Evaluation and Design Development of an Advanced Alarm Management System for Submarine Applications

Rebecca Liss and Klara Wiklundh

DEPARTMENT OF DESIGN SCIENCES
FACULTY OF ENGINEERING LTH | LUND UNIVERSITY
2024

MASTER THESIS

Evaluation and Design Development of an Advanced Alarm Management System for Submarine Applications

Rebecca Liss
re0274li-s@student.lu.se

Klara Wiklundh
kl3433wi-s@student.lu.se

September 27, 2024

Master's thesis work carried out at
the Department of Design Science, Faculty of Engineering, Lund University.

Supervisors: Günter Alce, gunter.alce@design.lth.se
Joakim Davidson Truuberg, joakim.truuberg@saabgroup.com
Anders Malmberg, anders.malmberg@saabgroup.com

Examiner: Joakim Eriksson, joakim.eriksson@design.lth.se

Evaluation and Design Development of an Advanced Alarm Management System for Submarine Applications
From a User Centered Perspective

Published by
Department Design Sciences
Faculty of Engineering LTH, Lund University
P.O Box 118, SE-221 00 Lund, Sweden

Subject: Interaction Design MAMM01
Division: Ergonomics and Aerosol Technology
Supervisor: Günter Alce
Co-supervisors: Joakim Davidson Truuberg and Anders Malmberg at Saab Kockums
Examiner: Joakim Eriksson

Abstract

This thesis aims to compare the user experience of an alarm system, which includes a conventional alarm list, with newly developed graphical ways to display alarms. In addition to these visual enhancements, new features requested by users, such as the ability to park alarms to minimise the overwhelming feelings during an alarm flood were integrated. Three iterations with usability tests were conducted, revealing that while the new graphical interfaces could not fully replace the traditional alarm list, they were highly appreciated as complementary tools. Some displays were better for analysing purposes rather than during active alarm floods. The study shows the value of combining traditional and modern approaches to alarm management to achieve a more effective and pleasant user experience.

Keywords: Alarm System, Alarm Flood, Alarm List Display, Design Evaluation, User Experience (UX), User Centered Design (UCD)

Acknowledgements

This master thesis study would not have been possible without Saab Kockums and our supervisor Joakim Davidson Truuberg. You have been super engaged in our project from start to finish and have always been eager to teach us anything we would like to know about Saab or submarines. We are deeply grateful to you for that and for being invited into your fantastic team. A huge thank you also to Anders Malmberg, for being our second supervisor at Saab, for always answering crazy fast and for giving us great advice when we needed it. Anders and Joakim, you gave us the opportunity to travel to Karlskronavarvet and conduct usability tests with Anders's former colleagues from Försvarsmakten (FM). Without the two of you, this and the once in a lifetime experiences that came with it would not have been possible.

To our supervisor at LTH, Günter Alce, thank you so much. We want to thank you for so many things, but above all else: for always giving us a good laugh, answering our calls, emails and being available when we needed your expertise or advice. Also for patiently waiting for us to finally start our master thesis, as we asked you to be our supervisor almost two years in advance. We are so grateful that you said yes!

We would like to thank our test participants from FM for helping us by being a part of this study. Hopefully, we have made a difference that will affect your work and make it easier in the future.

Thank you to our colleagues at Saab Kockums in Malmö for being able to participate in our last usability tests and above all for welcoming us warmly into your team.

Thank you to our family and friends for all the love and support during this master's thesis.

Lund, August 2024
Rebecca Liss and Klara Wiklundh

List of Acronyms and Abbreviations

A26 - Blekinge-class submarine
CENG - Chief Engineer
FM - (Försvarsmakten) The Swedish Armed Forces
FMV - (Försvarsmaktens Materielverk) The Swedish Defence Materiel Administration
Hi-Fi prototype - High-Fidelity Prototype
HMI - Human Machine Interface
Lo-Fi prototype - Low-Fidelity Prototype
MEO - Mechanical Engineer Officer
Mid-Fi prototype - Mid-Fidelity Prototype
P1 - Prototype 1
P2 - Prototype 2
P3 - Prototype 3
SUS - System Usability Scale
UCD - User Centered Design
UI - User Interface
UX - User Experience

Contents

Chapter 1
Introduction

This section aims to give an introduction to this master thesis within interaction design at Saab Kockums. A brief background will be presented as well as the purpose and goal, in addition to the research questions for this study. Furthermore, an introduction to the management of the internal systems used and alarm management systems will be explained. Since Saabs customers were also involved in the study, a brief presentation of them and their role in the puzzle is essential. Lastly, there will be sections dedicated to sustainable development and ethical considerations related to the subject.

1.1 Background

The defence industry has long been one of the most important industries in Sweden and in the world in general. Ever since 1937, when Saab was first established in Sweden, we as a country have enabled ourselves to keep up with the technological advancement when it comes to our defences and keeping our people protected and feeling safe [1]. The underwater craft that is called a submarine is used predominantly within the military nowadays, even though that was not the point with the invention from the start. In submarines, it is common to use control and monitoring systems to observe any inconsistencies or issues in any part of the submarine. This system is typically monitored by a single individual, responsible for ensuring that all issues are addressed. The system generates alarms and warnings for any problems that arise, and when multiple alarms occur simultaneously, it can be extremely overwhelming for the user [2].

1.2 Purpose and Goal

The purpose of this study is to enhance the User Experience (UX) in a monitoring and operating system used by submarine crew members. With a new submarine being developed at Saab, the amount of technical aspects and improvements of the systems have increased.

With that, the interface requirements have also increased, consequently, a UX study was considered necessary. The advanced Human Machine Interface (HMI) system to be tested and further developed in this study is the Ship Control and Monitoring System (SCMS) [3].

This usability study aims to aid Saab Kockums in improving the HMI in the area of alarm displays, primarily for usage in submarines, but could of course be implemented in other vehicles with the same type of operating system. One additional purpose is to help users recognize potential root causes of arising problems and take action when experiencing alarm floods in the system.

The following goals were defined for this master thesis.

- Examine the possibilities of improvements for the alarm handling interface
- Analyse possible solutions to potential problems of alarm management in the submarine system UX design
- Make it easy for the user to understand what actions to pursue when multiple alarms are being triggered simultaneously

1.3 Research Questions

The research focuses on the following key questions:

- In which way can multiple alarms that are being triggered simultaneously be presented in a simple way?
- How to present the alarms for the user to prioritize the severity of the alarms?
- How can the root cause (and correlating alarms) be displayed (during alarm flooding), to reduce cognitive load and give the user an effective way to solve the problem?
- How can the visualization be improved in order to give the user a better understanding of previous alarms and alarm floods?

1.4 Saab Submarine Introduction

The submarine is a complex invention, where everything needed must fit into a relatively small vessel. In the front of the submarine, there is a control room. In this room, there are a few crew members stationed in various roles, responsible for monitoring and operating the systems of navigation, combat control and also the control of internal systems, which is the focus of this study.

1.4.1 Management of Internal Systems

Many of the internal issues are monitored with alarms in an integrated system. The position handling the control of the internal systems is occupied by either the Chief Engineer (CENG) or Mechanical Engineer Officer (MEO). This specific workstation, in the submarine being built at Saab, has two large 26-inch touch screens which are also possible to manage with a mouse. There is also a smaller touch screen below to navigate these two screens. In addition to this, there are physical buttons and panels placed around the screens to complement the system with critical security instruments due to the need for redundancy in the system [4]. A visual representation of the workstation is presented in figure 1.1.

Figure 1.1: The station where SCMS will be used are the two seats closest to the left in the picture.

The interface that is used on these screens for the internal control of all systems on the submarine is called SCMS [5]. This system was developed by Saab and has been in development for 15 years. At the time of this study, the development and improvement of this system is continuously ongoing. The system itself handles all internal systems in the submarine that are to be monitored, excluding combat systems and navigation. For example, this could include temperature management in the fridge onboard or the operation of a Stirling engine [6].

1.4.2 Management of Alarms

In any monitoring system, it is common to have alarms or warnings that indicate when something is wrong. This is of course the case for SCMS, where an upgrade from earlier systems has been made in terms of how many alarms exist in the system. With the new technology development, there was a possibility to be more granular on each part, system or unit involved. For example, an alarm previously would alert that there is something wrong with the engine. In SCMS on the other hand, it could alert that there is a specific part, for example, an open valve, that should be closed, that connects to the engine. This is a huge improvement even though the amount of increasing alarms is both a blessing and a curse. In this case, there could be too much information to handle for a single person when many alarms show up on the screen simultaneously.

1.4.3 Submarine History

According to Beyer, the submarine was invented by William Bourne in the year 1578, though it was never built nor tested until later. The first-ever submarine being built was in 1620 by a Dutch inventor called Cornelis Drebbel. It was not until later in the year 1720 that the first military submarine was built [7]. From this point in history and forward, the submarine as we know it today was born.

1.4.4 Saab Kockums

This study was conducted at Saab Kockums, a part of the Saab Group, which is a number of companies originating from Saab AB. Saab has a total of 22 000 employees as of 2024. All companies have the common goal of *"Keeping people and society safe"* and Saab expresses that feeling and being safe is a human right [1]. In 2014 Saab made an acquisition of Thyssen Krupp Marine Systems AB (TKMS AB), which then became the unit called Saab Kockums. This to strengthen their capabilities to deliver all kinds of naval military systems. The possibility to deliver orders within Sweden in addition to export has increased with this acquisition and the continuous development that has followed [8]. According to Saab, the new submarine seen in figure 1.2, namely the Blekinge-class (A26) is the world's most modern submarine program [9]. However, this requires updated or even brand-new systems with new technological advancements, which Saab Kockums is responsible for delivering. Saab Kockums is also responsible for the maintenance of the Gotland-class in addition to the development of the new Blekinge-class.

Figure 1.2: The Blekinge-class submarine (A26) in the deep sea.

1.4.5 FMV and FM

This study was also carried out in collaboration with Saab's customers in Sweden, Försvarsmaktens Materielverk (FMV), which translates to The Swedish Defence Materiel Administration. FMV have a history of providing Försvarsmakten (FM), called The Swedish Armed Forces, with all the equipment and logistic services needed for each and every mission that they might have. FMV and FM are two important stakeholders for this study, as FMV are Saabs customers and FM will be the future users of the systems developed in the submarine.

1.5 Relevance for Sustainable Development

This study has a strong connection to The Global Goals for Sustainable Development. First and foremost, Goal 16: Peace, justice and strong institutions, pictured in figure 1.3 [10]. Working with any project within Saab, whether it is a master thesis like this one or a 6-10 year-long project, safety and peace are always the keywords laying the foundation for it.

Figure 1.3: Global goals number 8 and 16

The second one that this project affected was Goal 8: Decent work and economic growth, also pictured in figure 1.3 [11]. The submarine work environment alone is an abnormal situation for any type of personnel. Regarding alarm management, operators face significant pressure when using the alarm system, which can lead to stress. In a crisis or when many alarms are triggered, the operator must take quick and correct action to prevent fatal consequences. The current system interface tends to present large lists of alarms that can grow too quickly for the operator to be able to read. This can lead to worse performance and increased stress and cognitive load [12].

The overall work environment can also induce stress. The submarine has no windows at all and most of the time spent onboard will be below the surface with no chance to go out to take a breath of fresh air. With new technology, in this case, the Stirling engines, it is possible for A26 and its crew members to stay submerged for weeks [13]. This could be argued that it is not appropriate or according to decent working conditions. Even though the personnel work in shifts, similar to any other workplace operating around the clock, for example, a hospital, the hours could still differ or easily be extended to overtime. If there is a critical situation, the whole crew must step in and work to handle the situation until it is over. Who knows how long a critical situation can take? Furthermore, there is another problem that is very specific to submarine working environments. Since the operation is done under silence for long periods of time, the ability to have spare time when off the clock is not as free as with a normal job. The options for activities are limited, as there is no room for many activities but to sleep and read a book.

1.6 Ethical Considerations

When it comes to ethical considerations from a user-centered perspective, there are plenty within the naval field especially when it comes to the submarines. Firstly, there is a risk that every crew member takes when boarding the submarine, not only because of situations with enemies and war that could occur with it but also because the ocean itself is a dangerous place. If something were to happen onboard A26, for example, a fire, then the crew could be saved with the URF [14]. However, there is still a larger risk with the job when being far under the surface with no place to go if something happens.

In addition to the physical environment of the submarine, there are the ethical aspects of the technical environment making an impact on the mental state of the individual who works with any alarm system that is a bit more advanced. The general alarm systems in most environments similar to the submarine are often badly designed for the recommended amount to keep the cognitive load on a balanced level. This makes it impossible to execute a job requiring seamless decision-making amid such a diverse and heavy flow of constant new information.

Test persons were involved in the process of testing and developing prototypes for the thesis. In order to perform ethical tests and keep an ethical test environment, an informed consent form was used. The form included information on what the test would entail, the right to abort the test at any time and the option to sign and approve of the data collection. The selection of test participants was performed by a supervisor with connections to FM and FMV. Participating in the tests was completely voluntary.

Chapter 2
Theory and Methodology

The following section aims to provide an overview of the theoretical framework used in this study, focusing on a literature study including related work as well as UX design aspects and methods. It will cover key principles of user-centered design, including design thinking, usability and user experience. Additionally, it will outline core UX methods such as user research, prototyping and usability testing. Furthermore, the section will discuss how these methods contribute to creating effective and satisfactory user experiences.

2.1 Literature Study

Our project started with a Literature study, where we spent time searching for information on alarm systems on databases such as LubSearch and Google Scholar for peer-reviewed sources. We also searched through Saabs' own database for information or valuable studies on the subject but found many more related sources to the subject on the previously mentioned ones. The main terms used when searching were combinations of the following words: *alarm system, alarm flood, design, UX, usability, Interaction design, workload, Nuclear power plant* and *Control room*. Our findings laid the foundation for a literature study with a focus on related work.

To handle information in stressful situations is universally acknowledged as challenging, especially when the information is overwhelming. This is not specified to a single system, it is rather common in monitor and operating systems. In these systems, alarms are commonly used as notifications to indicate potential issues. The problem with these kinds of systems is the amount of alarms exceeding the users' information processing capability. According to Wu and Li, it lowers the users' ability to detect issues when there are too many alarms at the same time. It also hinders the user from taking necessary actions to deal with the origin of the alarm [15]. Wei et.al. agrees and states that an alarm flood significantly challenges the operators' ability to effectively monitor the system [16]. The definition of an alarm flood is according to Wu and Li, "A condition during which the alarm rate is greater than the operator/user can effectively manage" [15]. In other words, an alarm flood is when an overwhelming amount of alarms are triggered simultaneously.

Wei et.al also mention difficulties in understanding and interpreting the information during these stressful situations, where rapid comprehension and response are crucial [16]. Laberge et al. confirm this in their study on alarm management in process plants, stating that "when managing alarms these alarm floods is one of the most significant challenges facing process plant operators". What makes it so difficult to handle these situations is the continuous flood of alarms showing up on the screen. The authors further describe a typical alarm flood as a situation where hundreds of alarms appear within the first five minutes, followed by dozens more per minute over the next few hours [17].

According to Li et.al. the optimal manageable number of alarms per hour for any human mind is six alarms. In other words, one alarm per 10 minutes if they were to be spread out equally. However, in their study on a ship monitoring system, this number of alarms was reportedly exceeded by a lot [18]. However, Laberge et.al state that the response limit is closer to 11 alarms in a ten-minute period [17]. Wu and Li on the other hand published two recommendations for an average amount of alarms, depending on the situation. For normal operations the authors recommend one alarm per 10 minutes and during emergencies, no more than 10 alarms per 10 minutes [15].

There is no known or clear consensus on the frequency of occurring alarms that will allow a stable cognitive ability for the user. Wu and Li further explain that research on human performance indicates that the conditions of alarm flooding surpass the processing and response capabilities of a typical user [15].

In their study on alarm system design , Wu and Li identify three key functions that these alarm systems should incorporate. First, the system should notify the user of any system deviations, while also conveying the criticality and the root cause of the issue. Second, the system should guide the user's initial response to the deviation and confirm whether the action was effective in terms of correcting the issue. Finally, the system should assist in the post-analysis of any potential incidents. They also mention that the system's functions and roles differ between normal and emergency situations. During normal operations, the primary challenge is the lack of significant, informative data, whereas, in emergencies, the key issue is alarm flooding [15].

Given how extensively alarm systems are used today, what is it that makes the designs so poorly adapted to alarm management? In fact, it is overwhelming for the operator when an alarm flood occurs. In monitor and control systems, the design tends to not be in favour of the user when it comes to alarm floods. The usability problems are often the same in these systems since most of them also use a standardised chronologically sorted list when displaying the alarms [19]. In their study on alarm systems, Laberge et al. argue that there are three problems with these lists during alarm floods and these are presented in the list [17] :

1. Alarms appear and move through the list faster than the user can read them.
2. The alarm list quickly exceeds the screen capacity. This forces the user to either scroll or change pages to view all alarms, making it difficult to see everything at once. This consumes valuable time for the user and time is critical during an alarm flood.
3. Even high-priority alarms, often displayed in a red colour according to Berg et.al. [19], can be pushed off the list as new alarms appear. This increases the risk of missing critical alarms if the user is monitoring another display simultaneously [17].

Furthermore, Wu and Li explain some of the most crucial but general design problems that have come with the digitisation of alarm systems [15].

- **The keyhole effect** occurs when information is spread across multiple screens, each displaying only a fraction of the process of the system. The users' limited focus on the screens creates the illusion of viewing the system through a keyhole, increasing the risk of missing the overall picture.
- **Interface management issues** arise in large and complex systems with numerous screens to navigate. Users may become disoriented and struggle to return to previous screens, reducing their effectiveness in alarm management.
- **Disappearance of visual patterns** makes it harder to quickly identify system states when represented by digital numbers or text, compared to analogue controls intentionally placed in an operating room.
- **Decrement of teamwork transparency and communication** in the advanced alarm system's control room. The layout is less conducive to teamwork than the previous setup. Screen-focused work decreases visibility, hinders perception of team members' actions and minimizes communication.
- **Increased requirements of operators' knowledge and skill** are necessary due to the system's complexity and reduced teamwork transparency during normal operations.

In the best scenario, there would be solutions to avoid the occurrence of alarm floods completely. However, Laberge et.al. argue that trying to remove possible alarm floods entirely in monitor and control systems is impossible, stating that "even the best efforts do not eliminate the occurrence of alarm flooding". Besides trying to improve the design to make the user able to manage these alarm floods more easily, there is another aspect that the authors mention, namely for the user to train their ability and response strategy to critical situations such as alarm floods [17]. Although alarm management only accounts for about 10 per cent of the user's time, it is arguably the most important part of their work [15]. Hence, designing a usable and accessible system is vital.

2.1.1 Related Work

Several studies have attempted to enhance the usability of various alarm systems. Many of them have already been mentioned earlier in the literature study.

Alarm Tracker Summary Display

Among them is the solution presented by Laberge et.al. where the alarm list is compared to what the authors refer to as an "Alarm Tracker Summary Display". This solution involves creating a timeline-based alarm display, where the systems in the power plant are organized into separate timelines. The display is presented in figure 2.1. The created overview was a summary from the last 15 minutes. The results were contrary to the hypothesis, namely that orienting to alarms and performing a correct response was worse with the new summary display, even though the participants had been training for an average of four weeks with the new display. The participants even responded faster using the usual standard list instead. The expected increased effectiveness in the evaluation of alarms did not exceed any expectations either, as it was the same as the standard list. In this study they also mention the action of acknowledging alarms, meaning that the user has noted and addressed the issue. The authors mention several ways to implement the action of acknowledging alarms. One method involves using a touch screen where the user can mark and drag all alarms before acknowledging, while another allows the user to right-click with a mouse. They also mention that acknowledging alarms by equipment area or in a cluster was one of the most important aspects in their design [17].

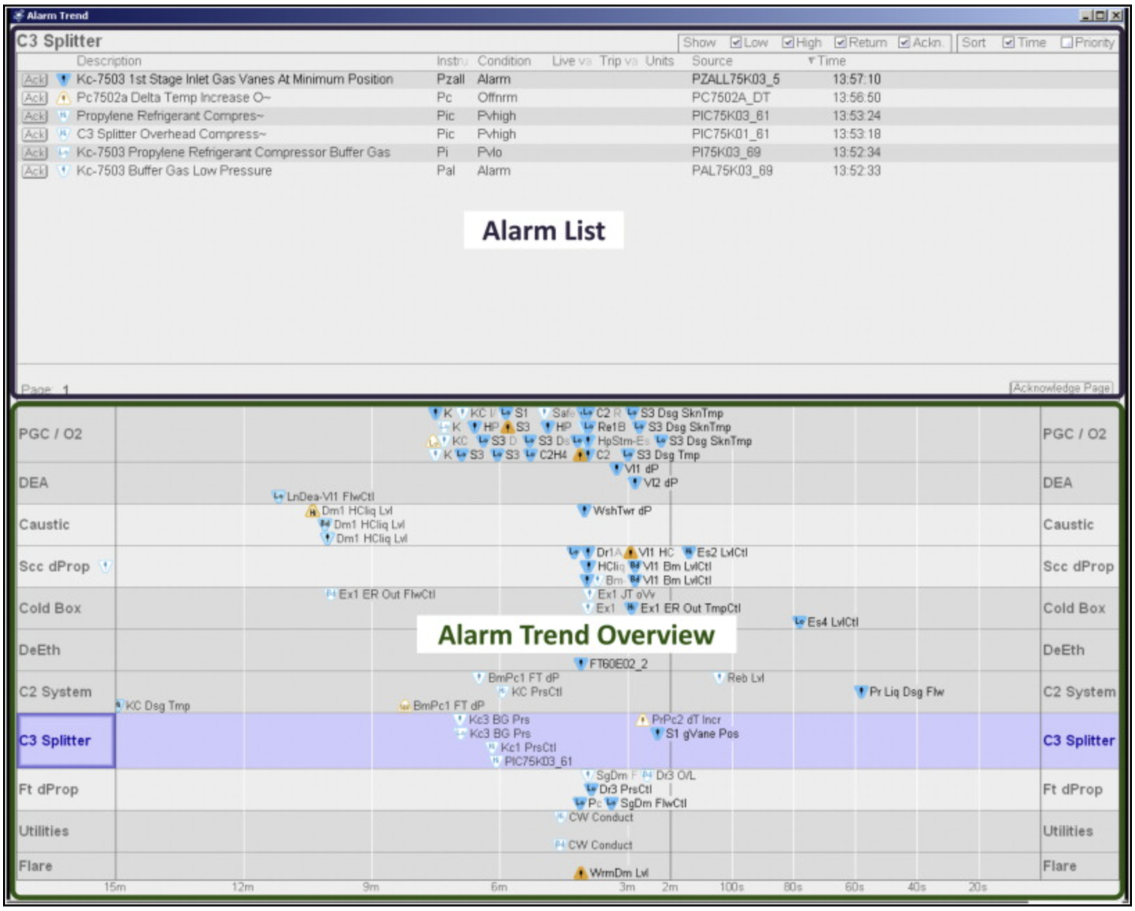

Figure 2.1: Picture of the new alarm summary display that Laberge et.al. evaluated in their study on alarm management. The alarms are generated solely for the illustration purpose [17].

Design Implementation of Ship Alarm Management System

A more recent study on alarm management onboard ships was published by authors Li et.al. in 2023. This specific study displays the problem of alarm flooding on a ship, where there is one responsible operator, namely the Chief Engineer. The three main reasons stated for alarm flooding to occur on the ship were described as following [18]:

- "Including many original alarms" - This means alarms on every individual part of a system that can be tracked.

- "Including many repeated alarms" - Repeated alarms are in this case described as many repeated alarms appearing after a single device defection. It also states the important fact that many liquid alarms being active can occur while the rough seaway is affecting the boat to askew.

