative usability testing and six participants were recruited. Two of the participants belong to the team who develop the station's current setup and the remaining four had never interacted with the station prior to the test. All of the participants were male and none of them have a class ‘C’ licence.

The setup for the test showing the station, facilitator and user can be seen in Figure

5.4. Having filled out the consent form outlining the ethical considerations of their participation, each participant was asked to perform a set of tasks while viewing the videos. The first test was to point out in which direction the different camera views were pointing for two different setups. The other three tests were more focused on performing tasks while driving the Pod, all the while participants kept to a Think Aloud protocol in order to give insight into their experience during interaction with the prototype. The videos were paused by the facilitator while the participant considered what to do next and subsequently resumed when they had performed their chosen action. Finally, participants were asked to reflect back on the different GUI versions after they had tested both of them. Notes were taken of uttered phrases and observations during the tests.

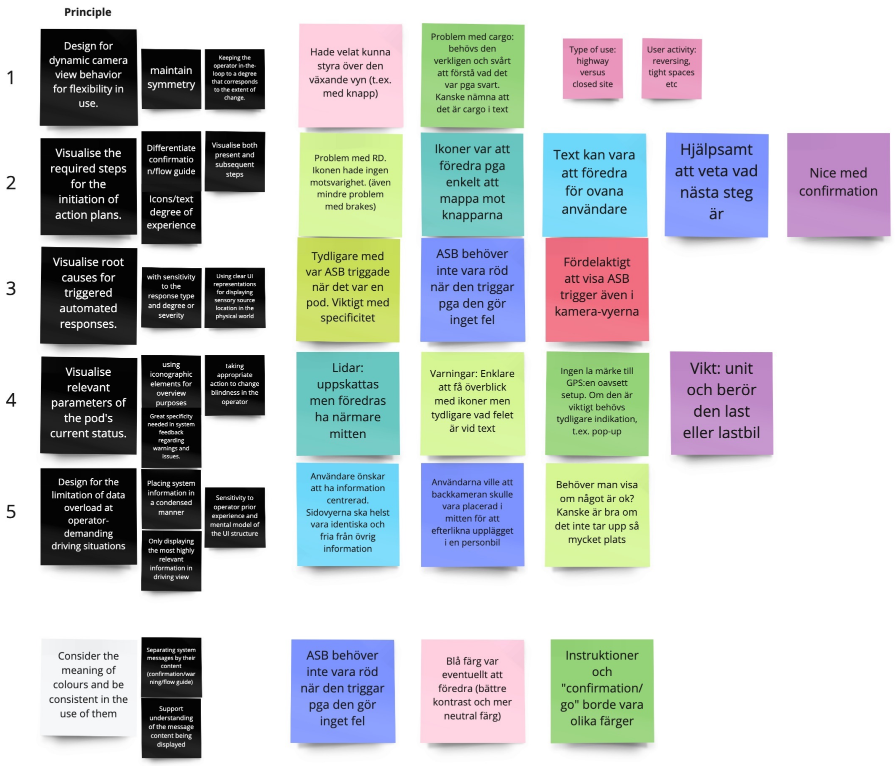
Figure 5.4: Setup for the comparative usability testing showing the station, facilitator and user.

5.2.4 Iteration 2

Following the comparative user testing of the first iteration of video prototypes, another round of affinity mapping and thematic analysis was conducted. The one difference in methodology this time around was that the content analysis was initially deductive, ordered according to the task and version of GUI solutions in conjunction with which they were observed. There were still elements of inductive analysis within not only these categories but also the summation of key insights that were collected at the tail end of the process. Examples of such key insights include participant closing statements such as appreciating reduced visual clutter or wanting to limit main screen real estate to include only the most important elements. This alternative approach to thematic coding was brought about in response to the format of the preceding user test, facilitating easier analysis and drawing of conclusions on a per solution basis. The result from this round of affinity mapping and thematic analysis can be seen in Figure 5.5.

Figure 5.5: Overview of resulting Miro board after the second round of thematic analysis.

The list of first iteration guidelines and established research questions were brought forward and compared to these new findings. As seen in Figure 5.6, post-its that were especially insightful or interesting were placed next to the guideline they related to. The guidelines were then updated to cover these new observations. Minor changes were done to the main guidelines, but the need for a higher degree of specificity was quickly highlighted and branching recommendations were, therefore, added to each guideline. As the number of identified themes grew, it became apparent that there were much more niche insights yet to be incorporated into the guidelines and so the branches were added accordingly. Examples of these new insights include recommendations for the visual differentiation between confirmation notices and other action prompts involved in the visualisation of action plans. Finally, the importance of colours was also highlighted, leading to the addition of a new guideline (the sixth guideline).



Principle

1 Design for dynamic camera view behavior for flexibility in use.
maintain symmetry
Keeping the operator in-the-loop to a degree that corresponds to the extent of change.
Hade velat kunna styra över den växande vyn (t.ex. med knapp)
Problem med cargo: behövs den verkligen och svårt att förstå vad det var pga svart. Kanske nämna att det är cargo i text
Type of use: highway versus closed site
User activity: reversing, tight spaces etc

2 Visualise the required steps for the initiation of action plans.
Differentiate confirmation/flow guide
Icons/text degree of experience
Visualise both present and subsequent steps
Problem med RD. Ikonen hade ingen motsvarighet. (även mindre problem med brakes)
Ikoner var att föredra pga enkelt att mappa mot knapparna
Text kan vara att föredra för ovana användare
Hjälpsamt att veta vad nästa steg är
Nice med confirmation

3 Visualise root causes for triggered automated responses.
with sensitivity to the response type and degree or severity
Using clear UI representations for displaying sensory source location in the physical world
Tydligare med var ASB triggade när det var en pod. Viktigt med specificitet
ASB behöver inte vara röd när den triggar pga den gör inget fel
Fördelaktigt att visa ASB trigger även i kamera-vyerna

4 Visualise relevant parameters of the pod's current status.
using iconographic elements for overview purposes
Great specificity needed in system feedback regarding warnings and issues.
taking appropriate action to change blindness in the operator
Lidar: uppskattas men föredras ha närmare mitten
Varningar: Enklare att få överblick med ikoner men tydligare vad felet är vid text
Ingen la märke till GPS:en oavsett setup. Om den är viktigt behövs tydligare indikation, t.ex. pop-up
Vikt: unit och berör den last eller lastbil

5 Design for the limitation of data overload at operator-demanding driving situations
Placing system information in a condensed manner
Only displaying the most highly relevant information in driving view
Sensitivity to operator prior experience and mental model of the UI structure
Användare önskar att ha information centrerad. Sidovyerna ska helst vara identiska och fria från övrig information
Användarna ville att backkameran skulle vara placerad i mitten för att efterlikna upplägget i en personbil
Behöver man visa om något är ok? Kanske är bra om det inte tar upp så mycket plats

Consider the meaning of colours and be consistent in the use of them
Separating system messages by their content (confirmation/warning/flow guide)
Support understanding of the message content being displayed
ASB behöver inte vara röd när den triggar pga den gör inget fel
Blå färg var eventuellt att föredra (bättre kontrast och mer neutral färg)
Instruktioner och "confirmation/go" borde vara olika färger

Figure 5.6: Miro board showing guidelines along with notes related to them.

Combining the previously established material with new findings produced a second iteration of design guidelines for remote operations:

1. Design for dynamic camera behaviour for flexibility in use.
    (a) Maintain symmetry for consistency in use.
    (b) Keep the operator in-the-loop to a degree that corresponds to the extent of the change.
2. Visualise the required steps for the initiation of action plans.
    (a) Differentiate confirmation elements from the rest of the flow guide.
    (b) Visualise both present and subsequent steps.
    (c) Balance the use of icons versus text in a way that matches operator experience.
3. Visualise root causes for triggered automated responses.
    (a) Use clear GUI representations for displaying sensory source location in the physical world.
    (b) Design notices with sensitivity to the response type and its degree of severity.
4. Visualise relevant parameters of the Pod's current status.
    (a) Use iconographic elements for overview purposes.
    (b) Apply great specificity in system feedback regarding warnings and issues.
    (c) Take appropriate action to combat change blindness in the operator.
5. Design for the limitation of data overload at operator-demanding driving situations.
    (a) Place GUI elements of system information in a condensed manner.
    (b) Design with sensitivity to operator prior experience and mental model of the GUI structure.
    (c) Display only the most highly relevant information in the driving view.

Another guideline was put forward as a candidate for potential addition:

6. Consider the meaning of colours and be consistent in the use of them.
    (a) Differentiate system messages by their content, such as confirmations, warnings and action steps.
    (b) Support rapid understanding of the message content being displayed.

Having established a second iteration of the design guidelines, work began on a new prototype to be used in their evaluation. Sketching and wireframing of an updated GUI layout was once again carried out in Figma. Alongside it, a flowchart of the video prototype showcasing the timing of events and the corresponding guidelines at play with specific user challenges matched below was developed (see Figure 5.7). These materials were then used as specifications in the production of two video clips in After Effects, once more applying new GUI elements on top of screen recordings of a Pod in action. These two video clips were then uploaded as unlisted on YouTube and had interactions applied to them via the Anima plugin in Figma (Anima, 2022). This enabled more high-fidelity interactivity during the user test via the Wizard of Oz method.

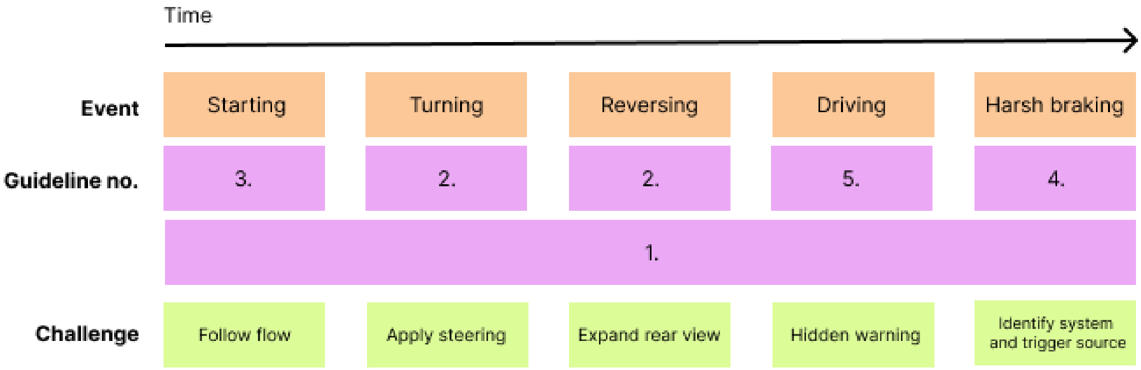


Figure 5.7: Flowchart of events in user testing of second iteration prototype.

