Wireless Helm Navigation for a Small Passenger Ferry

Master's thesis in Computer science and engineering

Emil Falk
Karl Loewe

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2025

Master's thesis 2025

Wireless Helm Navigation for a Small Passenger Ferry

EMIL FALK
KARL LOEWE

UNIVERSITY OF GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2025

Wireless Helm Navigation for a Small Passenger Ferry

EMIL FALK & KARL LOEWE

Supervisor: Aris Alissandrakis, Computer Science and Engineering
Advisor: Tobias Husberg, Cstrider AB
Examiner: Morten Fjeld, Computer Science and Engineering

Master's Thesis 2025
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Representation of the final high-fidelity prototype of the mobile application.

Typeset in LaTeX
Gothenburg, Sweden 2025

Wireless Helm Navigation for a Small Passenger Ferry
EMIL FALK
KARL LOEWE
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

This thesis investigates the design and development of a wireless mobile phone application intended to facilitate the safe and efficient navigation of a small passenger ferry. Conducted in collaboration with Cstrider AB, the project applies participatory design and Research Through Design (RTD) methodologies to address cognitive and usability challenges faced by ferry operators.

Through interviews, thematic analysis, and iterative prototyping in Figma and Flutter, the study identifies key user requirements, such as simplicity, intuitive controls, traditional visual metaphors, and minimal attentional demand. The final high-fidelity prototype emphasizes skeuomorphic design and streamlined functionality to support situational awareness and reduce cognitive workload.

While evaluations suggest that the application is generally perceived as intuitive, concerns persist around trust and reliability in safety-critical contexts. This work contributes both a validated interaction prototype and actionable insights into the design of mobile interfaces for maritime control systems, with implications for future development in wireless and remote vessel operation.

Keywords: Ferry, Mobile, Application, Mental Workload, Situation Awareness, Complacency, Mental Rotation, Memory, UI.

Acknowledgements

We want to thank our supervisor Aris Alissandrakis for all the valuable feedback and swift replies he has provided over the past few months.

A big thank you to Tobias Husberg from Cstrider for giving us the opportunity to carry out this project and for providing technical and maritime insights as well as productive feedback during our weekly meetings.

We are very grateful to Monica Lundh for assisting us in our search for interviewees.

Finally, we thank the mariners who participated in our interviews and user evaluations.

Emil Falk & Karl Loewe, Gothenburg, 2025-06-10

Contents

List of Figures

List of Tables

1 Introduction

This thesis investigates the design and implementation of a wireless mobile phone application to ensure the safe and efficient navigation and control of a small transport ferry. The research is conducted in collaboration with Cstrider, the developers of the transport ferry, to create a user interface (UI) that mitigates cognitive risks associated with marine navigation. The study focuses on UI design principles that enhance situational awareness, reduce information overload, and optimize information presentation for ferry operators.

Situational awareness is fundamental for safe maritime navigation, enabling operators to perceive, comprehend, and anticipate maritime conditions effectively. A lapse in situational awareness can lead to critical errors, particularly in confined or high-traffic waterways where swift decision-making is essential. Additionally, information overload—stemming from excessive or poorly structured data—can hinder an operator's ability to process critical information, increasing the likelihood of misjudgment and accidents. Effective UI design plays a crucial role in addressing these issues by structuring information in a way that enhances comprehension and usability.

Existing research on maritime navigation underscores that a well-designed UI can significantly improve an operator's ability to make informed decisions under time-sensitive conditions (Vu et al., 2019). UI design guidelines emphasize clarity, minimal cognitive load, and intuitive interaction, all of which are essential for reducing human error. By integrating these principles into a mobile application, the system can provide ferry operators with real-time data in an accessible and actionable format, thereby improving navigation safety and efficiency.

This thesis is structured as follows: The background chapter provides an overview of Maritime Autonomous Surface Ships (MASS) and examines the leading causes of marine navigation accidents, highlighting the cognitive challenges faced by operators. The theory section brings up relevant studies and theories pertaining to interaction design, cognition and UI design. The methodology section outlines the research approach, data collection methods, and system design framework. Subsequent chapters present the UI design process, implementation details, and evaluation of the application's effectiveness in steering operations, improving situation awareness and reducing information overload. The discussion brings up reflections of the process, the final design, ethical considerations, limitations, and aspects to consider for future work. Finally, the conclusion summarizes the project and brings up the potential relevance of the thesis in broader research.

1.1 Stakeholders

In this section the different stakeholders involved will be presented and their relation to the project.

1.1.1 Chalmers

Chalmers University of Technology is the institution where the project takes place and from where resources are available. Chalmers plays a key part in allowing the designers access to otherwise restricted research and study spaces. The interest Chalmers has in the project can be seen as the contribution that this work can bring to the field of maritime research and their existing database on students' theses.

1.1.2 Cstrider

One of the primary stakeholders in this project was Cstrider, the company that proposed this project. Cstrider had a direct interest in ensuring that the application aligned with their operational needs, brand identity, and usability expectations. Their role as stakeholders included defining high-level functional requirements and evaluating design proposals throughout the development process. Their involvement throughout the project ensured that the design remained relevant to real-world use cases and reflected the practical constraints and expectations of their vessel operation.

1.1.3 Ship Officers (Users)

Ship officers of at least class VIII education are the primary target group of the study. As they are the final users of the product they are also the most impacted. Should the solution be used in their work it is of great importance that it works well for them. Considering their part in the design of the solution they could have a vested interest in its success and implementation. They would also be the first group impacted by a bad design.

1.1.4 Designers (Students)

The designers have a vested interest in the outcome of this project. For the completion of this course the project needs to be adequately conducted. Furthermore the designers invests time and effort into the project, and therefore will be invested in the project outcome.

1.2 Aim

The practical outcome of this research will be a mobile application prototype that aligns with cognitive ergonomic principles and industry best practices. The application will be designed to streamline navigational tasks, provide critical alerts, and facilitate efficient decision-making for ferry operators. By leveraging insights from human factors research and real-world ferry operations, the study aims to develop a system that enhances operational safety and efficiency, contributing to the broader field of maritime digitalization.

1.3 Research Question

The primary research question guiding this study is: How can a wireless mobile phone application be designed and implemented to ensure safe and efficient navigation and control of a small transport ferry?

2

Background

2.1 Maritime Autonomous Surface Ships (MASS)

The term Maritime Autonomous Surface Ships (MASS) was created by the International Maritime Organization (IMO) and refers to commercial ships where human intervention is limited or non-existent. These vessels use advanced sensors, sophisticated software, and advanced communications technology to perform collision avoidance, navigation, and some shipboard tasks (McKie, n.d.). Before this definition came into existence, numerous terms like unmanned ships, intelligent ships, smart ships, and self-driving ships were used to define the same (Cheng, 2023). Once the IMO provided a definition for MASS, it became easier to address autonomous operations in the maritime space.

In December 2018, during the 100th session of IMO's Maritime Safety Committee (MSC100), the framework for the Regulatory Scoping Exercise (RSE) on MASS was approved. This framework categorizes maritime autonomy into four degrees (Cheng, 2023):

1. Degree one: Ship with automated processes and decision support: seafarers are on board to operate and control shipboard systems and functions. Some operations may be automated and at times be unsupervised but with seafarers on board ready to take control.
2. Degree two: Remotely controlled ship with seafarers on board: the ship is controlled and operated from another location. Seafarers are available on board to take control and to operate the shipboard systems and functions.
3. Degree three: Remotely controlled ship without seafarers on board: the ship is controlled and operated from another location. There are no seafarers on board.
4. Degree four: Fully autonomous ship: the operating system of the ship is able to make decisions and determine actions by itself.

The incorporation of MASS technology is especially important for small transport ferries that work in narrow water bodies and need high levels of precision. Based on the autonomy level, such ferries can either help the onboard crew with automated decision making (Degree One), be controlled through a remote command (Degree Two or Three), or operate without any human intervention (Degree Four). The integration of MASS on ferries has the potential to improve safety and reduce costs while making operational efficiencies better through lowering human resources, marine risk and fuel use (Cheng, 2023).

2.2 Safety concerns for marine navigation

The marine environment has numerous safety concerns, the majority of which are caused by human factors. These include fatigue, automation, situational awareness, communication, decision-making, teamwork, and overall health and stress levels (Hetherington et al., 2006). Although some cognitive and behavioral tendencies of individuals lead to maritime accidents, the overall safety climate on a vessel also determines whether people adopt safe working practices or not.

One of the most significant cognitive issues in maritime safety is loss of situational awareness (SA). Grech et al. (2002) discovered that the majority of human errors in the maritime sector fall into this category. Situational awareness is crucial in ship operation and navigation since it involves the perception of environmental elements, comprehension of their significance, and prediction of what they will become. Insufficient situational awareness can lead to poor decision-making, misinterpretation of navigational hazards, and increased collision or grounding risk.

In situational awareness, (Stratmann & Boll, 2016) state that Level 1 SA—detection of the primary elements in the environment—is the highest cause of human error in maritime accidents. What this means is that a failure to possess a simple perception of navigational data, threats, and system alerts can be very perilous. It is essential to ensure that marine operators are provided with simple, intuitive, and easily accessible information in order to reduce the frequency of accidents due to situational unawareness.

Fatigue is also a major factor in human errors at sea. Chen et al. (2015) indicate that more than 80% of shipping accidents have implications for human errors, of which fatigue is a prominent cause. Excessive working hours, poor rest, and routine operating conditions may impair cognitive abilities, leading to reduced vigilance and reduced reaction time. The shipping sector broadly acknowledges the influence of fatigue on operational safety, necessitating improved work-rest planning and fatigue management practices.

Despite improvements in seagoing technology, accidents are not decreasing by a commensurate margin (Caridis, 2001). Such long-standing concern results to a large extent from human factors, which are still the major cause of marine accidents (de la Campa Portela, 2005). Although human error is a common factor in maritime accidents, it is most likely an indication of systemic organizational failures and not personal lapses in judgment. According to Bhardwaj et al. (2022) organizational failures are the cause of most maritime accidents. Research has shown that approximately 80% of accidents are the result of organizational failures and not necessarily personal errors. This calls for well-structured operational procedures and effective management in minimizing risks. Despite the improvement in automatic systems and electronic interfaces, the human-technology interface is still a vulnerable area.