- "Unsuitable alarm classification" - It states that it is important to have a proper classification system, to not mix up alarms that have different levels of severity, as this can cause confusion. It is difficult to gather the right information to make a decision in these situations.

A Review of Alarm System Design for Advanced Control Rooms of Nuclear Power Plants

Another study by Wu and Li, where they have conducted a full literature study on the subject including findings from the one by Laberge et.al mentioned above. Their approach was to review alarm system design, especially alarm flooding, in Nuclear power plant control rooms. One of the most noteworthy issues with alarm visualization in the nuclear industry is that interfaces often replicate traditional mimic-based systems rather than exploring new innovative solutions. Therefore, the authors highlight some of the new alarm management designs, including the Alarm Tracker Summary Display from Laberge et.al., in addition to one in particular where the designers have modified the alarm list to instead show an event-driven timeline display. The alarms were presented by type and by temporal sequence [15]. Since alarm flooding is a common problem in advanced alarm systems many researchers have studied possible solutions to lessen the negative effects. Some of the most common ways according to Wu and Li, are summarised:

- **Alarm filtering** is defined as "eliminating the alarms that were determined to be less important, irrelevant, or otherwise unnecessary by processing techniques and making them not available to operators". It helps reduce the total amount of alarms, however, the users express disliking feelings towards this and they have two reasons for it. One is that they will always worry about important alarms being deleted and the other is that useful information might be removed from the system. Users would rather face more alarms than miss a single important alarm.

- **Alarm suppression** allows users to request the display of irrelevant or unnecessary alarms while keeping them hidden by default. This method has two drawbacks: the risk of suppressing an important, unexpected alarm and the added workload of retrieving suppressed data. Despite this, it is more effective than filtering and helps reduce the user's mental load.

- **Alarm prioritizing** involves ranking alarms based on urgency and safety, often using colour-coding to indicate significance. Prioritization can be static for simple systems or dynamic for complex ones. The recommended amount of priority levels ranges from three to six between sources and some declare a consensus of no more than four.

- **Group alarms** combine multiple low-level alarms into a single high-level alarm, highlighting the most severe condition and reducing alarm clutter. Operators can click to view detailed alarms. These alarms can be created using event-based, state-based methods, or the Abstract Hierarchy method, each with distinct advantages.

- **Dynamic alarm threshold** method builds on the static approach by factoring in multi-variable interactions and setting thresholds using historical data. It adapts alarm limits to match different operational conditions, cutting down on nuisance alarms and giving operators more relevant alerts.

- **Pre-alarms** serve as early warnings to help operators take preventative actions before critical failures occur, reducing the severity of failures. While they can reduce the frequency of critical alarms, they may also increase the total number of alarms, potentially leading to earlier alarm floods if not carefully managed.

2.2 Submarine classes and terminology

Submarines are complex crafts and people working on them use specific terms and language. Some of the background and classifications of Swedish submarines and the terms will be explained in the following sections.

2.2.1 Internal vs External Battle

Unplanned events on the submarine can be divided into two different categories, internal and external battle, as it is called within FM and also Saab. The internal battles include all the internal problems that could occur, such as a fire starting in the kitchen. The external battles, on the other hand, include everything happening and affecting the submarine from the outside. For example, if a submarine is hit and there is a breach where water is coming in, there would be an external battle to handle when avoiding getting hit yet again and possible combat. This is in addition to an internal battle to stop the water ingress into the submarine [4].

2.2.2 Saab Submarine Classes

The submarines being used in the Swedish military today are part of the Södermanland-class and Gotland-class. These have been used for almost 30 years and will continue to be used until the next generation of submarines is ready to be launched. Maintenance and upgrades have been made on the Gotland-class submarines and the new Blekinge-class (A26) will be added to the naval force and eventually replace the Södermanland-class [20]. Within Saab, the new submarine Blekinge-class is therefore being developed with a planned deliverable of two units for the Swedish Navy. The two of them are planned to be launched in 2027 and 2028 [21]. This is the new generation of submarines within Saab AB. Which is where this study was formed, within the new Blekinge-class submarine development. The two classes can be seen in figure 2.2. If nothing else is mentioned, A26 will be referred to whenever submarines are mentioned from now on.

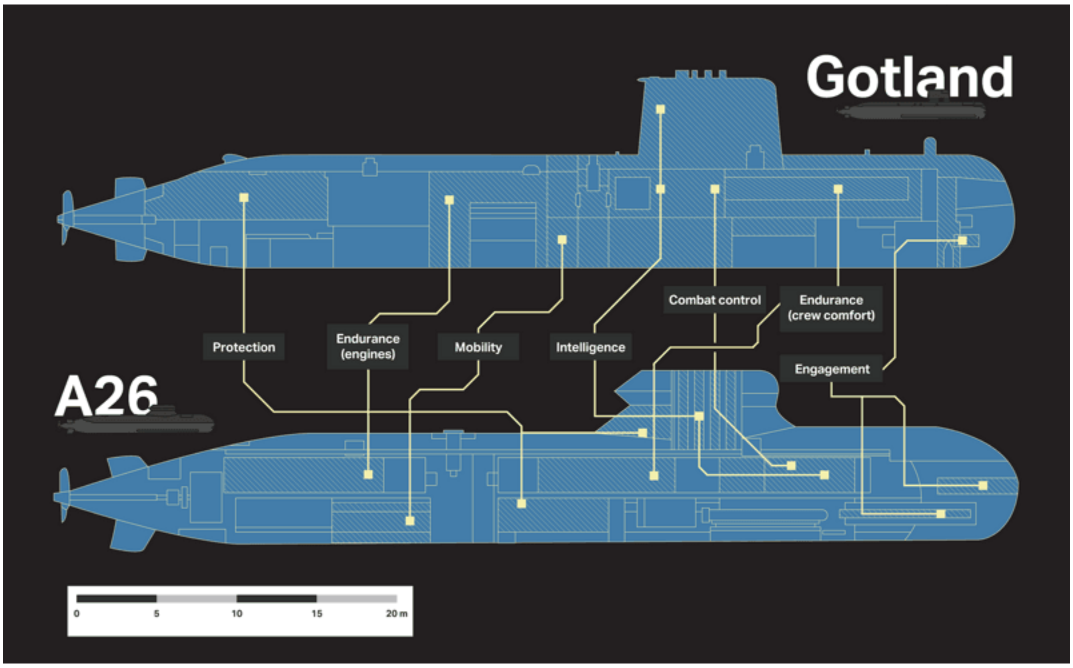

Figure 2.2: Pictures of the Gotland-class and Blekinge-class (A26) submarines in comparison.

2.3 Interaction Design Theories

Within interaction design, according to Preece et.al., there are four key steps in the Design process that any designer should be aware of [22]. The process is presented in figure 2.3 and the four activities are:

- Establish Requirements
- Design Alternatives
- Prototyping
- Evaluating

Within these four activities, there are plenty of steps, methods and terms that are useful to designers. The remainder of this chapter will describe some theory behind these steps and the concepts and methods needed to perform an interaction design process on any interface, system or product.

2.3.1 Design Thinking

Design thinking is a creative and analytic process to create innovative solutions to challenging problems. The process aims to give what the users need (desirable), be technologically possible (feasible) and generate profits (viable). Design thinking consists of five steps: *empathize, define, ideate, prototype* and *test*. The steps, also called phases, are not always sequential which can be seen in figure 2.4.

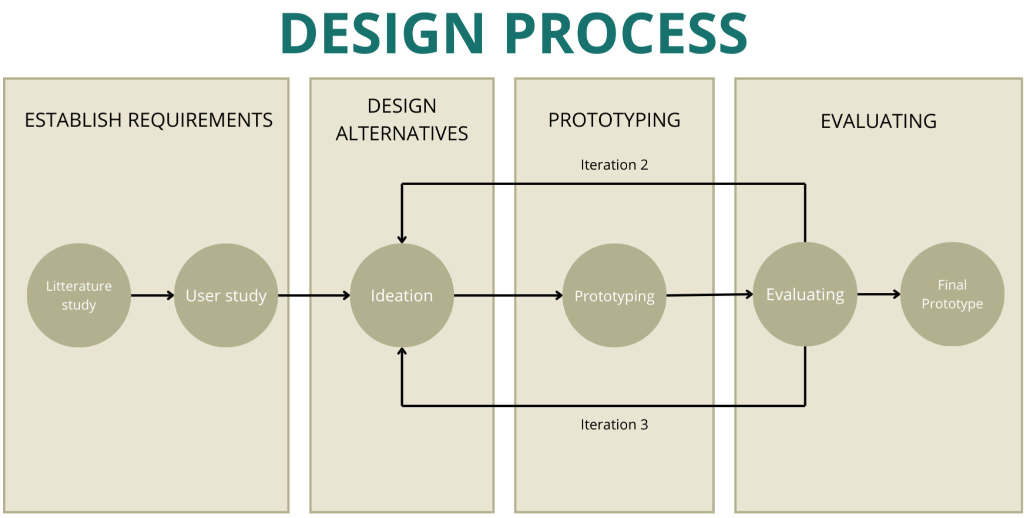

Figure 2.3: Interaction Design Process visualized

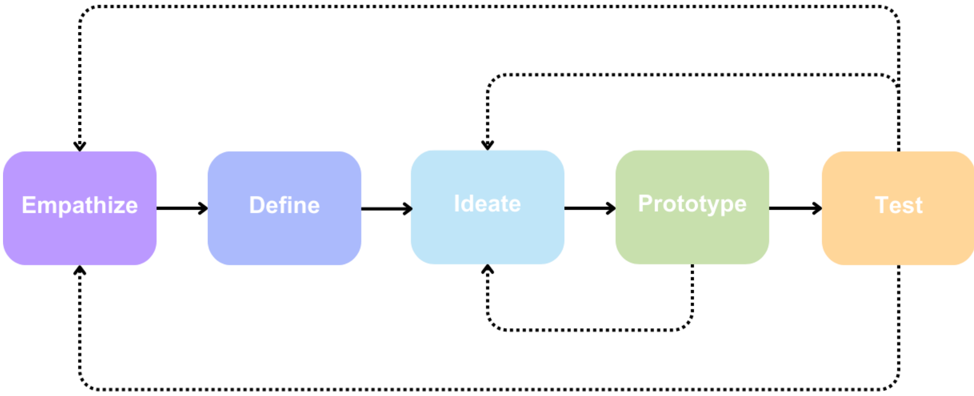

Figure 2.4: Design thinking visualized

Empathize encompasses the aim of understanding the user and the challenges they face. Designers must set aside their assumptions and preconceived notions to gain true insight into the user's needs. *Define* involves accumulating all the information, analysing it and formulating concrete definitions of the core problems. The *Ideate* phase focuses on generating innovative solutions and ideas to address the problems previously defined, often through methods such as brainstorming. *Prototype* refers to implementing the generated ideas. This phase is still experimental, so the prototypes should be simplified, scaled-down versions of the product or solution. Finally, the *Test* phase evaluates the prototypes, with results providing valuable insights. This builds the foundation for a new iteration, which can start at any point in the design thinking process.

Since design thinking is solution-focused it has the possibility to solve "wicked problems", meaning problems that are hard to define and solve [23].

2.3.2 User Centered Design

User Centered Design (UCD) is a term that explains the process of creating and designing a product in an iterative design process with a continuous focus on the user at every step. According to ISO standard 9241-210, User-Centered Design follows six principles [24]:

- *"The design is based upon an explicit understanding of users, tasks and environments."*
  This includes studying and identifying users in order to be cognizant of their needs, wants and demands. To what extent is the product usable and accessible for users?

- *"Users are involved throughout design and development."*
  By focusing on the user in each step of the development, the product is secure from misguided errors caused by the designers or the development team. Finding flaws and bugs early in the design process saves resources later. This minimizes the need to re-design key features or re-structuring a program when the product is close to finished or close to release.

- *"The design is driven and refined by user-centered evaluation."*
  User feedback from evaluations or testing is crucial to keeping the design process on track and ensuring that the focus remains on what the user truly needs and wants.

- *"The process is iterative."*
  An iterative design process consists of four steps: understand the problem, specify the user requirements, design a solution and evaluate, agreeing with Preece et.al. [22]. The cycle is repeated until the result is satisfactory. Due to the complexity of HMI, creating a perfect first version is impossible. However, through iterative development, the product can be refined to better meet user demands.

- *"The design addresses the whole user experience."*
  The UX is how a user interacts with and experiences a product or system. It includes efficiency, utility and "ease of use". However, usability does not solely refer to making products easy to use, it also takes into account the users' goals, the experience of using the product and feelings such as satisfaction.

- *"The design team includes multidisciplinary skills and perspectives."*
  By having a team built up of members with different perspectives and expertise, team-work and interaction can lead to creative and innovative ideas.

User Centered Design may seem like a resource-heavy process, but it pays off by creating a product that users are more likely to want for three reasons. First, reducing the risk of discovering bugs, faults or errors late in the process, which is both difficult and expensive to fix. Second, it creates empathy between the user and the designer. Last, this results in sustainable products in terms of both diversity and usability [25].

2.3.3 Usability and User Experience

Usability is a quality attribute that measures how easy a user interface (UI) is to use. The term usability is defined by five components [26]:

- Learnability - How easy is it to learn?
- Efficiency - How quickly can a user perform a task?
- Memorability - Is it easy to perform a task again after some time?
- Errors - How severe are errors? How many errors does the user make? How do they recover from them?
- Satisfaction - Is it satisfactory to use the system?

Usability is important because it helps the user and makes the product better. As Nielsen states about internal design projects, "... think of doubling usability as cutting training budgets in half and doubling the number of transactions employees perform per hour" [26]. Although the thesis technically does not focus on internal design projects, this quote still applies since the focus of the thesis is how the system will be used. Usability can be improved by usability testing or user testing.

The definition of User Experience according to Nielsen and Norman Group: User Experience (UX) encompasses all aspects of the end-user's interaction with the company, its services and its products. For a good UX, the experience must fulfil the user's needs and be enjoyable [27].

2.4 Design Guidelines

To improve the user experience there are certain design guidelines on how to apply design principles. Two of the guidelines are Norman's design principles and Jakob's 10 Usability Heuristics. Companies usually have their own design guidelines to ensure consistency in all products and systems. This often includes style, colours, layout, fonts, padding and much more [28].

2.4.1 Normans Design Principles

Donald Norman is a key figure in the space of human-centered design. He has defined several principles in his books and some of the most important are: affordances, constraints, conceptual models, feedback, mapping and signifier [29].

Affordances: "A term used to refer to an attribute of an object that allows people to know how to use it." In other words, it can be seen as a clue. The user should know what actions are possible with a single look at the object.

Constraints: "Refers to determining ways of restricting the kind of user interaction that can take place at a given moment." By giving the user limits on a product, the possible actions can become much clearer.

Conceptual models: "Mental simulations of devices that enable users to judge the means of operation and possible uses." It is a representation of a system based on concepts which are used to help the user understand.

Feedback: "Feedback is about sending back information about what action has been done and what has been accomplished, allowing the person to continue with the activity." It is a way to communicate to the user whether their action towards a goal was successful.

Mapping: "This refers to the relationship between controls and their effects in the world." It is a bridge between a user's mental model and the object they interact with.

Signifier: "... any perceivable indicator that communicates appropriate behaviour to a person." It communicates to the user where an action should take place. A few examples of signifiers are presented in figure 2.5.

Figure 2.5: Examples of symbols that can be used as signifiers.

An important comment on Normans' principles is that the implementation of the terms greatly depends on cultural context. Colours, symbols and placements can be perceived differently in different countries or regions. An example of this was when Amazon used the icon of a looking-glass, which was meant to signify "search" but was mistaken for a ping-pong racket in India [30].

2.4.2 Jakobs 10 Usability Heuristics for User Interface Design

Jakob Neilsen´s 10 Usability Heuristics for User Interface Design is a set of guidelines for creating a good system and user experience. They were created in 1994 and even though design and technology have evolved since then, the usability heuristics are still the same. The 10 usability heuristics are [31]:

- Visibility of System Status, which is to keep the user informed, with feedback, of what is happening in the system.
- Match Between the System and the Real World. The system should be easy to understand, "speaking" the users' language and follow real-world conventions (mapping).
- User Control and Freedom. Users need a way to undo or abort an operation if they make mistakes.
- Consistency and Standards. The design of the system should follow industry conventions so as to not cause confusion and increase users' mental load. The design should also be consistent throughout the system.
- Error Prevention. The system should be designed to prevent slips or mistakes. Slips can be avoided by providing constraints and default modes and mistakes can be prevented by supporting undo-functions and removing memory burdens.
- Recognition Rather than Recall. Important information should be visible and easy to remember to minimize the users' mental load. The user should not rely on memorization of the system.

- Flexibility and Efficiency of Use. The system should allow users to tailor their experience for themselves, like having shortcuts for experienced users.
- Aesthetic and Minimalist Design. There should not be irrelevant information presented in interfaces, this only leads to clutter and hides the relevant information.
- Help Users Recognize, Diagnose and Recover from Errors. The error message should be easy to notice and understand in addition to suggesting a solution.
- Help and Documentation. Provide help and documentation to guide the user if all nine factors above fail.

2.4.3 Saab UI Guidelines

In the Saab HMI/UI guidelines there are definitions and specifications on the design in SCMS. This includes spacing, fonts and sizes for the UI. In addition to this, the language and specific words and their abbreviations are specified. This HMI/UI guide is unfortunately not available to the public, hence we will not share anything more about this in particular. The guidelines at Saab were made to secure a consistent design and overall system recognition no matter which part of the system is being used. This is to improve the experience for the user and make the systems easier to use and learn [2].

2.5 User Study

As mentioned in UCD the user should always be in the focus when designing a product. Design thinking also acknowledges that empathizing with the user is important. The first step in the design process, according to Preece et. al., is to establish requirements and this is why a user study is imperative to this project/report [22]. In order to work with a user throughout the process the users need to be identified and information/data of the users must be collected [32].

2.5.1 Types of Users

Determining the different types of users is important in a UCD process since they are the main focus. There are three different kinds of user categories: the primary (users that frequently use the system), secondary (users that occasionally or indirectly use the system) and tertiary (users that would be affected by the system in some way). The primary and secondary users often want different things and it is important not to overlook the secondary users even though they are not the main focus [33].

2.5.2 Data Gathering

Data gathering can be conducted in different ways and through different methods. Primary research methods are the most effective way to collect user data and they include methods like interviews and focus groups. In short, primary research means interacting directly with the user about their challenges and constraints [34].

Interviews

Interviews are an essential research method that involves directly interacting with participants to gather firsthand information about their attitudes, experiences and opinions.

The format of interviews can be structured, with a script of questions, or unstructured, encouraging follow-up questions and detours around the topic at hand, or something in between. They both have different areas of usefulness, an unstructured interview can make the participant feel more comfortable and therefore offer up more information, but the conversation can also get off-topic and lead to gaps in the information wanted. The freedom and flexibility of conversation lends itself to exploring topics and brainstorming. The structured interviews may feel impersonal but is usually easier to predict and keep on time. This format is great for extracting specific information that is needed.

Key informant interviews, or expert interviews as it will be referred to in the report, are interviews with people who have expertise or specialized knowledge [35].

Focus Group

Focus groups consist of a group of users and a moderator to guide conversation and it is a qualitative research method. The unique property of focus groups, compared to interviews, is the group dynamics that form between the participants. When participants see each other as peers they are more likely to share opinions and experiences. They can build upon others' ideas, discuss freely and reach an exploratory phase, moving beyond the fundamental ideas [35].

Due to its unstructured nature, the group discussion can veer off topic. If that happens, it is important for the moderator to steer the conversation back on track. Another problem that can arise is that dominant participants who might have strong opinions are loud and/or stubborn. There is a risk of such participants persuading or silencing other participants, in that case, it is also the moderators' responsibility to step in [32].

2.6 Analysing Data

All the data gathered during the user study and literature study needs to be organized and analysed in order to be of use. Some of the methods used in this report are personas and affinity diagrams.

2.6.1 Personas

A persona is a general description (of features, traits and goals) of a typical user of a product. Personas are used to give personal insight into how a user might interact with a product and can give clues on preferences and dislikes. Since focusing on users is a key aspect of USD, personas are an important step in the design process. By creating several (3-5) personas a product can be designed to fulfil a "genre of users" needs and wants instead of trying to fulfil everyone's needs and wants and thereby failing to please anyone. Personas also help "humanize" the user by giving them names, a portrait and personalities. The persona is later used to test the scenario in which the prototype or product is used or tested [35].

A persona contains demographic information, like name, age, gender, living and family situations. They also contain personality traits, preferences, needs, wants and fears.

The steps in creating a persona are:

- Collect information about the users. The information the personas are based on must be true/accurate/realistic to be helpful in later steps.
- Organize and analyse the user information. Are there common features and traits among the studied users? Is there a few "typical" users?
- Create personas (3-5 depending situation). Write 1-2 pages of information and portrait per persona.
- Validate personas. Be sure to validate/confirm with users, or experts in the field, that the personas are realistic/true otherwise the personas will not be useful [32].

2.6.2 Affinity Diagram

Affinity diagramming is a process used to cluster data in meaningful ways. The data in question can be observations, information or insights from research. Affinity diagramming is done by writing down the data on Post-it notes and the team groups the Post-its based on similar attributes. This process helps visualize data, visualize data in context to other data and help draw connections between data groups. The groups can unearth trends and tendencies in the data set [35].

2.7 Ideation Methods

The ideation phase of the design process is where designers and developers make sense of the analysed data and create ideas to make new designs or solve problems. There are many methods that could be used in the ideation stage and two of them are brainstorming and bodystorming [34].

2.7.1 Brainstorming

Brainstorming is a process where a group of people try to generate possible solutions and/or ideas for a problem.

Key aspects of brainstorming are to value quantity over quality, postpone criticism, stimulate creativity and build on other participants' ideas. This is to generate as many ideas as possible and to encourage the participants to think outside the box. By building on other participants' ideas new solutions and new iterations/versions of the ideas can be created and new solutions can be revealed.

The first step of brainstorming is to formulate the problem that is to be solved. Then all participants get to write down or draw ideas on Post-it notes for a set amount of time. When the time is up all members pass the Post-it notes to their neighbour and they get to add or expand on the idea. This step is repeated. At the end all ideas are put up on a wall/whiteboard etc. and the ideas are organized and discussed [32].

2.7.2 Bodystorming

Bodystorming is a physical way to brainstorm. It is a qualitative and behavioural method to generate innovative ideas. Bodystorming is a combination of simulation and role-playing in order to experience what a real user might experience. By using props or a simulated environment active ideation can lead to new insights and solutions. The first-hand experience of the researcher invokes empathy for the users and leads to a connection between user and researcher as well as reflection and heightened awareness of the UX [32].

The props do not need to be replicas of what a user would use but are similar enough to help with immersion. Scenarios and personas can also be helpful in a bodystorming session [35].

2.8 Conceptual Design

Conceptual design is part of the second step in the design process, after establishing requirements, namely designing alternatives. The point is to create some kind of visual idea of the product, before deciding anything about colours, typography and themes. The conceptual design relies on a well-conducted user study, where the foundation and pillars of the future design are supposed to take form. Usually when doing conceptual design, methods such as making a user flow is common [22].

2.8.1 User Flow

A user flow is defined as: "a set of interactions that describe the typical or ideal set of steps needed to accomplish a common task performed with a product". User flows are tools for designers to understand and optimize the users' experience. It is often structured around a task/goal the user wants to perform and focuses on the specific and granular interactions and/or steps. User flows can be represented as flow charts, diagrams or wireflows (simple visuals over the views and arrows to represent navigation) [36].