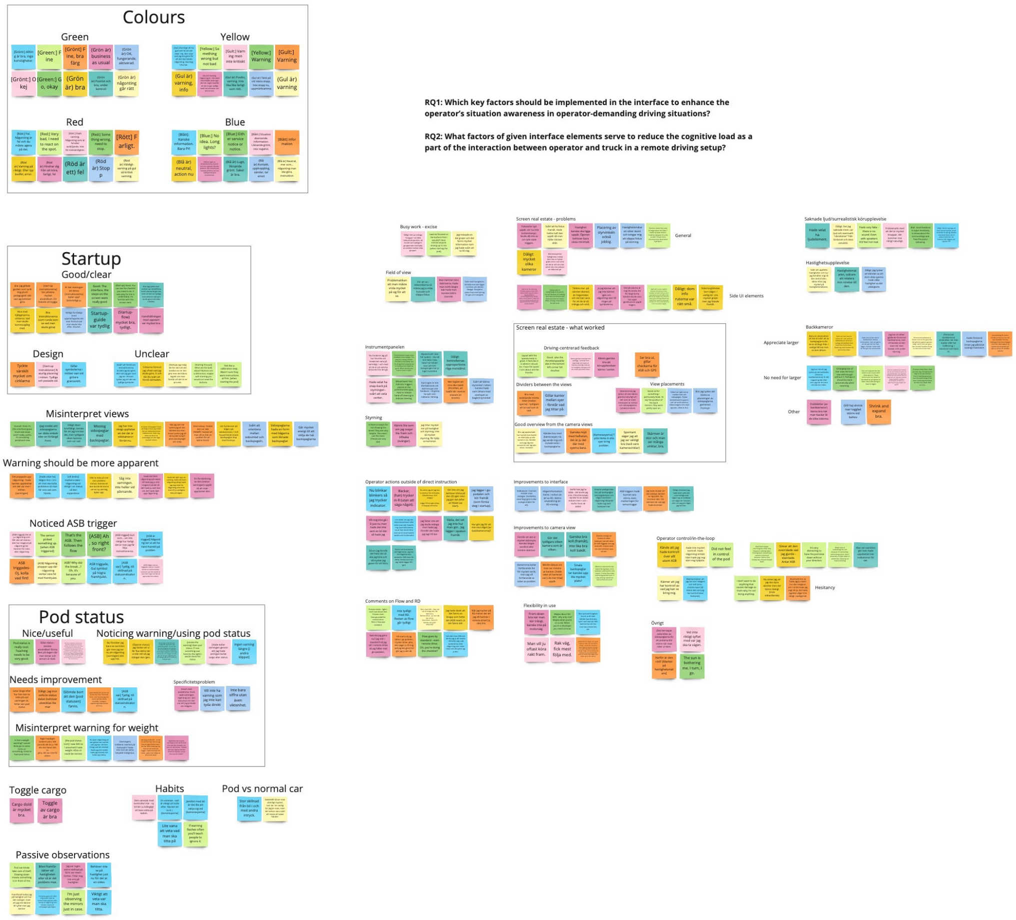
The second iteration usability test was of a more free-form format than the comparative user testing that had preceded it. Ten participants were recruited from within Einride to partake in the test, 30% of whom were women. The participants' prior interaction with the station was varying as it ranged from not having seen it before to having used it for long periods of time. Two of the participants had a class 'C' licence and two additional participants had used the station previously. Being more exploratory and open-ended in nature, participants only received instructions at the beginning of the test and during the swap to the second clip. After finishing the test, they were asked reflective questions regarding their experience, but they were otherwise largely unrestricted in their interaction with the prototype. Having filled out the prerequisite consent form, participants were explicitly encouraged to interact with the controls and interface as they saw fit, all while thinking aloud. The facilitator had the responsibility of providing the instructions as well as carrying out the Wizard of Oz illusion via the Anima interactions. The other member was instead responsible for taking notes of observed behaviours, Think Aloud-utterances and question responses while also ensuring that the session was being properly recorded. To secure further reflections from participants, each was sent a follow-up inquiry the day after their testing session. This delayed reflection asked them to name things about the interface that stood out to them as positive and negative. The question sent to the participants was:

- Thinking back on it, was there something you remember standing out as needing to improve the most and is there something you particularly liked in the interface?

Nine out of ten participants provided a response and their feedback was noted down. Most replies were things they had already said during the evaluation, but there were also some comments regarding things the participant had forgotten to mention the day before. One such response started with "I forgot to mention..." highlighting that this feedback would not have been received if the delayed reflection method had not been used. The comments that restated things the participants said during the evaluation indicated what they found most memorable and important. These comments were, therefore, considered to be more important.

5.3 Implementation

Having thus concluded another usability test, the session notes from each participant were considered through the lens of the established thesis research questions. Every listed observation with relevant ties to said questions was pooled together onto individual notes on a Miro board for thematic analysis via affinity mapping. Content analysis was largely inductive with only the commentary on colours being categorically separated from the rest of the response material. Individual observations were once again clustered into groups, seen in Figure 5.8, and ordered according to emerging themes to be compared against the guidelines from which the prototype had first been conceived.



RQ1: Which key factors should be implemented in the interface to enhance the operator's situation awareness in operator-demanding driving situations?

RQ2: What factors of given interface elements serve to reduce the cognitive load as a part of the interaction between operator and truck in a remote driving setup?

Figure 5.8: Overview of the clustered notes on the Miro board as a result of the third thematic analysis.

The guidelines were summarily updated in light of new user feedback, trimming a few branching bits of specified recommendations. An example of this is that branch (c) of the previously second guideline, *Balance the use of icons versus text in a way that matches operator experience*, was removed. The reason for the deletion of this branch was that no one had trouble understanding the iconographic instructions even though several of the evaluators had no prior experience. The listed order of the guidelines were also changed for this iteration. This was mainly done in order to bring the overarching guideline regarding data overload to the beginning of the list so that it can be kept in mind for the duration of the reading.

1. Design for the limitation of data overload in operator-demanding driving situations.
    (a) Place GUI elements of system information in a condensed manner.
    (b) Display only the most highly relevant information when driving.
2. Design for dynamic camera view behaviour for flexibility in use.
    (a) Maintain visual symmetry of camera layout for consistency in use.
    (b) Keep the operator in-the-loop to a degree that corresponds to the extent of the change.
3. Visualise the required steps for the initiation of action plans.
    (a) Differentiate confirmation elements from procedural action steps.
    (b) Visualise both present and subsequent steps.
4. Visualise root causes for triggered automated responses.
    (a) Use clear GUI representations for specifying the source location in the physical world.
    (b) Design notices with sensitivity to the response type and its degree of severity.
5. Visualise relevant parameters of the Pod's present operational status.
    (a) Use iconographic elements for overview purposes.
    (b) Take appropriate action to combat change blindness in the operator.
    (c) Use great specificity in system feedback regarding warnings and events.
6. Consider the meaning of colours and be consistent in the use of them.
    (a) Differentiate system messages by their content, such as confirmations, warnings and action steps.
    (b) Use established colour conventions to direct attention and enhance understanding of system messages.

Other themes, seen in Figure 5.9, focused on the physical input devices, such as the Stream Deck, potentiometer and throttle. The design of these elements was originally considered to be outside the scope of this thesis, but due to the high number of comments regarding them, it was decided to highlight these insights. One evaluator mentioned that it was hard to reach the button for putting on the indicator without letting go of the potentiometer, making it impossible to steer and press the button simultaneously. Another person did not like the high resistance in the throttle and several evaluators reported difficulties in knowing how much they were turning the potentiometer. Others also reported having problems understanding how the input on the potentiometer affected the Pod's steering. Based on this feedback, another potential guideline was put forward:

7. Design input devices to enhance usability.
    (a) Augment inputs with visual feedback of the requested effect.
    (b) Design tangible affordances to support user understanding of control orientation.

Instrumentpanelen

Nu funderar jag på hur throttle och break kan vara på samtidigt. I och med att de är på samma så borde inte det gå.

Potentiometer needs more feedback, more weight. Too easy for a heavy vehicle. Throttle is a bit better but it's tiring (from experience) so having a lock might be good. When remote you have to have it open.

Mycket kraft i den här spaken - ska väl inte hålla i hela tiden? Jo? Då blir man väl trött. Svängningen har ju inget motstånd.

Dåligt: kontrollernas oppproportione rliga motstånd.

Hade velat ha snap-back på styrningen - svårt att veta center.

[Would want the] Indicator triggers placed on the potentiometer grip. Hard to release hand off steering to indicate steering.

Styrningen är inte återfjädrande och skärmen ger inte feedback - vill gärna ha spår som indikerar riktning.

Mer logiskt att röra den bakåt [throttlen, att bakåt blir reverse snarare än broms]

Svårt att känna styrvinkel i ratten - kanske hade varit lättare med overlayet av begärd synvinkel

Styrning

Is there a reason for not showing the lanes? They are very helpful. The ones I use the most. [These] are very bad.

Känns lite som om jag svajar lite fram och tillbaka (svänger).

Jag tittar mycket ner på hastighet och styrning. Inte så mycket styrning, får hjälp av kameran.

Figure 5.9: Post-it notes forming the theme regarding physical input devices.

5.3.1 Evaluation of Guidelines

As the guidelines had been updated, evaluations were held to understand how the intended interpretation of each guideline matched the interpretation made by evaluators. Different methods were considered for this evaluation and a decision was made to use a combination of the unstructured and semi-structured interview methods for different parts of the session. It was decided not to use the AGREE II framework as it is a more quantitative and structured method and the goal of this evaluation was to allow them sufficient freedom to make reference to a broad spectrum inquiry as a part of the pairwise discussion.

During this evaluation, the penultimate version of the proposed design guidelines was presented to three pairs of Interaction Design Master students from Chalmers University of Technology to evaluate the legibility and composition of each recommendation. A total of six participants partook in the evaluation, all of whom were female and currently in the last term of their degree. Upon filling out the prerequisite consent form, each pair was shown part of a local news segment showcasing Einride's remote operating station to introduce them to the context at hand. A link to the news segment can be found in Appendix D. One member of the design team facilitated while the other took notes and managed the audio recording. The evaluation itself took the form of an unstructured interview during which a few initial questions were asked, followed by a discussion between the participant pair and the facilitator covering each of the proposed guidelines and their branches. Following this, a final pair of questions were asked before the session was over. The questions asked during the interview can be seen in Appendix E.

This set of interview discussions allowed for modifications to be made to the guidelines based on the perspective provided by the Interaction Design students. The notes taken during each session were compiled into a new document according to the question answered or the guideline to which they were referring. Every guideline and its branches then had their intended meaning compared with the interpretation offered by the other designers. Following this, prospective changes to wording and structure was discussed between the thesis authors. In the end, select phrases were edited for clarity and specific words exchanged for ones with more suitable definitions. Example excepts of changes made for the final iteration of guidelines can be found below.

A phrase that was edited for clarity (the first point is the previous phrasing and the second point is the final result):

- **Keep** the operator **in the loop** to a degree that corresponds to the **extent of the change**.
- **Inform** the operator **of changes** to a degree that corresponds to the **impact of said changes**.

A phrase with changed wording (before/after):

- Place GUI elements of system information in a **condensed** manner.
- Place GUI elements of system information in a **well-grouped** manner.

6

Results

In this chapter, the results achieved throughout the Research through Design process are presented in an attempt to answer the established research questions. To reiterate, the questions to be answered are as follows:

- RQ1: What are key interface factors for enhancing the operator's situation awareness in demanding driving situations?
- RQ2: What factors of given interface elements serve to reduce the cognitive load as a part of the interaction between operator and truck in a remote driving setup?

The presented findings are ordered according to their place in the design process and sectioned according to the corresponding iteration. Design guidelines are discussed as they appear and evolve throughout the design cycles. Finally, the final iteration of said guidelines is presented and expounded upon via short summaries and mention of example use cases.

6.1 First Iteration of Prototypes

The video prototypes created in the first iteration of the design process were all based on a video recording of an operating Pod and have different interface overlays on top of the footage. Each pair of prototypes showcase two solutions for the same guideline, using an identical sequence of events, with the difference being the interface proposal itself. The prototypes are described in the following sections in conjunction with the guideline from which they were designed.

6.1.1 Guideline 1

Design for dynamic camera behaviour for flexibility in use.

The first guideline (G1), which sought to achieve flexibility in use via dynamic camera view behaviour, was embodied through the design of an alternate structure of camera views. The first solution of the prototype was the layout currently in use for the station with no changes made to it. This original solution can unfortunately not be illustrated in this report due to the material being confidential. For the second solution, the camera views had been reorganised to create a new layout. This new layout can be seen in Figure 6.1 which also highlights the changes differentiating it from the original layout. The second version also featured a dynamic rear camera view that grew in size at times where it was speculated that the operator would require a good view of the back part of the Pod. This behaviour was envisioned to be automatic in times such as shifting gear into reverse.

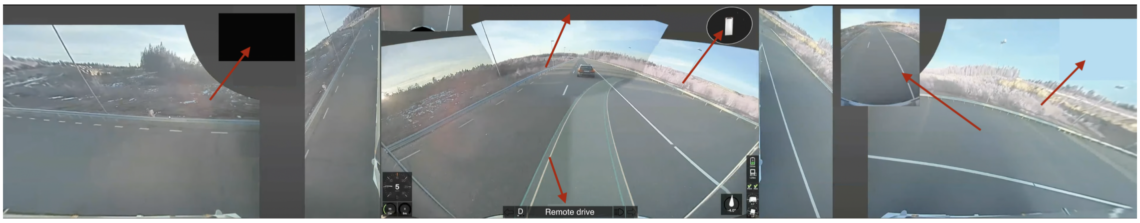


Figure 6.1: Camera layout of the second solution for G1 with red arrows pointing out the parts that differs from the original layout.