Poorly designed user interfaces, vexingly ambiguous or unclear warning systems, and substandard information display tend to enhance human error rather than minimize it.

The UI must be optimized to enhance situational awareness, reduce cognitive workload, and facilitate good decision-making. As loss of SA is one of the main reasons for maritime accidents, the interface must display clear real-time information about the location of the ferry, traffic around it, environmental conditions, and navigation risks.

Fatigue-caused errors can be prevented by keeping the UI straightforward and not demanding excessive mental effort. An unvarying interface with strong visual hierarchies, warning anticipations, and streamlined workflows can prevent information overload, reducing the likelihood of misinterpretation and sluggish response. Additionally, the application of adaptive automation—where the system assists but never overrides human choice-making—can help in workload balancing and sustaining the operator attentiveness.

2.3 Influencing factors for marine navigation

Vu et al. (2019) describe several different factors that can influence the operation of an user interface for marine vessels. Some of these are: individual factors, experience and habits; traffic conditions, density and complexity; and geographical location, ocean waters or coastal waters. The main usability issue reported was information overload, where the concerns lie with distractions and issue identification. This study focused on providing input for the guidelines of Standardised modes of operation (S-mode) for the International Maritime Organization (International Maritime Organization. Maritime Safety Committee, 2019)(IMO, 2019). This study and S-mode is not targeted at ferries specifically and it has not focused on mobile applications. From this we can sufficiently conclude that there exists a knowledge gap in regard to control and navigation of a ferry through wireless mobile applications.

3
Theory

In this chapter, different relevant theories and research will be presented. The design methodology will be briefly explained and presented. Followed by a short section detailing working- and long-term memory systems. Mental Workload and Situation Awareness are then explained, followed by automation complacency. Lastly guidelines for UI design and control systems will be presented.

3.1 Design Methodology

3.1.1 Wicked Problems in Design

Wicked problems are messy, interconnected issues that cannot be tackled in a straightforward way in design research. Contrary to typical problems, which can be clearly stated and systematically solved, wicked problems are indeterminate in type, ever-changing, and situated in the very dense social, cultural, and technological contexts. Their subtlety will often result from the interplay of diverse user needs, contradictory stakeholder agendas, and the increasing pace of technological evolution (Buchanan, 1992; Wong, 2023). Among the characteristics that differentiate wicked problems in design research is that no problem statement can be easily discerned. The diversity of the requirements makes it impossible to identify a single, "correct" solution. As Rittel and Webber explain, 'Solutions to wicked problems are not true or false (right or wrong); they can only be good or bad.' (as cited in Buchanan, 1992; Wong, 2023).

Another characteristic of wicked problems is that there is no definitive stopping point (Buchanan, 1992; Wong, 2023). In design research, usability testing and user feedback tend to uncover new problems, leading to iterative design adjustments. Unlike technical bugs that can be fixed once and for all, user experience issues can remain as subjective or context-dependent issues. Although best practices and design guidelines can provide starting points, solutions need to be adapted to particular contexts. Moreover, the integration of design elements makes wicked problems more complex. A change to one aspect of the interface, say, the implementation of a new model of navigation, will resonate in other parts of the user experience. In research, this requires systems thinking, where one watches out for the overall impact of the work and not get fixated on individual features (Buchanan, 1992; Wong, 2023).

Lastly, there exist stakeholder tensions that occur within the design of intricate systems. Researchers and designers have to balance among user demands, technical realities, and company objectives. As an illustration, a feature meant to improve the user experience might be resource-intensive and expensive to build, leading to confrontations among engineering teams and design teams.

Wicked problems are dynamic, multifaceted challenges that require ongoing research, iteration, and collaboration. By embracing ambiguity, adaptive design solutions can tackle complexity and ultimately end up with user-centered interfaces that evolve with changing user needs and technological advancements.

3.1.2 Research Through Design

Research through Design (RtD) is a design-led research approach in which the process of design itself becomes a way of creating new knowledge. As opposed to traditional research, where one begins with a hypothesis and a linear workflow, RtD is an iterative and exploratory process. Designers develop artefacts such as prototypes or user interfaces as a method of investigating tricky research questions. These artefacts not only stand for solutions to problems but also function as platforms for questioning, reflection, and learning.

RtD is suited to wicked problem-solving in design, where wicked problems are not certain in definition, and solutions are never actually complete. With the uncertainty and dynamic character of wicked problems, traditional research designs that rely on controlled environments and static variables may be insufficient. RtD is strongest in situations where seeking out and iterating can give rise to meaningful insights (Zimmerman et al., 2010).

In the context of UI design and research, RtD is very good for understanding user behavior, testing new interaction patterns, and addressing complex design challenges. For example, when designing an interface for a health application or educational platform, user needs can be diverse, and technological constraints may continually shift. Through iterative prototyping and user testing, designers can explore different solutions, gather feedback, and refine interfaces to better meet user and stakeholder expectations. The artefacts created in RtD are not just functional products but also serve as research outputs that capture insights about user preferences, interaction flows, and design strategies. As Gaver (2012, p. 941) puts it "... research through design should be appreciated for its proliferation of new realities, and its theory considered as annotation of the artefacts that are its fundamental achievement".

A key aspect of RtD in UI research is its emphasis on reflection-in-action. Designers continuously evaluate and adapt their work based on observations and user interactions. This iterative reflective practice generates knowledge that extends beyond the immediate design context, contributing to broader design principles and guidelines (Zimmerman et al., 2010). RtD also encourages a participatory approach, where users and stakeholders are actively involved in the design process. This collaborative aspect is suited for tackling wicked problems, as it ensures that diverse perspectives are considered and that solutions are grounded in real-world contexts.

3.2 Memory

3.2.1 Working Memory

Working memory is a cognitive system responsible for temporarily storing and manipulating information. Baddeley (1986, 1995)'s model identifies three core components: the phonological store, the visuo-spatial sketch pad, and the central executive, with a later addition of the episodic buffer (Baddeley, 2007). The phonological store processes verbal and auditory information, allowing rehearsal through a phonological loop. In contrast, the visuo-spatial sketch pad is responsible for handling visual and spatial information, such as images and dynamic spatial displays. The central executive plays a crucial role in managing attentional control, allocating resources, and resisting distractions. The episodic buffer serves as a temporary storage system that integrates information from the working memory components, perception, and long-term memory, facilitating conscious awareness.

Research has shown that working memory capacity is linked to various cognitive functions, including reading comprehension, multitasking, problem-solving, and decision-making. It has also been found to decline with age, which can impact cognitive performance, as observed in tasks such as remembering and executing instructions (Taylor et al., 2005). Another important aspect of working memory is its role in attentional control. Higher working memory capacity is associated with a greater ability to sustain attention and resist distractions, reducing the likelihood of mind-wandering and task neglect (McVay & Kane, 2009).

The distinction between different components of working memory has several practical implications. The phonological store and visuo-spatial sketch pad operate independently and are susceptible to different types of interference, which is relevant for multitasking and task design. The central executive, responsible for control and management functions, can also be affected by interference, influencing the ability to perform concurrent tasks. Furthermore, the relationship between different types of memory codes and display modalities has implications for the design of auditory versus visual interfaces, as well as verbal versus spatial presentations (Wickens et al., 2013).

3.2.2 Long-Term Memory

Long-term memory (LTM) stores information in various forms, primarily categorized as procedural and declarative knowledge. Procedural knowledge, or implicit memory, involves skills and actions that individuals may not be able to verbalize but demonstrate through behavior (Reder, 1996). Declarative knowledge consists of semantic memory, which pertains to general facts and concepts, and episodic memory, which involves personal experiences (Poldrack & Packard, 2003; Tulving & Schacter, 1990). These three LTM systems, implicit, semantic, and episodic, are believed to exist independently in the brain. LTM is not merely a passive storage system but an active perceptual process (Barsalou, 2008). According to the grounded cognition approach, recalling conceptual knowledge activates areas in the brain that simulate sensory and motor elements.

3.3 Mental Workload and Situation Awareness

3.3.1 Attention and Workload

Parasuraman et al. (2008) defined mental workload as “the relation between the function relating the mental resources demanded by a task and those resources available to be supplied by the human operator”. Vidulich and Tsang (2012) explain that there are two main determinants of mental workload: exogenous task demands as specified by factors such as task difficulty, task priority, situational contingency, and endogenous supply of attentional or processing resources to support information processing such as perceiving, updating memory, planning, decision making, and response processing. They continue by explaining that this endogenous supply is modulated by individual differences such as one’s skill set and expertise.

A widely accepted theoretical framework for attention in workload research is the energetics model (Hockey et al., 1986; Wickens, 2001). This model suggests that attentional resources are finite and can be strategically allocated. Performance generally improves with increased resource allocation until capacity limits are reached (Norman & Bobrow, 1975). Furthermore, Kahneman (1973) proposed that attention is managed through a feedback loop based on task priorities, personal relevance, and performance monitoring.

According to Vidulich and Tsang (2012) the notion of a single, undifferentiated attentional resource pool was challenged in the late 1970s by studies suggesting the existence of multiple specialized resources. Wickens (1980, 2008) developed a multiple resource theory, proposing that attention is divided along three primary dimensions: (1) stages of processing (perceptual vs. response), (2) processing codes (spatial vs. verbal), and (3) input/output modalities (visual vs. auditory). This model explains multitasking performance by highlighting competition for similar resources, where tasks demanding the same type of resource lead to higher workload and reduced efficiency. Greater similarity in resource demands among task components leads to more intense competition for the same resources, reduced spare capacity, and consequently, a higher resulting workload (Vidulich & Tsang, 2012).

Research supports the view that individual differences, such as expertise and skill level, influence the supply or availability of processing resources. Skilled individuals demonstrate more efficient resource utilization, leading to reduced cognitive strain and improved performance (Just et al., 2003). Studies using neurophysiological methods indicate that practice reduces cognitive load by enhancing procedural knowledge, thereby freeing up resources for other tasks (Haier et al., 1992; Parks et al., 1988).

In applied settings, understanding workload and SA dynamics is critical for optimizing human performance in high-stakes environments such as aviation, healthcare, and automation. By incorporating workload and SA assessments, researchers and practitioners can design better training programs, improve system usability, and enhance overall safety and efficiency in complex operational settings (Durso & Alexander, 2010; Wickens, 2008).