2.9 Prototyping

To be able to test the usability of a new system it is necessary to create a prototype of the design. This part of the design process is called prototyping. Since testing can and should be initialised early on, the prototype does not have to be finished or even possible to interact with in a usual manner when first tested. In fact, testing concepts early in the design process allows designers to make adjustments based on user behaviour more effectively. As the design progresses, it becomes increasingly difficult to implement changes due to a stronger attachment to the existing prototype. According to Walker et al., the level of fidelity in a prototype is described as how far from the final product the design is when testing [37]. This divides the prototype into three different categories, Low Fidelity (Lo-Fi), Mid Fidelity (Mid-Fi) and High Fidelity (Hi-Fi).

2.9.1 Low Fidelity Prototype

A Low Fidelity (Lo-Fi) prototype is the furthest version from a finished design. This form of a prototype often includes the basic parts of the interaction or design that are to be tested. However, it may lack details such as images and logotypes, exact padding and margins or correct information in different paragraphs on the page. The Lo-Fi prototype is often made on paper since this is the quickest way to create a fast prototype sketch. The paper prototype, along with simple digital prototypes in computer programs are two commonly used mediums within design[37].

It is recommended in a design process to create and test a Lo-Fi prototype before making a Mid-Fi or Hi-Fi prototype. This is an essential part of the process to find structural or major flaws. Thereafter it is much easier to concentrate on improving the prototype and design those details that were not necessary during the Lo-Fi prototype[37].

2.9.2 Mid Fidelity Prototype

Mid Fidelity (Mid-Fi) prototypes are not as common or as well-known as Lo-Fi and Hi-Fi since they lie somewhere on the scale between the two versions.

There is no conclusive definition of what a Mid-Fi prototype looks like or contains but it can be explained as a prototype with limited functionality and that may include clickable areas and navigation. It is more realistic than a Lo-Fi and less realistic than a Hi-Fi.

Unlike a Lo-Fi paper prototype, a Mid-Fi prototype is more representative of the final product design and can therefore present plots and visuals better. Due to the more "finished" appearance of the prototype, user-testing with the prototype will generate other kinds of feedback, like design flaws or preferences [38].

2.9.3 High Fidelity Prototype

The High Fidelity (Hi-Fi) prototype is supposed to represent the finished design, both functionality and design details. This includes the interactions to be tested and design aspects such as padding and margins, all in a digital prototype. Since the design process is all about iterations, the Hi-Fi prototype could be tested and iterated until the finalized design is set and ready for development [37].

Due to the high standards, of the Hi-fi representing the final product, the prototype usually demands greater resources both in investment and in time creating it. Since Hi-fi creation and testing are at a later stage of the design process and more realistic looking, user feedback from user testing is mostly focused on surface-level flaws or design aspects. The expectation is that the larger flaws with functionality or navigation would be found earlier in the process [39].

2.10 Evaluating

Evaluating the prototype or prototypes by conducting usability testing is a fundamental part of the design process. The result of this could increase the amount of users or customers to a platform and with that, the possible sales and profits earned will also increase. For this reason, many companies make sure to test their websites, applications and other products nowadays [40].

2.10.1 Usability Testing

Usability testing is a research methodology when a moderator (researcher) asks a participant to perform a specified task using a prototype or interface, then observes the participant's behaviour and collects feedback [41].

Usability testing has three components: working with representative users, asking them to perform representative tasks and observing what they do. According to Nielsen, elaborate usability tests are a waste of time and resources. The best results come from running several smaller tests with no more than 5 users [42]. As seen in figure 2.6 about 80% of the problems can be found with only 5 users.

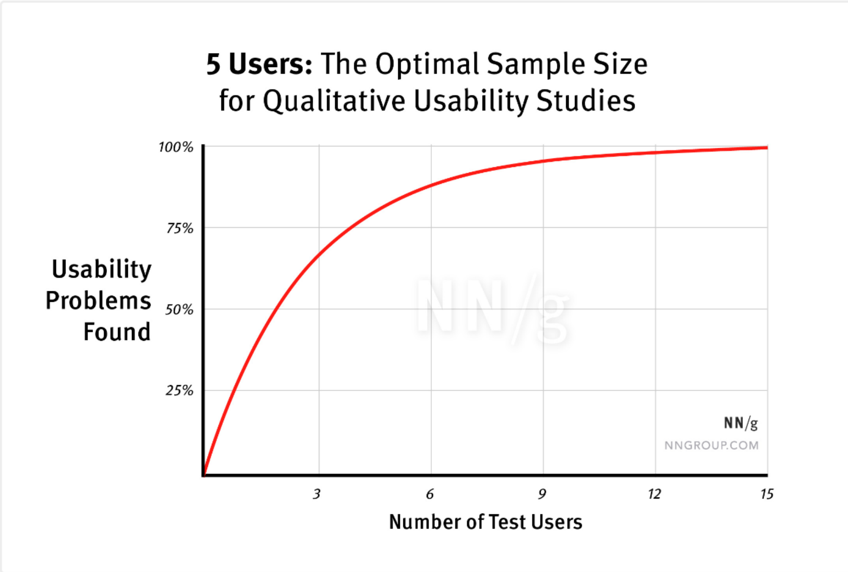

Figure 2.6: Why you only need to test with five users by Jakob Nielsen [42]

There is a simple yet important methodology when conducting a usability test for a prototype. The most important part, according to VanArsdall in his *Handbook of usability testing*, before actually testing is to conduct a test plan [40]. A fully developed test plan should include the following parts:

- Purpose and goals
- Research questions
- Participant characteristics
- Methods
- Task list (Scenarios)
- Test environment, equipment and logistics

- Test moderator roles
- Data to be collected and evaluation measures
- Report contents and presentation

After developing a test plan, it is recommended to conduct at least one pilot test. A pilot test ensures that the scenarios and test plan will result in the necessary responses or actions. It also verifies that the tasks are properly formulated and understood by the test participant [40].

Quantitative vs Qualitative Data

There are two concepts called quantitative and qualitative data. In the book named *Introduction to Data Analysis: Quantitative, Qualitative and Mixed Methods* by Tiffany Bergin, quantitative data is described as numerical, or represented by numbers [43]. The key method for extracting quantitative data according to the author is experiments and in this study evaluations and surveys are far more relevant methods in usability testing. Qualitative data on the other hand, is described as textual or visual data. Bergin describes a qualitative collection of data with three key words, in which they are often combined, namely experiencing, enquiring, examining. The most important part in collecting and analyzing qualitative data is to find themes and patterns. This is where quantitative and qualitative data are completely different. A number can be compared to a number quite easily, while thoughts and behaviours could be more challenging to compare to one another [43].

Think Aloud

To "think aloud" is a technique used in many usability tests, where the moderator asks the user to openly share their thoughts and actions throughout the test. This is carried out while the user perform the tasks at hand. The "think aloud" technique then helps designers comprehend both the user's understanding of the designed prototype and the user's desired design or functionality with their product. There is also a disadvantage to the think aloud technique, where the user gets distracted by their own talking, which leads to them not displaying the same user behaviour as they usually would. However, in most cases think aloud is useful and a necessary part in the usability tests [40].

Scenarios

In usability testing the tasks the user performs during tests are often called scenarios. Scenarios are a way to create an environment where the user can use the system as they would in a real situation. They are structured to minimize the need for the test leader to intervene during the test. The scenarios set the scene for the user to be able to do the task that is expected of them during the test. This means that the scenario will provide the user with information on the results expected of them. Usually, real names and data are provided to make it seem real to the users. The task scenarios could be either read out loud by the test leader or given to the users to read instead [40].

Questionnaires

Questionnaires are tools used in surveys to gather information from participants about themselves, their opinions or their behaviour, in a written format.

The most important aspect of creating a questionnaire is the formulation of the questions, they must be easy to understand to prevent misunderstandings. Additionally, the structure/format of how the participant can respond to the question is critical, the answer format must match the question. An open-ended question, with a free-text field option to answer, can provide depth to the response. A close-ended question, with a scale, ranking, single or multiple choice question, is easier to analyse numerically.

Questionnaires are most useful together with other methods, for example, can observation data be a good supplement to the self-reported text-based responses [35].

Questionnaires can be used in several parts of user testing, both before the test, during and after the test is completed.

The last type of questionnaire included in the usability tests is the post-test questionnaire. These are by many seen as the most important ones to include, as they are used to get more information from the user about their preferences. In a debrief purpose there is an importance in asking the same questions in the same way to every test participant. This is one of the reasons why many use the already standardised prepared questionnaires such as System Usability Scale (SUS) [40].

System Usability Scale

The System Usability Scale (SUS) is a questionnaire that aims to get the users' honest opinions of the usability of the system. The SUS is a part of the ISO standard and it uses a likert scale from 1, meaning strongly disagree to 5, meaning strongly agree, as seen in figure 2.7 [40].

Figure 2.7: Pictures of the Likert scale used in System Usability Scale

The SUS is a questionnaire based on ten questions seen in Appendix A.1.2. Half of them, the uneven-numbered questions, are positively stated and the other half, the even-numbered ones, are negatively stated [44]. The questions are intentionally arranged in a mixed order to encourage users to carefully consider their responses. Another reason to alternate the order is to avoid the behaviour of providing identical answers to each question throughout the questionnaire.

The SUS is calculated into a score where the number answered on questions 1, 3, 5, 7 and 9 are subtracted with one before being summarised. On the remaining questions 2, 4, 6, 8, and 10, the answer is subtracted from the number five. The SUS score for each test person is calculated using the summarized points from the even and odd questions, where both are added together and then multiplied by 2,5. The final result will range from 0 to 100 points for each user after the multiplication [44]. The equation and formula for calculating the score can be found in Appendix A.1.2.

To get the total score for the prototype the average for all users points need to be calculated. A higher score indicates higher usability, while a lower score means that there could be faults in the system design. According to Sauro, the average score is 68 and anything above is considered a good product.

Chapter 3
Design Process

In the third chapter, we will present the design process and results of this study. This will include digital illustrations, drawings and images. Photos from the actual events will not be displayed due to photography prohibition in all buildings owned by Saab. Instead, a simplified illustration was created to replace the image that otherwise would have been displayed.

3.1 User study

Due to limited direct access to users during the study, an alternative approach was required. We acquired knowledge about submarines and SCMS through conversations and interviews with system developers and experts. This was to understand further the potential issues the user might encounter. During the project planning phase, valuable information and feedback was received from a user. Based on this feedback, insights from the literature study and expert interviews, we compiled personas and a comprehensive list of system challenges.

3.1.1 Saab Submarine Mini-courses

To find out more about our users from FM and their environment we needed general information on their workplace, the submarine and the systems they use, SCMS. We received a generous introduction to submarines, our users and SCMS's role in this during three mini-courses with our supervisors from Saab. The first one was to inform us how the submarines work and any technical information that we were either curious about or that was necessary for our project. The second one gave more detail on the systems in the submarine that our team is responsible for, in this case, the alarm handling system SCMS. The third mini-course was to introduce us to the roles onboard and ranks within FM. The roles and ranks were specified for a naval ship, as one of our supervisors has experience and expert knowledge in this area. Although it might differ a bit on a submarine in terms of names on the roles, they have similarities and helped us a lot during our user study. All of the information from the mini-courses was documented in order to remember and to be in place in case of uncertainty. Apart from the submarine specifics we did Saab onboarding and e-learning courses.

3.1.2 Expert Interviews

Submarine personnel from FM have a tight and strict schedule, therefore we had no access to our users during the user study. Since the process could not include a survey or any interviews with the users, we had to get the information needed for this user study from other sources.

The team we were assigned to at Saab are incredibly talented and knowledgeable and they were all aware of several of the existing user problems. Therefore, we collected all data about our users' problems by interviewing and asking for information from the developers and experts of SCMS. The people we interviewed [4][2][12] knew the details of SCMS design issues in great depth and many had previously tested the system several times with FMV present, though mostly in developer functionality aspects rather than UX.

With the experts' knowledge of this very complex and large system, we received valuable information that we could not access in any other way. They explained aspects of the system in detail that they knew could lead to problems and even shared ideas on how to prevent them. Some of the issues mentioned were corroborated by the results of the literature study, such as issues of alarm flooding and the "keyhole effect". The team also had insight into who our users would be. Different user data such as age ranges, education, roles and professions were discussed to paint a picture of a typical user.

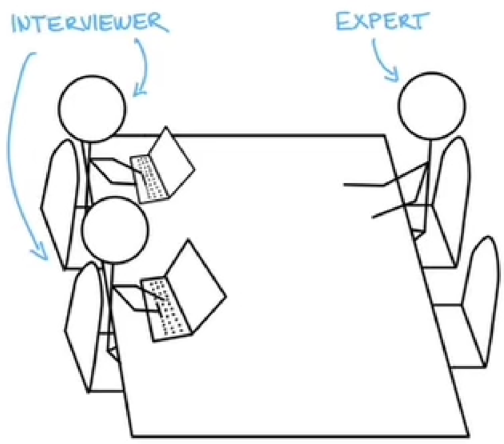

Figure 3.1: Drawing of the scenery during the expert interviews

The interviews were conducted with an open structure since the goal was to explore the topic at hand and acquire as much knowledge about the users as possible. Different problems were discussed and eventually, a clearer picture was formed. As usual, we took notes to document all the information. A follow-up interview or dialogue was held whenever additional information was needed. Figure 3.1 displays how the interviews were conducted.

The data collected during these expert interviews provided ground for parts such as finalizing our research questions, creating personas, creating scenarios and identifying the main issues to be solved during the brainstorming session. The results of the interviews will be presented throughout the report.

3.1.3 User Feedback

In addition to the expert interviews held, we received unexpected information and feedback on the system from an FM test person when booking the usability tests. The user was passionate and motivated to share their thoughts on the current control and monitoring system used on ships within FM. The system at hand is similar to SCMS in many ways, even though it is much less modern, used on a much smaller screen and not developed by Saab. The feedback included requests for features and improvements and all of the information received was summarised in a list.

- There are too small alarm windows and it is hard to read all the alarms.
- There is a risk of prioritized important alarms disappearing at the bottom of the list when new alarms appear.
- A function where one could swipe alarms to a list they could handle later, to reduce the amount of un-important alarms in the list.
- A better filter of importance to be able to prioritize alarms.
- Add reminders of critical alarms that have not been dealt with in due time, to help the user not forget or miss critical alarms.
- Add a help window to guide the user to take the next steps or to explain the possible consequences of the alarm.

3.1.4 Usability Problems in SCMS

The questions and issues originating from the expert interviews, literature study and user feedback were mapped out and examined. The result was presented in table 3.2.

3.1.5 User Group for SCMS and Test Participants

All users for SCMS will be submarine personnel from FM or similar units from other countries. When conducting our user study we needed to book our upcoming usability tests with FM early on in the process, even before creating our personas. When booking, we were unable to select specific types of users for testing. Therefore we kindly asked our supervisor for users that would closely match the ages, genders and equivalent roles of either a MEO or CENG from other parts of the naval flotilla. The gender statistics from all of our usability tests combined are displayed in figure 3.3. The statistics of our tests were compared to the gender statistics of the whole FM personnel in the same figure [45]. Moreover, the specified age group for test participants in comparison to the age group of SCMS users according to

Table 3.2: Usability problems in SCMS

Problem	Description	Source
Learning and mastering the system takes a long time.	This may be due to design flaws. A more intuitive design would likely reduce the training duration. If the submarine is delivered to another country with shorter mandatory training periods, issues may arise.	Expert interviews
There is a difficulty in transitioning from previous systems, which had an insufficient amount of alarms, to the new system featuring a large number of highly detailed alarms.	With too few and less detailed alarms, identifying and resolving issues becomes challenging. It is crucial to find the right balance between an excessive number of detailed alarms and insufficient alarms with inadequate detail.	Expert interviews Literature study
Alarm flooding is an existing problem in most digital alarm monitoring systems.	The design in alarm monitoring systems is usually not built for alarm floods, rather than a few occasional alarms. The question remains on how the information could be displayed in order to not overwhelm the user in a stressful situation such as alarm floods.	Expert interviews Literature study
Difficulty in prioritizing alarms during alarm floods.	When faced with hundreds of alarms, the user needs to know what to prioritize. Currently the system displays a long list of alarms, with new alarms appearing at the top while the user is reviewing the old alarms. This setup can overwhelm the users. Moreover, the system lacks an indication of which alarms have triggered other alarms.	Expert interviews Literature study
Difficulty in getting an overview of the system. The users are experiencing the "Keyhole effect".	In a highly complex system with many components, it can be challenging to view the entire system screens. This effect of not seeing the full picture is similar to looking through a keyhole. The users usually struggle to get an overview of the alarms in the system.	Expert interviews Literature study
The filtering of alarms requires improvements.	There are existing filtering options available on different pages, however they are still not providing sufficient assistance to the user during an alarm flood.	Expert interviews
Difficulty to determine root causes in alarm floods.	Users frequently struggle to identify the root cause of alarms during alarm floods. Currently, the system does not show any connections between alarms triggered by the same event.	Expert interviews Literature study

Saab [4] are displayed in figure 3.4. This statement of the users being 30-45 years old was made based on the roles of a MEO or CENG.

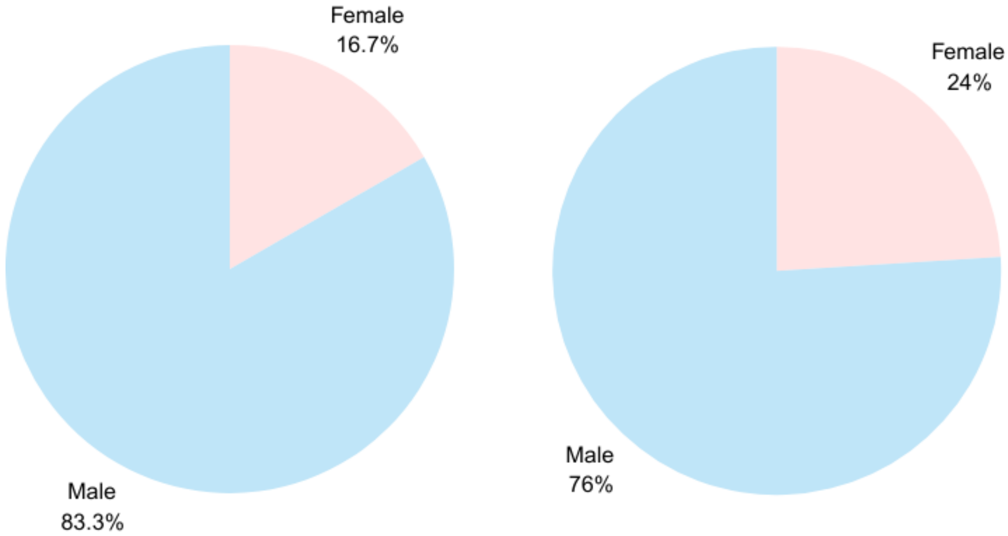

Figure 3.3: Gender distribution from the usability tests (Left) and FM personnel statistics from 2023 (Right) [45]

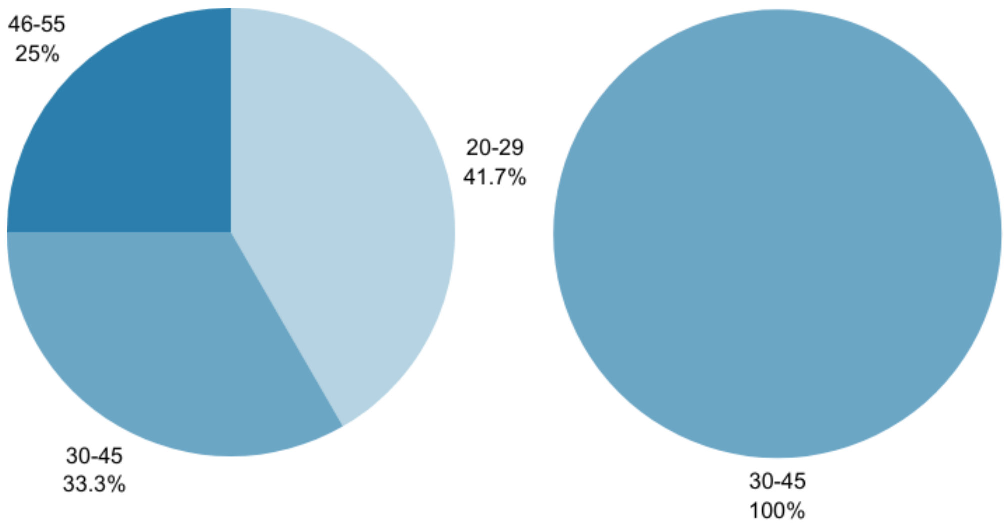

Figure 3.4: Age distribution from the usability tests (Left) and the assumed specified age group for SCMS made by Saab (Right) [4]

3.1.6 Personas

Based on the information from the expert interviews, three different personas were created. The three personas were ranked according to the military rank within FM. With a few adjustments after our supervisors reviewed the personas, they were summarising our user group well.

Since SCMS will only be used by submarine personnel, our target users were all quite similar. Similarities were for example that they will all have the same type of baseline education, the system will be used in the same environment for everyone and the users are required to be free from any disabilities or any type of functional variant such as colourblindness etc.

The key difference between the personas was the experience with the system at hand. The older generation will have used multiple variations of the system up to this point. They might have used the system when it was not as digitized and consisted mostly of physical buttons, switches and annunciators. The younger generation will not be as used to the current system and might therefore be more open to change. They will also be more used to digital systems in general due to them growing up using computers or digital devices daily.

The age range of the primary users lies between 30-45, they operate the system first-hand and have had extensive eductaion. The secondary users are generally younger, around 20 years old, and observe the primary user operate the system. We classify Lars-Erik and Oscar as primary users and Filip as a secondary user.

We used AI-generated pictures for the persona profile to make the personas feel as real as possible.

The personas are presented in the figures 3.5, 3.6 and 3.7.

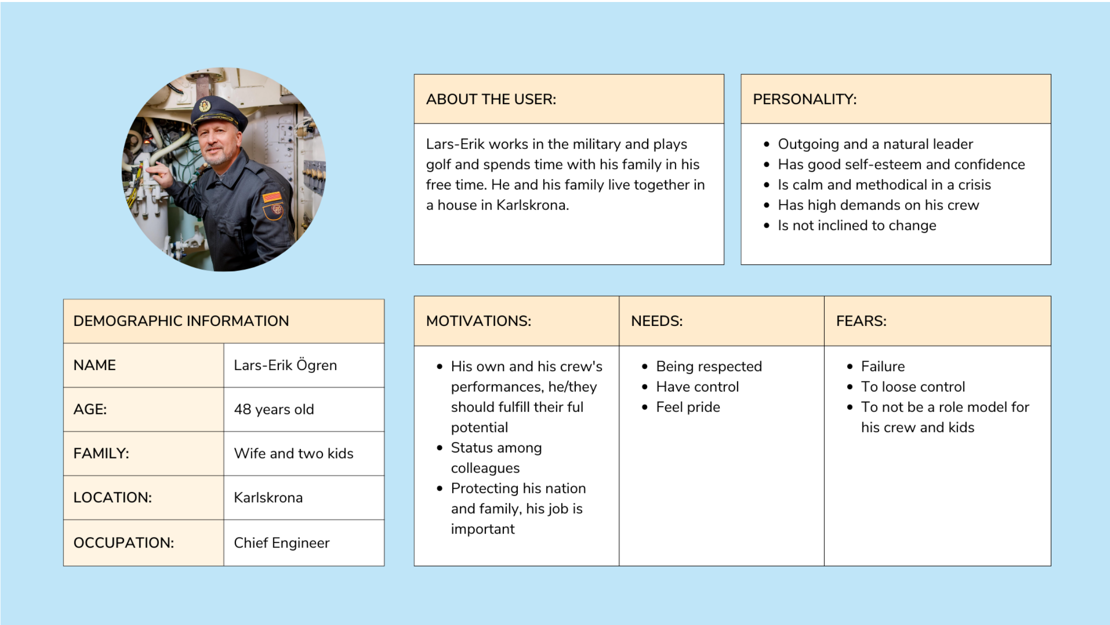

Figure 3.5: Persona 1

3.2 Ideation

The methods and results of the ideation phase will be presented in this section.