6.1.2 Guideline 2

Visualise the required steps for the initiation of action plans.

Two different solutions, using the context of a vehicle startup process, were produced for the second guideline (G2). Each proposed solution featured a distinct style of guiding the operator through this flow, the first one being text-based and the other icon-based. The first solution featured rectangles with written instructions centred on the screen (seen in Figure 6.2). It used opacity and movement to give feedback to the operator regarding the timing, completion and most importantly, order of each required action input. The second functioned much the same in terms of visual feedback but with the primary difference being the replacement of written prompts with iconographic ones (seen in Figure 6.3). The animated flow was also placed slightly more to the left-hand side of the screen as opposed to being centred. The reason for this was mainly to test out another type of animation for progressing through the steps. The instructions in this solution moved in a circular motion as opposed to falling down as they did in the first solution.

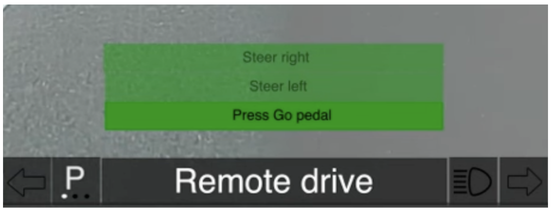


Figure 6.2: Text-based instructions used in the first solution for G2.

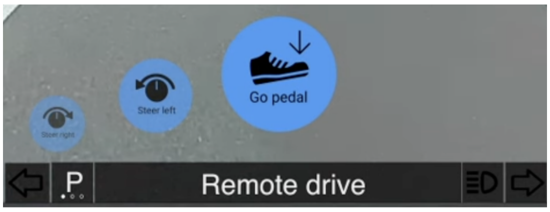


Figure 6.3: Icon-based instructions used in the second solution for G2.

6.1.3 Guideline 3

Visualise root causes for triggered automated responses.

Guideline number 3 (G3) was instantiated in the form of an enhanced notification style for the automatic safety brake system (ASB) with the purpose of comparing design solutions for the visualisation of root causes for triggered automated responses. The two proposed solutions each featured an interface element signalling the status of the ASB system with the additional functionality of displaying an area or direction in which the ASB system had detected an obstacle, as well as a guiding action flow reminiscent of that posited in G2 in order to resume driving post-trigger event. The first solution, seen in Figure 6.4, again featured a text-based action guide matched with a rectangular ASB status element that used colour and iconography to signal whether it was active, dormant or triggered. This proposed solution used a circular element placed on the periphery of the ASB rectangle in order to display the area in which the sensor had triggered. The circle used colour and size to give visual feedback of the direction and proximity of said trigger. The second proposed solution, seen in Figure 6.5, featured a square element with a coloured frame that similarly signalled the status of the ASB system, augmented with icons to match. Upon being triggered, the initial element would expand to display a top-down view of a miniature Pod with a circle element showing the precise location of the sensor that was triggered. The unfolded element retracts automatically as the truck disengages the ASB.

Put direction in drive
Change driving mode to remote drive
Disengage ASB

Figure 6.4: The first solution for G3 with a text-based action guide and a red ASB status element to signal that it has been triggered.

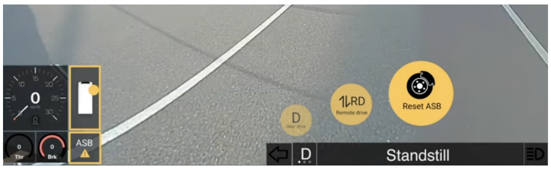
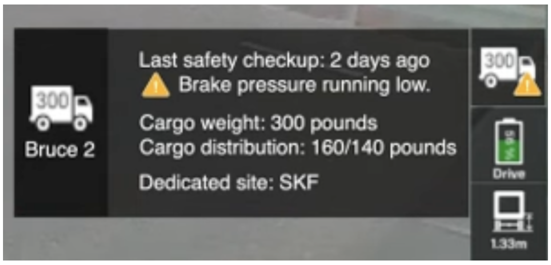


Figure 6.5: The second solution for G3 with an icon-based action guide and two status elements with yellow borders to signal that the ASB has been triggered.

6.1.4 Guideline 4

Visualise relevant parameters of the Pod's current status.

Guideline 4 (G4) focuses on ways through which relevant parameters of the Pod's current status could be visualised for the operator. In the proposed solutions, this was realised in the form of a new interface window where one solution showcased a more text-based style and the other a more icon-based one. In both cases, the window was manually controlled by the operator through the use of a dedicated button on the Stream Deck. The first solution used text to inform the operator of parameters such as cargo load and warning notices (see Figure 6.6). The other solution used two rows of icons dedicated to giving an overview of vehicle-related and cargo-related factors respectively (see Figure 6.7). G4 also visualised different solutions for showcasing the magnitude and strain on brakes applied: one in the style of a brake pressure gauge using the colour and fill of said gauge for visual feedback while the second solution was integrated into the already existing brake element. This other solution used the expansion and opacity of a circular element filling the space within the gauge's arch to signal the same strain and magnitude of applied brakes.



Last safety checkup: 2 days ago

⚠ Brake pressure running low.

Cargo weight: 300 pounds

Cargo distribution: 160/140 pounds

Dedicated site: SKF

Figure 6.6: Text-based visualisation showing the Pod's status as the first solution for G4.

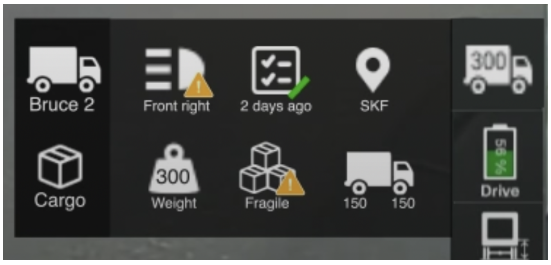


Figure 6.7: Icon-based visualisation showing the Pod's status as the second solution for G4.

6.1.5 Guideline 5

Design for the limitation of data overload at operator-demanding driving situations.

Guideline 5 (G5) differs from the rest in that it did not have a single example task or solution by which to evaluate it, but was instead made a part of the evaluation of all preceding design solutions. Each was considered through the lens of limiting the risk of information overload, especially in situations considered to be more demanding for the operator.

6.2 Second Iteration of Prototypes

The prototypes in the second iteration of the design phase were fewer in number and only showcased a single variant of the interface elements up for evaluation. The videos were longer in duration and made for a more linear experience while also featuring increased interactivity via Anima animations. Like in the previous section, this part covers the theoretical background to the tasks inlaid into the prototype in order of the guideline from which it originated. The guidelines themselves remained largely unchanged in the transition into the second iteration but became more specific through the inclusion of branches offering more specialised advice under each main guideline.

6.2.1 Guideline 1

Design for dynamic camera behaviour for flexibility in use.

- *(a) Maintain symmetry for consistency in use.*
- *(b) Keep the operator in-the-loop to a degree that corresponds to the extent of the change.*

G1 still primarily influenced the layout of the camera feeds themselves and resulted in a similar structure to what was seen in the first solution presented in the previous prototype. A few alterations were done to match feedback given during the comparative usability test. Instead of a single rear-view feed on just one side of the setup, this prototype showcased the feed on each side of the middle screen, providing symmetry for consistency in use and to meet operator expectations. The expansion functionality was made a toggle using a button as opposed to automatic in order to better keep the operator in the loop and animated using Anima. The lidar view (seen in Figure 6.8) was also placed in the top-right corner of the middle screen and the cargo view was adjusted to the upper corner of the far right screen. Said cargo view was also kept hidden until the operator hit a dedicated button - a change from the previous iteration. A screenshot showing the camera layout can be seen in Figure 6.9.

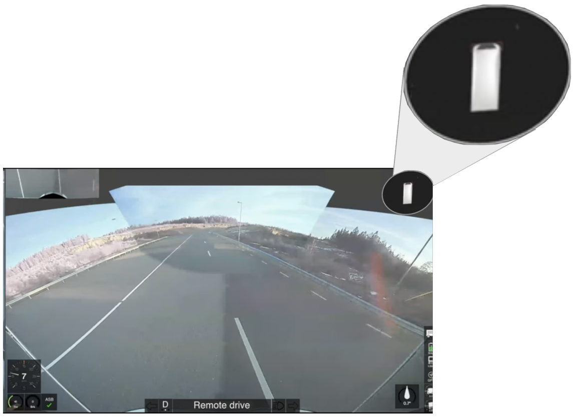


Figure 6.8: The lidar view, i.e. a live feed of the Pod's lidar sensors, highlighted in the interface.

Figure 6.9: Camera layout for the prototypes in the second iteration with red arrows pointing out differing parts from the layout in the first iteration.

6.2.2 Guideline 2

Visualise the required steps for the initiation of action plans.

(a) Differentiate confirmation elements from the rest of the flow guide.

(b) Visualise both present and subsequent steps.

(c) Balance the use of icons versus text in a way that matches operator experience.

G2 expounded upon the iconographic version of the action plan prompts with another set of alterations born from the user test (see Figure 6.10). The colour of the standard prompts was changed to blue and the confirmation element "Ready to drive" was given a green background to differentiate its status of a confirmation notice - a prompt that did not require any additional input from the operator. A few icons were also redesigned to better match those on the Stream Deck as it was observed that users heavily relied on matching symbols to figure out which action was being requested of them. The action prompts retained their use of opacity and movement to give visual feedback of which step was the current one as well as which action would be requested next.

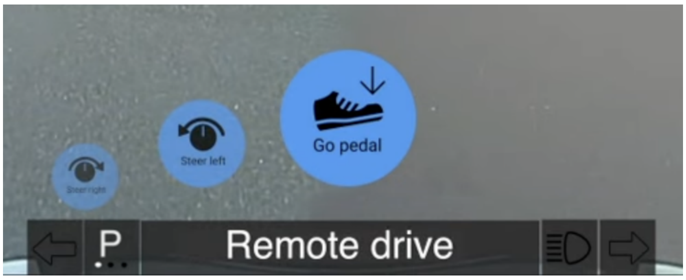


Figure 6.10: Iconographic action plan prompts.

6.2.3 Guideline 3

Visualise root causes for triggered automated responses.

(a) Use clear GUI representations for displaying sensory source location in the physical world.

(b) Design notices with sensitivity to the response type and its degree of severity.

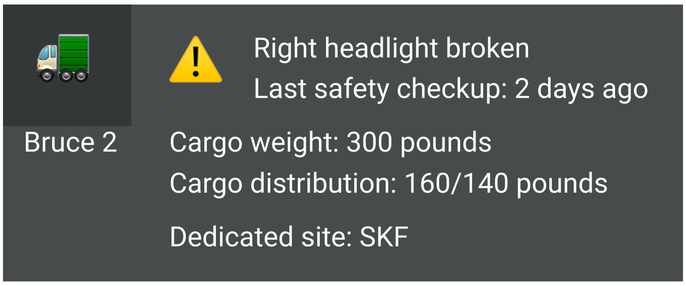
G3 employed the second of its two solutions to better visualise the context surrounding the triggering of the ASB system with minor alterations made (see Figure 6.11). Similarly to in the preceding section, the "Ready to drive" confirmation notice was given a green background to better differentiate it from the actionable prompts of the startup procedure following the automated braking. None of the braking visualisations were included in this prototype as it was deemed difficult to evaluate.

Figure 6.11: Screenshot of the middle screen in the video prototype when the ASB has triggered.

6.2.4 Guideline 4

Visualise relevant parameters of the Pod's current status.
(a) Use iconographic elements for overview purposes.
(b) Apply great specificity in system feedback regarding warnings and issues.
(c) Take appropriate action to combat change blindness in the operator.

G4 had its initial proposed solutions synthesised into a hybrid of the two, employing a text-based foundation with iconographic elements for an easy overview of information and direction of operator attention (see Figure 6.12). This approach was chosen due to an identified need for specificity in regard to warning notices. Bright colours were used in icons of special importance, such as that of the warning triangle, in an attempt to combat change blindness in the operator. As was the case in the previous prototype, the status window was manually operated through a dedicated button that was animated using Anima in order to give it interactivity via the Wizard of Oz method.