3.3.2 Memory and Situation Awareness

Situation Awareness (SA) is a concept that, like mental workload, is difficult to define precisely. Endsley (1990) provided one of the most widely accepted definitions, describing SA as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future”. SA is closely linked to perception and working memory processes, as information must not only be available but also perceived, processed, and integrated with prior knowledge stored in long-term memory (LTM).

Adams et al. (1995) distinguished between the process and product of SA: the product refers to the operator’s state of awareness, while the process involves the perceptual and cognitive activities needed to build and update that awareness. SA varies between individuals based on their knowledge and experience, just as mental workload can differ depending on attentional resources and skill levels.

According to Vidulich and Tsang (2012) a key factor influencing SA is expertise, which is developed through extensive practice and leads to improved perception, memory organization, and decision-making. Experts possess a vast amount of domain-specific knowledge, which is structured more efficiently than that of novices (Reingold et al., 2001). Research has shown that experts can recognize patterns, categorize information, and respond to situations more quickly and accurately than novices (Charness & Tuffiash, 2008). Furthermore, they demonstrate superior procedural knowledge, allowing them to execute complex tasks efficiently (Druckman & Bjork, 1991).

Ericsson and Kintsch (1995) proposed the concept of long-term working memory (LTWM) as a defining feature of expertise. Unlike standard working memory, LTWM has a larger capacity and persists for longer durations, enabling experts to store and retrieve relevant information more effectively. Expert problem-solving involves more than simply recalling stored solutions to past problems. It also requires effectively applying extensive knowledge and reasoning skills to tackle new challenges (Charness, 1989; Horn & Masunaga, 2000).

Beyond knowledge and memory, experts also exhibit metacognitive abilities, such as self-monitoring, planning, and efficient resource allocation (Glaser, 1987). These capabilities allow them to quickly assess situations, anticipate future events, and retrieve appropriate responses from memory, leading to superior problem-solving and decision-making (Charness, 1995).

3.3.3 Mental Workload and Situation Awareness

Mental workload and situation awareness are closely related but distinct concepts (Vidulich, 2003; Wickens, 2002). Wickens (2001) differentiates them by describing mental workload as an energetic construct concerned with "how much" cognitive effort is required, while SA is a cognitive concept focused on the "quality" of an operator's understanding of the situation. In practice, researchers assess workload by measuring the amount and type of mental effort involved, while SA is evaluated based on its depth, accuracy, and scope (Vidulich, 2003).

According to Vidulich and Tsang (2012) both the level of workload and quality of SA are influenced by exogenous (external) and endogenous (internal) factors. Exogenous factors include task complexity, uncertainty, and demands, while endogenous factors relate to an individual's ability and skill level. A given task may impose different workloads on different individuals depending on their expertise. As mentioned earlier, a high skill level effectively equates to having a greater supply of processing resources.

Vidulich and Tsang (2012) further explain that, since SA and workload rely on many of the same cognitive resources, they often compete. High workload can reduce SA by demanding excessive attention, but in some cases, greater effort (e.g., frequent information sampling) is necessary to maintain SA (Endsley, 1993). This means that high workload can be associated with either good or poor SA, depending on whether the operator actively engages in maintaining awareness. However, poor SA does not always lead to higher workload—an individual may simply fail to recognize a deteriorating situation and take no corrective action. Ideally, high SA supports more efficient resource allocation, reducing workload while maintaining performance (Vidulich & Tsang, 2012).

To balance SA and workload, strategic management (executive control) is essential. This includes cognitive functions such as task coordination, resource allocation, planning, information chunking, and inhibition of irrelevant details (Vidulich & Tsang, 2012). Neurophysiological research supports the idea that executive control is a distinct cognitive function consuming processing resources. The prefrontal cortex, responsible for executive control, is more active when tasks require greater planning (Baker et al., 1996), involve larger amounts of information (Braver et al., 1997), or demand multitasking (D'Esposito et al., 1999). Moreover, individuals with frontal lobe damage struggle with higher-order cognitive tasks (Shallice, 1988), further emphasizing the role of executive control in balancing workload and SA.

3.3.4 Complacency (Over trust)

Automation complacency refers to the reduced monitoring of automated systems due to an over-reliance on their reliability. This phenomenon has been extensively studied in aviation, industrial automation, and supervisory control systems (Casey, 1988; Wiener, 1981). While automation can enhance efficiency and reduce workload, it can also lead to reduced vigilance, making it difficult for operators to detect failures when they do occur (Endsley & Kiris, 1995; Kaber et al., 1999).

Automation complacency is particularly problematic when failures are rare, unexpected, or occur after long periods of error-free performance (Bainbridge, 1983; Parasuraman, 1987). Research indicates that when operators are engaged in multiple tasks, their attention is often diverted away from monitoring automation, leading to a decline in failure detection (Parasuraman & Manzey, 2010; Parasuraman et al., 1993). Studies have shown that operators who rely on automation are less likely to detect system failures, particularly in high workload conditions where manual tasks demand more cognitive resources (Wickens & Dixon, 2007).

Empirical studies support these findings. For example, pilots using cockpit automation systems detected fewer engine malfunctions compared to when they manually performed flight tasks (Galster & Parasuraman, 2001). Similarly, air traffic controllers relying on automation to detect aircraft conflicts showed reduced visual scanning of radar displays and were less likely to notice automation failures (Metzger & Parasuraman, 2005). The phenomenon has also been observed in driving simulations, where automated steering failures were often overlooked, leading to delayed or absent manual interventions (de Waard et al., 1999).

Several strategies can help mitigate automation complacency. Research suggests that trust and dependence in automated systems can be calibrated by providing users with automation reliability information (Neyedli et al., 2011; Wang et al., 2009). Another effective strategy for mitigating over-trust and complacency is exposing users to automation failures during training, before real-time use. Studies indicate that when individuals experience a failure early in their interaction with automation, they are more likely to develop appropriate skepticism and verify automated recommendations. This approach contrasts with merely informing users about potential failures, which has been found to be significantly less effective (Bahner et al., 2008).

3.4 UI Design

3.4.1 UI Design Considerations in Mobile Applications

Ensuring the effectiveness of user interaction (UI) design is instrumental in optimizing ease of use and overall output. A good quality UI is designed to greatly reduce the amount of work users have to put in and ease their interaction with digital services. An overly complex interface could cause frustration and disengagement. Increasing the cognitive effort required to interact with an application goes against a key principle that should not be compromised (Banga & Weinhold, 2014, p. 43). One of the primary concerns is the type of touch interaction. Instead of a conventional mouse pointer, users now engage with mobile interfaces using their fingers. Fingers have a greater range of contact of roughly 40 to 50 pixels thus making precise delicate interactions challenging. Hence, larger buttons and touch-friendly features need to be implemented. Additionally, some users may find it hard to perform some advanced gestures while some users will find it intuitive. This demonstrates the importance of simple interactions in mobile devices (Banga & Weinhold, 2014, p. 47).

It is also important to minimize interface friction. Focusing on "touch points," regions where users interact with the content and interface of the system, is one of the most promising ways of doing so (Banga & Weinhold, 2014, p. 97). By optimizing these interactions, users will have a smooth experience which will, in turn, motivate them to keep using the application. In addition, iconography is an important part of an application's usability as it can provide immediate understanding of an application's user interface regardless of the presence or absence of language. This is especially useful in a worldwide information technology market where users come from different countries and languages. (Banga & Weinhold, 2014, p. 97).

Skeuomorphic interfaces also known as skeuomorphism have the potential to enhance usability further by employing design cues mimicking real-world objects, such design will most often lead to more intuitively understood interfaces. Recognising the function of such objects allows for a faster grasp of the purpose and interaction of an application, thereby also reducing the need for explicit instructions (Banga & Weinhold, 2014, pp. 104–105). Modern UI trends, as in, flat design often overlook the importance of skeuomorphism, but it has its benefits when it comes to user control comprehension.

Another common denominator across all these concerns is simplicity. A simple design boosts usability because it decreases a user's mental effort and cognitive load while also enabling a more natural navigation of the interface (Banga & Weinhold, 2014, p. 168). This is in line with the broader goal of creating interfaces that need little to no instructions in order to be used, reducing the need for possible intervention (Banga & Weinhold, 2014, p. 171). Another critical aspect of usability is consistency in UI elements, familiarity with established patterns will allow for a smooth transfer of knowledge from one application to another. There are platform conventions and common design patterns that one can adhere to in order to create interfaces that feel familiar and intuitive (Banga & Weinhold, 2014, p. 174).

3.4.2 Navigation, Mental Rotation, and Cognitive UI Design

Effective mobile app navigation and UI design rely heavily on the cognitive principles of human spatial awareness, mental rotation, and perception. If one understands how individuals perceive spatial relationships and manipulate frames of reference (FORs), it is possible to create intuitive interfaces with less cognitive load and higher usability. This chapter discusses important cognitive factors in mobile application navigation, such as human-centered frames of reference, mental rotation, spatial awareness, and display optimization techniques.

In a human-centered, egocentric, or ego-referenced frame, spatial dimensions are perceived primarily as left-right, front-back, and up-down (Franklin & Tversky, 1990; Previc, 1998). Among these dimensions, there is higher perceptual salience of some axes. For example, the up-down axis is highly differentiated due to environmental cues such as gravity and the sky-ground difference. Similarly, the front-back axis is relevant to hazard perception since what is in front is visible, whereas what is at the back requires special attention (Previc, 1998). Differentiation between left-right is not as salient, and this leads to confusion.

When frames of reference are not aligned, users must perform a frame of reference transformation (FORT), which increases cognitive load, prolongs task completion time, and raises the likelihood of errors (Wickens et al., 2013, p. 125). A common example is the difficulty of navigating with a north-up map while traveling southward, requiring frequent mental rotation for alignment between map and environment

(Wickens et al., 2013, p. 125). To enhance usability, mobile applications should offer both "track-up" (heading-up) and "north-up" orientations, as track-up displays reduce mental transformation effort, whereas a fixed north-up orientation may help to facilitate a broader spatial understanding (Wickens et al., 2013, p. 132).