3.2.1 Bodystorming

The current system setup for the submarine was available for tests anytime at Saab. It is a simulation of SCMS (complete with screens and touchpad) developed for test purposes while the submarine is being built.

Thanks to this valuable asset, a bodystorming session was conducted. We got to sit in a seat in front of the two screens, click around and monitor the alarm list impersonating our personas. We noted things that we did not understand and things that were confusing. From this, we tried to spot potential weaknesses in the design as well as think of possibilities for improvement.

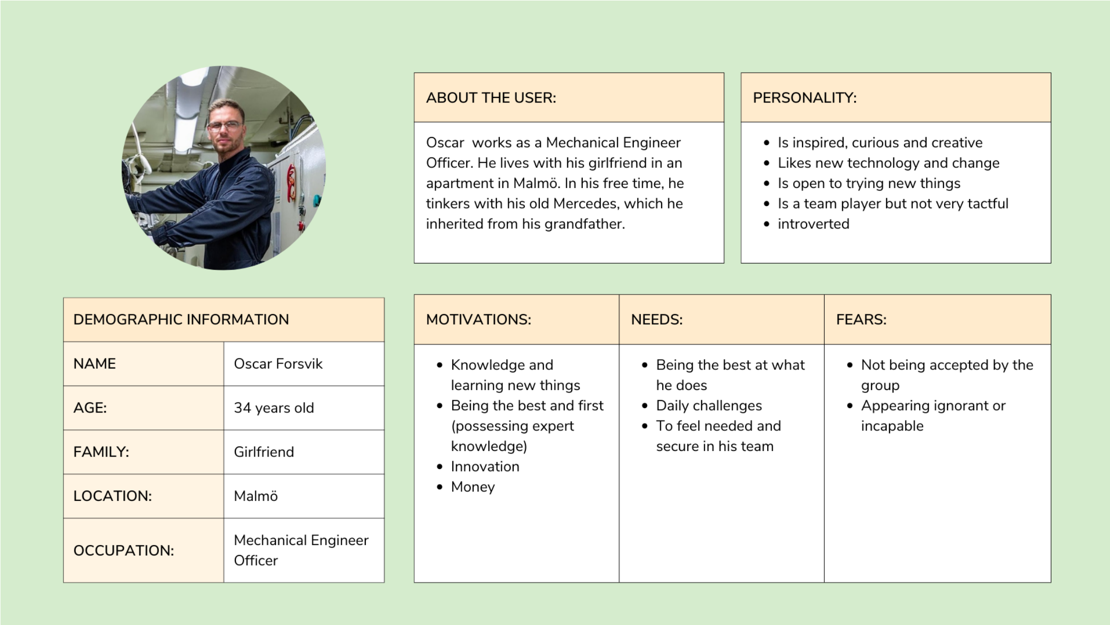

Figure 3.6: Persona 2

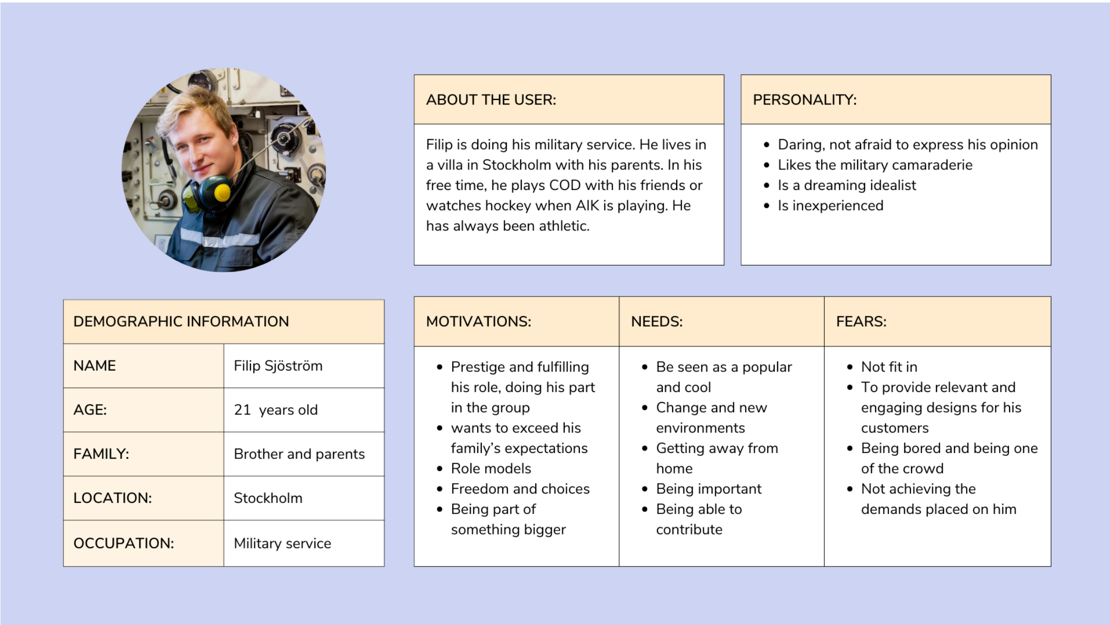

Figure 3.7: Persona 3

For an earlier project, we had the opportunity to visit Malmö Museum, where we made sure to check out the submarine exhibition and got to enter an old submarine. This experience helped us imagine the small operating space and work environment even though it was not the same submarine class.

General impressions of the systems were that the small alarm list on the home-screen was too small and was difficult to read when many alarms were generated. The maximised alarm list grew very long in a short period, overwhelming us with information and there was no way to get an easy overview of the alarms or historic alarms.

3.2.2 Brainstorming

With the information and knowledge gathered in the early stages of the study two brainstorming sessions were conducted with the two of us as participants. The first brainstorming session was conducted by setting a timer for twenty minutes and separate brainstorming was performed in silence, not showing each other any of the Post-its. When the timer was up we switched Post-it notes and started another timer for twenty minutes to continue brainstorming with a new base of the others' ideas. This resulted in both of us building and further developing the separate ideas. After this whole process, the last and final stage of our brainstorming was the joint discussion where all the Post-its were explained in more detail and new ideas were collected and saved on additional Post-its.

The second brainstorming session was carried out a few days later and repeated the earlier format. There were a lot fewer ideas in the second session. However, this second session resulted in a few new design options. During the brainstorming sessions, we did not restrict our ideas to any specific pages or area of the system, since alarms are handled, filtered and seen in many different views. The result of the brainstorming is presented in the affinity diagram section.

3.2.3 Affinity Diagram

The brainstorming results were organised using an affinity diagram, grouping related ideas from the Post-its into categories. All the Post-its were divided into four groups which are presented in figure 3.8.

In the top left: General ideas for system improvements, including creating an easier filtering system, enhancing the management of multiple alarms simultaneously, sorting or organising alarms in a tree structure, clearly linking related alarms and providing information about the potential root cause of each alarm.

Bottom left: Various methods to visualise alarms in a graphic overview instead of a traditional list. The idea was to implement either a timeline, bubble chart, treemap, spiral graph or heat map of the submarine.

Bottom right: Concepts for designing and formatting alarm lists, including navigating between pages based on criticality level, organising alarms in different tree structures, using dropdown menus to hide following alarms, incorporating signifiers in the alarm status bubble, providing help text and implementing a priority list function.

Top right: General GUI techniques to enhance the visibility of alarms in lists, such as multimodal alarms with sound or vibration, highlighting the small alarm list when new alarms appear, keeping a longer alarm list displayed as a side panel and changing the colours of status symbols.

Some ideas were discarded after careful consideration. One example was the idea of providing a help text for each alarm. This would be a good idea according to users and Jakob´s 10 Usability Heuristics "help users recognize, diagnose and recover from errors". However, the design would demand expert knowledge, be incredibly hard to calculate and the development would take up too much time, not matching the time frame of the project.

Figure 3.8: Brainstorming Post-its notes arranged in an affinity diagram

3.3 Prototyping

In normal cases where a new design is created, a mood board or other ways of displaying the possible style for the different prototypes is a good way to start the prototyping phase. However, in this case, the design guidelines were already established by Saab for SCMS. According to Jakob´s 10 Usability Heuristics "consistency and standards" following guidelines increases users' mental load [31]. As a result, there were fewer decisions to be made regarding colours, margins, and sizes. This allowed us to concentrate more on UX and User Interface (UI), exploring new ways to present the same information while ensuring its usability.

3.3.1 Scenarios

To ensure the test scenarios were based on realistic situations onboard a submarine, we received help from a colleague. Together we focused on two main problems that could trigger alarm floods that our personas would have to manage during their shift. The first scenario was a fire in a specific area of the submarine and the second scenario was fluid-level alarms triggered by heavy seaway. Both created hundreds of alarms in just a few minutes, which was perfect for testing. The scenarios from the Lo-Fi and Hi-Fi tests are provided in the appendix A.1. Some of the specific tasks in the scenarios were inspired by Jakob´s 10 Usability Heuristics, especially "visibility of system status" and "recognition rather than recall". We wanted to test if the user could understand the state of the system and also learn functionality and navigation easily [31].

3.3.2 Views to Redesign and Test in SCMS

SCMS is a large and complex system with numerous views and for the purpose of redesign and testing in this project, the most relevant views were the *maximised alarm list*, the *history view* and the *alarm list view*. These were the only views considered when deciding which parts of the system would be redesigned and evaluated.

Firstly, this study focused on analysing the maximised alarm list, which is accessible from any page in the system and can be displayed on either of the two screens for the user to see.

Secondly, the page displaying older alarms will be tested, as it is designed for use in less stressful situations. During the expert interviews, one issue raised was the need for a better overview of alarms, particularly when multiple alarms occur simultaneously. In comparison to the normal alarm list, we hoped to get a better understanding of this page might be enough to get that overview. If not, we hoped to be able to display it more graphically.

The third page, namely the alarm List, was the only page specifically designed in the HMI as an alarm-focused page. We decided to exclude it from our prototype. This decision was due to the almost identical resemblance to the History alarm page and similar functionality to the maximised alarm list. Our decision was also based on the conclusion that the same new design ideas for the maximised alarm list could be implemented for this one later on.

3.3.3 Lo-Fi

For the Lo-Fi phase, we decided to create three different prototypes, naming them P1, P2 and P3. The first was a replica of the current system and the two other variations included new displays based on the brainstormed ideas. The focus of these ideas was to provide the user with a clearer overview of the alarms and reduce the overwhelming feeling caused by the maximised alarm list during an alarm flood.

Making one of the prototypes with a design closely resembling the current system was a crucial step for both our study and for Saab. The opportunity to usability-test this type of alarm management case in SCMS with relevant personnel from FM had not yet been executed prior to this study. The vital part in the first prototyping phase was therefore to compare the original design of SCMSin P1, with the new ideas in P2 and P3 and compare the feedback for all three prototypes. Our hypothesis was that the feedback would answer whether something was good with the original system, what could be improved and what additions we should focus on.

Voting on Ideas for Lo-Fi Prototype P2 and P3

From the ideation phase and especially our brainstorming session we had several ideas and small sketches we wanted to implement and test in a prototype. To determine which variations to include in the design of the different prototypes, we used a voting system inspired by the Eurovision Song Contest. The different ideas and variations were 12 in total and we used what we call the *Eurovision voting system*.

Our Eurovision voting system worked as follows. Each participant gets an array of numbers, which equals the number of ideas the participants are voting on. In this case, there were 12 ideas and each participant got 12 Post-its. Each Post-it had a number from the array, 1-12 in this case. Then each participant got to vote on the ideas with one Post-it each. In the

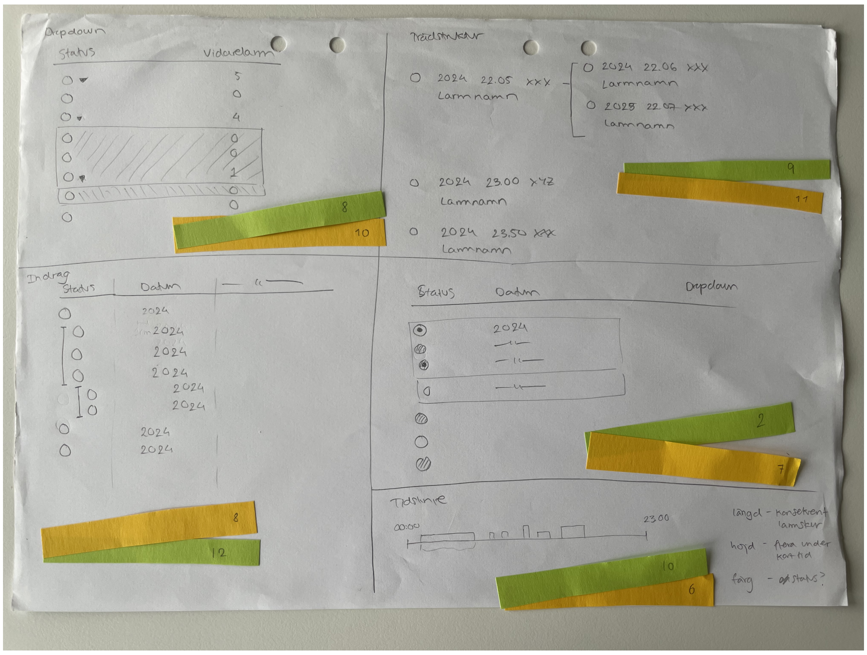

Figure 3.9: Eurovision Voting System part 1

Figure 3.10: Eurovision Voting System part 2

end, the results were summarised for each idea and the total points were compared in a list of the lowest total points to the highest. This is a type of weighted dot-voting system which allows the participants to prioritise the ideas they valued highest and lowest, respectively, without any required dialogue with other participants. It is a fast and effective way to prioritise which ideas participants cared for and believed in the most. In the end, the points were added together and the design variations were ranked in order from the most popular, three ideas with a total of twenty points, to the least popular, with two points. After this voting

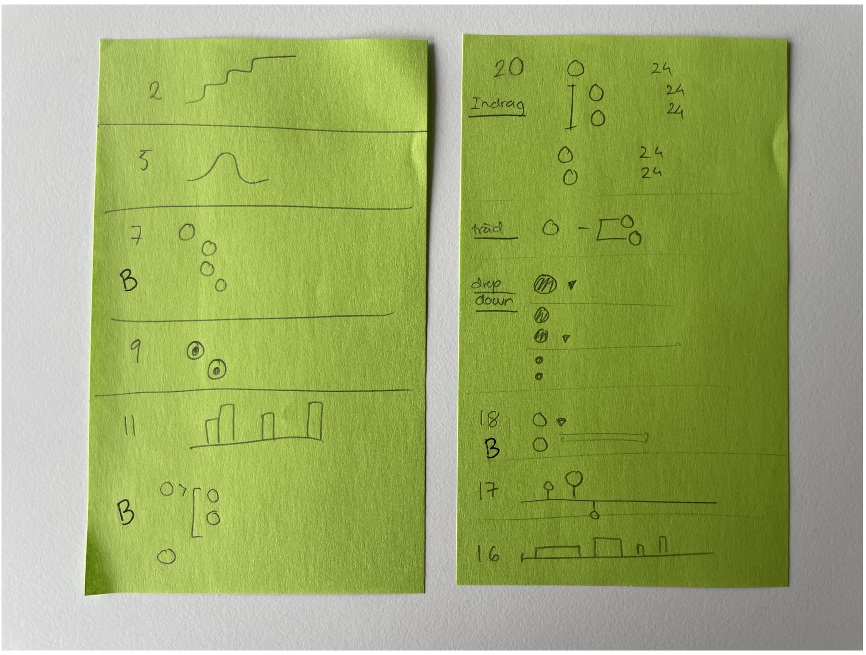

Figure 3.11: Results from Eurovision Voting System

session, the prototype design variations were settled. The different design variations and the voting with Post-its are displayed in the figures 3.9, 3.10 and 3.11.

Lo-Fi Prototype P1, P2 and P3

We decided to redesign two of the original types of views for every prototype. The views we focused on were the maximised alarm list and the history page.

To create the prototypes we started by recreating the current system in a 1:1 scale on paper. By cutting and putting two A3 papers together we got a 26-inch screen and then measured everything we needed in the current system on the interface in the lab. The original base design was used as a base for all three versions of the prototype.

To save time and minimise the risk of one prototype being destroyed, we scanned the prototypes after each step in order to be able to reuse and save the designs.

We created two copies of all three prototypes because of the testing circumstances, which will be further described in the following sections.

The home-screen consists of an upper panel with six buttons, a short alarm list with three alarms at the bottom with corresponding buttons to handle the list. The middle of the screen is intentionally left blank. This space in the original system is usually filled with the systems, components and subsystems of the submarine to visualise the status or state being monitored, which we are unable to display for the public eye during any of our iterations in the design process. We did not include that information since it would not be necessary for the usability tests. Instead, we concluded that this would cause confusion or distract the users and be time-consuming during the usability tests. The prototype is presented in figure 3.12. This is the base for all three prototypes P1, P2 and P3.

The differences between the prototypes are:

- In P1, the alarm list is a traditional list, presented with different alarms depending on

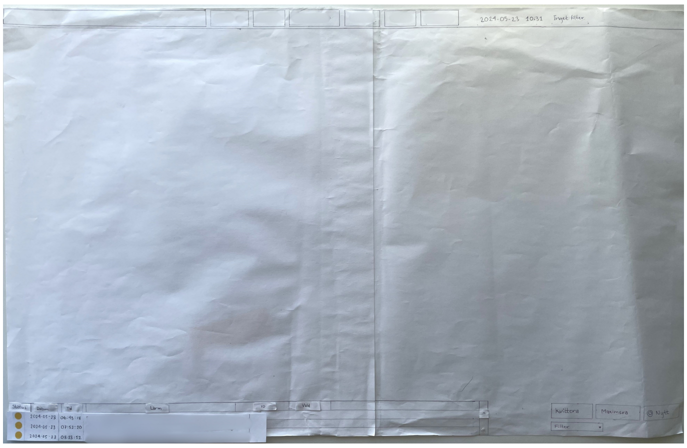

Figure 3.12: Lo-Fi Home-screen view used for all the Lo-Fi prototypes

the scenario. The history page is the same as in the original design of SCMS.

- In P2, the consequence alarms and the common and less critical alarms and warnings in the list are indented, in other words, tabbed in to indicate that there is a parent alarm. The history page is changed to an overview page with a map of the alarm positions in the submarine. The *submarine view* is presented in figure 3.13.

- In P3, the alarms are presented in a drop-down list and the parent alarm has a signifier displayed as an arrow, to signify that more alarms are hidden. P3 also included two versions of a time-based overview page instead of the history page, named the *time-line view* and the graphical *time-circle view*. Even though these were two different ideas from the start, they were implemented into the same prototype and considered as one page. The seven day time-circle view is presented in figure 3.14.

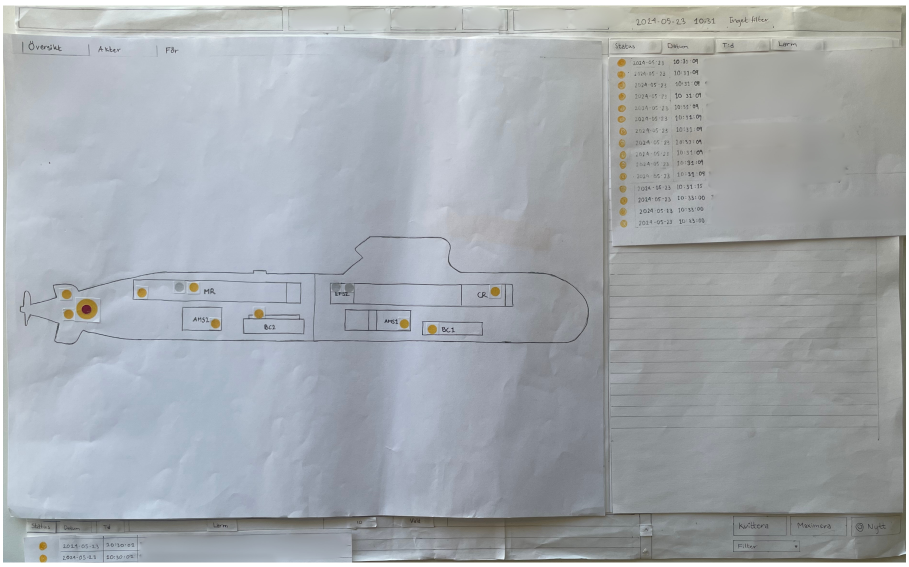

Figure 3.13: Lo-Fi P2 Submarine View

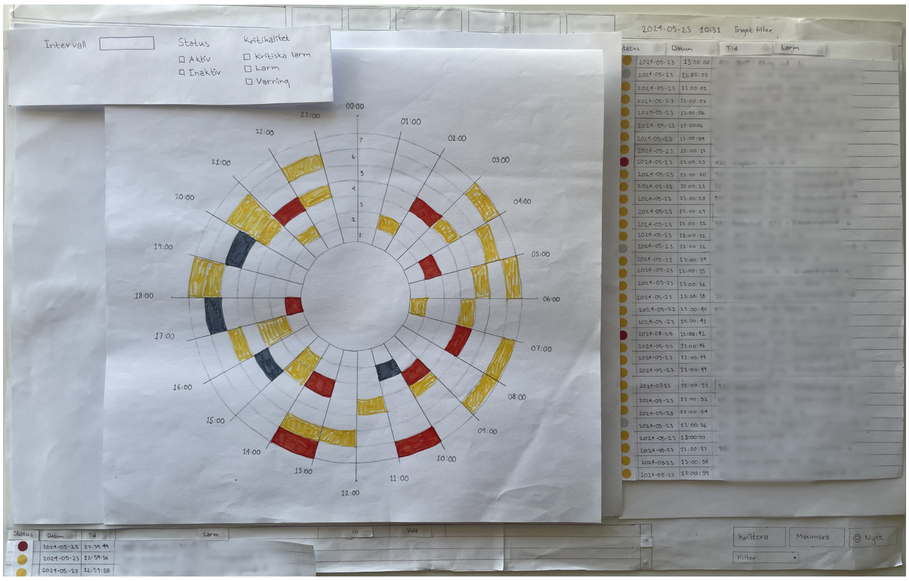

Figure 3.14: Lo-Fi P3 seven day Time-circle View.

The goal of the new design ideas was to group and categorise alarms to make them easier to handle for the user in P2 and P3. The different alarm list designs from all three prototypes are presented in figure 3.15.

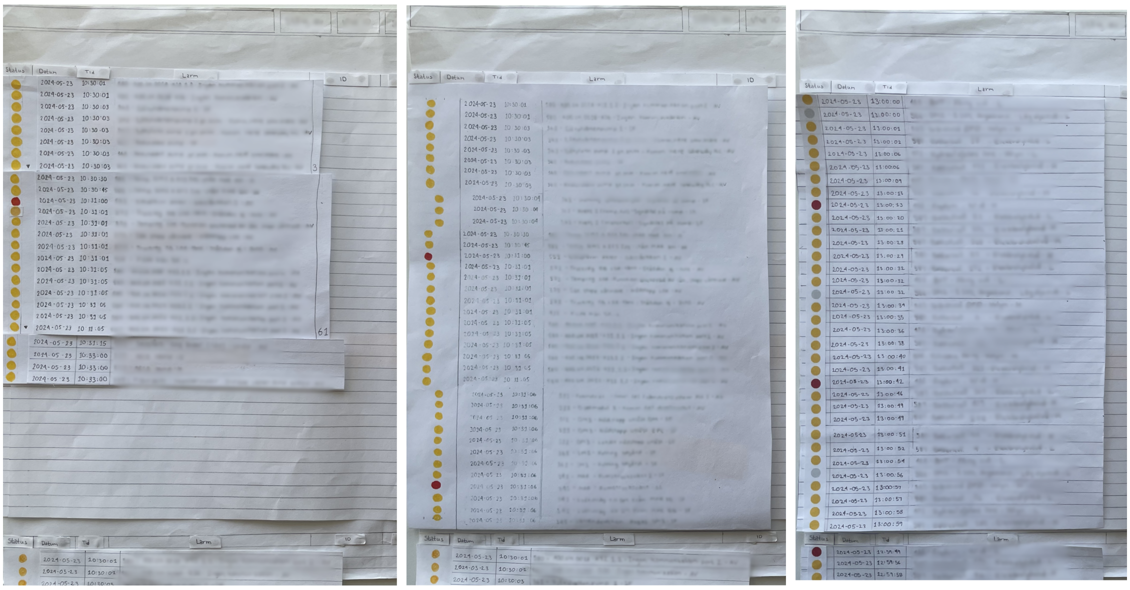

Figure 3.15: Different alarm list versions for each prototype. From left to right: P3, P2 then P1

3.3.4 Mid-Fi

Since this study at Saab was carried out in collaboration with FMV, who provided users from FM for the usability testing, we had limited options of when we could test our different prototypes. With their hectic schedules, we were offered a second opportunity to do a usability test, however, this test would be approximately two to three weeks prior to the planned Hi-Fi usability tests. This resulted in a Mid-Fidelity (Mid-Fi) prototype with another iteration of the design.