⚠ Right headlight broken
Last safety checkup: 2 days ago

Bruce 2

Cargo weight: 300 pounds
Cargo distribution: 160/140 pounds

Dedicated site: SKF

Figure 6.12: Window with information about the Pod's status.

6.2.5 Guideline 5

Design for the limitation of data overload at operator-demanding driving situations.

(a) Place GUI elements of system information in a condensed manner.

(b) Design with sensitivity to operator prior experience and mental model of the GUI structure.

(c) Display only the most highly relevant information in the driving view.

G5 was once more incorporated as more of an overarching principle evaluated throughout the driving experience as opposed to one embodied in one or more functionalities in the interface. However, the need for placing system information GUI elements in some proximity to one another, as opposed to being spread out across the screen real estate, was pointed out in the preceding usability testing pointed out. As such, the second iteration video prototype showcased three clusters of GUI elements at the bottom of the middle screen and was a reason for the adjustment of the lidar feed's positioning onscreen. Further efforts were made to adjust the appearance of the physical button and the action prompt icon for easier matching, indirectly supporting less experienced operators into taking a prompted action and reducing their cognitive load. A few elements such as the GPS status indicator was redesigned into a simplified icon for quicker and easier interpretation following similar reasoning.

6.2.6 Guideline 6

Consider the meaning of colours and be consistent in the use of them.

(a) Differentiate system messages by their content, such as confirmations, warnings and action steps.

(b) Support rapid understanding of the message content being displayed.

Finally, a new guideline regarding the meaning and consistency in the use of colours had emerged following the analysis of user feedback on the first iteration of video prototypes. This was a largely spontaneous finding that nevertheless sparked discussion when inquiries were made regarding the efficacy of the two design proposals from G2. Colours proved to be very effective visual signals for the differentiation and interpretation of various interface elements. This factor might also serve to reduce cognitive load and thus made it highly relevant to the given research question. This was incorporated throughout the second iteration prototype - primarily as part of interface elements connected to G2, G3 and G4, as mentioned earlier in this section.

6.3 Design Guidelines

Concluding and combining the findings brought about through the evaluation of two consecutive iterations of interface prototypes as well as feedback given by interaction designers, the following section lists guidelines and best-practice recommendations for the design of a remote operations interface for remotely driven trucks. All guidelines are presented alongside their respected branches and their meaning is expounded upon in subsequent sections.

1. Design for the limitation of data overload in operator-demanding driving situations.
    (a) Place GUI elements of system information in a well-grouped manner.
    (b) Display only the most highly relevant information when driving.
2. Design for dynamic behaviour in the display of camera feeds for flexibility in use.
    (a) Maintain visual symmetry of camera layout for consistency in use.
    (b) Inform the operator of changes to a degree that corresponds to the impact of said changes.
3. Visualise the required steps for the initiation of action plans.
    (a) Differentiate call-to-action elements from confirmation notices.
    (b) Visualise both present and upcoming steps of the action plan.
4. Visualise root causes for triggered automated responses.
    (a) Use clear GUI representations for specifying the source location in the physical world.
    (b) Design notices with sensitivity to the response type and its degree of severity.
5. Visualise relevant parameters of the truck's condition.
    (a) Use iconographic elements for overview purposes.
    (b) Take appropriate action to counter change blindness in the operator.
    (c) Use great specificity in system feedback regarding warnings.
6. Consider the meaning of colours and be consistent in the use of them.
    (a) Differentiate system messages by their content, such as confirmations, warnings and action steps.
    (b) Use established colour conventions to direct attention and enhance understanding of system messages.

6.3.1 Guideline 1

Design for the limitation of data overload in operator-demanding driving situations.
(a) Place GUI elements of system information in a well-grouped manner.
(b) Display only the most highly relevant information when driving.

The first guideline focuses on data overload and highlights the importance of limiting it. The operator should not be presented with too much information at once since that can be overwhelming, as mentioned in branch (b) (G1b). After all, data overload can not only cause distraction but also bait user attention through misplaced salience (Endsley and Jones, 2014). G1b further states that the information shown in the interface should be highly relevant to the operator. This is especially true when driving and additional information can, for example, be shown when the vehicle is standing still and the operator has not yet started driving.

The (a) branch (G1a), instead focuses on the arrangement of GUI elements in the interface. There should not be many different groups of GUI elements scattered on the screen as that requires the operator to move their gaze and head much more often. Having many things compete for the user's attention is considered visual excise and should, according to Cooper et al., be avoided whenever possible (Cooper et al., 2014). The GUI elements in the prototypes were placed in three groups, all placed at the bottom of the middle screen, which evokes the Gestalt theory of proximity by forming perceptual clusters (Fekete et al., 2008). Having them placed on the middle screen and not too far away from each other makes it easier to quickly glance at one of them.

6.3.2 Guideline 2

Design for dynamic behaviour in the display of camera feeds for flexibility in use.
  (a) Maintain visual symmetry of camera layout for consistency in use.
  (b) Inform the operator of changes to a degree that corresponds to the impact of said changes.

Allowing the display of camera feeds to be dynamic is the focus of the second guideline and dynamic, in this case, mainly refers to the feeds being able to change size. Buttons can be used to resize some views and toggle the visibility of others allowing for an interface that can be adapted to best suit the current driving situation. There are, however, some things to consider when allowing this type of behaviour and they are stated in the branches (G2a and G2b).

G2a points out the importance of keeping a visual symmetry of the camera feeds. This means that the layout of the camera feeds should be the same even if their size or visibility is toggled. Keeping the interface layout consistent facilitates the operator's navigation of the interface and limits potential confusion (Cooper et al., 2014). A camera feed that can be hidden should have some type of symbol indicating where it will be shown and feeds that grow in size should have the placement even after the size has changed. G2b states that it is important to inform the operator of any changes that happen. Consistency in use means that the layout should be mindful that regardless of dynamic changes made to camera views, the operator should be able to recognise each view and be able to predict their positioning.

6.3.3 Guideline 3

Visualise the required steps for the initiation of action plans.
  (a) Differentiate call-to-action elements from confirmation notices.
  (b) Visualise both present and upcoming steps of the action plan.

The third guideline describes the importance of visualising the steps an operator needs to take to reach a goal. An example of a situation where this is useful is when starting the Pod. Before being able to start driving, there are a set of actions the operator has to do, in a specific order. These steps can be hard to remember and visualisations can, therefore, help the operator. The visualisations should, as stated in branch (a), show both the current and upcoming steps so that the operator is prepared for what happens next. When the operator has completed all the steps, a confirmation notice can be shown and that should, as mentioned in branch (b), differentiate from the visualisation of the steps that call to action. All of these recommendations serve to reduce the likelihood of active failures occurring, specifically mistakes and procedural violations (Reason, 2000).

6.3.4 Guideline 4

Visualise root causes for triggered automated responses.

(a) Use clear GUI representations for specifying the source location in the physical world.
(b) Design notices with sensitivity to the response type and its degree of severity.

An automated vehicle has automated responses and the fourth guideline states that when these are triggered, it is important to visualise the root cause of that response. An automated response can for example be triggered if an object appears in front of the vehicle and is picked up by a sensor that causes the automatic safety brakes to trigger. In this type of situation, it is important to both use visualisations to show that the system has been triggered and the reason for why it was triggered. The (a) branch states that the source location, e.g. where the object is located, should be visualised in the GUI. It is also important to consider what type of trigger it was and how severe it is, which is stated in branch (b). Keeping the operator in the loop ensures that the operator updates their mental model of SA accordingly and increases the chances of them successfully identifying potential issues (Endsley and Jones, 2014). It also serves to mitigate one major cause of accidents within the field of HAI by ensuring that the operator is supplied with sufficient feedback from the system (Sheridan and Parasuraman, 2005).

6.3.5 Guideline 5

Visualise relevant parameters of the truck's condition.

(a) Use iconographic elements for overview purposes.
(b) Take appropriate action to counter change blindness in the operator.
(c) Use great specificity in system feedback regarding warnings.

Operators have no way of checking the truck's physical condition by examining its outside and it is, therefore, important to use visualisations to show the current state of the truck. An example of information that can be of relevance for the driver is the cargo's weight, type and if it has been anchored correctly.

There are different ways that relevant parameters of the truck's condition can be visualised, but there are some things to consider. The first branch, G5a, states that iconographic elements are good for overview purposes but branch (c) points out that system feedback regarding warnings should use great specificity. The reason why those need more specificity is that it is important to fully understand the warning as that can affect the operations and potentially become dangerous. Similarly to what was presented in the summary of the previous guideline, visualising indicators of the truck's condition mitigates the risk of the operator receiving insufficient feedback, one of three major causes of bad outcomes in automation (Sheridan and Parasuraman, 2005). Furthermore, it bridges the Gulf of Evaluation by providing the operator with information they can use to better plan and problem solve (Behymer and Flach, 2016).

Branch (b) points out the need to take appropriate action to counter change blindness in the operator. What an appropriate action is relates to the importance of the operator noting that change. If it is important to inform the operator about something, then it should be obvious to them that something has changed. A small change in colour or iconography may, for example, be difficult for the operator to notice. By paying attention to how interface elements are presented to the operator, the amount of extraneous cognitive load experienced by the operator can be reduced (Sweller et al., 1998).

6.3.6 Guideline 6

Consider the meaning of colours and be consistent in the use of them.

(a) Differentiate system messages by their content, such as confirmations, warnings and action steps.

(b) Use established colour conventions to direct attention and enhance understanding of system messages.

The focus of the sixth guideline is colours and how they should be used. People usually associate colours with a specific meaning and it is important to consider those meanings. It is also important to be consistent in how colours are used throughout the interface. Branch (a) delves deeper into this by stating that system messages should be differentiated based on their content. An example of this is when blue was used for the steps in the action plan, but green was used for the confirmation notice. This better highlights the difference between the two different types of messages. Furthermore, it presents yet another way to reduce the amount of extraneous cognitive load placed on the operator of this interface (Sweller et al., 1998).

The other branch, G6b, points out that using established colour conventions is a great way of aiding the operator. One common use of colours is green, yellow and red that often is interpreted as a scale from good to bad. Having something change from green to red can, for example, let the operator know that something has changed and need attention. Incorporating rich modeless feedback in the design of GUIs allows the operators to internalise information much quicker, perhaps just by a glance (Cooper et al., 2014). However, it is also important to consider colour blindness and that a colour change never should be the sole indicator of change (Story, 1998b).

7

Discussion

In this chapter, the methods used throughout the project are discussed along with the result and how general the resulting guidelines are. Potential future work is also discussed and ethical considerations are considered.

7.1 Result Discussion

This section expounds on results produced through user testing as they pertain to the theoretical background of the thesis. Starting with discourse regarding the established research questions, presented findings from both iterations of prototype evaluations are then brought up to give further context for each separate discussion.

Before delving into the specifics of the outcome, how the finished guidelines look to answer the established research questions merits some additional deliberation, especially regarding RQ1. This research question states that the thesis work seeks to identify key interface factors for the enhancement of the operator's situation awareness, but arguments can be made whether or not this is achieved via the produced guidelines. The design of interface solutions can be considered a wicked problem. As mentioned previously, wicked problems tend to become more understandable as the process closes in on a solution (Plattner et al., 2010). This is mirrored in the thesis work, the scope changing in time with design solutions being produced as part of the Research through Design process.

This shift lead to an increased focus on the practical application and iteration of new insights as opposed to the identification of specific interface and usability factors. As a consequence of this, the guidelines developed through this design effort do not, in fact, encompass an exhaustive list of key factors for interface design as it pertains to either RQ1 and RQ2. Instead, they provide design insight that underpins the broader discussion surrounding those same factors. It is generally considered difficult to ask the right question from the very start of a project and as such, allowing for appropriate design methods and iterative work to lead the way have answered the research questions in a manner that fits the process. The extensibility criterion for the quality assurance of Research through Design projects states that new research needs to be based on the contributions produced in the process (Zimmerman et al., 2007). The design work carried out as part of this thesis fulfils this criterion by maintaining a high sensitivity to the products of the process and thus leads to the research questions being attended to in a less direct manner.