There exists substantial individual differences in mental rotation ability, which has an effect on how users process navigation interfaces and map orientation (Wickens et al., 2013, p. 127). Since mental rotation demands cognitive resources, interfaces should minimize unnecessary transformations. Navigation applications (maps, radars, GPS, etc.) benefit from dynamic map rotation that aligns with user movement, reducing mental workload.

A narrow perceptual focus is another influencing factor that can limit attention to objects directly in front, thereby reducing awareness of landmarks on the periphery (Wickens et al., 2013, p. 130). Consequently, navigational UI design should emphasize situation awareness (SA) by incorporating landmark similarity and visibility cues. This can be achieved by displaying both the immediate route and a broader area of the map to improve users' understanding of relative and absolute locations (Wickens et al., 2013, p. 131).

The effectiveness of a navigation display depends on scale and information density. Smaller scale maps facilitate global situation awareness by allowing users to view multiple landmarks simultaneously, whereas larger scale maps enhance navigation precision by highlighting route details (Wickens et al., 2013, p. 133).

Small display mobile phones are prone to showing miniaturization effects, where increased item density "scrunches" the items in close proximity so that it is difficult to identify useful information (Wickens et al., 2013, p. 134). Clutter can be prevented by employing highlighting techniques like pre-attentive differences in features (e.g., color coding), which aid the user in efficiently processing data (Wickens et al., 2013, p. 135). However, users' preferences vary as one's mess is another's essential data (Wickens et al., 2013, p. 135). Thus, having customizable presentation settings allows for different cognitive styles.

Consistency of orientation and elements across displays aids user performance by reducing cognitive switching costs (Wickens et al., 2013, p. 136). Abrupt discontinuities in UI layouts cause disorientation, and therefore transitions should be smooth and natural (Wickens et al., 2013, p. 144). Also, directional judgments are made more rapidly when they align with compass headings, pointing to the need for standard orientation cues (Wickens et al., 2013, p. 136).

In navigation tasks, users will often be engaged in closed-loop tracking, wherein they provide ongoing corrections of their inputs based on system feedback (Wickens et al., 2013, p. 146). Tracking success is a function of bandwidth—the frequency at which corrections are required—and system gain, a measure of the ratio of control movement to cursor movement (Wickens et al., 2013, p. 147). Touch screens inherently have a gain of one, since the user's finger serves both as cursor and as control, ensuring natural interaction (Wickens et al., 2013, p. 147). Transmission lag between control input and system response could, nonetheless, disrupt tracking performance, particularly for high-speed navigation tasks (Wickens et al., 2013, p. 147).

Effective navigation and cognitive user interface design for mobile applications necessitate consideration of spatial perception, mental rotation, and display optimization. Transformations of frame of reference should be avoided, dynamic map orientations must be provided, and mental load diminished through clear visual presentation. Consistency, perceptual prominence, and adaptive display characteristics are crucial in order to maximize user experience. Adhering to these concepts, mobile applications can simplify navigation and improve cognitive spatial awareness for a broad range of users.

3.4.3 Icons, Legibility, Recognition, Display Instructions, and Redundancy in UI Design

Icons are an integral part of mobile application UI design for quick and efficient recognition of functions. Research has identified that representational pictures or icons are just as fast to recognize as words and are therefore useful in UI design (Camacho et al., 1990). Good icon design, nonetheless, calls for ample care for distinctiveness and clarity to avoid confusion by users (Wickens et al., 2013, p. 172).

One of the key characteristics of icon usability is 'concreteness', and it refers to how well an icon represents an object or concept from the real world. For first-time users concrete icons are more intuitive, but the strength of concreteness diminishes with familiarity (Isherwood et al., 2007). This does not apply to mobile phone icons, however, as their small size can conceal critical information (Schröder & Ziefle, 2008). Similarly, 'visual complexity' also affects usability; although highly detailed icons can enhance recognition by means of prior knowledge, they do increase search times, particularly on small handheld screens (Huang, 2008; McDougall et al., 2000). One way to decrease the search time is to have a high contrast ratio between the icon and the background upon which it is presented (Huang, 2008).

Another important consideration for icon design is semantic distance, or the closeness of the relationship between an icon and what it is meant to represent (e.g. a printer for 'print'). Isherwood et al. (2007) has found that it is easier to understand icons with a close semantic relationship, particularly for new users. Similarly, 'familiarity' plays a part in usability, where users who have prior experience with an icon or object it depicts are better able to recognize it (Isherwood et al., 2007). In addition, 'aesthetic appeal' can also play a role in usability by reducing the negative impact of visual complexity, shown in a study where more aesthetically pleasing icons produced faster recognition times (McDougall et al., 2009).

Legibility is yet another essential element of UI design, especially for mobile apps with limited display area. It is found through studies that capital letters are more efficient to process when used in short labels, but lowercase letters facilitate better readability for longer blocks of text (Baker & Grether, 1954). Therefore, UI components like navigation buttons and labels may benefit from capital letters, while body text and directions should employ lowercase for better readability (Wickens et al., 2013,

p. 167). Abbreviations and acronyms of text should be avoided, since research has indicated that words are read and perceived quicker than their abbreviations (Norman, 1981). However, when there is not enough space, designers must make a trade-off between bottom-up sensory quality (using larger text for legibility) and top-down processing (using smaller text in order to fit more information) (Wickens et al., 2013, p. 168). Furthermore, high-probability messages (those which are expected by the user) should be short, and low-probability messages should be more informative (Wickens et al., 2013, p. 170).

Display instructions should be designed in order to enhance cognitive processing in relation to environmental constraints. Users are better at the detection of the presence of an unexpected element than the absence of an expected one and, therefore, salient information must always be clearly displayed (Wickens et al., 2013, p. 179). UI guidelines also recommend redundancy, that users are supported by using several different media to aid understanding and accommodate a variety of user needs (Wickens et al., 2013, p. 181). Text and pictures both have complementary functions in UI design, where pictures are useful for spatial location and objects and words to convey abstract actions and ideas (Wickens et al., 2013, pp. 183–184). Voice control can be useful where people's visual attention elsewhere is required, e.g., in conditions of poor visibility. For high-noise levels environments, textual control is preferable. Where user context is in doubt, redundancy, both in voice and text, can assist with comprehension and usability (Wickens et al., 2013, pp. 183–184).

In mobile applications, UI elements can display either status (indicating states) or command (implying actions) information. While command displays can lead to faster decision-making in high-stress contexts (Wickens et al., 2007), there has been little difference in usability between the two formats in some studies (Barnett, 1990; Sauer et al., 2009). As there is inconsistency in the results, a conjunction of both types of displays would give the most ideal outcomes, allowing users to see status information and also obtaining action-oriented instructions (Wickens et al., 2013, p. 178).

Effective UI design for mobile applications entails a strategic compromise among icon comprehensibility, readability, presentation of instructions, and redundancy. Familiar icons with appropriate levels of concreteness must be the primary option for designers, display text in a readable format based on the content length, and implement redundancy where required to achieve understanding. Status and command displays also need to be tailored as per user needs in order to achieve optimum usability under varied circumstances. By leveraging these design principles, mobile applications can provide intuitive and efficient user experiences.

4

Methodology

This chapter goes over the various techniques and methods that will be utilized during the course of the project. Table 1 outlines the general methodological structure divided up into the phases of: Research; Ideation; Prototyping; and Evaluation. These phases will be described in the next chapter.

Table 4.1: The methodological structure for the course of the project

Phases	**Methods**
Research	Literature Review Interviews Thematic Analysis
Ideation	Sketching Crazy 8 Participatory Design
Prototyping	Low Fidelity High Fidelity Rapid Prototyping
Evaluation	NASA-TLX System Usability Scale Interviews

4.1 Qualitative

4.1.1 Interviews

Interviews are a form of qualitative research and is perhaps the most commonly used method for gathering qualitative data (Alsaawi, 2014). Interviewing allows researchers to gain insight into respondents perceptions, understandings and experiences in one or more given subjects (Ryan et al., 2009). There are several different forms of interviews, two of which will be presented below.

1. Semi-structured interview: This is one of the more common methods of conducting an interview. It is a mix of unstructured and structured interviews.

This entails pre-planned questions but they are inherently open-ended allowing for elaborations, explanations, follow up questions and possibly more in-depth responses. There is a requirement that researchers have a level of overview knowledge within the subject matter beforehand (Alsaawi, 2014).

2. Focus group interviews: This method seeks to have the respondents engage in a session of brainstorming moderated by an interviewer asking questions and sometimes structuring the discussion. The number of participants can range from 4 -12. Focus group interviews can generate more in depth rich data as a result of debates between subject experts but it will be harder to transcribe and take notes (Alsaawi, 2014).

Blomberg et al. (1993) make sure to point out that interviewing can not be considered as recording fact. Respondents' answers are inherently subjective, not objective, meaning it is with great care that researchers will have to derive useful insights from the data.

4.1.2 Thematic Analysis

Thematic analysis (TA) is a method for extracting information from one or several sets of qualitative data (Clarke & Braun, 2016). TA does this by identifying, analyzing and interpreting patterns of meaning within qualitative data. The method stresses the active role of the researcher in the process. The building blocks that will result in patterns of meaning can be divided into two parts: codes and themes. Codes are the smallest units of analysis that capture interesting features of the data (potentially) relevant to the research question." (Clarke & Braun, 2016). While themes are larger building blocks that consist of codes who share a central organizing concept.

TA can be used for both small and large data sets and can be either inductive (data-driven) or deductive (theory-driven) (Clarke & Braun, 2016). There are six phases to TA (Braun & Clarke, 2006):

1. Familiarizing yourself with the data: Transcribing data (if necessary), reading and re-reading the data, noting down initial ideas
2. Generating initial codes: Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
3. Searching for themes: Collating codes into potential themes, gathering all data relevant to each potential theme.
4. Reviewing themes: Checking if the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic 'map' of the analysis.
5. Defining and naming themes: Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
6. Producing the report: Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research

question and literature, producing a scholarly report of the analysis

4.2 Quantitative

4.2.1 NASA-TLX

The NASA Task Load Index (NASA-TLX) is a widely used (Vidal-Balea et al., 2024) subjective workload assessment tool developed by Hart and Staveland (1988) to quantify workload across various tasks. It provides a multidimensional measure of workload by considering six contributing factors: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration.