The Mid-Fi prototype was a step between Lo-Fi and Hi-Fi prototypes. The prototype was created in Figma [46] with static views and no interaction or functions implemented. This was due to time constraints and the navigation functionality was deemed less important at this stage in the process. The goal was to highlight details that were missed or iterated from the Lo-Fi testing. Furthermore, it was important to verify the feedback from the Lo-Fi tests and the usability of the newly designed views, not the navigation between views.

This time one single Mid-Fi prototype was created, as the key features and views of the previous three prototypes P1, P2 and P3 were combined into one.

The prototype consisted of a home-screen view with a small and maximised alarm list, a submarine view, a time-line with 24 hours and a time-circle view with 7 and 14 days, respectively. To make the vision clear we made a simple sketch of all the views needed which can be seen in figure 3.16.

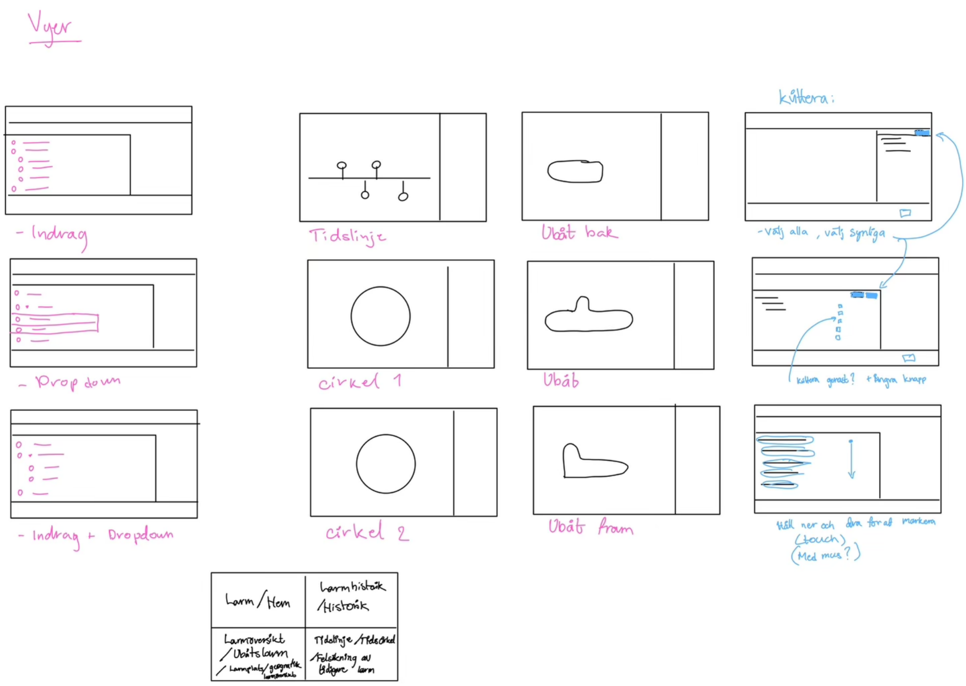

Figure 3.16: Sketched views for the planned Mid-Fi prototype

Home-screen

The design of the home-screen view is similar to the current design in the developed SCMS and is presented in figure 3.17. There is an upper panel with several buttons which will not be in use during our usability tests except for the select navigation button, in the figure named "Hemskärm" with a drop-down arrow. At the bottom of the screen, there is, regardless of what is displayed on the rest of the screen, a small alarm list accompanied by buttons, to help manage the alarms. The alarm list has a panel header to mark the colons: status, date, time,

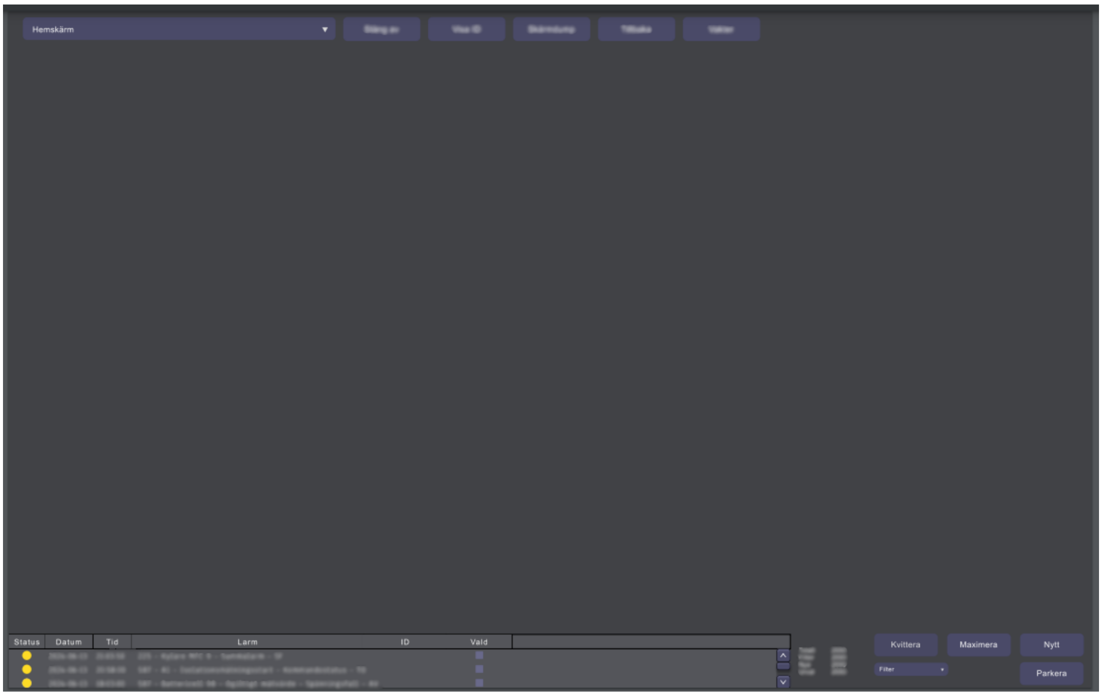

Figure 3.17: Mid-Fi prototype Home-screen view

alarm, ID and chosen. All alarm texts and other sensitive or unnecessary information will be blurred in the images. The buttons next to the small alarm list are: acknowledge, maximise, new, filter and park. The large middle part of the screen is intentionally left blank once again, precisely as the Lo-Fi prototypes. The maximised alarm list, both the ordinary one and the dropped-down combined with tab-in version, can be seen in figure 3.18.

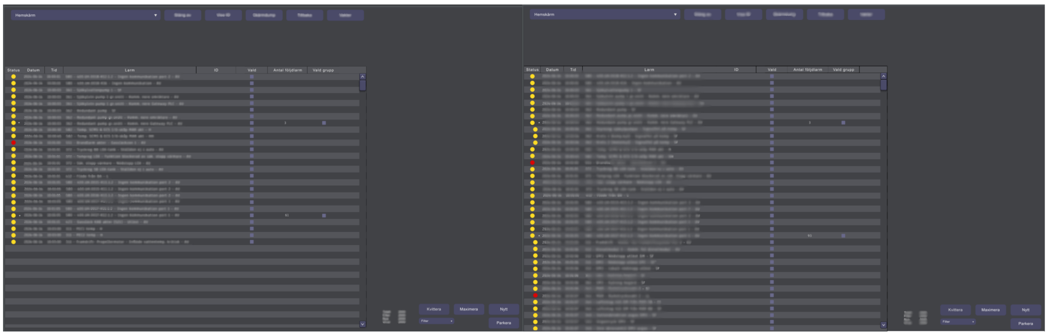

Figure 3.18: Mid-Fi prototype, from left: Maximised Alarm List View and right: Dropped-down in combination with Tab-in Alarm List View

The maximised alarm list has two extra panel headers: "antal följdlarm" meaning following alarms to root causes, and choose group. The root cause alarms that has following alarms have a small arrow next to the status symbol. If the root cause alarm row is clicked on, a list of alarms would appear with the status symbol slightly indented to the right to indicate that they are connected to the main alarm row.

Submarine view

The submarine view is presented in figure 3.19. The design of the overview submarine page is similar to our Lo-Fi prototype. In the middle of the screen there is a model of a submarine with some of its main components in boxes. Different coloured and sized circles are placed as representatives of the places of triggered alarms. There is also a side panel with an alarm list showcasing these active alarms.

Some of the differences compared to the Lo-Fi prototypes are the filtering menu and the tabs to navigate between the different submarine parts. We changed the menu so that the content would be placed in order of size. The tabs were highlighted so that the user would know which view they are currently looking at.

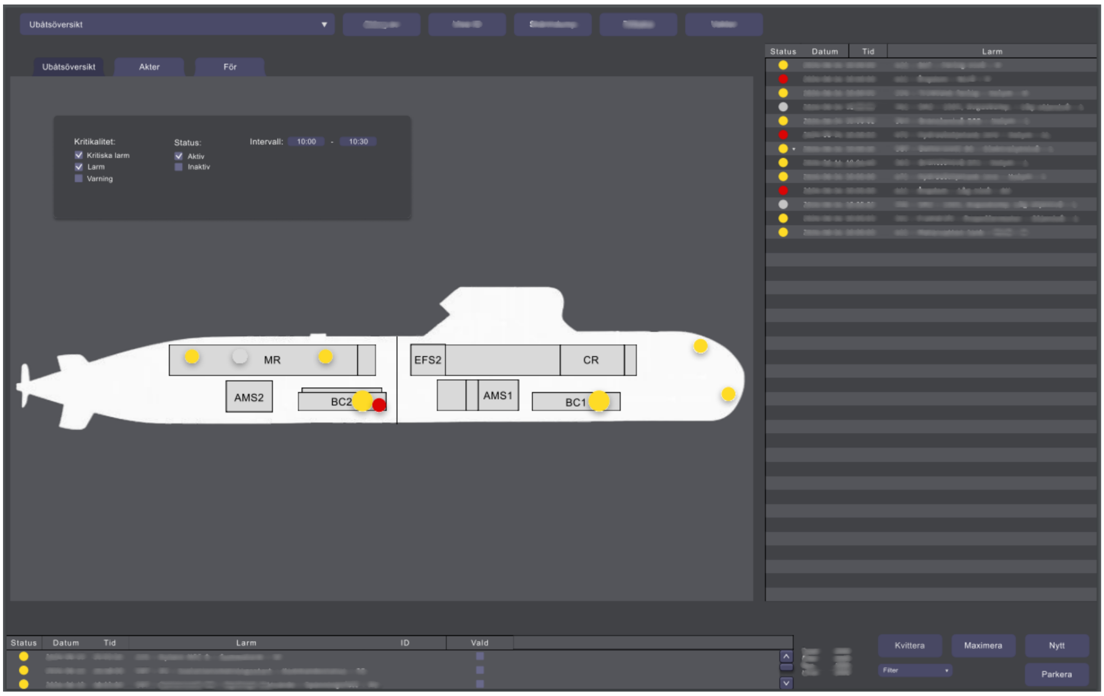

Figure 3.19: Mid-Fi prototype Submarine View

Time views

The 14-day time-circle seen in figure 3.21, the 7-day time-circle and the balloon time-line in figure 3.20 used the same tab-navigation as the one implemented in the submarine view. These three views had their main graph in the middle of the screen, a side alarm list and a type of filtering menu as well.

The time-line had the same menu as the submarine view and the two time-circles had only the criticality filter and a new search function, translated to type of alarm.

The time-circles had dates in falling order starting from the inside and out, as well as a helping text "today" next to today's date. The circle symbolises a 24-hour clock with slots for alarms in each hourly interval. The colour of the slot references the criticality level of the alarm or alarms being triggered at that time slot. We also created a view for when a certain liquid level alarm had been searched. The view presented a 7-day time-circle with alarms and critical alarms selected to be shown. The alarm list next to the circle was also filtered.

For the balloon time-line view, we kept the placement of the grey warnings underneath the time-line, to be sure that the users would know where to focus and make it clear which alarms to prioritise.

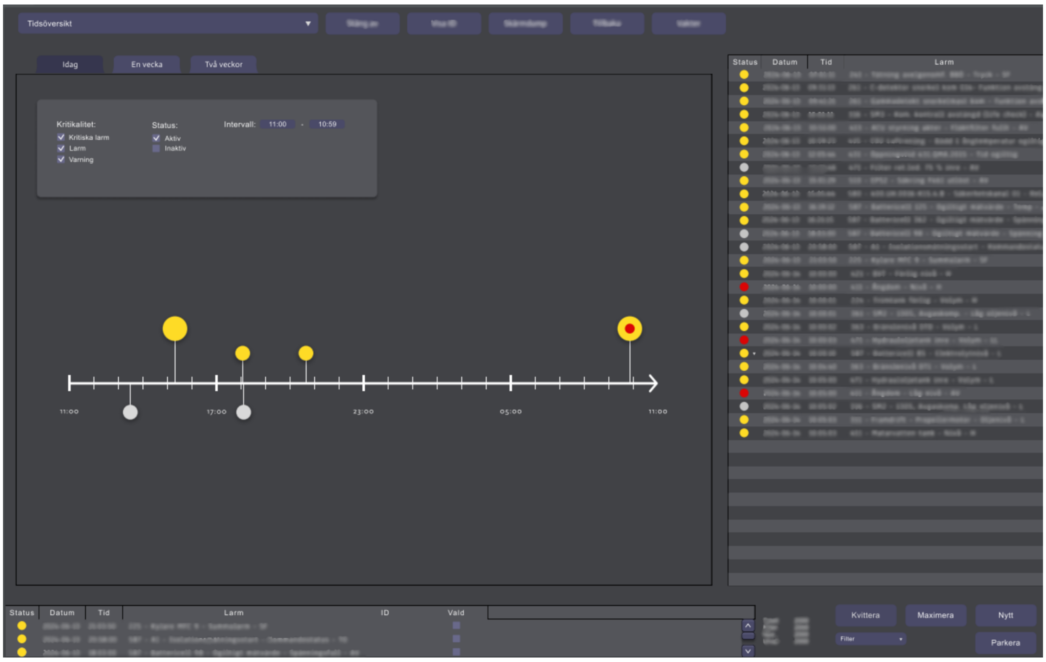

Figure 3.20: Mid-Fi prototype Time-line View

Figure 3.21: Mid-Fi prototype 14-day Time-Circle View

Park and Acknowledge functions

In addition to the alarm list and overview pages, the beginning of a park function was added to the prototype. The park function was explored due to feedback from the first usability tests on the Lo-Fi prototypes. The park function would be a way to set aside alarms that the user did not want to see in the current alarm list. This in order to lessen the amount of alarms and to help the user to prioritise the important ones more easily. The difference between acknowledge and park would be that acknowledge removes inactive alarms when pressed or checks active alarms to signify that the user has taken action and has acknowledged the alarms. The park function would allow the user to customise what alarms would be in the alarm list which is positive according to Jakob´s 10 Usability Heuristics "flexibility and efficiency of use" [31].

3.3.5 Hi-Fi

After the Mid-Fi prototype, the foundation for the Hi-Fi prototype was already prepared. This way the Hi-Fi prototyping went faster than expected. The design updates in general were kept simple according to Jakob´s 10 Usability Heuristics "aesthetic and minimalist design" [31]. We did not want to add new functions or buttons unless it was deemed necessary.

For the Hi-Fi to be able to be tested, a prototype flow was created with considerations to the stated time in each scenario and to be able to include every view in a suitable order. All the views and the flow are presented in figure 3.22.

To make the flow as simple as possible, navigation menus on repeated views were created as components in Figma. The navigation select menu placed top left on every view, has six options for navigation. Making the menu a component lessened the amount of flow arrows from 28x6 to only six arrows. This makes it easier to further scale the prototype and easier for new creators to understand the flow.

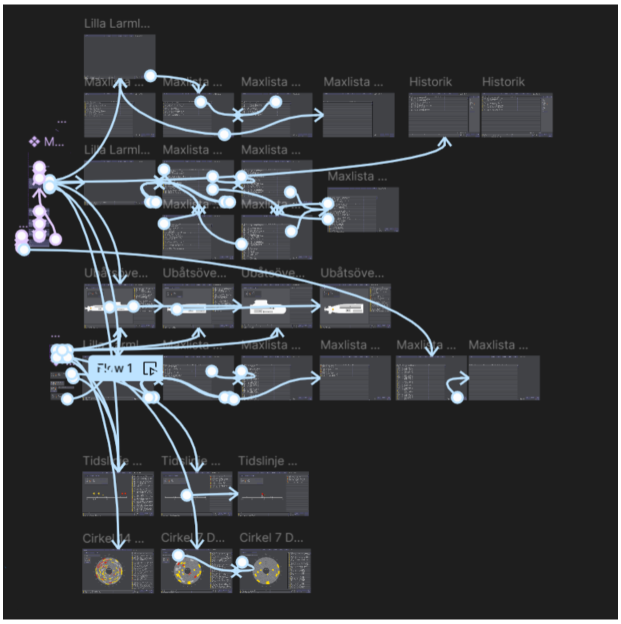

Figure 3.22: Screenshot of the prototype flow in Figma.

It was not only the navigation that was added, but also functionality such as making checkboxes clickable and filter functions. Especially the park and acknowledge functionalities were developed to let the user perform the task with as few clicks as possible. First, click on the panel header "vald" or selected group checkbox, then the button "parkera" or acknowledge, in Swedish translated to "kvittera". Some of the new design features can be seen in figure 3.23.

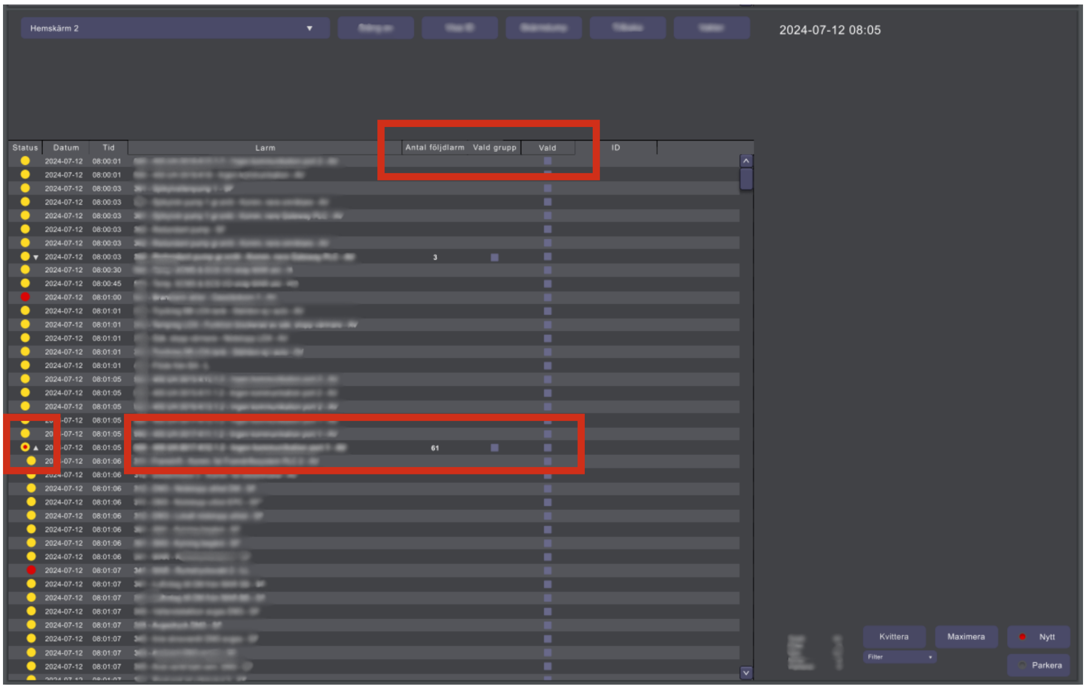

Figure 3.23: Hi-Fi prototype Maximised Alarm List View

Status symbols for alarms and group alarms were updated after feedback. The group alarm status symbol was previously single coloured, but changed to two colours to show if a group alarm contains any other more severe criticality. The time-line status symbol containing multiple alarms of different criticality were also switched from yellow-red to red-yellow to make the most critical colour most visible for the user. Status symbols were added next to the criticality-filter. Some of the new design features can be seen in the figure 3.24.

3.4 Testing

There were three usability testing phases in the design process, one for each prototype. Both places and participants resembling the users were different each time.

3.4.1 Pilot-Testing

Due to secrecy within Saab, a simple pilot test could not be conducted without relevant personnel. Therefore, pilot tests have been conducted on people with access, but unfortunately not a real user of the system.

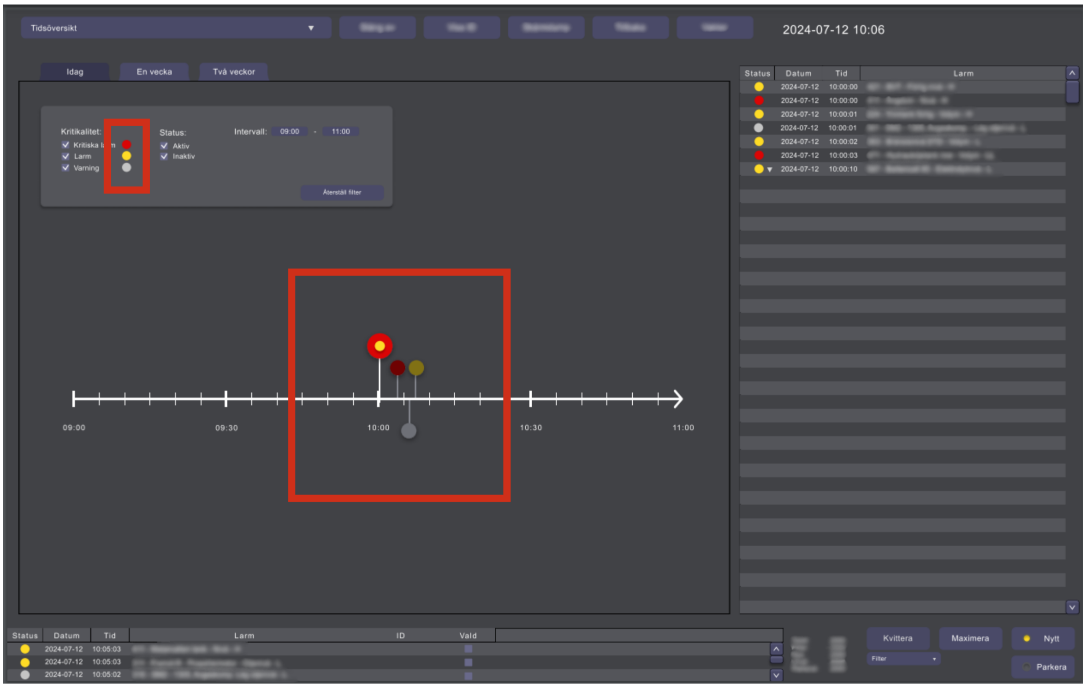

Figure 3.24: Hi-Fi prototype Time-line View

The pilot testing was a helpful practice for us to conduct the test. It also spotlights issues both in the prototypes, in the scenarios and tasks etc. There was no time to fix the prototype bugs in all iterations but the formulation and order of tasks were always updated and clarified.