7.1.1 Situation Awareness

Looking back at the results achieved by way of practical design work through the lens of established academic theory, five particular situation awareness (SA) risk factors stood out as particular threats to remote operators within the established context. These five are the SA demons of attentional tunnelling, data overload, complexity creep, errant mental models and out-of-the-loop syndrome.

As mentioned in earlier chapters, attentional tunnelling is the phenomenon where a user fixates on one set of information to the exclusion of others (Endsley and Jones, 2014). It is largely a natural consequence of selective attention, a cognitive phenomenon that implies the withdrawal of attention from some source of information to focus on another (James, 1890). This was especially observed during the second of the two user tests. Several participants commented on their tendency to fixate their attention on one of the three clusters of GUI groups; more specifically the cluster at the bottom left-hand corner of the central screen featuring, among other things, the speedometer and the new ASB display. A few also reported an awareness of this becoming a potential issue in terms of maintaining awareness of GUI changes within other interface clusters. Branch (a) of G1 (G1a) largely addresses this issue by recommending that GUI elements be placed in some amount of proximity to each other while G5b also highlights the need to mitigate the risk of change blindness in the operator. Exactly how to best combat this cognitive phenomenon was not investigated and may be subject to further study.

Data overload occurs when the intake of necessary information outpaces the user's ability to process it (Endsley and Jones, 2014). This was enough of a concern early on in the design process that it was among the first guidelines to be established. Apart from G1, which in its entirety is dedicated to this factor, there is also G3a which points to the need for easy differentiation between notices of different natures. Furthermore, G4b ensures the effective communication of the degree of severity inherent to sudden system events. Furthermore, the entirety of G6 carries a similar message in that the conscious application of colours may further facilitate the internalisation of these factors. This adheres to the three-level model of situation awareness acquisition in that the guidelines listed serve to reduce the complexity inherent to given tasks and interface design overall. G3 in itself tackles the complexity of tasks, such as the Pod startup procedure, which interview participants previously listed as challenging to remember due to the sheer number of steps required. The visualised solution presented to users was very well-received in both iterations. Concern remains whether the amount of visual data and screen real estate presented to the operator may not, in fact, still be a cause of overdue load as it relates to the next risk factor: complexity creep.

Complexity creep is the adverse effect an abundance of features may have on the user's ability to form a mental model of the system itself (Endsley and Jones, 2014). This was certainly observed as a part of the struggle for users in the second usability test. In the first usability test, the first task set before each participant was to identify each camera feed presented to them and what field of view around the truck they corresponded to. All participants managed to correctly identify each view, but only after considerable amounts of deliberation and looking around the screens presented to them. There were no such tasks presented to the participants of the usability test for the second iteration, and several of them reported struggling to comprehend where exactly to look for specific information about the environment around the vehicle. This further relates to the phenomenon of errant mental models where the user is implied to use an incorrect mental model when interpreting incoming information, leading to misunderstandings, mode errors and representational errors (Endsley and Jones, 2014). Apart from struggling to form a mental model of the functionality of the remote operating station, a few participants reported mistaking the rear-view feeds for the side mirrors and vice versa. The guidelines address certain parts of this concern via recommendations such as G1b, G2 and G4 through the mitigation of information overload. However, the thesis work did not investigate the best ways to onboard a user or differentiate feeds from each other. Another suggested functionality for enhancing operator understanding of the camera layout would be a tool through which the operator could select a feed to get further details of the specific field of view, potentially using a top-down or AR-visualisation.

Finally, out-of-the-loop syndrome occurs when the operator of an automated system has their situation awareness undermined by keeping them ignorant of system performance (Endsley and Jones, 2014). This is where the guidelines to bridge the physicality gap between vehicle, system and operator come into play. G4 bridges the gap by looking to quickly inform and reconnect the operator with a sense of control of the truck after an automated system response springs into action. The example situation used in the evaluations was that of the ASB system triggering, and subsequent GUI solutions were well-received by participants of both iterations of user evaluations. G5 seeks to bridge the gap by presenting relevant information that would otherwise be lost by the operator's physical displacement from their vehicle. A solution suggestion such as the status window addresses this issue by supporting the operator in carrying out knowledge-based behaviours, as described by the SRK framework. Knowledge-based behaviours are actions carried out during decision-making in novel situations (Behymer and Flach, 2016). This term can be said to intersect with that of the SAS of working memory - the conscious part of the central executive of working memory used to make strategic assessments (Baddeley et al., 2015). An application such as the status window presents prudent information to these systems, keeping the operator in the loop regarding its performance. Prudent information, in this case, might refer to characteristics such as cargo weight, brake pressure or sensor status. It also acts to bridge the Gulf of Execution inherent to the maintenance of the mental model of situation awareness. The status window supplies information that is subsequently perceived and used by the operator to form an action plan. The key here is that the task of formulating action plans should remain the responsibility of the human operator, mitigating not only the complacency problem of automation but also playing into the strengths of the human-machine team by leaving the task of practical implementation to the machine. This ties back to the four stages of large-scale automation systems presented in earlier chapters.

7.1.2 Cognitive Load

Moving beyond the scope of situation awareness and its relation to the guidelines presented, there is also the matter of cognitive load to consider in order to properly address the second of the two research questions. Observations and results achieved as part of the design process can be tied together with studies in human cognition and contemporary HAI principles, which the following section will expound upon.

Theories surrounding cognitive load largely agree on the concept of finite cognitive resources being the underlying reason for which the adverse effects of the phenomenon may manifest (Sweller et al., 1998). As stated previously, the complacency problem of automation is less of a visual issue and to a larger extent a matter of attention (Metzger and Parasuraman, 2001b). As such, visual solutions may prove insufficient if they do not employ strategies to manage the operator's cognitive resources. G1b presents one way to do so by stating that only the most important visual elements should be presented to the operator when driving. Other guidelines and subsequent branches previously listed in Section 7.1.1 as serving to mitigate data overload can be said to work in a similar manner. As stated previously, due to the pool of cognitive resources being a shared source for three distinct types of cognitive load, the reduction of the extraneous load type remains the most effective way of preventing adverse effects (Sweller et al., 1998). Extraneous cognitive load is exactly what these data-limiting guidelines are dealing with when streamlining and directing the presentation of interface elements.

The formation of mental models is generally stated to cause germane cognitive load due to the high resource cost required to internalise a process into long-term memory (Sweller et al., 1998). As was mentioned previously, the structure of the camera feed caused significant confusion for several participants in the second iteration of user tests - enough so that one admitted to being unable to attend to other things. Meeting user expectations, as mentioned in G2a, is thus an indirect prompt to consider the operator's mental model as they navigate the visual feedback to accomplish various tasks - meeting their expectations by keeping views consistent wherever possible. Meeting user expectations facilitates them forming habitual skills and schema for navigating the interface, increasingly supported by the automatic mode of the attentional controller of working memory and thus freeing up cognitive resources to be used in other matters. One such matter could involve maintenance of the user's mental model of the system, ensuring its accuracy. Mismatching mental models of operation is one of three prevalent causes of bad outcomes in HAI (Sheridan and Parasuraman, 2005), making adherence to G2a a safety matter as well as a usability concern. G3 and G4 reduce additional cognitive load by presenting actionable steps and helping the operator maintain awareness of their place in time. This was represented in the interface using a combination of confirmation notices as well as upcoming and presently requested action prompts.

Maintaining a HCD and user-centred approach to designing a cooperative automated system will need to take human cognition and mental models into consideration. Be it investing resources in mental schemas that will benefit the operator in the long run or making adjustments to better support working memory through interface design, each task needs to be tailored according to the specifications of the relevant memory system. The discussion comes full circle with the consideration that most recommendations lifted from the proposed guidelines boil down to keeping the operator continually informed of system performance while minimising operator cognitive load through the limitation of excise on-screen. It is a trade-off, a balancing act that needs to be investigated further from the point at which these proposed guidelines left off. As mentioned in G5, specificity and ease of overview both have their parts to play in the interface, entirely dependent on the needs of the operator in the moment. It was observed throughout the two user tests how participants intuitively matched the symbols on-screen with the options presented to them on the Stream Deck - an example of interaction design that supports working memory and ended up being highly popular among user test participants.

7.1.3 Design of Input Devices

While the design of input devices connected to the remote operating station was not considered a relevant topic within the scope of this thesis, some interesting learnings eventually coalesced into its own potential guideline following the usability test at the end of the second iteration. As described in earlier chapters, this potential guideline presented a set of usability considerations inspired by comments made by participants from various stages of user testing. The guideline itself ranges from the application of AR elements onto the road to the design of tangible user interfaces and did not go through the same amount of refinement that it may have deserved. Nevertheless, it was deemed interesting enough to warrant mention as part of the results achieved through the Research through Design process.

The guideline itself targets the enhancement of usability through conscious affordances made to the physical controls as well as a suggested improvement to the visualisation of input feedback. The proposed branch (a) suggests that AR elements on-screen are to be recommended to support the operator in their understanding of the practical impact of their input. Examples of this being put into use would be a visual AR projection of the truck's stopping distance or requested steering. The generation of future states belongs to the third level of situation awareness and involves the creation of mental models forecasting events to come - a phenomenon that purposeful design of AR feedback might serve to facilitate. Branch (b) of this proposed guideline instead focuses on more physical matters concerning the tangible affordances of the controls themselves.

A few participants in each usability test had comments regarding the instrumentation, citing difficulties in understanding whether the steering knob was centred enough to have the truck drive in a straight line. They also pondered the significant difference in manual resistance between the throttle and the steering knob, describing it as slightly unnerving. Some expressed concern about incurring strain injuries as a result of keeping the throttle engaged enough to keep the truck in motion. Another user observed how engaging the indicator lights forced them to let go of the steering knob entirely, which is a decidedly unsafe situation in any form of vehicle. In light of these findings, branch (b) of the input design guideline recommends physical affordances to understand the orientation of the controls - effectively removing the need for the operator to seek visual confirmation and lose focus on the road or some other matter. The design of input devices specialised for remote trucking operations is key in the development of a comprehensive user interface, and this minor contribution to the discussion can only be said to scratch the surface of the issue.

7.2 Methodological Discussion

This project had an iterative process where some methods were used in several iterations. One such method was thematic analysis, which was used to analyse the results from the expert interviews and both user evaluations. As for many qualitative research methods, there is a risk of research bias when using thematic analysis (Mackieson et al., 2019). To limit potential biases, both writers took part in the thematic analysis. One person coded half of the data and the other person coded the remaining data. When all data had been coded, a switch was made between the writers, allowing both to go through all the data and highlight parts that they found interesting or important. Both writers also partook in identifying patterns in the created codes. The bias was likely lowered by having two people's opinions during each step of the process, but a potential problem, in this case, is that both writers had similar backgrounds. No one had any prior experience of driving a truck and very little time had been spent observing the remote operating station prior to the first interview. Having similar backgrounds might result in opinions that are quite unified which might result in bias.

Another method where bias is common is interviews which were held on several occasions during the process. The formulation of questions can heavily affect the answer given by the interviewee and it is, therefore, important to consider the phrasing (Goh, 2020). The aim was to not ask any leading questions and a script was used to make sure that the information and questions were phrased the same way for all interviewees. One thing to note is that interviews were held in both Swedish and English and the questions were likely phrased slightly differently between the two languages, which could have affected the answers. The person that did not ask the questions was instead notetaking, and to limit potential bias from the notetaker, all interviews were recorded. These recordings were later if something seemed unclear or if one person remembered something else being said that was not noted down. To limit potential bias even further, it would most likely have been beneficial to transcribe all recordings. However, as this would take a much longer time to do, it might result in less time being left for other parts of the project.

Bias can also be present for evaluators when doing user tests. In the comparative usability testing, performed during the first iteration of the process, the users had to compare two pairs of solutions. These solutions visualised the same thing in different ways and having seen one solution made it easier to understand the other.

It was, therefore, easier for the evaluators to understand the second solution. To limit the bias for a specific version of the solution, the tasks were performed in alternating order. Every other person started with the first version of the solution and the others started with the second version. This probably means that the overall result was more accurate.