The methodology involves two primary steps: (1) a weighting procedure and (2) workload ratings (Hart & Staveland, 1988). In the weighting procedure, participants compare the six factors in a series of pairwise comparisons to determine their relative impact on workload for the specific task. The number of times each factor is chosen is used to calculate individual weighting coefficients. In the second step, participants rate each factor on a scale, typically ranging from low to high demand. The final workload is computed by multiplying the ratings by the corresponding weights and averaging the results.

NASA-TLX has been validated in multiple studies and is preferred over single-dimensional workload measures due to its sensitivity to task variations and its ability to identify specific sources of workload (Jorgensen et al., 1999; Moroney et al., 1995). It is applicable across diverse operational environments, including aviation, healthcare, and human-computer interaction research (Hart, 2006).

4.2.2 System Usability Scale (SUS)

The System Usability Scale (SUS) is a standardized instrument used to assess the perceived usability of a system, product, or service. Originally developed by John Brooke in 1986 at Digital Equipment Corporation, SUS provides a reliable, low-cost method for obtaining a global measure of usability based on user feedback (Brooke, 1996; Lewis, 2018).

The SUS questionnaire consists of ten items, each rated on a five-point Likert scale (Likert, 1932) ranging from "Strongly Disagree" to "Strongly Agree." These items alternate in tone, with odd-numbered items expressing positive statements about usability (e.g., "I think that I would like to use this system frequently") and even-numbered items expressing negative statements (e.g., "I found the system unnecessarily complex"). This alternating structure helps minimize response bias. The SUS is typically administered immediately after users interact with a system, ensuring that responses reflect their immediate perceptions without the influence of prolonged reflection or external discussion.

SUS produces a single score that reflects a composite assessment of the system's overall usability. Scores are calculated as follows:

1. For positively worded items (1, 3, 5, 7, 9), the response value is subtracted by 1.
2. For negatively worded items (2, 4, 6, 8, 10), the response value is subtracted from 5.
3. The resulting values are summed and multiplied by 2.5 to normalize the score to a 100-point scale.

Sauro (2011) analyzed data from 446 studies and over 5,000 SUS responses, applying a logarithmic transformation to normalize scores and computing percentile ranks. Sauro and Lewis (2012, 2016) used these ranks to develop the curved grading scale (CGS), where the average score (68) corresponds to a C grade. The CGS provides more granularity than an adjective scale and aligns with industry practices, where a score of 80 or higher (A-) indicates an above-average user experience.

Table 4.2: The Sauro–Lewis CGS

SUS Score range	**Grade**	**Percentile range**
84.1-100	A+	96-100
80.8-84.0	A	90-95
78.9-80.7	A-	85-89
77.2-78.8	B+	80-84
74.1-77.1	B	70-79
72.6-74.0	B-	65-69
71.1-72.5	C+	60-64
65.0-71.0	C	41-59
62.7-64.9	C-	35-40
51.7-62.6	D	15-34
0.0-51.6	F	0-14

SUS has demonstrated high reliability and validity across various domains, with a typical Cronbach's alpha reliability coefficient of approximately 0.91, indicating excellent internal consistency (Bangor et al., 2008). It has been widely used in both industrial and academic contexts to compare the usability of different systems, track improvements over iterative designs, and benchmark usability against industry norms (Lewis, 2018).

4.3 Participatory design

Participatory design is an approach that involves the end-users in the design process. By incorporating the future users, and potential stakeholders, in the design process the aim is to better meet needs and expectations (Interaction Design Foundation - IxDF, 2023). Robertson and Simonsen (2012, p. 7) underscore the importance of participatory design by highlighting how users' genuine participation in the design process gives insight into their world that might differ from second hand accounts or what ideal theory specifies. It is also important in the sense that the end products to be used in everyday activities will shape how users go about those same activities (Robertson & Simonsen, 2012, p. 7). There are several benefits to participatory design. It can improve user satisfaction by leading to products better suited for the users. User engagement will increase by them having a more vested interest in the final product because of their participation. It can save time and money by finding flaws in the design process early on and also improve innovation by including unique and more perspectives. Lastly it also ensures a better social inclusion by broadening the spectrum of users and keeping the design accessible (Interaction Design Foundation - IxDF, 2023).

4.4 MoSCoW Method

The MoSCoW method is one of the most commonly used prioritization techniques for managing requirements. The method can help with determining the importance and order of development through categorizing requirements into a hierarchical structure. The four hierarchical categories are *Must Have*, *Should Have*, *Could Have*, and *Won't Have (this time).* By categorizing requirements into these categories, it helps to streamline a project, to allocate time and resources more efficiently. This method facilitates clear communication between stakeholders and developers by aligning expectations and supporting scope management, especially in iterative development environments. Some further advantages of this method is its flexibility and adaptability, clear prioritization, and focus on core functionalities (Kostev, 2023; ProductPlan, n.d.). The four different categories are described below:

- Must Have - Requirements that are critical to the function and success of the system.

- Should Have - Requirements that are important but not more so than the Must Have's and can be postponed if need be.

- Could Have - Requirements that would improve upon the system but are mostly nice-to-have features and would have a small impact on the system.

- Won't Have (this time) - Requirements that fall outside of the scope of this project, these will not be implemented for now or not at all (Kostev, 2023; ProductPlan, n.d.).

Some drawbacks of the MoSCoW method is the inherent bias if used with too few or not the correct stakeholders. There is also risk for ambiguity within the categories if some requirements are valued slightly differently within the team (Kostev, 2023; ProductPlan, n.d.). To mitigate such drawbacks it is important to evaluate requirements separately and collectively, to include all relevant stakeholders early on, and to make sure team members understand each requirement.

4.5 Software Tools

This section will go over the software tools used when working on this project, what the different tools were and how they can be used.

4.5.1 Figma

Figma is a digital design software that allows cooperation within the same file. The tool is mostly used for designing UI/UX prototypes such as interfaces, graphic visuals, wireframes, sketches, and mock-ups. The advantage of Figma is its easy-to-use functionality and ability to include interactivity within the prototypes (Figma, n.d.).

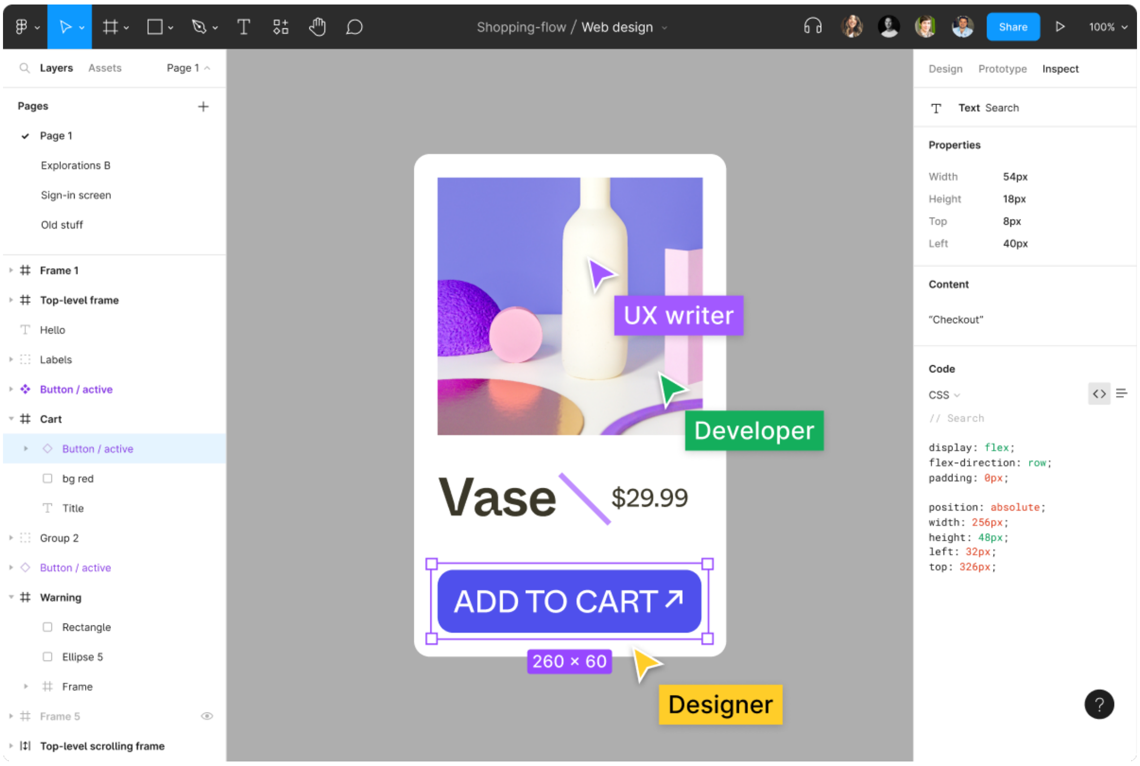


Figure 4.1: Figma Design

4.5.2 Flutter

Flutter is Google's open source user interface (UI) toolkit for building high-performance applications that run on Android, iOS, Web, and desktop from a single codebase (Amadeo, 2018). The toolkit packages each application with a built-in rendering engine that outputs pixel data directly to the display, rather than relying on the native UI libraries of its host operating system. This makes it so that the same UI code can be used for all intended platforms, which simplifies multi-platform support and facilitates a consistent look and feel across multiple platforms. Flutter's hot reload mechanism speeds up development by instantly reflecting code changes in a running app without restarting or losing state (Amadeo, 2018).

The application logic and UI definitions are written in Dart, an object-oriented language designed by Google. Dart facilitates readable, maintainable code, fast startup and smooth runtime performance.