3.4.2 Lo-Fi

The Lo-Fi testing was performed on a single day in Karlskrona with five users from FM. All five of them were involved in the Swedish Navy, within the 3rd Naval Warfare Flotilla.

First there was an introduction to what the day would bring, then the one by one tests began.

Before the test could start the user signed an informed consent form, since the tests would be recorded. The tests were conducted in two different rooms so that the tests could run in parallel. This was necessary since we only had a short time span for performing both testing and a focus group. A visualisation of how the test session was performed is presented in figure 3.25.

Since three prototypes would be tested a schedule of which order the prototypes would be presented to the users was followed to decrease the risk of bias. The schedule is displayed in table 3.26. The tests took around 45-60 minutes.

Focus Group

After all tests were conducted a focus group session was held with four of the five users, as one of them had to leave ahead of schedule. However, this was a highly valuable session, since the participants had several opinions and benefited from discussing the different prototypes with each other. Before they were asked any specific questions, they all had the opportunity

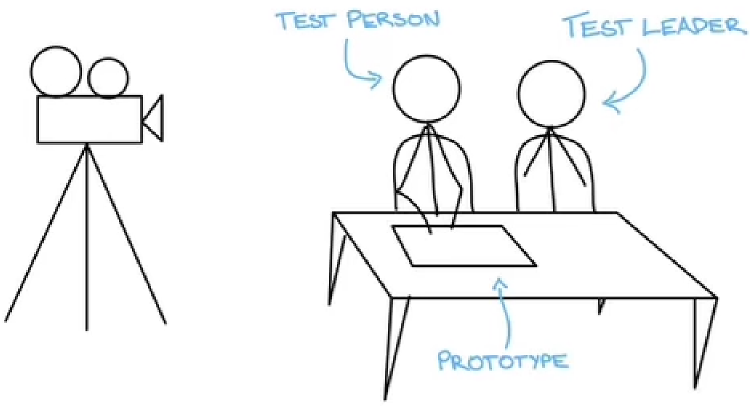

Figure 3.25: Drawing of the scenery during the Lo-Fi tests

Table 3.26: Table of the scheduled order of prototypes during tests.

Test Persons (TP) 1-5	First	Second	Third
TP1	Prototype 1 (P1)	Prototype 2 (P2)	Prototype 3 (P3)
TP1	P2	P3	P1
TP3	P2	P3	P1
TP4	P3	P1	P2
TP5	P1	P2	P3

to share their thoughts openly on the separate prototypes or problems they have with their current system used on ships today. The prototypes were put out in front of the participants on the table to be able to get an overview. The differences could be displayed easily, and they could point at the different prototypes when referencing while they were speaking.

Many questions regarding the different design aspects were asked and what the users thought of the different views in regards to their experience. They were also asked to share specific situations they might have use for the newly designed views and whether they think they would actually use it instead of the traditional alarm list. Lastly, we followed up with questions about examples of situations where alarm flooding might occur.

Due to the circumstances, this was the first day we could speak with users. Therefore, this session was also meant to create a solid ground for the upcoming iterations, as the decisions were now based on first-hand user-centered data. The session took a little over 30 minutes in total, then the session was completed.

3.4.3 Mid-Fi

The second test was performed by four test participants. It was originally planned that five users would participate in the tests but one of them cancelled since he could not make it. The tests were approximately 45 minutes and consisted of a consent form, a brief introduction to the test, the performance of the test and afterwards, filling out a SUS questionnaire. The tests were conducted on Saab's premises in Karlskrona by users from FM within the Swedish Navy and all tests were recorded for our benefit.

Since the prototype was made in Figma, the prototype was displayed on a large screen and the user was seated in a chair in front of it. To make the testing process more efficient both of us were present at each test, one was the test leader and the other took notes and observed the user. The test setup is presented in figure 3.27.

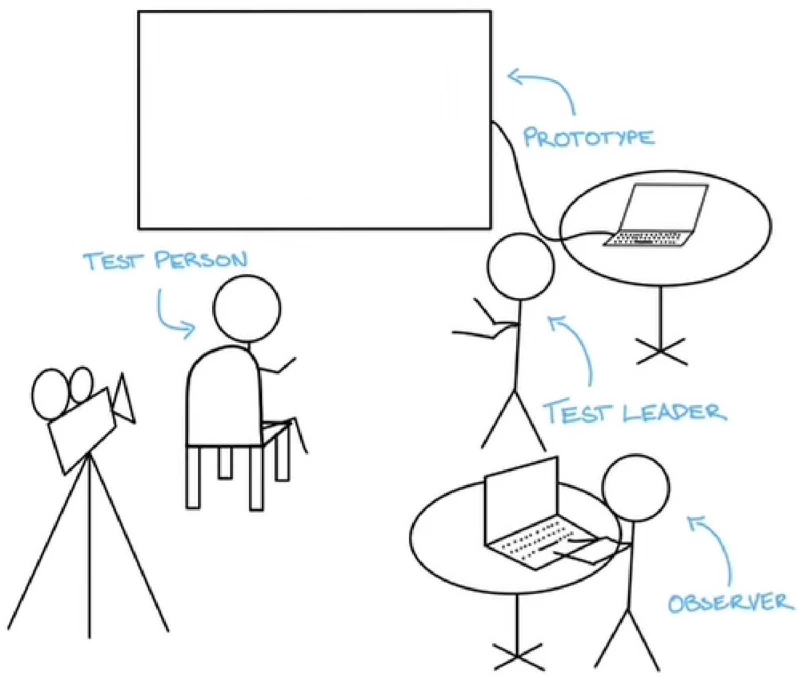

Figure 3.27: Drawing of the scenery during the Mid-Fi tests

The test leader explained different scenarios about the submarine and asked the user questions. The scenarios were similar to the ones from the Lo-Fi test session but not in the same order, slightly shorter and somewhat modified. For example, the dates and times were changed and the views were presented in a different order.

The user answered what their impressions of the design were and explained how they would interact with the still images. The test leader switched between views on the prototype and most comments were noted by the observer.

When the test was finished the users filled in a SUS questionnaire in a Google form.

3.4.4 Hi-Fi

In contrast to the first two usability test phases conducted in this study, this one was not planned in coordination with FMV and FM. This was due to their schedules not matching the time frame set for the Hi-Fi tests. Instead, it was carried out in collaboration with relevant Saab colleagues at the office in Malmö.

Among the five people testing the prototype during this phase, there was one who had previously been stationed on a submarine by FM in the early parts of his career. The other were colleagues from our team, some who worked more with SCMS and some who had never used the system. This was also the first time that women have been testing our solutions. As there is usually a majority of men within FM, we wanted to include at least one woman in our usability study, but managed to get two to participate.

The test included reading and signing a consent form and this time we were only recording the screen and voices. There was no need to record the test with a video camera anymore since the users could interact with the prototype in this test.

The Hi-Fi prototype was shown on a large TV screen, similar to the previous test. The users were seated in front of it by a table and could navigate using a mouse. See figure 3.28 of the setup for the Hi-Fi tests.

During the test, one of us was the test leader and one was the observer, constantly taking notes of the user behaviour and comments about the system. The scenarios were still the same as in earlier stages. When the test was completed, a quick debrief using a SUS questionnaire was filled out by the users.

3.5 Evaluation

All data from the tests were compiled and analysed from the recordings and questionnaires. The first test session generated the most qualitative data since it tested three versions and the questions and tasks were open in nature. This was valuable since general user feedback was the most important feedback in the early stages of prototyping. The second two test sessions, Mid-Fi and Hi-Fi, gave more quantitative data. This was good for evaluating the chosen designs and how to fine-tune aspects of the design that did not work. The Hi-Fi test specifically generated quantitative data in the form of the number of clicks to complete a task. Measuring time was an optional way of collecting data and was quickly disregarded since tasks during the test were mixed with opinion-based questions. This led to test participants answering questions and simultaneously clicking around in the prototypes and in some cases figuring out how to achieve later tasks.

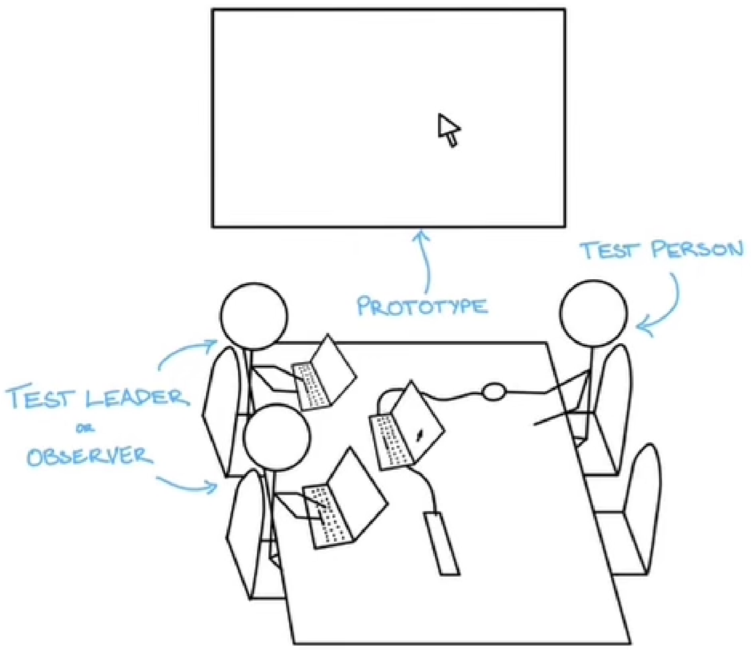

Figure 3.28: Drawing of the scenery during the Hi-Fi tests

3.5.1 Lo-Fi

All test sessions were influenced by the confidentiality required by Saab and FM. This means that transcribing the videos filmed during the tests with any tool not approved by Saab was out of the question. This led to several days of work in the Lo-Fi test phase, where the videos of the tests and the focus group session were manually transcribed. In the end, the most valuable feedback and quotes from the tests were summarised in tables where they were categorised.

From this categorisation, a prioritization list of the users' problems, new feature ideas and improvements was created, which later led to the prototypes Mid-Fi and later Hi-Fi.

Prototype 1, P1

Feedback on P1 is presented in table 3.29. The observations (in the table) in the bolded caption are specifically regarding the original alarm list. The other observations regarding the design of the system in general are not bolded, which will apply to the later prototypes as well.

In short, the minimised alarm list with three alarms was too small and the maximised version contained too many alarms. This led to the participants having a hard time understanding the sequence of events and placement of the origin of the alarms. Four participants explained that they wanted specific filter functions to deal with the large amount of alarms.

All members wanted to acknowledge all alarms at once and learned to use the maximise button quickly. Three participants noted and complimented the order of appearance of the alarms and four mentioned and explained how the colours in the systems worked.

In addition to the feedback in the table, all participants wanted a go-to-page or button. They wanted to see the state of the systems, but this feedback was excluded from our official results since it was out of the scope of the report.

Table 3.29: Test results from P1 summarised

Observation	Test Persons (TP) 1-5	Feedback from TP
The 3 alarms- alarm list is to small	TP1, TP2, TP4	They all agreed on a larger list
Order of the alarms	TP1, TP3, TP4	Usually the newest alarms is at the top and will burry older (important) alarms, they mention that they are worried to miss something
Too many alarms in the alarm list	TP1, TP2, TP3, TP4	There is too much information, it must be filtered in some way, they are afraid they will miss something
Understand the meaning of the colors	TP1, TP2, TP4, TP5	They do not know submarine systems but understand that red is important and grey is not Note: TP3 "I do not look at the colors"
Do not know where in the boat the alarm was triggered	TP1, TP2, TP4, TP5	They try to read the text and mentions that "that would be a good function (to know where...)"
"Acknowledge" alarms	TP1, TP2, TP3, TP4, TP5	They want to "acknowledge" all alarms at once
Understand (learns) to use maximise	TP1, TP2, TP3, TP4, TP5	
Do not understand sequence of events	TP1, TP2, TP4, TP5	Reads the alarm list, but have a hard time to understand
Asks specifically for a filter function	TP1, TP2, TP3, TP5	Want to group alarms, want to filter a certain type of alarm and want to set a delay in showing alarms

Prototype 2, P2

Feedback on prototype 2 is presented in table 3.30.

All participants had trouble navigating to the new submarine view and all but one instantly explained that they understood the colour, criticality and placement of the alarm symbols on the submarine. They understood where the alarms had been triggered, which was not clear in P1. Four participants wanted to click on the circles to see corresponding alarms in the alarm list or to see more information on the alarms portrayed. Four participants tested the front and back zoom of the submarine and two of them shared that they would like to zoom in even further.

The three participants that did not start with P2 did not notice or comment on the new design of the alarm list and two participants still felt that there were too many alarms in the list. There were also some issues with understanding the sequence of events for three participants.

Table 3.30: Test results from P2 summarsed

Observation	Test Persons (TP) 1-5	Feedback from TP
Can not find the submarine view	TP1, TP2, TP3, TP4, TP5	They all need a hint to find the view, one TP mentioned that "I would like to have the page open all the time..."
Understands the alarm symbols on the submarine	TP1, TP2, TP4, TP5	They understand the color, criticality and placement.
Click on the alarm symbols	TP1, TP2, TP3, TP5	They would like to click on a circle and want the corresponding alarms to appear in the alarm list
Submarine front and back zoom	TP1, TP2, TP4, TP5	They like the zoom-in feature, two participants would like to zoom-in more
Still to many alarms	TP1, TP2	There are still too many alarms and too much information
Does not notice/comment on new alarm list design	TP1, TP4, TP5 note: TP2, TP3 started with P2	One liked the idea when it was explained
Do not understand sequence of events	TP1, TP2, TP3	Reads the alarm list, but have a hard time to understand

Additional feedback that was expressed concerned the learning curve of these complex systems. These thoughts were briefly mentioned during P1 but explicitly mentioned during P2. Some translated quotes:

"The system must be simple, XX weeks of training and you must know the system"

> *"It should not take 2 years for someone to feel confident in (their ability in) the system"*

> *"If you did not have experience, this (referring to the alarm list) would probably just become a list of anything"*

The last quote explains that the significance of the alarms gets lost in the amount of information. They also mentioned that a large amount of alarms can make the user nonchalant towards the alarms.

Prototype 3, P3

Feedback on P3 is presented in table 3.31.

Four participants needed hints to recognise the drop-down menu but they liked the feature, one in particular mentioned that they wanted to click up the drop-down feature directly after they clicked down. All participants had a hard time locating where the alarm had been triggered and three asked for the previous view of the submarine. Only two participants could navigate and find the time-line view. This view helped three participants to understand the sequence of events. The circles on the time-line were clicked on by all participants and they all wanted corresponding alarms to appear in the alarm list.

The opinions on the 7- and 14-day time-circles were split between the participants. All participants said that the graph was confusing and hard to read. After consideration, they mentioned that it could be used more effectively with filters, like searching for a specific alarm. When probed four participants could find trends and tendencies in the graph and three could make up scenarios on why the graph looked like it looked. Although the graph grew in popularity the longer participants looked at it and got used to its appearance, the opinion on whether it was a good addition to the system was split.

Additional notes were that two participants started to notice that one scenario was meant to be a non-critical alarm flood. This may be because of the format of the alarm list or just because they had experienced the scenario at least one time prior.

Focus group

The summarised takeaway from the focus group is that they would like a combination of the three versions because the submarine, time-line and time-circle help provide context to the alarms in different ways. Pictures and graphs can help visualise text and help the user take more accurate action to prevent critical incidents.

The submarine view was good for current alarm supervision and helps provide an understanding of the location of origin of alarms which then can be handled faster. The time-line view was best for understanding the sequence of events during the evaluation of an alarm flood directly after the event. The time-circles were no one's favourite, but were a better way to evaluate and analyse alarm systems over a longer time. An alarm list or a time-line would not provide the same overview. They give an example of an alarm going off every third hour, which would be easy to spot in the time-circle and difficult in a list format.

Most participants did not notice the difference in alarm list design. One participant noted that the drop-down list was the best and easiest to get an overview of the alarms. Another noted the danger of having hidden alarms that can go unnoticed.

Table 3.31: Test results from P3 summarised

Observation	Test Persons (TP) 1-5	Feedback from TP
Do not click on drop-down	TP1, TP2, TP4, TP5	They needed a hint to click on the drop-down, one TP said that they wanted to click up immediately
Do not know where in the boat the alarm was triggered	TP1, TP2, TP3, TP4, TP5	Three asked for the submarine view
Do understand sequence of events	TP1, TP2, TP3,	Mentions that they understand the sequence of events better with the time-line
Can not find the time-line view	TP1, TP4, TP5	Note: TP2 and TP3 finds the time-line
Click on the alarm symbols	TP1, TP2, TP3, TP4, TP5	All want to click and wants alarms to appear
7 day and 14 day time-circle	TP1, TP2, TP3, TP4, TP5	All said that the graph is confusing, they also mention that it would be more useful together with filters, at the end can four participants find trends in the graph and three make up possible scenarios

A function missing currently was the park function or "deal-with-later"- function, as the users described it. They wanted a way to filter out non-critical alarms to create a smaller list of the highest prioritised alarms.

They also highlighted how important it is to learn the system fast and to feel secure using it. With enough time, a user can learn a badly designed system and the experience will cover it. There should be a tool to help users make educated decisions if unsure.

The participants discussed the use of colour and how it compares to their current system.

Note: they discussed how to present states of systems as well as using warnings before alarms, but this is outside of our scope.

Takeaways from Lo-Fi Test

After all the data, information and feedback were analysed we compiled a list of features and aspects to change for the next iteration. The original SCMS design had many flaws, but so did the other prototypes.

For the next iteration, we decided to:

- Make a single prototype containing a submarine view, a time-line and a time-circle. This came from the focus group, that each view was useful in different scenarios. A simple history of previous alarms after an alarm flood did not provide any insights into the sequence of events or placement of the origin of the alarm in the submarine. We decided to keep the time-circles even though they said they did not like it at first. It was still better than an alarm list of seven days and the participants could see trends in the data.

- Combine the drop-down list and tab-in list. We could not identify which version improved the design and performance the most. The participants focused mostly on the alarm text and one explicitly said they thought the format was an error, saying "it is hard to draw straight lines".
- Make a function to acknowledge all alarms. In the test, we asked how they wanted to acknowledge alarms and all participants agreed on this matter.
- Make a function to park alarms. Many participants said that there were too many alarms and wanted a specific function to park or "deal-with-later" to make the alarm list shorter and prioritised.

3.5.2 Mid-Fi

Because we both were present during each test and the tests were shorter, with only one prototype to be tested, the transcription was quicker for the Mid-Fi test. Each task the test person executed was numbered, questions and answers were written down as well as specific body language. The significant feedback on the design and relevant functions was highlighted and collected in a separate document. All the information was prioritised in order of most to least important, and a category of things to write about in the report was also created. The categories were: implement now, need to be discussed how to implement (still important), nice to have and write about in the report.

Prototype

Feedback on the Mid-Fi prototype is presented in table 3.32.

Some of the general feedback on the design was that, as in previous tests, all participants agreed that the small 3 alarms- alarm list was too small and one participant suggested that it could contain at least four to five alarms. Only one participant needed a hint to first use the maximise button and all participants understood the use of color. The participants were given a task to remove alarms and no one understood the park function. Instead, they wanted to acknowledge alarms, two wanted to acknowledge all at once and two wanted to be able to choose to acknowledge all or specific alarms.

Specific feedback on the new alarm list design was that all participants noticed the following alarms and two liked it. All participants but one found the drop-down function and liked that all information was not visible and that the alarms were grouped. Two participants did notice a critical alarm hiding in the drop-down list and addressed that there needs to be some sort of signifier so as to not miss any critical alarms. When observing the maximised list, all participants mentioned that they wanted to sort or filter to get the most important alarms. Three participants wanted to click on the header panel "status" to sort the alarms based on criticality, like in an Excel sheet. Two participants still felt that there were too many alarms in the list.

Feedback on the submarine view and time-line were both positive, all participants liked them and they wanted more details, specifically on the submarine. The navigation between

Table 3.32: Test results from Mid-Fi prototype summarised

Observation/Task	Test Persons (TP) 1-4	Feedback from TP
The 3 alarms- alarm list is to small	TP1, TP2, TP3, TP4	They all agreed on a larger list, one suggest 4-5 alarms
Understand (learns) to use maximise	TP1, TP2, TP3	TP4 needed a hint, wanted to click on the alarm-panel
Understand the mening of the colors	TP1, TP2, TP3, TP4	TP4 greatly disliked the gray color
"Acknowledge" or park alarms	TP1, TP2, TP3	Two want to "acknowledge" all alarms at once, and two wanted to be able to choose all or specific alarms. No one understood parkera
Maximised list	TP1, TP2, TP3, TP4	They all noted "Following alarms", two liked it and one did not
Find drop-down	TP1, TP2, TP3	Two found the drop-down immediately and one was slower. The first two noted that a red alarm was hiding in the list
Filter critical alarms	TP1, TP2, TP3	They all wanted to click on the status header-panel to sort the column in the list after criticality ("like excel"), TP4 wanted a "prioritize"-button
Too many alarms in the alarm list	TP1, TP4	Even with the drop-down there was too many alarms
Does not notice/comment on filter-menu on the submarine view	TP1, TP4	TP2, TP3 can filter
Can navigate between views	TP1, TP2, TP3	
Submarine as an overview	TP1, TP2, TP3, TP4	They all liked it and wanted more details on the submarine
7 day and 14 day time-circle		TP1 and TP4 did not understand or like it, TP2 and TP3 understood and said it would be good for analysis
7 day search-function	TP1, TP2, TP3, TP4	They all used the search-function
Time-line as an overview	TP1, TP2, TP3, TP4	They all understood the time-line
Interacting with submarine, time-circle and time-line	TP1, TP2, TP3	They wanted to click on the circles, squares and balloons and see corresponding alarms highlighted in the alarm list

the views was easy for the three participants. The 7-day and 14-day time-circle was not liked by TP1 and TP4 and they did not understand what it portrayed or what it would be useful for. TP2 and TP3 were more positive, they understood the graph and said it would be good for analysis. In all three overviews, submarine, time-line and time-circle, three participants wanted to click on the alarms in the views to see the corresponding alarms be highlighted in the alarm list.

Some additional feedback was that one participant mentioned that the date often is unnecessary in a current alarm list since the alarms are studied second by second. Another participant mentioned that the time-line would be specifically good for analysing alarm floods and that the time-circle makes analysis further back more efficient (no need to look in ledgers to find data).

By observing the test participants it was clear that some were more comfortable than others using the system. It presented itself in quotes like "I am not as good at computers as the young people are" and talking more about their old system than the prototype at hand. Even though some feedback on test results may be due to them being reluctant to change, the feedback is still valuable since it is reflective of users of the system overall.

Debriefing SUS Results

After each test in the Mid-Fi process, the test person got to fill out a SUS questionnaire. These questions are specified in Appendix A.1.2. The individual results from the questionnaire are presented in figure 3.33. Additional details on the results of each question answered by every user are available in the appendix A.1.2. Each test person has a final SUS score and the average and total score for the Mid-Fi prototype was calculated to be 60 points. This indicates that the design still needs some work til the usability is improved. The optimal score would of course have been higher than 68, but as one of the test persons, TP1, wrote in the open section for any last comments when filling out the SUS, the interaction part was missing in this prototype with pictures. The quote mentioned by TP1 is stated in Swedish below.

> TP1 - *"Would be nice to be able to click around in the system myself"*

Another quote that is worth mentioning from the SUS is the following.