The evaluators participating in the first round of usability tests were all male engineers that did not have a class 'C' driver's licence. Their opinions might have been more alike, considering their similar backgrounds. Their backgrounds were also similar to those who developed the original interface and may, therefore, have been more suited to them. The developers mainly consist of male engineers from the same university and it is likely that their culture might have impacted the design. The original setup of the station might be perceived as more "geeky" as a result of the culture shared by the developers. One example of this is the chair used for the station, which resembles a typical gaming chair.

Not having much experience in driving trucks made them relate more to driving cars and several of them mentioned that they wanted the setup to imitate that of a car. However, the experience of driving a car is quite different from driving a truck and that was one of the reasons why it was decided to consider demographics more when recruiting participants for the second usability evaluation. As the user group was limited to employees at Einride, it was difficult to recruit a diverse group. Retrospectively, it had probably been possible to assemble a less homogenous group for the first evaluation and demographic factors were, therefore, considered to a greater extent for the second evaluation. The participants in the second round of user evaluations included both males and females. The participants had different levels of experience driving trucks and were of different national background. The feedback given during this evaluation gave new insights, which might be a result of having a more diverse user group.

Another difference that was noted between the two user tests was the understanding of the views. During the second evaluation, several evaluators asked about what the camera views were showing during the course of the test. This question was, however, not asked during the first evaluation. The reason for this might be the first part of the first user test where the participants were asked to identify each camera view by drawing on a paper. This gave them time to carefully look at all the views and better understand what they were showing. The second evaluation did not have this element as part of the test and the participants were instead asked to start driving earlier. The GUI elements were explained before this test but the camera views were not and the first good look they got of them was when they were actually driving.

Seeing the participants perform the exercise where they had to draw on a paper to mark what field of view the camera views showed inspired the writers to use more hands-on methods that involve users in the future. A method where users got the chance to sketch solutions together or with the writers would likely have been a good method to use at the beginning of the process. Sketching and discussing possible solutions with others could, most likely, have resulted in a broader variety of sketches and ideas. If the process could be redone, this type of method would have been used instead of the sketching method that was used at the beginning of the process. A co-design workshop as part of the Inspiration phase might have injected the project with some additional creativity by getting formative input which could have provided a counterweight to a more implementation-focused narrative.

7.3 Generalisation

The development of the guidelines has focused on the station used for operating an Einride Pod. However, the feedback received in the last evaluation pointed out that it is likely that these guidelines also can be used when designing an interface for remote operation stations for other types of vehicles. Another given suggestion of usage for these guidelines was video games that include driving. This shows that the guidelines can be perceived as quite general even though the process focused on one specific setup and vehicle. The reason for this might be that the basics of driving vehicles are quite similar; the driver has to keep track of the surroundings, status of the vehicle and instruments. The instruments and surroundings may be different for different vehicles, but the basics of it are quite similar. One example that most people can relate to is driving different cars. The driving experience of driving two different car models can be quite different, but if you know how to drive one, you most likely know how to drive others.

7.4 Future Work

The evaluations done in this project have mainly been based on video prototypes where the user could perform a set of predefined actions. The users could, however, not drive the Pod and control how it moved. Using instruments such as the throttle and steering knob did not affect how the Pod behaved in the video and actions related to that were, therefore, difficult to evaluate. If given more time, the next step would be to conduct user evaluations with a mockup of the interface that allowed the users to control the Pod's movement. This type of testing would also allow for the development of guidelines related to the physical instruments.

The evaluation methods used in this project have only produced qualitative data and it would be interesting to continue this work by using a quantitative evaluation method. One such method that could be used is eye-tracking which is a method that can be used to measure the cognitive load for a user that interacts with a system (Zagermann et al., 2016). Qualitative and quantitative methods have different pros and cons and they can, therefore, give different types of insights (Rahman, 2020). The qualitative data that has been generated has been analysed and there is a risk that some insights have been missed. It would, therefore, be interesting to see if quantitative data could support the result or if it would result in new insights.

Future work for this project would also include the investigation of how aural and haptic feedback can be used to aid the user. It was mentioned during several interviews and user evaluations that the users missed sound from the Pod's surroundings as well as sound related to the interface (e.g. sound for the indicator). It was also mentioned that haptic feedback likely would help the user get a better understanding of which type of road surface the Pod is driving on. Both haptic and aural feedback would likely make the driving situation more similar to one where the driver is placed in the vehicle. This might create a more intuitive driving experience and it would be interesting to evaluate how this is best achieved.

7.5 Ethical Considerations

Intelligent transportation systems aim to help users in transport management by increasing safety and productivity while decreasing emissions (Gohar and Nencioni, 2021). These benefits, however, also come with new safety challenges and vulnerabilities. Core functions of such systems have been successfully hacked on several occasions and there are instances where hackers have been able to control a vehicle remotely over the internet. Larger vehicles can cause severe damage if they are operated by people with bad intentions which proves a need for increased safety precautions for remotely controlled and autonomous vehicles.

Cowger and Alfred describe that the decisions an autonomous vehicle makes while driving are decided based on algorithms created by humans (Cowger and Alfred, 2018). They further describe that there are situations in which a collision is impossible to avoid but exactly how the collision will unfold is always impacted by the algorithms. This, according to Cowger and Alfred, implies that there are situations in which a decision has to be made regarding what or who the vehicle will collide with. These decisions have to be decided beforehand when implementing the algorithms which entail that the decisions have to be made by people. It can, however, be very difficult to decide on who or what to hit as that decision can have extensive consequences involving the death of others. Cowger and Alfred mention that the decision made regarding who is selected to be the victim of an accident is most likely an innocent person. They further identify that there are not sufficient laws covering who is liable in those types of situations. When innocent people are affected, it is important that they legally are entitled to compensation for the damages, according to Cowger and Alfred.

There are more aspects to consider when identifying how innocent people can be affected by autonomous or remotely controlled vehicles. A vehicle that can be remotely controlled typically has several cameras mounted on the vehicle to cover different angles of the surroundings to give the operator a good picture of the vehicle's location (Liu et al., 2017b). The videos are live-streamed at the remote driving station but it is most likely possible to save the recordings as well. A person can be captured on video if they are close to the vehicle, no matter their angle to the vehicle, which means that they could be captured on video without them noticing. A similar problem is, according to Rakower, faced regarding the situations captured by Google Street View (Rakower, 2011). The pictures found on Google Street View are available for the general public to view and capturing a person in a situation might, therefore, have tremendous consequences for that person. Rakower concludes that a global threat is posed by Google Street View regarding the right to privacy. As images taken of people can have bad outcomes for them, it is important to consider how videos captured by vehicles will be used to ensure people's privacy.

Designing for autonomous or remotely operated vehicles can also impact the drivers of the vehicles. Lugano mentions that AI-based virtual assistants might result in professions such as truck drivers being outdated and unneeded (Lugano, 2017). The same thing might happen if the interface of remote operation stations would be significantly easier to interact with, especially if the driving is combined with more autonomous driving. Having an interface that allows for quick context switching would make it much easier for one operator to be responsible for several trucks. The trucks could then drive autonomously and only request the operator's attention in certain situations, allowing one operator to oversee the operation of many trucks. This would entail that just a fraction of the current truck drivers would be needed and many would lose their jobs.

Remote operation stations with an interface that makes it easy to operate the vehicle could likely be used to operate many other types of vehicles as well. It could, for example, be used to operate taxis, ships, buses, aeroplanes and many other types of vehicles. This would entail a shift in the job market as many people would lose their jobs due to reduced demand for manual drivers. It would also mean that most vehicles would have to be rebuilt since a driver's cabin would no longer be needed. This space could instead be used to fit more passengers in the vehicle.

When considering the operators, it is important to make sure that the design is inclusive since all drivers are different. Some of the guidelines that have been presented in this project mention the use of colours and it is important to also consider colour blindness when using incorporating colours in the design. To accommodate users with colour blindness, the meaning of data should never be conveyed using only colours (Cooper et al., 2014). Colours were never used as the only indicator of something in the prototypes but it is hard to know how it would have been perceived by someone who is colourblind as no colour blind person took part in the evaluations.

When creating an inclusive design, it is also important to consider people's different sizes and body types to offer an ergonomic working environment. Adjustability is an important factor when creating an ergonomic design that accommodates a wide range of people (Oyewole et al., 2010). The height and angle of the chair that is currently used for the station can be altered, but the table can not be lowered. This creates a situation that is not ergonomic for the shorter users as they have to choose between having a good position relative to the floor or the table. The physical instruments on the table are mounted to the table which makes it impossible to alter their position. This can be particularly difficult for people that are left-handed or people with smaller hands. These two examples clearly show that more work needs to be put into making the station more adjustable.

When considering the ergonomic effects the station can have, there is also the aspect of hours spent sitting down. This new working environment for truck drivers might impact the hours they spend sitting down. Prolonged sitting can have a bad effect on blood circulation and can cause back pain (Oyewole et al., 2010). The regulations in the EU state that a truck driver cannot drive longer than four and a half hours before needing to take a break (Transportstyrelsen, 2021b). Taking a quick break might, however, be easier when sitting at a remote operating station instead of in a truck. A truck driver cannot leave the vehicle and can, therefore, not stand up when on the road. Shorter breaks in the driving, such as when stopping for roadworks, might let the operator stand up for a few minutes. There is, however, a risk that the station causes the driver to sit down more instead. When taking a break while driving a truck, it is common that the driver has to leave the truck to walk to a restroom or restaurant. Being in an office setting does not require the operator to go outside and the distance to the restroom and kitchen is, therefore, most likely shorter.

8

Conclusion

The aim of this Research through Design project was to answer the research questions by producing a set of guidelines that can be used when designing an interface for a remote operation station. The two research questions aimed to be answered remain the following:

What are key interface factors for enhancing the operator's situation awareness in demanding driving situations?

What factors of given interface elements serve to reduce the cognitive load as a part of the interaction between operator and truck in a remote driving setup?

The 3i Model of Design Thinking, with the three core activities Inspiration, Ideation and Implementation, laid the base for the process of this thesis. The second activity, Ideation, was performed in an iterative manner to better incorporate feedback from the users. Two different video prototypes were created during these iterations and used to evaluate the guidelines. The guidelines were updated after each evaluation to better suit the observations and feedback. As the last step, the understanding of the guidelines was evaluated by fellow interaction designers to make sure that the guidelines are interpreted as intended. In the end, the final iteration of the guidelines appear as follows:

1. Design for the limitation of data overload in operator-demanding driving situations.
    (a) Place GUI elements of system information in a well-grouped manner.
    (b) Display only the most highly relevant information when driving.
2. Design for dynamic behaviour in the display of camera feeds for flexibility in use.
    (a) Maintain visual symmetry of camera layout for consistency in use.
    (b) Inform the operator of changes to a degree that corresponds to the impact of said changes.
3. Visualise the required steps for the initiation of action plans.
    (a) Differentiate call-to-action elements from confirmation notices.
    (b) Visualise both present and upcoming steps of the action plan.
4. Visualise root causes for triggered automated responses.
    (a) Use clear GUI representations for specifying the source location in the physical world.

    (b) Design notices with sensitivity to the response type and its degree of severity.
5. Visualise relevant parameters of the truck's condition.
    (a) Use iconographic elements for overview purposes.
    (b) Take appropriate action to counter change blindness in the operator.
    (c) Use great specificity in system feedback regarding warnings.
6. Consider the meaning of colours and be consistent in the use of them.
    (a) Differentiate system messages by their content, such as confirmations, warnings and action steps.
    (b) Use established colour conventions to direct attention and enhance understanding of system messages.

The resulting guidelines highlight important factors to consider when designing an interface for a remote operation station. These guidelines address both how to enhance the operator's situation awareness and reduce the cognitive load. Some guidelines, such as G5, are especially relevant when considering situation awareness and others, such as G3, are more focused on cognitive load. However, most guidelines are relevant for both subjects.

User evaluations were an important part of the process and it would be beneficial to incorporate user evaluations where the user could control an actual Pod in possible future work. It would also be interesting to look into how haptic and aural feedback could be incorporated into the interface to help the operator.