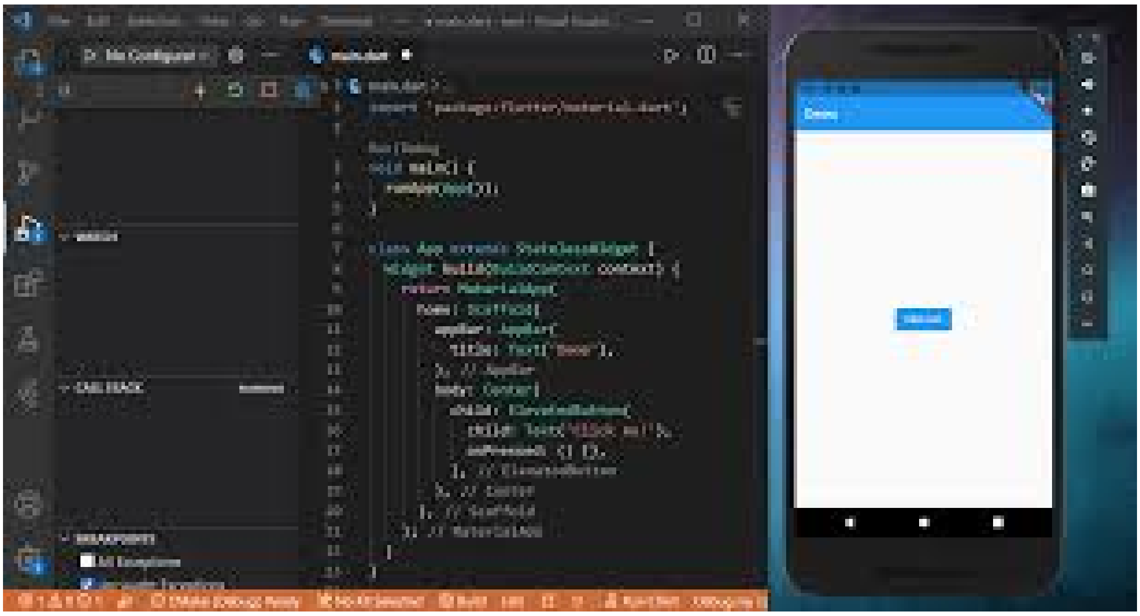


Figure 4.2: Flutter

5
Process

The purpose of this chapter is to detail the various processes and methods used for the project. Iteration was a key aspect of this design process, in which multiple steps were taken more than once throughout the project. The following sections describe the processes involved for each phase: Research, Ideation, Prototyping, and Evaluation.

5.1 Research

This section will describe the processes and methods used to collect and understand the data. The following subsections will include: Literature Review, Interviews, and Thematic Analysis.

5.1.1 Literature Review

The literature review began with a systematic literature search aimed at refining the research question and building a foundational understanding of the relevant knowledge domains. This step was essential for identifying theoretical frameworks, existing technological solutions, and cognitive considerations relevant to the development of a mobile application for wireless ferry control.

A range of keywords was employed to locate pertinent studies across disciplines, including: maritime, ferry, design, user interface (UI), application, wireless control, mobile phone, cognition, situational awareness, attention, information overload, cognitive workload, mental rotation, navigation, and automation. These terms were often combined to narrow the search scope and increase specificity—for example: "wireless control UI ferry", "mobile phone application navigation ferry UI", or "ferry navigation cognitive workload".

Literature was sourced from multiple academic and professional platforms, including Google Scholar, Scopus, the Chalmers Library portal, ACM Digital Library, and general academic databases. Sources were selected based on their relevance to the research questions, the credibility of the publication outlet, and how directly the work addressed the intersection of design, cognition, and maritime control systems.

This review informed the theoretical and practical direction of the project by highlighting challenges in maritime human-machine interaction, cognitive limitations under operational stress, and the importance of designing for situational awareness and reduced information overload in safety-critical environments.

5.1.2 Interviews

The purpose of the interviews in the research phase was to collect information on typical ferry navigation and end users' views on steering with a mobile phone application. The participant profile consisted of maritime experts and operators with at least a Class VIII classification of ship officers. Participants were recruited by contacting several different shipping companies in the Västra Götalandsregionen and throughout Sweden. Recruiting participants with the right background was crucial because only domain experts possess the specific knowledge and hands-on experience needed to evaluate the idea for realistic circumstances.

The interviews were conducted in a semi-structured manner. Several open-ended questions were prepared, allowing for elaboration, follow-up questions, and discussion. The questions were divided up in categories based on the specific field or aspect that they touch (See Appendix A).

There were a total of four interviews with potential end users (n = 4). Three meetings also took place with lecturers and professors in fields such as maritime sciences, ship officer, and human factors engineering. The interviews took place either via a Microsoft Teams meeting or in person in a private setting and were all in Swedish. Before each interview there was a consent form informing the participants about the purpose of the interview, short description of the project, and information about data gathering and management. Three interviews were recorded for future transcription and each participant orally gave their agreement for recording. During each interview there were two interviewers and one respondent. One or both of the interviewers took notes during the interview. While one interviewer had more focus on taking notes, the other asked questions and led the discussion, but the other interviewer intervened when appropriate to ask questions or help.

Transcribing the interviews was done through either the Microsoft Teams transcription function or *VS code*. For the interviews that were not transcribed the data was manually reviewed and read multiple times and the important information was organized. The information was deemed important if the respondent stressed the point, how often it was repeated, how strongly it correlated to the project focus.

5.1.3 Thematic Analysis

In this project, thematic analysis (TA) was employed as the primary method for analyzing qualitative data, following the approach outlined by Clarke and Braun (Braun & Clarke, 2006; Clarke & Braun, 2016). TA was selected for its flexibility in handling both small and large datasets and its capacity to support either inductive (data-driven) or deductive (theory-driven) analysis. The method facilitated the identification, interpretation, and analysis of recurring patterns of meaning within the data, with a strong emphasis on the active and interpretive role of the researcher. The goals of this process were to identify components and functions to be included in the interface, as well as determine the cognitive demands put on ships officers during operation of a vessel.

The analytical process was guided by the conceptual distinction between codes and themes. Codes, understood as the smallest meaningful units of analysis, were used to capture notable features of the data relevant to the research question. In order to identify codes, the transcripts and notes taken by the interviewers were split question by question or section by section and summarized. These summaries were then analyzed and codes were extracted from them. The codes were then gathered, and similar codes were grouped into themes. The themes that were identified pertained mainly towards different aspects of usability and user preferences, but also the nature of navigating on water and the users' aversion towards certain technologies.

Table 5.1: Themes identified through a Thematic Analysis of the Interviews in the research phase

Themes
Intuitive Controls
Simplicity
Redundancy and Safety
Experience and Learning
Steer with sight
Traditionality
Feedback and Responsiveness
Accessibility and Configurability
Environmental Circumstances
Joystick
Digital Advantages

5.2 Ideation

This section will describe the processes and methods used to generate ideas and translate the data into clear design directions. The following subsections will include: Sketching, Crazy 8s, Participatory Design.

5.2.1 Sketching

Sketching is an important step when visualizing and externalizing the insights gathered during the research phase. It requires the designer to produce simple, rough, low-effort sketches that communicate an idea based on stakeholder needs and domain knowledge. Sketching is iterative and informal in nature, allowing for divergent paths and supports multiple alternatives.

In this project sketching served as a critical tool for analyzing the limitations and design requirements that arise when aligning user expectations with the functionality of the mobile application platform.

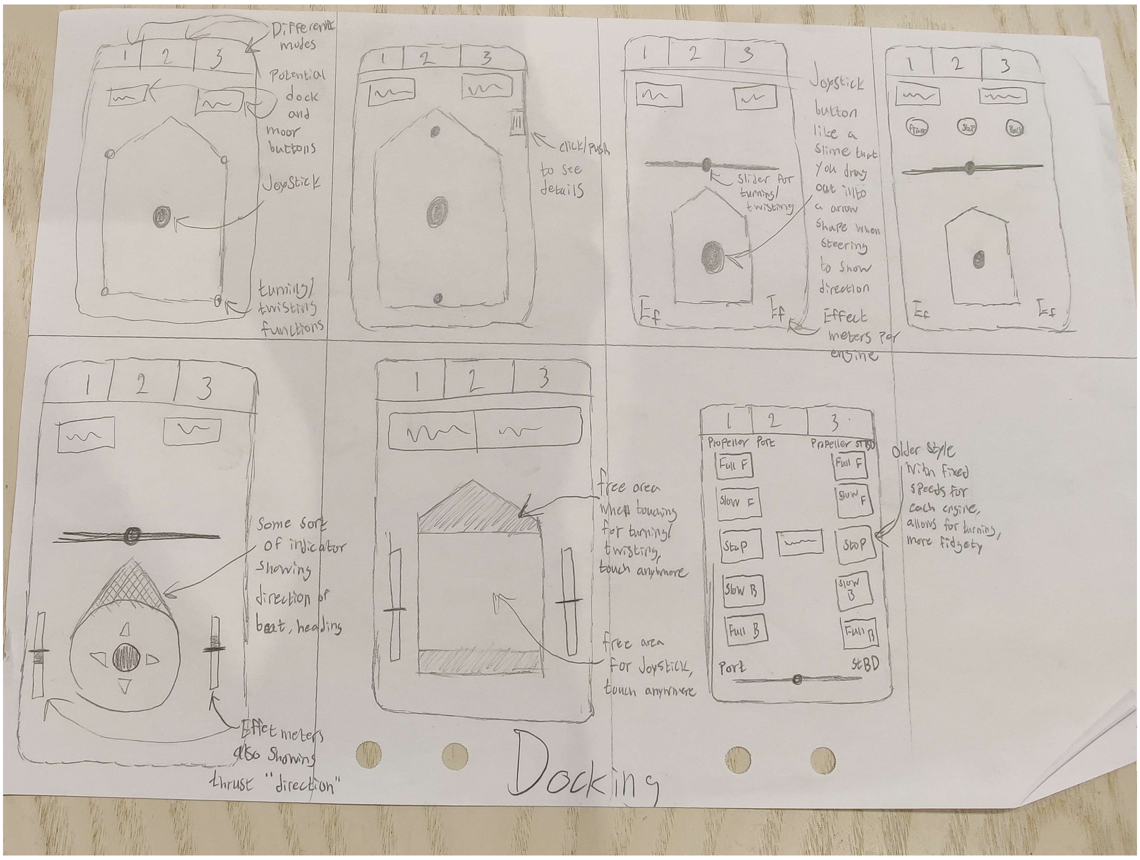


Figure 5.1: Simple Sketches with a focus on Joystick steering

The sketches were made with pen and paper. The pens were both regular pencils and ballpoint pens. The paper used was standard printing paper. After the interviews it became clear that the preferred way of holding the phone was in an upright position. Therefore the sketches proceeded from that information and were drawn in an upright position. The initial sketches help to understand the available space for buttons, functions, font size, and layout.