> TP2 - *"The system is easy to use and so far seems to have the most important functionalities. I believe it is important with great overviews of alarms and subsystems. Make sure to focus on making the main functionalities good and not implementing too many new functions that are not used in the end."*

This quote is interesting since it mentions the importance of getting the fundamental and most important functions to work properly in the system before too many functions are implemented. The statement about there being too many unused functions in several systems the test person has used before is as important to have in mind when designing a large and complex system as SCMS.

The two remaining test persons had the following comments. The first one confirms the design choices previously made by Saab, with alarms represented by a yellow colour and critical alarms by a red colour. The second comment is a general recommendation to make the system simple.

TP3 - "*Good colours. Clear with red and yellow circles that the eye is drawn to.*"
TP4 - "*Aim for simplicity.*"

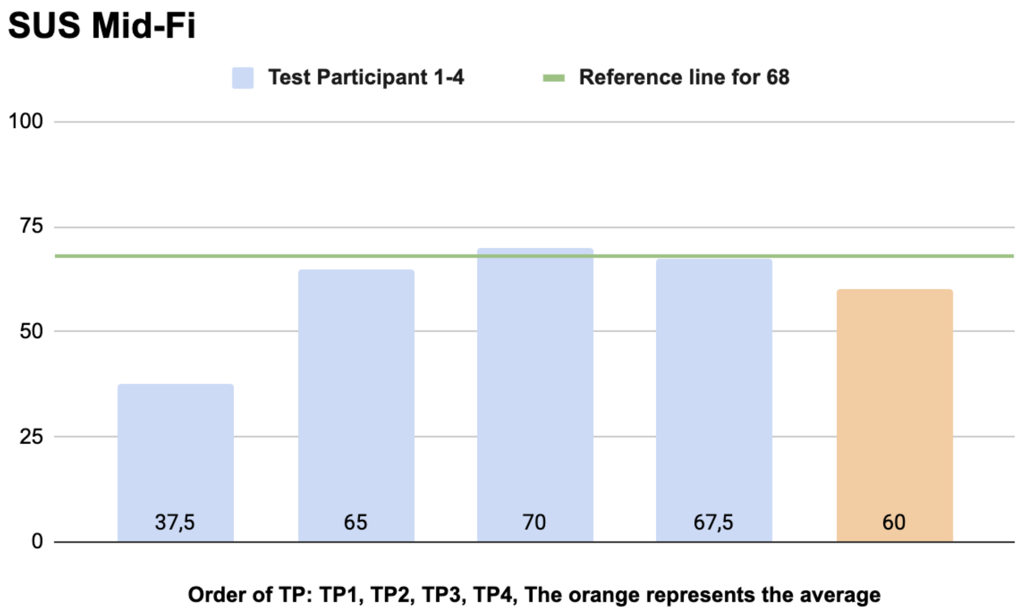

Figure 3.33: Results from the System Usability Scale questionnaire for the Mid-Fi test phase. The total score for the Mid-Fi is summarised in the orange staple to the right.

Takeaways from Mid-Fi Test

The changes to be made for the next iteration are listed in the following lists. As mentioned before, all feedback was categorised into different categories to prioritise quick changes and more difficult ones.

Implement now (quick changes):

- Change the following alarm or header of the alarm group to something to include a group of similar alarms. The following alarm caused confusion for one participant and it is bad design to assume users will connect following alarms with similar alarms when it is a misnomer. Another change is to make the title more prominent and stick out from the list to prevent users from mistake a single alarm for a group alarm. And to make the drop-down arrow larger, if a signifier can not be seen it will not be used.

- Remove dates when it is not necessary. There is a lot of information already in the alarm lists and it would be interesting to test if results changed without a column of data.

- Make a new status symbol for a group alarm with both critical and normal alarms. This is very important and an obvious improvement of the alarm list design.

- Add criticality status symbol to the filtering menus. This is not necessary for the function of the system but a simple way to improve the users' mental model of the importance of the alarms.

- Make alarm lists so the user can interact with certain alarm indicators on the submarine view or time-line. Add functionality to the prototype participants have been asking for.

- Make sort function like in Excel. Many participants asked for this feature and it is a good design to take advantage of users' mental models and already known design elements/functions.

Need to be discussed how to implement:

- Make park a much clearer function. Many participants have asked for this function but it will not be used if it is not intuitive how to use it. They did not understand what a park button could do.

- Submarine zoom, several ideas were emerging from the tests on how they wanted to zoom in on the submarine. Either use a signifier like a magnifying glass, click on the submarine or select with a computer mouse.

- Acknowledge all alarms, it is important that this function can be performed with as few clicks as possible to save time during an emergency.

- Time-circle was difficult to understand but we did not want to give up on it. The date needed to be more distinct and easy to understand.

- Navigation between views needed to be implemented. In the original system, there was an additional touchscreen mostly used for navigation, but in this case, we needed another way to be able to test the next prototype.

Prototype

3.5.3 Hi-Fi

The screen recordings in the Hi-Fi tests were similar in workload to the tests from the Mid-Fi testing phase. While evaluating the Hi-Fi prototype and the feedback from the participants, the data was once again categorised.

The test result of the Hi-fi prototype is summarised in two tables as seen in table 3.34 and table 3.35. The tasks are listed in test order and the table is split into the two scenarios that were presented to the participants.

All participants found the maximise button within three clicks and the second time they were asked to explore the alarm list they remembered the button. Two participants wanted to test a new way to maximise, click on "new" and click on the alarm panel. Later in the test, we wanted to see if the new button would be used to explore the alarm list and only one did.

Four out of five managed to acknowledge all alarms with only two clicks, but it was harder to find the acknowledged alarms in the history view. Participants looked again in the

Table 3.34: Test results from Hi-Fi prototype summarised part 1

Observation of task	Result/feedback from 5 TP	Note
Can find and use maximise (click on maximise button)	First attempt: 3 TP Second attempt: 1 TP (+1 click) Third attempt: 1 TP (+2 click)	
"Acknowledge" alarms (click on select all and then "acknowledge" button)	First attempt: 4 TP Second attempt: 1 TP	
Can find "acknowledged" alarms (navigation-select-button and then history)	First attempt: 3 TP (+1 click) Second attempt: 1 TP (+2 click) Third attempt: 1 TP (+6 click)	Most TPs either clicked on maximise (to se that the alarms were really gone) or on filter first
Comments on appearance on history-view	TP1, TP2, TP3, TP4, TP5	They all note the filtering system, 3 want to know the status of the alarms
Have learned to use maximise button (click on maximise button)	First attempt: 3 TP Second attempt: 2 TP (+1 click)	One wanted to click on the panel and another tested the new-button
Can find drop-down function (click on drop-down list)	First attempt: 5 TP	They all note the "following alarms" header
Can sort the alarms by status (click on the status-header-panel)	First: 2 TP Second: 2 TP (+1 click) Third: 1 TP (+2 click)	3 TP clicked on filter first
Comments on appearance on the maximised alarm list	TP1, TP2, TP3, TP4, TP5	They all understand the colors and the group-alarm-symbol, 2 say they like the drop-down function
Can navigate to the submarine view (navigation-select-button and then submarine overview)	First: 4 TP Second: 1 TP (+4 click)	
Can navigate between views (two ways: using tabs or click on boat)	Click on tabs: 2 Click on boat: 3	
Comments on appearance on submarine view	TP1, TP2, TP3, TP4, TP5	They all liked it and they understood the color, criticality and placement of the symbols.
Comments on alarm filter on submarine view	TP1, TP2, TP5	They all wanted to click and interact with the filter-function

maximised alarm list or the button "filter" and then found the navigation to the history view. All participants wanted to interact with the filtering system unprompted.

The navigation between views throughout the prototype worked well after the first task. The drop-down-function was found by all participants and they all noticed the following alarm. At first, two participants said they liked the drop-down function and later during the test, two additional participants agreed. The next time participants were asked to examine the alarms, all participants knew the drop-down function. When asked to sort the alarms according to criticality, two participants clicked on status in the header panel and the rest clicked on the filter button at first.

The submarine view was popular with the participants and easy to understand. Two participants used the tabs to navigate the boat and the rest clicked on the boat where they wanted to zoom in.

All participants knew how to park alarms, similar to acknowledge and four participants could find the parked alarms on the first try. Some of the participants had clicked on the filter button previously but this result was still surprising.

The time-line view was intuitive and the participants understood the size, placement and colours of the balloons. An important note, from one participant, was that the intervals time function could make users prone to make mistakes since there is no date on either time-box. All participants could use the interval function.

The 7-day time-circle was still contentious. Two participants liked it instantly, two participants liked it over some time and one did not like the time-circle. All participants could search for a specific alarm and when asked, in the 14-day time-circle, all could read trends in the data presented.

Debriefing SUS Results

In the Hi-Fi usability tests the SUS questionnaire was used once again to conduct a proper debriefing. The same questions were asked during this test as in the Mid-Fi SUS, hence they are found in Appendix A.1.2. The individual results from the Hi-Fi questionnaire are presented in figure 3.36. The table is constructed in the same way as for the Mid-Fi tests, although this time the total score for the Hi-Fi prototype was calculated to 83 points. This indicates that the usability of the design has improved significantly since the previous tests.

The following comments were written by each test person in the open section during the SUS. One test person actually had a previous career working on a submarine within FM and knows very well that the alarms and alarm floods are a problem in SCMS, wrote:

> *"I think the system has captured the need to be quick and easily comprehend the amount of information. Some of the visualisation, specifically the one with the weekly displays, could be done in another way that is easier to understand."*

TP1 stated that the system information was easy to understand and navigate. However, the one and two-week history pages were less intuitive and could benefit from an alternative design. On the other hand, the second test person really liked the idea of the weekly displayed history pages.

> TP2 - *"Everything looks good, I like the weekly views of the alarms!"*

The remaining three answered the questionnaire's open question with the following statements. TP3 describes the perspective in which he answered the questions in the SUS.

> TP3 - *"I have answered from the perspective of an inexperienced submarine personnel, with a background in programming and automation."*

TP4 praised the system with the quote "Well done!". TP5 was in agreement with TP4 on this and additionally noted that the system was easily managed, easy to navigate and featured excellent functionality.

TP4 - *"Well done!"*
TP5 - *"Good work! Relatively easy to navigate, good functionalities and easy to use."*

Takeaways from Hi-Fi Test

If we (or Saab) were to make a fourth version of the prototype we would focus on:

- Switch the order of dates in the time-circle. Two participants mentioned that it was more intuitive to look from the outside and into the middle of the circle.
- Add more detail to the submarine when it is zoomed in. There was also an idea to make a 3D model of the boat to increase understanding of placement in the boat.
- Find a clustering algorithm to make the time-line more bullet-proof. When testing the time-line several participants have asked what is the best way to "group" the alarms is, we have eye-balled the amount of alarms and estimated placements.
- Investigate the colouring of the alarm groups, both in the alarm list and on the time-line. What is the difference in perception if the alarm is yellow with a red center position or red with a yellow center position?

Table 3.35: Test results from Hi-Fi prototype summarised part 2

Observation of task	Result/feedback from 5 TP	Note
Can maximise (two ways: click maximise or new)	Click on maximise: 4 Click on new: 1	2 who clicked maximise also mentioned that new was an option
Have learned the drop-down function (click on drop-down list)	First attempt: 5 TP	
Can "park" alarms (click on select all and then "park" button)	First attempt: 5 TP	It took longer time for all TPs to perform this task. 3 TPs noticed change in the park button after clicking on it
Can find the "parked" alarms (click on filter and then show parked)	First: 4 TP Second: 1 TP (+5 click)	3 TP had looked at the filter-button previous, and most TPs would like a more obvious button
Can navigate to the time-line (navigation-select-button and then time-overview)	First: 3 TP Second: 2 TP (+1 click)	2 TPs first clicked on history
Can filter time-line by time (click on interval in filter-menu)	First attempt: 5 TP	
Comments on the appearance of the time-line view	TP1, TP2, TP3, TP4, TP5	They understand the size, placement and color of the balloons. One comment on the time for a 24-hour period
Understands that balloons are clickable	TP1, TP2, TP3, TP4, TP5	
Can navigate back (three ways: clear filter, tab or navigation-select-button)	Clear filter: 2 Tab: 2 Navigation-select: 1	Note: they also wanted to click on interval
Can navigate to 7 day time-circle (click on tab)	First attempt: 5 TP	
Can search for specific alarm (click on search-function "type of alarm")	First: 5 TP	They question whether it should be a search function or drop-down
Comments on the appearance of the 7 day time-circle	TP1, TP2, TP3, TP4, TP5	2 TPs said easy to understand, 2 TPs needed some time to understand, 1 TP did not understand
Can navigate to 14 day time-circle (click on tab)	First: 5 TP	
Can read trends in the graph (few alarms 01-02, many critical alarms 20-23)	TP1, TP2, TP3, TP4, TP5	They all can read trends

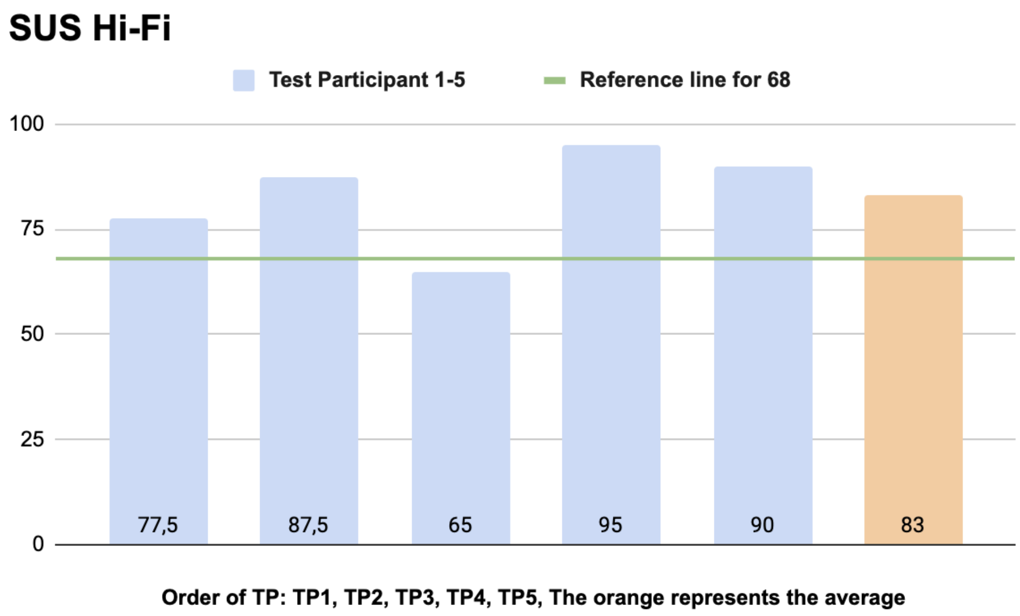

Figure 3.36: Results from the System Usability Scale questionnaire for the Hi-Fi test phase. The total score for the Hi-Fi is summarised in the orange staple to the right.

Chapter 4
Discussion

The fourth chapter will dive into a discussion about the different processes, methods and results for this study. The results will also be presented with the literature study from Chapter 2 in mind and discuss the test participants' relevance for the usability testing. The concluding section of the discussion will focus on further development, offering recommendations for Saab and the teams working with SCMS.

4.1 Limitations

With an already advanced system being continuously developed, there could have been a much larger study or many more minor ones made on the SCMS alone. This project, however, was limited to studying and improving the interface pages and functionality mainly involving alarms. The specific interface pages the study was limited to test and further develop in UX design aspects are specified in Chapter 3, in the section called Prototyping. The study did not focus on alarm handling systems using multimodal effects like sound or vibration, since it is not conducive to the silent missions of submarines.

Another limitation is that the scenarios will focus solely on internal battles, as SCMS is designed primarily for monitoring internal systems and therefore mostly internal battles in general.

4.2 Process and Methods

Overall, the process of this project has progressed quickly, with significant effort to meet deadlines and complete tasks, especially prototypes on time. The necessary deadlines impacted the overall project plan leading to the conduction of the originally unplanned Mid-Fi test.

4.2.1 Reflection from Lo-Fi Testing

During the Lo-Fi test, we noticed a general flaw with the Lo-Fi design. The different formats of the alarm lists in the three prototypes were too similar to actually get comparable data. Since we knew what we were testing, the alarm lists, we deemed that the changes were apparent and a key feature of the design. The test participants on the other hand got overwhelmed by buttons they could not use, the alarm text they did not understand and they did not know what features that would be tested.

The alarm list variations, original, drop-down and tab-in, should have been tested in another format since they did not benefit from a minimal design Lo-Fi prototype. A Lo-Fi prototype generally benefits broad stroke features and design aspects, which is why the submarine view, time-line and time-circles accumulated more feedback and opinions.

4.2.2 Mid-Fi Addition

In our opinion, the Mid-Fi test was a significantly successful part of our project. Especially since scaled-down tests and above all, fast but insightful iterations, in general, are more common and display the reality of the professional UX work. While the Lo-Fi tests and focus group provided valuable insights by allowing us to engage with users of similar systems for the first time, the Mid-Fi test was conducted with more information and a better understanding of our users and their needs. The Mid-Fi test offered crucial insights on how we could improve the Hi-Fi further. This resulted in some major changes such as the introduction of the new function for parking unwanted active alarms.

4.2.3 Overall Process Reflection

In the ideal and best possible version of this study, we would have preferred to engage with our users from the start. In a usability study, this could have been achieved through either a survey, interview or direct conversations with users about the system. This approach would have provided a user perspective that was not yet fulfilled prior to the Lo-Fi tests.

This was the first usability study conducted on SCMS and its alarm management at Saab, making it challenging to identify and address existing problems from a user perspective. Our team consisted of many talented engineers and programmers, but we lacked a fully dedicated UX designer for SCMS. This is why we chose to do a comparative study with the existing system in the Lo-Fi tests rather than solely designing new pages and features. The system needed to be tested with real users for us to begin to comprehend what actions were necessary.

4.2.4 Usability Test Participants

The participants in the user tests were not that far from our personas in terms of gender, age, occupation and rank within FM. At least not during the Lo-Fi and Mid-Fi tests. Since the participants were contacted directly through our supervisor at Saab, the participants were all voluntarily participating as well. Even though FM are incredibly occupied with their hectic schedules they made time for our usability testing and some even kept asking during the project about the progress and results. This shows that the test participants were highly motivated, even though they participated during working hours. The participants were in the correct target group since they were all men, within the navy and with the same type of rank and roles as our personas and our real target group, but there is still a few aspects that differ.

The first one is that the real target group has the expertise of submarines that all participants except the first one from the Hi-Fi test were lacking. During the usability tests, the majority even mentioned it themselves. They could see that they were missing key knowledge and probably would have known more about the scenarios and alarms displayed in each one if they were submarine personnel.

The second one is the lack of the younger generation among the test participants. Even though our primary users are usually between 30-45 years old, the secondary user who is sitting next to the primary user at all times when using the system will still have to learn, navigate and use SCMS. This role is usually occupied by the younger generation doing their mandatory military service, meaning they are around 20 years old and have no prior experience working in SCMS or any other similar monitoring system. This part of the target group, even from the crew on ships was difficult to get a hold of and therefore we chose to test our Hi-Fi prototype on a young colleague in his early 20s.

Although we never met our main target users in real life and never got to test our prototypes on them directly during this study, we did get plenty of results and great feedback that led us in a direction where the prototype was improved.

4.2.5 A Certain Type of User

An interesting aspect noticed during the first two test sessions with the Lo-Fi and Mid-Fi prototypes was that most users seemed apprehensive at first. Due to the participants' background in military training, they were all strict and professional. A conclusion from observing the participants was that they seemed determined not to make a mistake and therefore put high pressure on themselves to read the alarm text to try to understand the situation/scenario. This led to a few participants not noticing the bigger picture, like not noticing the change in the alarm list format.

Since the prototype testing sessions were supposed to be explorative and to generate new ideas we needed to intervene and encourage the users to "think outside the box". Participants were probed several times to share their opinions on the prototypes and were heavily encouraged when they shared both positive and negative feedback. When asked what they would like in a system or how they would like to perform a task without any constraints or limits, they were astounded.

This might have influenced the results during the test sessions but since the test participants share background and personality traits, with actual users it is still valid.

The only note is to consider that the user group might give harsher feedback since a single mistake could be the difference between life and death in their perspective.

4.2.6 SUS questionnaire

During the Mid-Fi and Hi-Fi test sessions, several participants asked what perspective they should fill out the questionnaire, referring to the fact that they did not have complete knowledge of submarines and their systems. The question: "I think that I would need the support of a technical person to be able to use this system" was especially difficult to answer. We told the participants that asked the question, to imagine that they were submarine personnel and disregard the lack of understanding of specific submarine knowledge, for example, the alarm text. We also asked them to note in the last open text field if they answered the questions from another perspective. However, we did not receive any specific comments on this matter, so we are not entirely sure if they answered the questions from the mentioned perspective or from their own. There is a possibility that the lack of submarine knowledge might have influenced the result of the SUS score to a more negative score than it could have been with the correct knowledge. Nevertheless, we are highly satisfied with the clearly increased score from Mid-Fi in comparison to Hi-Fi.

In addition to the SUS questionnaire, we could have added questions about the user experience, especially during the Hi-Fi test. This is to gain even more insight into the users' thoughts and opinions but alas we did not think it would be necessary. If there would be another iteration of testing more questions would be added.

4.3 Results

The result of the study performed during this thesis is presented in the following section. First a brief summary of the final prototype and discussion of the controversial time-circle. Then all research questions are answered by drawing insights and conclusions from the design process and literature study.

4.3.1 Prototype and Time Circle

The final result of the prototype is the Hi-Fi prototype and it still has elements in need of improvement. The alarm list design with drop-down and tab-in was a good solution to the problem of displaying too many alarms (and thereby overwhelming the user) and displaying root cause and correlating alarms. The submarine and time-line provided an overview directly and just after the alarm floods, but the time-circle was divisive.

The time-circle usually split the test participants, some liked it and some did not, but most time participants grew to like it after some time. The presentation of the data seemed at first to be intimidating, but at the same time, participants could draw conclusions and see trends in the data. Some could even imagine scenarios where data would look like the way it was presented in the graph.

Even though there was some negative feedback we decided to keep the time-circle format because the result was interesting and the design accomplished what it was meant to do. The time-circle is the part of the prototype that needs to improve the most and be tested comparatively to other graphs to find the best and most intuitive way to display several days of alarm information.

4.3.2 Research Questions

The study had four research questions which are answered in this section.

In which way can multiple alarms being triggered simultaneously be presented in a simple way?

This question has been central to our project from the very beginning. The results from all three tests highlight the need for alternative methods to display alarms when many alarms are being triggered simultaneously aka during alarm floods.

The main problem is that too much information appears and the user can not tell what is important or what to prioritise, which is mentioned in Jakobs 10 heuristics. The user can not get an overview of the situation during or after the alarm flood. The first step to solve the problem is to minimize the amount of information, this can be done by grouping or hiding information.

Among the approaches tested, the drop-down list was the most effective in relieving the user of the overwhelming workload from the list on the screen. However, a few testers initially failed to notice the hidden alarms, but once they did, they agreed that this was the optimal display method. The drop-down list minimises the amount of information on the screen by hiding groups of alarms.

Displaying many alarms or combating the amount of alarms have been analysed in previous studies like Li et.al's "Design Implementation of Ship Alarm Management System" where they mention alarm filtering. They define alarm filtering as: "eliminating the alarms that were determined to be less important, irrelevant, or otherwise unnecessary by processing techniques and making them not available to operators". The key difference between their definition of filtering and our solution to use the drop-down list is that filtering alarms remove them from the list permanently. This was not something users or the test participants were interested in and in the study, they mentioned that important information can be removed when filtering an alarm list. Information can get lost in drop-down lists but it is not removed.

Another aspect of presenting multiple alarms being triggered simultaneously is our time-line view. During testing it became clear that the time-line view was much less intimidating than the alarm list even though they were representing the same amount of alarms.