Teachable moments from this thesis work include a sense of fascination with how easily operators lacking any sort of experience driving remotely took to said task. It may be that experience driving a personal car carries over comprehensively enough to benefit in a use context such as the Pod, or perhaps that experience born from operating screen-based devices, in general, played a part. Furthermore, colour proved to be a much more effective communicator than expected. It seems that perhaps the automotive field has the benefit of an established convention of colours thanks to the traffic light metaphor - something that might prove very helpful when designing for the visual language of the interface.

Both chosen research questions were answered through a user-centred RtD process. Although the finished results present more of a discussion surrounding specific key factors as they pertain to the two questions, they nevertheless qualify as relevant findings of a Research through Design process. The resulting range of guidelines highlights the recommended approach to designing for a set of key factors identified as pertinent to the problem space. However, instead of simply pointing out which of these are relevant for the inquiry in question they, in accordance with the project scope shifting with learnings achieved, discuss and prescribe advice for how to work with these selfsame factors to provide the best interface for remote operations.

This outcome is a consequence of the design problem's nature as a wicked problem, revealing itself as more understandable the closer one gets to a solution. It is also a product of the design process adhering to the extensibility criterion of the Research through Design process - that new findings must be based on the contributions produced as part of the work. Aspects such as these did affect how the research questions were answered.

In conclusion, these guidelines aim to support developers, designers and researchers in their efforts to realise or study interface solutions in the remote operation sector. At present, much remains to be explored in the way of established convention and methodology and it is hoped that this thesis work might inform and inspire future endeavours in the field.

References

Ackerman, E. (2015). Ford developing cross country automotive remote control. *IEEE Spectrum*. https://spectrum.ieee.org/ford-developing-cross-country-automotive-remote-control

AGREE Next Steps Consortium. (2017). The agree ii instrument [electronic version]. http://www.agreetrust.org

Al-Khafajiy, M., Kolivand, H., Baker, T., Tully, D., & Waraich, A. (2019). Smart hospital emergency system. *Multimedia Tools and Applications*, *78*(14), 20087–20111.

Anderson, N. (2021, February 18). Comparative usability testing: A key method for actionable design feedback. https://dscout.com/people-nerds/comparative-usability-testing

Anima. (2022, April 26). *Anima: Design to code*. https://www.animaapp.com

Apple. (2022, April 25). *Voice memo on the app store*. https://apps.apple.com/us/app/voice-memos/id1069512134

Aylward, K., Dahlman, J., Nordby, K., & Lundh, M. (2021). Using operational scenarios in a virtual reality enhanced design process. *Education Sciences*, *11*(8). https://doi.org/10.3390/educsci11080448

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2015). *Memory* (2nd). Psychology Press.

Baerkahn, J. (2020, October 31). Update #16: A hole lot to think about. https://graffito.com/update-16-retail-restaurant-preparedness-copy/

Bargas-Avila, J. A., & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: A critical analysis of empirical studies of user experience. https://doi.org/10.1145/1978942.1979336

Behymer, K. J., & Flach, J. M. (2016). From autonomous systems to sociotechnical systems: Designing effective collaborations. *She Ji: The Journal of Design, Economics, and Innovation*, *2*, 105–114. https://doi.org/10.1016/j.sheji.2016.09.001

Bergstrand, F., & Landgren, J. (2009). Information sharing using live video in emergency response work. *International Journal of Emergency Management*, *6*. https://doi.org/10.1504/IJEM.2009.031567

Bergstrand, F., & Landgren, J. (2011). Visual reporting in time-critical work: Exploring video use in emergency response. *Mobile HCI 2011 - 13th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 415–424. https://doi.org/10.1145/2037373.2037436

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chalmers University of Technology. (2021a, June 24). Chalmers maritime simulators. https://www.chalmers.se/en/researchinfrastructure/cms/Pages/default.aspx

Chalmers University of Technology. (2021b, June 24). Full mission bridge simulator. https://www.chalmers.se/en/researchinfrastructure/cms/Pages/Full-Mission-Bridge-Simulator.aspx

Chong, M. K., Whittle, J., Rashid, U., & Ang, C. S. (2015). Cue now, reflect later: A study of delayed reflection of diary events. *Proceeding 15th IFIP TC 13 International Conference*, 367–375. https://doi.org/10.1007/978-3-319-22698-9_24

Chung, G. (2016, August 2). Systematic innovation through ideo's 3i model. https://www.linkedin.com/pulse/systematic-innovation-through-ideos-3i-model-cplp-cssbb-pmp-/

Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About face: The essentials of interaction design* (4th). Wiley Publishing.

Cowger, J., & Alfred, R. (2018). Liability considerations when autonomous vehicles choose the accident victim. *J. High Tech. L.*, *19*, 1.

Dominguez, C. (1994). Can sa be defined? In M. Vidulich, C. Dominguez, E. Vogel, & G. McMillan (Eds.), *Situation awareness: Papers and annotated bibliography* (p. 11). Wright-Patterson Airforce Base.

Dow, S., MacIntyre, B., Lee, J., Oezbek, C., Bolter, J. D., & Gandy, M. (2005). Wizard of oz support throughout an iterative design process. *IEEE Pervasive Computing*, *4*(4), 18–26.

Eden, G., Sharma, S., Roy, D., Joshi, A., Abdelnour Nocera, J., & Rangaswamy, N. (2019). Field trip as method: A rapid fieldwork approach, 1–7. https://doi.org/10.1145/3364183.3364188

Einride. (n.d.-a). Einride pod. https://www.einride.tech/pod

Einride. (n.d.-b). Making earth a better place. https://www.einride.tech/about

Einride. (2021a, November 2). Pod in ny. https://storage.googleapis.com/einride-press-kits/pod-in-ny.zip

Einride. (2021b, December 7). Remote operation and the future of trucking. https://www.einride.tech/insights/remote-operation-and-the-future-of-trucking

Einride. (2021c, November 2). Remote pod operation. https://storage.googleapis.com/einride-press-kits/remote-drive-station.zip

Embrey, D. (2005). Human error understanding human behaviour and error. *Human Reliability Associates*, *1*.

Endsley, M. R., & Jones, D. G. (2014). *Designing for situation awareness: An approach to user-centered design* (2nd). CRC Press.

Endsley, M., & Kiris, E. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*. https://doi.org/10.1518/001872095779064555

Endsley, M., & Robertson, M. (2000). Situation awareness in aircraft maintenance teams. *International Journal of Industrial Ergonomics - INT J IND ERGONOMIC*, *26*, 301–325. https://doi.org/10.1016/S0169-8141(99)00073-6

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32–64.

Fekete, J.-D., van Wijk, J. J., Stasko, J. T., & North, C. (2008). The value of information visualization. In A. Kerren, J. T. Stasko, J.-D. Fekete, & C. North (Eds.), *Information visualization: Human-centered issues and perspectives* (pp. 1–18). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70956-5_1

Figma Incorporated. (2022, March 29). *Figma* (Version 1). https://www.figma.com/community

Fridsma, L., & Gyncild, B. (2017). *Adobe after effects cc classroom in a book (2018 release)*. Adobe Press.

Giacomin, J. (2014). What is human centred design? *The Design Journal*, *17*(4), 606–623. https://doi.org/10.2752/175630614X14056185480186

Gnatzig, S., Chucholowski, F., Tang, T., & Lienkamp, M. (2013). A system design for teleoperated road vehicles. *ICINCO 2013 - Proceedings of the 10th International Conference on Informatics in Control, Automation and Robotics*, *2*.

Goh, W. (2020, May 11). 5 types of biases in user interviews and how to mitigate them. https://uxdesign.cc/5-types-of-biases-in-user-interviews-and-how-to-reduce-them-6b863fb65af1

Gohar, A., & Nencioni, G. (2021). The role of 5g technologies in a smart city: The case for intelligent transportation system. *Sustainability*, *13*(9). https://doi.org/10.3390/su13095188

Gröndahl, T. (2022, March 1). Personal communication [Personal interview].

Hanington, B., & Martin, B. (2019). *Universal methods of design expanded and revised* (1st). Quarto Publishing Group USA.

Hartson, R., & Pyla., P. S. (2012). *The ux book: Processes and guidelines for ensuring a quality*. Elsevier.

Hooijdonk, C. (2022). Explorations in multimodal information presentation.

Hosseini, A., Richthammer, F., & Lienkamp, M. (2016). Predictive haptic feedback for safe lateral control of teleoperated road vehicles in urban areas. *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, 1–7. https://doi.org/10.1109/VTCSpring.2016.7504430

IDEO. (n.d.). Secondary research. https://www.designkit.org/methods/secondary-research

IDEO. (2015). *The field guide to human-centered design* (tech. rep.). Designkit.

IDEO.org. (n.d.). Human-centered design sits at the intersection of empathy and creativity. https://www.ideo.org/tools

ISO. (2019). *Ergonomics of human-system interaction - part 220: Processes for enabling, executing and assessing human-centred design within organizations (iso 9241-220:2019)* (tech. rep.). International Organization for Standardization.

ITS ConGlobal. (n.d.). About its conglobal. https://itsconglobal.com/about/

ITS ConGlobal. (2020, December 10). Its conglobal partners with phantom auto. https://itsconglobal.com/2020/12/10/its-conglobal-partners-with-phantom-auto/

ITS ConGlobal. (2021, November 29). About its conglobal. https://itsconglobal.com/2021/11/29/its-conglobal-announces-remote-trucking-operations-at-the-alliancetexas-mobility-innovation-zone/

James, W. (1890). *The principles of psychology* (2nd).

Kaber, D., & Endsley, M. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, *5*, 113–153. https://doi.org/10.1080/1463922021000054335

Kaber, D. B. (2018). Issues in human–automation interaction modeling: Presumptive aspects of frameworks of types and levels of automation. *Journal of Cognitive Engineering and Decision Making*, *12*(1), 7–24. https://doi.org/10.1177/1555343417737203

Kaufmann, E., Cappé, O., & Garivier, A. (2014). On the complexity of a/b testing. *Conference on Learning Theory*, 461–481.

Liu, R., Kwak, D., Devarakonda, S., Bekris, K., & Iftode, L. (2017a). Investigating remote driving over the lte network. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 264–269. https://doi.org/10.1145/3122986.3123008

Liu, R., Kwak, D., Devarakonda, S., Bekris, K., & Iftode, L. (2017b). Investigating remote driving over the lte network. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 264–269. https://doi.org/10.1145/3122986.3123008

Lugano, G. (2017). Virtual assistants and self-driving cars. *2017 15th International Conference on ITS Telecommunications (ITST)*, 1–5.

Mackieson, P., Shlonsky, A., & Connolly, M. (2019). Increasing rigor and reducing bias in qualitative research: A document analysis of parliamentary debates using applied thematic analysis. *Qualitative Social Work*, *18*(6), 965–980.

Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies*, *55*(4), 587–634. https://doi.org/10.1006/ijhc.2001.0503

Mead, L. (2021). The road to sustainable transport. *IISD Earth Negotiations Bulletin.*

Metzger, U., & Parasuraman, R. (2001a). Automation-related "complacency": Theory, empirical data, and design implications. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*, 463–467. https://doi.org/10.1177/154193120104500442

Metzger, U., & Parasuraman, R. (2001b). The role of the air traffic controller in future air traffic management: An empirical study of active control versus passive monitoring. *Human Factors*, *43*(4), 519–528. https://doi.org/10.1518/001872001775870421

Miro. (2022, March 29). *Miro online whiteboard* (Version 0.7.7). https://miro.com/online-whiteboard/

Norman, D. (1990). The 'problem' with automation: Inappropriate feedback and interaction, not 'over-automation'. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *327*, 585–93. https://doi.org/10.1098/rstb.1990.0101

Olsson, E., & Jansson, A. (2006). Work on the bridge – studies of officers on high-speed ferries. *Behaviour & Information Technology*, *25*(1), 37–64. https://doi.org/10.1080/01449290500066711

Organization, W. H. et al. (2015). *Framework for a public health emergency operations centre*. World Health Organization.

Oyewole, S. A., Haight, J. M., & Freivalds, A. (2010). The ergonomic design of classroom furniture/computer work station for first graders in the elementary school. *International Journal of Industrial Ergonomics*, *40*(4), 437–447.