5.2.2 Crazy 8's

Crazy 8s is a rapid ideation technique commonly used in early-stage design processes to encourage creative thinking and generate a wide range of ideas in a short amount of time. The method involves sketching eight distinct ideas in eight minutes, pushing participants to think beyond their initial, most obvious concepts. The emphasis is not on the quality or detail of the sketches, but rather on quantity, speed, and originality.

This method is particularly useful for divergent thinking, where the goal is to explore a broad solution space before narrowing down. It lowers the barrier for participation by requiring minimal materials, typically just a sheet of paper folded into eight sections and a pen, though digital tools can also be used.

In this project, Crazy 8s was used during the ideation phase to rapidly explore interaction models and interface layouts. It proved useful in breaking out of design fixation and promoting a more experimental mindset. By having team members sketch individually before discussing ideas, the method ensured a wider range of perspectives and reduced the risk of groupthink. The outcomes from this session informed the development of early low-fidelity prototypes and set the stage for further iteration and evaluation.

5.2.3 Participatory Design

To ensure that a design aligns closely with the needs of end users, it is essential to involve them directly in the design process. Participatory design facilitates the inclusion of experience-specific insights that designers alone may not have access to. It also fosters a sense of ownership and investment among users, which can enhance long-term engagement and acceptance of the final product.

Given that this project targets a specific user group and is being developed for a particular company, early understanding of both user preferences and organizational expectations was crucial. Actively involving stakeholders helped align the product's direction with practical needs and constraints from the outset.

Throughout the project, weekly evaluations and brainstorming sessions were held with a representative from Cstrider. These collaborative meetings were instrumental in deepening the mutual understanding of the problem space and in generating new ideas through the exchange of diverse perspectives. In addition, user interviews were conducted in which participants were invited to evaluate and critique early design concepts. This feedback quickly highlighted which ideas were viable and which were less likely to succeed in practice.

5.3 Prototyping

This section will describe the processes and methods used to elaborate on the ideation phase. The following subsections will include: Low-Fidelity Prototyping, High-Fidelity Prototyping, Rapid Prototyping.

5.3.1 Low-Fidelity Prototyping

Low-fidelity prototyping plays a vital role in the design process by enabling designers and stakeholders to visualize and discuss the concept and direction of a product before investing significant resources. Sketching, when sufficiently concrete, is a common form of low-fidelity prototyping that supports early exploration of ideas.

Following the ideation phase, iterative development and ongoing feedback are essential. Starting with low-fidelity rather than high-fidelity prototypes provides flexibility and promotes a step-by-step process that ultimately results in a more refined and user-validated product. Investing too early in high-fidelity design can be inefficient if the solution fails to meet user or business requirements (Interaction Design Foundation - IxDF, 2018b).

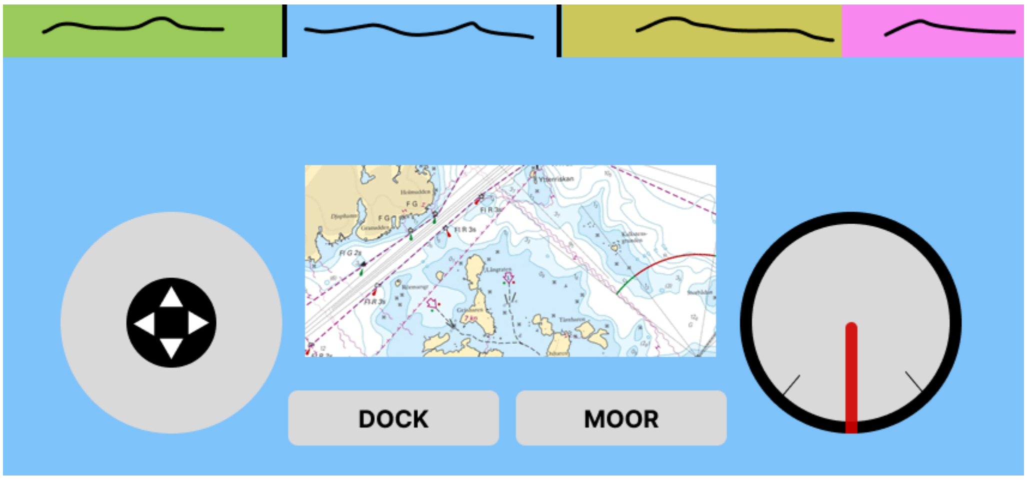


Figure 5.2: Simple wireframe with circular joystick and seamap.

In addition to paper sketches, low-fidelity prototypes were also developed using Figma. While initial sketches helped define the basic structure, Figma allowed for further exploration of layout and interaction. Design elements such as color, iconography, legibility, navigation, and button size were more clearly articulated and could be evaluated in context. Wireframes were used to map the flow between views and modes, providing a more holistic understanding of the application.

User feedback highlighted several important design insights. For instance, the circular joystick function in Docking Mode (See Figure 5.2) was identified as problematic due to issues related to mental rotation and spatial orientation. Additionally, the seamap feature was considered unnecessary, as users reported a preference for keeping their attention on the environment rather than looking down, and seldom used seamaps during river navigation.

To support varying user preferences and improve usability, the design incorporated multimodal interaction, allowing users to adjust values using either a slider or increment/decrement buttons. This approach acknowledges the importance of adaptable interfaces in accommodating different interaction styles.

5.3.2 High-Fidelity Prototyping

High-fidelity prototyping is best used when there is substantial theory and feedback grounding the design choices. A high-fidelity prototype can yield very valuable information that might be hard to come by from earlier iterations. When the prototype resembles the final product, something that looks more polished and has some degree of functionality, user-testing can give deeper more applicable results. It is also more likely that the feedback received from users on high-fidelity prototypes will focus on superficial details rather than the content and the whole (Interaction Design Foundation - IxDF, 2018a). In this project the high-fidelity prototypes were developed using Figma and Flutter in VS Code.

The prototype developed in Figma focused on realistic design visuals such as typography, color schemes, icons, and near-final layouts. Interactivity in Figma is limited

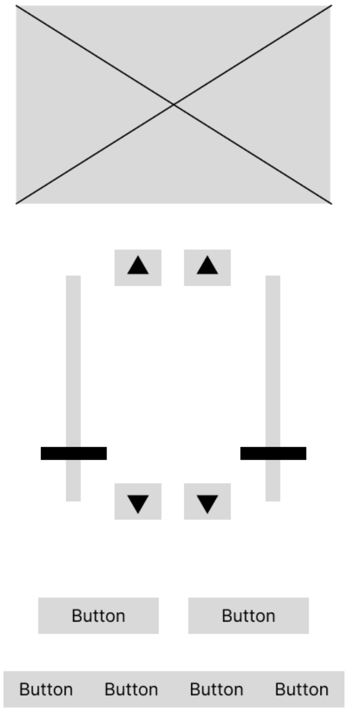


Figure 5.3: Simple wireframe in upright position.

in terms of what can be simulated, but this prototype focused on navigation flows, input controls, and modal views. This level of detail and interactivity ensures that the user can get a near-final experience of what the product would look like and function.

The prototype developed using Flutter in VS code focused on programming a working prototype in terms of functionality. There was less focus on visual design elements because of the limited time available and the existence of a visually clear prototype in Figma.

5.3.3 Rapid Prototyping

The core idea of rapid prototyping boils down to fast iterations and continuous evaluation. This is not a standalone method but rather an approach to the design process. By receiving feedback on low-fidelity prototypes, sketches included, it can validate the design direction and choices. Using the feedback gained from the stakeholders the design is iterated upon and progresses further. Then the process keeps repeating itself until a satisfactory high-fidelity prototype has been validated. The continuous evaluation from designers and stakeholders creates a product that more closely aligns with user needs and expectations, minimizing potential dissonance between users and final product.

In this project rapid prototyping was used throughout the design process in the ideation phase and the prototyping phase. In the early phases, paper sketches and Crazy 8s exercises were employed to explore diverse concepts. These were followed by interactive low- and mid-fidelity prototypes created in Figma, which enabled testing of interface layout, navigation flow, and key features such as the joystick control.

Prototypes were regularly reviewed in collaboration with a representative from Cstrider, facilitating quick feedback cycles and informed decision-making. The ability to revise designs in short intervals made it possible to adapt to user needs, such as removing the sea map or simplifying controls for better situational awareness during docking.

5.4 Evaluation

This section outlines the methods and processes used to evaluate the high-fidelity prototypes. The NASA Task Load Index (NASA-TLX) and the System Usability Scale (SUS) were employed as structured reference points to gather insights into perceived workload and usability. However, the results are not interpreted as absolute measures for benchmarking. Rather, these tools were used to support qualitative understanding and comparative evaluation within the scope of the project.

The rationale for using these methods as reference points lies in the exploratory nature of the research and the highly context-dependent use cases. The interaction design is shaped by specific environmental constraints and individual user preferences, which limits the applicability of standardized metrics in drawing broad or final conclusions. Nevertheless, the methods offer a useful framework for identifying patterns, validating design assumptions, and guiding further iteration.

Each of the methods described below were used in conjunction with each other for each evaluation session. The session started with NASA-TLX, followed by SUS, and ended with the interview. The following subsections will include: NASA-TLX, System Usability Scale (SUS), Interviews.

5.4.1 NASA-TLX

The NASA-TLX is typically used as a structured tool to gain subjective assessments of the overall workload across various tasks. To this end the result of using the NASA-TLX is usually a final overall workload score. This project used a think aloud approach to the evaluation. Rather than focusing on the overall score given by the scales used, the main benefits of using these methods was to identify potential pain points in the UI.