The time-line view groups alarms depending on when in time they were triggered. Most detailed information is removed and it only provides the most important information: are there many alarms? Are there critical alarms? When did they happen?

These considerations tie back to our research question by addressing how multiple alarms can be presented simply and effectively. It highlights the complexity of creating a solution that meets diverse user needs while maintaining clarity and usability. The testing results indicate that while the drop-down approach reduces clutter, it also introduces new challenges related to the visibility of critical alarms, thus necessitating further refinement.

How to present the alarms for the user to prioritise the severity of the alarms?

The severity of alarms can be regarded as both criticality of an alarm, but also if it will lead to other alarms. In other words, an alarm can have critical direct consequences and alarms can be non-critical but start a chain reaction of other alarms that could be severe.

Criticality is often measured on a scale. Li et. al mentions in "Design Implementation of Ship Alarm Management System" that there is no census on how many scales but approximately three to five, but no more than six. From the different criticality levels, it is easy to colour-code each scale. The alarms were already coloured according to the severity in SCMS before this study, with red being the most critical alarm, just like Berg et.al. recommends [19]. The colours used in the system were grey (background), red (critical alarm), yellow (alarm) and grey (warning). The colours red and yellow carry meanings of importance.

Additionally, presenting alarms so the user can prioritise alarms based on severity can be done by adding a signifier explaining the criticality but also by providing information on how many following there have been or will be. By adding the panel header "following alarms" and the number of following alarms in the drop-down alarm list, users can easily see the root cause and take action.

A significant challenge with the drop-down alarm lists was how to prioritise and display consequential critical alarms. Some users preferred that the critical alarm should always remain visible or at least at the top instead of the root cause, while others emphasized the importance of displaying the root cause as the first alarm. This significant divergence in user preferences might be a reason to test these two designs further.

How can the root cause (and correlating alarms) be displayed (during alarm flooding), to reduce mental load and give the user an effective way to solve the problem?

At the beginning of the project, we researched different ways to display information in a tree structure, in order to display the root cause and correlating alarms but were quick to realise that it was difficult to scale up.

The submarine has thousands of possible alarms and an alarm flood can generate hundreds of active and inactive alarms. Most examples of a tree structure had less than ten nodes which would be alarms in this case.

During an alarm flood where multiple alarms trigger other alarms, it is important to act fast and act on the critical alarms that may lead to new alarms. If a user can stop a chain reaction of alarms the situation will be under control, less stressful and more manageable.

Since the tree structure would not scale up, we decided to group correlating alarms in a two-layered tree structure. Alarms could only have one parent, this would lead to larger groups but still provide the function of hiding following alarms. If an alarm has triggered 60 other alarms during an alarm flood, the parent alarm would be the header for the drop-down list. This makes it easier for the user to know which of the maybe 70 alarms in the alarm list to start with since the parent alarm is highlighted. If the parent alarm is solved, some of the 60 following alarms might be solved as well since they depend on the parent alarm. This way of grouping alarms reduces the amount of information in the list, groups information and helps the user prioritise alarms and thereby reduce the mental load.

To display the alarm floods in other ways than an alarm list was important and with the view of the submarine, this was possible. This was a way to display where in the submarine the flood was occurring. This view was extra appreciated by the users from FM, since they felt a lack of knowledge about submarines. When asked about the fire in a scenario, almost every user in all tests immediately pointed at the large circle of where the fire started.

How can the visualisation improve to give the user a better understanding of previous alarms and alarm floods?

This question gave us the most creative freedom, allowing us to explore alternative methods for helping users analyse and visualise alarms using graphical diagrams instead of relying solely on the traditional alarm list. Based on user feedback and insights from our literature review, the existing alarm list in SCMS is deemed necessary and irreplaceable, as the users value the ability to review past alarms in a list format.

The challenge, therefore, was not to replace the alarm list but to complement it with additional views that could enhance the user's understanding of alarm trends over time. The visualisations we designed, such as the timeline and weekly circles, were intended to provide a quick overview of the severity of alarms over the past hours, days, or weeks. Initially, we hypothesised that these graphical tools would be valuable even during alarm floods. However, the results from the Lo-Fi tests revealed that users preferred to use the timeline and circles from a historical or analytical perspective, rather than in real-time during alarm floods.

In the Hi-Fi tests, there was a scenario where users were supposed to search for the occurrence of a specific type of alarm, which was then displayed in the time-circle for the last week, rather than viewing all alarms. The users responded positively to this feature and all five test participants were able to quickly understand and do it effectively. This feedback suggests that while the timeline and time circle diagrams are effective for post-event analysis, their utility during active alarm floods may be limited. Moving forward, it will be important to refine these tools to better align with the real-time needs of users while maintaining their value for historical and analytical purposes.

4.4 Future development

As for further development, there are a lot more suggestions and work to do on the system in general.

The first recommendation for further development would be to keep testing the system with users during the development process. We got many insights during the test sessions and if we could continue, test sessions with submarine personnel would be next. A similar study could expand on the result from this thesis with more focus on the cognitive load a user might experience. Tools like eye-tracking can enhance understanding of a user's experience. By measuring the reaction time of task management, different design solutions can also be weighted against each other.

The second recommendation, for development in SCMS, we suggest to implement the following things from this study:

- An overview of the current alarms displayed on a map of the submarine. Participants have suggested that it would be helpful to have the submarine view open on one of the screens at all times or as a home screen. It helps the user with spatial awareness and provides an overview.
- A park function to remove alarms in the alarm list. This would help the user to customise their alarm list and make prioritisation easier. A safeguard, as to not forget alarms, could be to set all parked alarms on a 20-minute timer to return. There is

also ground for studying if the function should be called something else, a few test participants did not understand the term park.

- A overview to be able to analyse the root cause of an alarm flood after the event, an example would be a time-line. The time-line format was intuitive and helped put alarms into perspective.

The last recommendation, from a UX perspective, there are a few things that we did not have time to test that we suggest for future usability studies:

- A more detailed map of the submarine and making the design responsive to zoom-in to increase the level of detail. A separate project could be to create a 3D model of the submarine.

- Making the submarine view clickable on each room or area in the submarine and displaying the results, either in an alarm list or as a go-to page function. This is so the user can retrieve even more details about where the alarms are coming from.

- Test different formats or graphs to present seven or fourteen days of alarms. The time circle might not be the best way to present the data.

- Further develop and test where the "show parked alarms" function should be placed if not in the filter part. The function should be easy to perform with as few clicks as possible.

Chapter 5
Conclusion

This last chapter concludes the study, bringing up the most important aspects and points discussed. This thesis aimed to answer the four research questions, to develop, design and evaluate a prototype encapsulating the insights learned. The prototype was inspired by SCMS. Three iterations of design, Lo-Fi, Mid-Fi and Hi-Fi prototypes were tested.

- *In which way can multiple alarms being triggered simultaneously be presented in a simple way?* The issue with multiple alarms being triggered simultaneously is that the information is hard to comprehend and a solution to the problem is to lessen the amount of information presented. This was accomplished with a drop-down alarm list, to hide and group information and a time-line view to reduce details and simplify a sequence of events.

- *How to present the alarms for the user to prioritise the severity of the alarms?* Severity includes the criticality of an alarm and the amount of following alarms. A method to prioritise the alarms is to use signifiers, the status of criticality on symbols in different colours and to highlight alarms leading to the following alarms, done in the drop-down alarm list.

- *How can the root cause (and correlating alarms) be displayed (during alarm flooding), to reduce mental load and give the user an effective way to solve the problem?* The root cause and correlating alarms can be presented in a tree structure, but since the advanced alarm system in SCMS contains too many alarms a typical tree structure does not work. Instead, the drop-down list works as a two-layered tree structure by highlighting the root cause so the user can prioritise and solve the root cause before it leads to more alarms.

- *How can the visualisation improve to give the user a better understanding of previous alarms and alarm floods?* Comparing the alarm list to the different overviews, the submarine, time-line and time-circles, it became clear that the overviews served other purposes. The overviews enhanced the user's understanding of alarm trends over time and place.

References

[1] Saab AB. Company in brief: About: Saab.

[2] Rebecca Liss, Klara Wiklundh, and Colleague 1 Saab Kockums. Expert interview scms, Apr 2024.

[3] Jan 2022.

[4] Rebecca Liss, Klara Wiklundh, and Colleague 2 Saab Kockums. Expert interview scms, Apr 2024.

[5] Saab AB, 2022.

[6] Saab AB. The secret to the world´s most silent submarine: Stories: Saab, Mar 2015.

[7] Greg Beyer. The history of submarines: From science to stealth, May 2023.

[8] Saab AB. Saab completes acquisition of tkms ab (kockums), Jul 2014.

[9] Saab AB. Oceanic: Saab.

[10] Goal 16: Peace, justice and strong institutions, Jan 2024.

[11] Goal 8: Decent work and economic growth, Jan 2024.

[12] Rebecca Liss, Klara Wiklundh, and Colleague 3 Saab Kockums. Expert interview scms, Apr 2024.

[13] Saab AB. The world´s most modern submarine program: Stories: Saab, Jun 2015.

[14] Saab AB. A submarine in space: Stories: Saab, Jul 2020.

[15] Xiaojun Wu and Zhizhong Li. A review of alarm system design for advanced control rooms of nuclear power plants. *International Journal of Human-Computer Interaction*, 34(6):477 – 490, 2018.

[16] Liuhong Wei, Wenxin Guo, Fushuan Wen, Gerard Ledwich, Zhiwei Liao, and Jianbo Xin. An online intelligent alarm-processing system for digital substations. *IEEE Transactions on Power Delivery*, 26(3):1615–1624, July 2011.

[17] Jason C. Laberge, Peter Bullemer, Mischa Tolsma, and Dal Vernon C. Reising. Addressing alarm flood situations in the process industries through alarm summary display design and alarm response strategy. *International Journal of Industrial Ergonomics*, 44(3):395 – 406, 2014.

[18] Guobing Li, Hongling Zhao, Mengyin Ma, Yang Zhao, and Peiwen Zhang. Design and implementation of the ship alarm management system for chief engineer. *2023 IEEE 6th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC), Information Technology,Networking,Electronic and Automation Control Conference (ITNEC), 2023 IEEE 6th*, 6:1039 – 1042, 2023.

[19] Øivind. Berg, Magnhild. Kaarstad, Jan Erik. Farbrot, Christer. Nihlwing, Tommy. Karlsson, and Belen. Torralba. *Simulator-based Human Factors Studies Across 25 Years. The History of the Halden Man-Machine Laboratory.* Springer London, 2011.

[20] Saab AB. Saab delivers upgraded gotland-class submarine to fmv, Dec 2015.

[21] Saab AB. Saab receives further order for the a26 submarine for sweden, Aug 2021.

[22] Jennifer Preece, Yvonne Rogers, and Helen Sharp. *Interaktionsdesign : bortom människa-dator-interaktion.* Studentlitteratur, 2016.

[23] Design thinking (dt), Jun 2024.

[24] Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems (ISO 9241-210:2019). Standard, International Organization for Standardization, Geneva, CH, May 2019.

[25] Apr 2024.

[26] Jakob Nielsen. Usability 101: Introduction to usability, Jan 2024.

[27] Don Norman and Jakob Nielsen. The definition of user experience (ux), Jan 2024.

[28] Mar 2024.

[29] Donald A. Norman. *The design of everyday things.* Basic Books, 2002.

[30] Ian Batterbee. Don norman's seven fundamental design principles, Feb 2024.

[31] Jakob Nielsen. 10 usability heuristics for user interface design, Feb 2024.

[32] SEVENTH FRAMEWORK PROGRAMME. User study guidelines, 2008. https://www.certec.lth.se/fileadmin/certec/publikationer/HaptiMap_d12.pdf.

[33] Eason K.d. Person. Information technology and organisational change: K.d. eason: Taylor, Apr 2014.

[34] IDEO. *The Field Guide to human-centered design: Design kit_2015.* IDEO, 1 edition, 2015.

[35] Bella Martin and Bruce M. Hanington. *Universal Methods of Design: 100 Ways To Research Complex Problems, develop innovative ideas, and Design Effective Solutions.* Rockport Publishers, 2012.

[36] Kate Kaplan. User journeys vs. user flows, Jan 2024.

[37] Miriam Walker, Leila Takayama, and James A. Landay. High-fidelity or low-fidelity, paper or computer? choosing attributes when testing web prototypes. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46(5):661 – 665, 2002.

[38] Peter Martinez. Everything you need to know about mid fidelity prototype, Jun 2024.

[39] Kara Pernice. Ux prototypes: Low fidelity vs. high fidelity, Feb 2024.

[40] J. Rubin, D. Chisnell, and J. Spool. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests.* Wiley, 2011.

[41] Kate Moran. Usability testing 101, Jan 2024.

[42] Jakob Nielsen. Why you only need to test with 5 users, Feb 2024.

[43] Tiffany Bergin. *Introduction to Data Analysis: Quantitative, Qualitative and Mixed Methods. [Elektronisk resurs].* SAGE Publications, 2018.

[44] Katharine Betteridge. What every client should know about sus scores.

[45] Försvarsmakten. Personalsiffror, Feb 2014.

[46]

Appendices

Appendix A
Usability Testing

A.1 Scenarios

In this section the scenarios and focus group questions for both the Lo-Fi and Hi-Fi test are presented in detail.

A.1.1 Lo-Fi

Presented in the figure A.1 are the questions asked for each scenario read to the user during the test.

Focus Group

From the test we had several questions prepared for the focus group session.

“Vi kommer gå runt i en cirkel, ni får gärna bygga på andras svar. Ni får gärna utgå från prototyperna ni fått testa idag, men berätta gärna även om era problem ni känner att ni har med systemet ni använder idag. Har ni lösningar, idéer så kör på! vi är här för att vi vill hjälpa till och förbättra.”

Problemspecifika (15 min):

- Vad är problemet idag för er?
- Hur hade ni velat förbättra systemet?

Prototypspecifika (15-20 min):

- kan ni beskriva er upplevelse med det “riktiga systemet”, aka. den med bara listor?
- vad funkar bra/mindre bra?
- vilka potentiella utvecklingsmöjligheter har det?

Nr	Scenario	Vad säger/gör vi	Vad ska de göra	Tid
1	Lösa larm	Har de läst igenom lappen och signerat? Läs högt: "Observera att det inte är din förmåga eller dina svar som testas idag, utan endast systemets användbarhet. Du har alltid rätt att avbryta testet när du vill utan att behöva ange din anledning. Om du har några frågor under tiden så är det bara att fråga. Vi vill även att du tänker högt under testets gång" "Beskriv gärna varje ny vy du får se"		8
2	Lösa larm	Du har precis gått på ditt skift och vill få en överblick över larmen. Interagera med pappersprototypen för att ta reda på mer om larmen som redan finns.	Hitta maxlistan	3
3	Brand	Klockan slår 10:30 och det dyker nu upp massor med oroande larm på skärmen, samtidigt som du hör/ser brandlarmet gå i det externa brandlarmssystemet. Du vill nu veta mer om orsaken till larmet och studerar dina larm som dykt upp. Vad tror du har hänt?	Hitta orsak: brand i skåp	5
4	Brand	Vilken del av ubåten som är drabbad? Se om du kan hitta information om detta i SCMS.	Förstå var i ubåten händelsen skett - historik	5
5	Brand	Kan du se om några larm har utlöst varandra?	Hitta rotorsaker till andra larm	3
6	Brand	Kan du följa ett händelseförlopp eller kan du dra några sutsatser om larmen?	Öppna tankar	5
7	Brand	Din besättning har jobbat hårt och lyckats släcka branden och du vill ta bort de aktuella larmen för branden i systemet. Hur hade du förmedlat i systemet att du sett och hanterat larmen?	kvittera (acknowledge) larm	5
8	Nivå	En tid har gått sen den senaste larmskuren, klockan är nu ca 13:00. Ni är på väg upp till ytan när det återigen dyker upp en stor mängd larm samtidigt. Hitta orsaken till larmen, eller åtminstone vad för typ av larm som kommit upp. Hur bedömer du situationen? (Hur gör du för att få en översikt och bedöma problemet?)	Hitta nivålarm eller orsak sjögång i maxlistan Öppna tankar	3
9	Nivå	Vilken del av ubåten som är drabbad? Se om du kan hitta information om detta i SCMS	Förstå var i ubåten händelsen skett - historik	8
10	Nivå	Kan du se om några larm har utlöst varandra?	Hitta rotorsaker till andra larm	3
11	Nivå	Kan du följa ett händelseförlopp eller kan du dra några sutsatser om larmen?	Öppna tankar	5
12	Nivå	Sjögången är över och du vill nu hantera nivålarmen. Hur hade du förmedlat i systemet att du sett och hanterat larmen?	kvittera (acknowledge) larm	4

Figure A.1: Scenarios from Lo-Fi tests

- övergripande tankar om prototyperna?
- kan ni beskriva er upplevelse med ubåts-prototypen (heatmap och indrag)?
- vad funkar bra/mindre bra?
- vilka potentiella utvecklingsmöjligheter har det?

- specifikt heatmap specifikt indrag kan ni beskriva er upplevelse med tidslinje och tidscirkeln (den med olika tidslinjer och dropdown menyer)?
- vad funkar bra/mindre bra?
- vilka potentiella utvecklingsmöjligheter har det? specifikt tidslinjer, dropdown menyer

A.1.2 Hi-Fi

Presented in the figure A.2 are the questions asked for each scenario read to the user during the test.

A.2 System Usability Scale Questionnaire

The eleventh question was added as a final chance to write down thoughts now that the test was finalized.

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome (besvärlig/opraktisk) to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.
11. Do you have any feedback or thoughts on the tested prototypes? Please feel free to share them here or on the provided papers beside you.

The SUS is calculated into a score where the number answered on questions (q) 1, 3, 5, 7 and 9 are subtracted with one before being summarised [44].

$$R(q) = UserRating$$

$$1 \leq R(q) \leq 5$$

$$q \in \{1, 3, 5, 7, 9\}$$

Nr	Scenario	Manus	Task
1	Lösa larm	Vänligen gå till hemskärm! Kl är 07:00 och du har precis gått på ditt pass. Du ser att din kollega har lämnat några inaktiva larm kvar i larmlistan, så du vill kika närmre på vad det är för larm som finns i listan. Du kollar med en kollega och larmen är redan åtgärdade. Du vill därför ta bort alla larm i larmlistan på en och samma gång. Nu när larmlistan blev tom vill du försäkra dig om att det inte var några viktiga larm som du kvitterat (acknowledge), gå in på larmhistoriken och se efter hur det ser ut. Ledning: Finns det kanske ett sätt att välja alla larm på en och samma gång för att sedan kunna kvittera (acknowledge)?	• Hitta maximera • Kvittera (acknowledge) lösa larm • Kolla på historik
2	Brand	Vänligen gå till hemskärm 2. Kl 08:00, Det dyker upp en stor skur med aktiva larm och du ser ett brandlarm blinka på nöd-panelen. Du ser att det är 91 larm i listan, (men du har ännu inte riktigt fått koll på allt,) se om du kan få reda på vilka larm det är. Hur upplever du larmlistan? Hint/ledning: Undersök rubrikerna i larmlistan och hitta att det finns "grupper" Du tycker att det var lite rörigt med så många larm, så du stänger de följdlarm som du öppnat. Du vill nu istället snabbt få koll på alla kritiska larm högst upp i listan, hur hade du gjort det? Hint/ledning: Det finns en filtreringsknapp men se om det finns något annat sätt, tänk excel!	• Hitta maximera • Hitta drop-down funktion • Sortera på status
3	Brand	Du inser att det är en brand och du vill nu veta hur omfattande larmen i ubåten är samt vart i båten larmen finns. Finns det något sätt att se detta? Kan du se var i båten larmen har utlösts? Du vill nu titta närmare på/undersöka vad som hänt i den akterliga delen där det indikerar att det finns rött larm. Hint/ledning: Kolla vad som har hänt i aktern/hur många larm finns det i aktern? Säg att det hade funnits många fler larm utspridda över ritningen på båten. Hur hade du gjort för att endast få se det röda larmet i listan?	• Hitta ubåten • Se kartan och förstå cirklarna • Klicka/ navigera runt • Klicka på den röda markören eller på filtrering
4	Nivå	Vänligen gå till hemskärm 3. Klockan är nu 10 och du ser nya larm på skärmen. Du vill nu veta vilka alla nya larm är. Om maximera, tror du att det finns något annat sätt att se de nya larmen på? Ta reda på vilka följdlarm du fått och delge gärna om du misstänker någon orsak till larmen. Du inser att alla larm med battericeller är något du inte vill se i larmlistan just nu, men inte heller vill kvittera (acknowledge) innan du har kontroll på situationen i båten. Du vill därför markera att du vill ta hand om dessa larm lite senare. Hur gör du det? FRÅGA: Hade du velat ha en prioritera funktion istället för parkera funktion? Hint/ledning: Hur hade du gjort för att lägga undan eller med annan formulering: dölja eller markera larm för att kunna titta på dem senare? Du vill nu se endast de larm du precis har parkerat, hur tror du att du kan se det? Hint: Du behöver hitta ett sätt att filtrera din maximerade larmlista	• Hitta Nytt eller maximera • Hitta drop-down • Parkera larmen i en grupp • Hitta parkerade larm
5	Nivå	Nu har en liten tid gått sen senaste skuren och du vill visa för din Chief hur senaste dygnets larm-översikt sett ut. Din Chief undrar om du kan visa vad som hänt specifikt där mellan 9-11, där det varit som mest larm. Undersök vilka av larmen som kom först i den här larmskuren. Hur hade du gjort då? Ni har diskuterat klar larmskuren och du vill nu återgå till att visa ett dygn. Hint: Du vill nu återställa	• Hitta tidslinje • Visa kl 9-11 • Tryck på ballongen • Gå tillbaka till ett dygn
6	Nivå	Ni vill nu undersöka hur larmen sett ut en vecka tillbaka. Din chief vill även veta hur ofta nivålarmen för battericellerna dykt upp. Kan du se om du kan få upp endast dessa i cirkeldiagrammet? Hint: Sök på battericell grupp	• Klicka på en vecka • Sök efter battericeller
7	Nivå	Du vill även visa hur den allmänna larmsituationen sett ut två veckor tillbaka för din chief. Kan du se något mönster eller trender i cirkeldiagrammet?	• Hitta 2 veckor • Berätta att det finns många röda larm på kvällarna

Figure A.2: Scenarios from Hi-Fi tests

$$SumOdd = \sum (R(q) - 1) \tag{A.1}$$

On the remaining questions 2, 4, 6, 8, 10, the answer is subtracted from the number five [44].

$q \in \{2, 4, 6, 8, 10\}$

$$SumEven = \sum(5 - R(q)) \tag{A.2}$$

$$UserScore = (SumOdd + SumEven) \cdot 2,5 \tag{A.3}$$

Table A.3: Results from the System Usability Scale questionnaire for the Mid-Fi test phase. The total score for the Mid-Fi is summarised in the bottom right corner.

PARTICIPANT	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	SUMMARY
P1	1	5	3	2	2	3	3	3	3	4	37.5
P2	3	3	4	1	3	2	3	2	3	2	65
P3	4	1	4	4	4	2	5	2	4	4	70
P4	5	4	4	2	4	2	4	2	3	3	67.5
SUMMARY											60

Table A.4: Results from the System Usability Scale questionnaire for the Hi-Fi test phase. The total score for the Hi-Fi is summarised in the bottom right corner.

PARTICIPANT	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	SUMMARY
P1	5	2	4	2	3	2	5	2	3	1	77.5
P2	4	2	5	2	4	1	5	2	5	1	87.5
P3	4	2	4	5	4	2	4	2	4	3	65
P4	5	1	5	1	4	1	4	1	5	1	95
P5	4	1	4	1	5	1	4	1	4	1	90
SUMMARY											83