Phantom Auto. (n.d.). Solution. https://phantom.auto/solution

Plattner, H., Meinel, C., & Leifer, L. (2010). *Design thinking: Understand–improve–apply*. Springer Science & Business Media.

Prinzel, L., & Pope, A. (2000). The double-edged sword of self-efficacy: Implications for automation-induced complacency. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *44*, 107–107. https://doi.org/10.1177/154193120004401338

Rahman, M. S. (2020). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: A literature review.

Rakower, L. H. (2011). Blurred line: Zooming in on google street view and the global right to privacy. *Brook. J. Int'l L.*, *37*, 317.

Reason, J. (2000). Human error: Models and management. *MBJ: British Medical Journal*, *320*(7237), 768–770. https://doi.org/10.1136/bmj.320.7237.768

Riesberg, D. (2016). *Cognition: Exploring the science of the mind* (6th). W.W. Norton & Company, Inc.

Ritchey, T. (2013). Wicked problems. *Acta morphologica generalis*, *2*(1).

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, *4*(2), 155–169.

Ross, J. (2017, March 6). Conducting qualitative, comparative usability testing. https://www.uxmatters.com/mt/archives/2017/03/conducting-qualitative-comparative-usability-testing.php

Rouse, B., William. (1988). Adaptive aiding for human/computer control. *Human Factors*, *30*(4), 431–443. https://doi.org/10.1177/001872088803000405

Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, *100*, 349–363.

SAE International. (2021, May 3). Sae levels of driving automation™ refined for clarity and international audience. https://www.sae.org/blog/sae-j3016-update

Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, *9*(4), 297–323. https://doi.org/10.1080/14639220701561775

Scerbo, M. (1996). Theoretical perspectives on adaptive automation. *Automation and Human Performance: Theory and Applications.*

Sharp, H., Preece, J., & Rogers, Y. (2019). *Interaction design : Beyond human-computer interaction.* (Vol. Fifth edition). Wiley. https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=2097340&site=eds-live&scope=site&authtype=guest&custid=s3911979&groupid=main&profile=eds

Sheridan, T. B., & Parasuraman, R. (2005). Human-automation interaction. *Reviews of Human Factors and Ergonomics*, *1*(1), 89–129. https://doi.org/10.1518/155723405783703082

Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors*, *37*(1), 137–148. https://doi.org/10.1518%2F001872095779049444

Spence, R. (2014). *Information visualization: An introduction* (3rd). Springer Publishing Company, Incorporated.

Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual review of psychology*, *44*(1), 453–495.

Story, M. F. (1998a). Maximising usability: The principles of universal design. *Assistive Technology*, *10*(1), 4–12. https://doi.org/10.1080/10400435.1998.10131955

Story, M. F. (1998b). Maximizing usability: The principles of universal design. *Assistive technology*, *10*(1), 4–12.

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296. https://doi.org/10.1023/A:102219372820

Transportstyrelsen. (2021a, October 1). Driving licence categories. https://www.transportstyrelsen.se/en/road/Driving-licences/im-going-to-take-my-driving-licence/driving-licence-categories/

Transportstyrelsen. (2021b, December 3). Regler om kör- och vilotider. https://www.transportstyrelsen.se/sv/vagtrafik/Yrkestrafik/Kor--och-vilotider/regler-om-kor--och-vilotider/

Tubik. (2018, May 4). Ux practices: 8 handy tips on cta button design. https://uxplanet.org/ux-practices-8-handy-tips-on-cta-button-design-682fdb9c65bc

Wang, S., Yu, B., Ma, Y., Liu, J., & Zhou, W. (2021). Impacts of different driving automation levels on highway geometric design from the perspective of trucks. *Journal of Advanced Transportation*, *1*. https://doi.org/10.1155/2021/5541878

Wiegmann, D., Rich, A., & Zhang, H. (2010). Automated diagnostic aids: The effects of aid reliability on users' trust and reliance. *Theoretical Issues in Ergonomics Science*, *2*, 352–367. https://doi.org/10.1080/14639220110110306

Wu, X., Cao, J., & Douma, F. (2021). The impacts of vehicle automation on transport-disadvantaged people. *Transportation Research Interdisciplinary Perspectives*, *11*. https://doi.org/10.1016/j.trip.2021.100447

Yeager, P. (2020, August 4). What is intermodal shipping...and why should shippers care? https://www.up.com/customers/track-record/tr080420-what-is-intermodal-shipping.htm

Zagermann, J., Pfeil, U., & Reiterer, H. (2016). Measuring cognitive load using eye tracking technology in visual computing. *Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization*, 78–85.

Zhang, Y., & Wildemuth, B. M. (2009). Unstructured interviews. *Applications of social research methods to questions in information and library science*, 222–231.

Zimmerman, J., & Forlizzi, J. (2014). Research through design in hci. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of knowing in hci* (pp. 167–189). Springer. https://doi.org/10.1007/978-1-4939-0378-8_8

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in hci, 493–502. https://doi.org/10.1145/1240624.1240704

A

Interview Questions

1. Can you quickly describe how you have been involved with the Pod? For example, how much have you been driving it?

Remote Driving Questions

2. Have you had previous experience driving a truck physically?
3. If so, what are the main differences - good or bad?
4. What was your experience like when you first started driving remotely?

Situation Awareness

5. Is there an aspect of the driving task that you find more difficult to monitor? For example, surrounding traffic or the speedometer?
6. Is there one that has a tendency to draw unnecessary attention to itself?
7. Have you driven remotely for an extended period of time before?
8. If so, to what extent does fatigue affect your remote operation of the Pod?
9. What aspect of remote driving are you most likely to forget checking? On the GUI or otherwise? Gradient and angle of truck.
10. Can you give an estimate of how often this might happen?
11. Can you recall ever misunderstanding what the driving station was showing you?
12. Can you recall an instance where the system/Pod responded to you in a way that you didn't expect?

Automation

13. When actively driving the Pod, what are the main aspects of the drive/GUI that you monitor?
14. Have you supervised the Pod in autonomous drive mode before?
15. If so, when in a supervisory mode of the Pod, what aspects of the drive are you continuously monitoring?
16. Do you generally feel that the system keeps you appropriately informed of what it's currently doing? In autonomous mode mainly.
17. Why/Why not?

18. When in a supervisory mode of the Pod, what are some tells/events/system messages that would make you resume active control outside of the mandatory validation (such as at pedestrian crossings)?
19. When faced with a novel situation where you find yourself needing to manually operate the Pod, what sort of assistance would you like to receive from the Pod itself?

Engineer-specific Questions

20. Regarding the current GUI: how did you pick which elements to display?
21. How did requirements from Trafikverket play in, if at all?
22. Why does it look the way it does?
23. Is there something from a previous setup that did not work as intended or was misinterpreted by the users?

B

Questions for the First Iteration of User Testing

We'll start each test by putting on a video recording of an Einride Pod out in the field and ask you to complete a few tasks. I will pause the recording while you decide what actions to take to complete said task, and resume it again after you've carried out your plan. We encourage you to use the controls in front of you. You use the lever to apply throttle or brakes, the potentiometer next to it to turn left and right and under the table you'll find a pedal you may need to press at certain times.
You'll be trying out two different design solutions to a single task, of which the test has a total of four. We strongly encourage you to give feedback on both solutions and how they compare to one another.

- Do you have any questions so far?

Intro

The recording is turned on. You'll have a couple of minutes to look through the interface and make sure that you understand what the different icons and elements represent. Feel free to ask questions if anything is unclear.

G1: Camera views

You'll now get to see two different camera layouts. The paper in front of you has the same views printed on it and we'd like for you to draw in which direction you think each camera is pointing.

- *For both setups:*
  - Can you identify each view?
- Which of the two solutions do you think works best for its purpose? Why?

G4: Pod status

The second task is to take stock of the Pod you're operating.

- *For both setups:*
  - What's the current cargo weight at this moment?
  - Are there any active safety warnings?
  - What does it signal?

  - How long has it been since a safety checklist was conducted on this Pod?
  - What do you think about the brake visualisation?
  - Did you lose GPS connection at any point? [If yes] Which one did you lose?
- Which of the two solutions do you think works best for its purpose? Why?

G2: Flow

The next task is to get the Pod ready for driving and get it moving. A safety check has already been conducted. [Play/pause video in time with participants following each step in flow]

- *For both setups:*
  - Get the Pod ready for driving.
- Which of the two solutions do you think works best for its purpose? Why?

G3: ASB

You will now drive the Pod on a straight road for a short while. You may press the Go pedal and apply throttle.

- *For both setups:*
  - What happened?
  - Did you see where the obstacle supposedly was?
  - Make the Pod ready to keep driving.
- Which of the two solutions do you think works best for its purpose? Why?

Final words

- Is there anything else you would like to say?
- Are there any changes you'd like to make to one or several solutions?

C

Questions for the Second Iteration of User Testing

- What did you think about the layout of the camera views?
- *Did he/she use the button for expanding/shrinking the rear-views?*
  - **Yes:**
    - What made you use the button to shrink/expand the rear-view camera?
    - Do you think it ended up being helpful?
  - **No:**
    - Was there a reason for you not using the expand/shrink button for the rear-view cameras?
- How was your experience starting the Pod? Did you understand what to do? Why did it work/not work out?
- What do you think happened when the Pod suddenly stopped?
- Why did it stop?
- Did you see the warning?
  - *If no, point it out.*
- *Did he/she open the status window?*
  - **Yes:**
    - What made you open the Pod status window?
    - What was the issue?
    - How would you have resolved the problem?
  - **No:**
    - What would you have done to investigate the warning?
    - Why did you decide not to do it?
- Do you feel that you could keep track of everything?
- Do you feel that you were in control of the Pod?
- Is there anything you'd like to change about the interface?
- Was there anything you liked in particular?
- What is your take on colours following this?
  - What does yellow mean to you?
  - What does green mean to you?
  - What does blue mean to you?
  - What does red mean to you?
- Any other comments?

D

Article from SVT

An article was published by Sveriges Television (SVT) on the 8th of April 2022 (Norén, 2022). This article includes a video, created by V. Cakic, that showcases an operator at Einride operating a Pod from the station. This video clip was shown to the participants in the last evaluation to give them some context on the subject. The article and video can be found here:

https://www.svt.se/nyheter/ekonomi/svenska-el-uppstickare-utmanar-lastbilsjattarna

References

Norén, A. (2022). Svenska uppstickare tror på eldrivna lastbilar. *Sveriges Television*. https://www.svt.se/nyheter/ekonomi/svenska-el-uppstickare-utmanar-lastbilsjattarna

E

Questions for Evaluation of Guidelines

Introduction

Shows video of the station being used.

Now that you have seen a short video of the station being used, do you remember seeing anything similar before or does it remind you of something?

Here are the current guidelines that we have produced and we will go through them one by one. For each guideline, we'll ask you to state how you interpret it and what you think it means.

Current guidelines:

1. Design for the limitation of data overload in operator-demanding driving situations.
    (a) Place GUI elements of system information in a condensed manner.
    (b) Display only the most highly relevant information when driving.
2. Design for dynamic camera view behaviour for flexibility in use.
    (a) Maintain visual symmetry of camera layout for consistency in use.
    (b) Keep the operator in-the-loop to a degree that corresponds to the extent of the change.
3. Visualise the required steps for the initiation of action plans.
    (a) Differentiate confirmation elements from procedural action steps.
    (b) Visualise both present and subsequent steps.
4. Visualise root causes for triggered automated responses.
    (a) Use clear GUI representations for specifying the source location in the physical world.
    (b) Design notices with sensitivity to the response type and its degree of severity.
5. Visualise relevant parameters of the Pod's present operational status.
    (a) Use iconographic elements for overview purposes.
    (b) Take appropriate action to combat change blindness in the operator.
    (c) Use great specificity in system feedback regarding warnings and events.
6. Consider the meaning of colours and be consistent in the use of them.

   (a) Differentiate system messages by their content, such as confirmations, warnings and action steps.
   (b) Use established colour convention to direct attention and enhance understanding of system messages.

Finishing questions

- Do you feel like something is missing in the guidelines?
- Can you imagine these guidelines being applied to something else?