Before the NASA-TLX was administered, the instructors gave the participants walk-throughs of both high fidelity UI prototypes, showing them its various features and functions. The participants were also shown a simple diagram of how the boat would behave when using the different joysticks in the UI (see Figure 5.4). The participants were then instructed to perform some tasks using the high-fidelity prototypes and to verbalize their thought process while interacting with the UI. The tasks were formulated in a way that ensured that the user interacted with each feature of the UI:

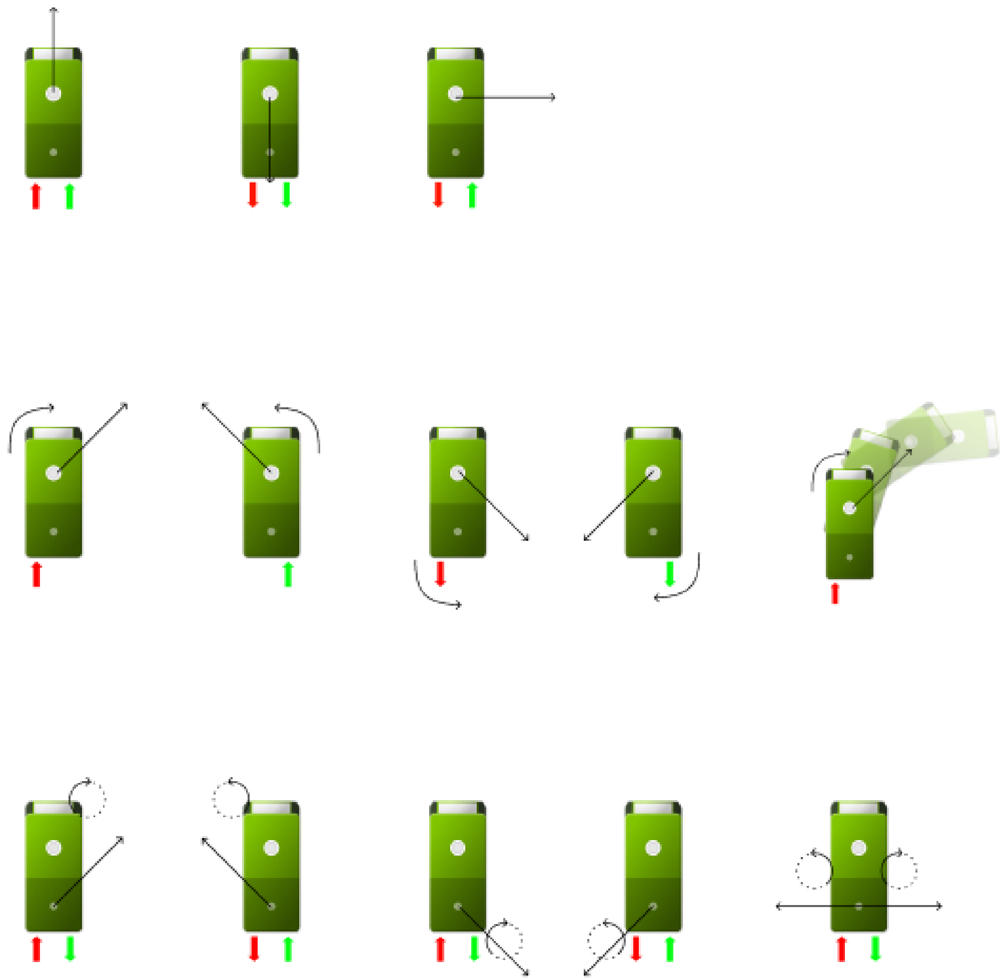
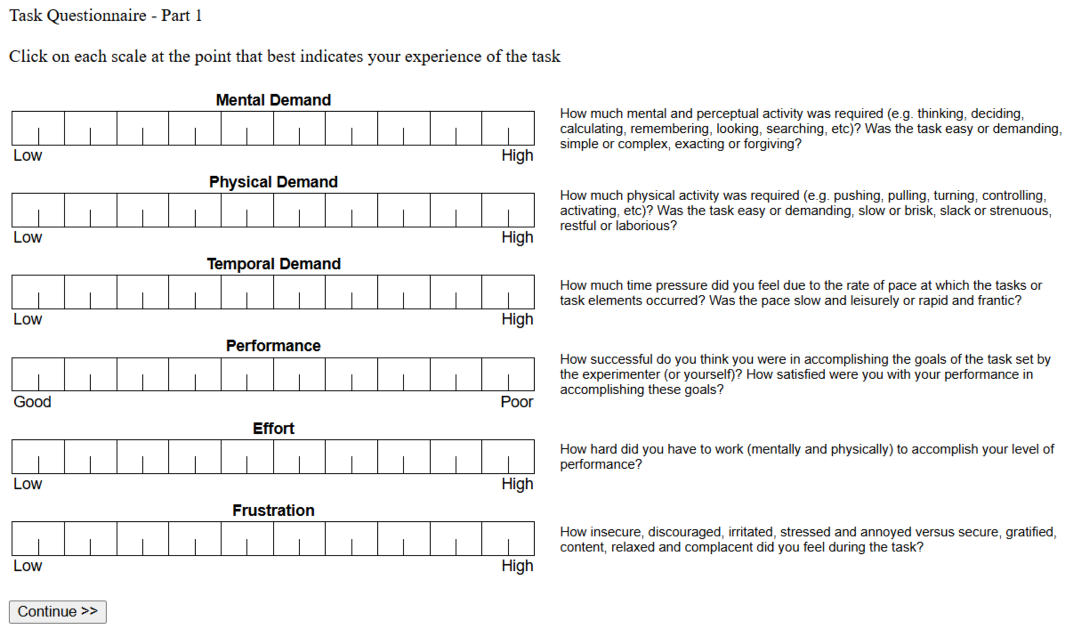


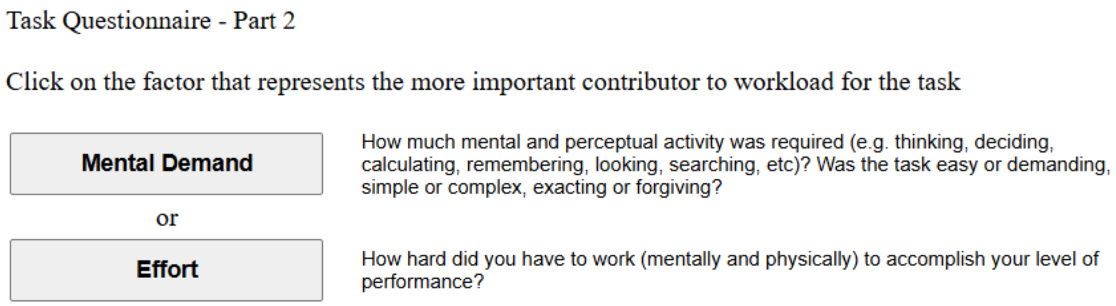
Figure 5.4: Vector and Force representation of steering in Docking mode

1. With the Flutter prototype: Steer full speed forward in docking mode and then perform a U-turn with the stern.
2. With the Flutter prototype: Set the port power to 60% and the starboard power to 55%, also set a course of 30 degrees.
3. With the Flutter prototype: Set the rudder angle to 25° port and speed over ground to 4.6 knots.
4. With the Figma prototype: Find and turn on the auto pilot with the predetermined route Lindholmspiren to Stenpiren.

Immediately after completion of the tasks, the NASA-TLX was administered using a web app (see Figure 5.5). The web app features the typical workload scales and pairwise comparisons, with the addition of scale descriptions next to each scale. The descriptions being taken verbatim from the original pen and paper documentation.



(a) Workload rating scale



(b) Pairwise comparison

Figure 5.5: The NASA-TLX web application (https://www.keithv.com/software/nasatlx/nasatlx.html)

5.4.2 System Usability Scale (SUS)

The SUS was administered immediately after the NASA-TLX in the form of a printed out paper questionnaire (see Appendix B). When filling out the SUS, participants were allowed to interact with the Figma prototype. Any questions posed regarding functions in the UI were answered, but any questions about the rating scales were avoided.

5.4.3 Interviews

The purpose of the interviews in the evaluation phase was to collect feedback on specific design aspects and an understanding on the applicability of the product in the work place. Participants were recruited by contacting several different shipping companies in the Västra Götalandsregionen and throughout Sweden. Interviewing participants with the right background was crucial because only domain experts possess the specific knowledge and hands -on experience needed to evaluate the prototype for realistic circumstances.

The interviews were conducted in a semi-structured manner. Several open-ended questions were prepared for each view of the application, thereby categorizing them according to mode of travel. The interviews were held at the end of each evaluation session with a participant to ensure that the application had been fully explored and sufficiently understood. Any questions the participant had during the interview were answered and they had access to the application to check any aspect again if needed. During the interview one of the interviewers were more responsible for asking questions and directing the conversation while the other interviewer had more focus on taking notes.

6

Results

In this chapter the results from the research phase will be described. The final design will be shown and the different choices will be explained. Lastly the results from the prototype evaluations will be presented.

6.1 Findings from Research Phase

In this section the findings from the interviews and TA (Thematic Analysis) will be presented and collated to create a list of requirements.

6.1.1 Interviews

Early on in the project, several semi-structured interviews were conducted in order to gain insight into the field of ship navigation and the needs and expectations of ship officers. There were several keypoints that were repeatedly stressed by several participants. Below are the findings from the interviews collated and categorized into the different themes found in the TA, which can be found at section 5.1.3.

Intuitive Controls
It was clear during the interviews that the respondents wanted controls that are intuitive and easy to understand. They mentioned *Recognition factors* and wanted the application to *"... to be consistent with reality..."*. One participant said *"Weight should be on function, and not flashy UI."* and clarified that they wanted *"Clear with what is there (in application) and what you can do."*. There was a desire for the controls to look analog but were skeptical of the realizability *"Interviewer: It might be more that one wants to imitate the real controls as much as possible at least. Respondent: If it is possible."* and *"...difficult to resemble analog functions perhaps?"*. They wanted to be able to maneuver in the app the same way as in current practice.

Simplicity
Participants mentioned how the current controls in many ships often have too many notifications and sounds and would like to avoid that *"I would want as few sensors as possible."* and *"... if anything should beep or buzz it should be because of something catastrophic."*. One participant was very clear with that *"Everyone who drives boat professionally are tired of sounds from instruments."*. From this the discussion continued towards controls and one participant said *"I want to fiddle as little as possible when driving."* and continued *"Minimize controls as much as possible when*

docking, only the essentials.". There were skepticism of bundling all the different controls into one application *"A screen with a lot of different controls is hard to debug."* and *"Do not like it when all controls, navigation and steering is in the same unit."*. The phones limited size, they said, meant that not all the valuable controls could fit.

Redundancy and Safety
One major concern from the participants was the issue of safety and alternative ways of steering the boat. They were glad to know that alternative steering methods exist on the boat in case of losing connection on the application *"Feels safe to be able to go to fixed stationary controls and easily switch off the wireless one."*. One participant also said that *"Controls that do not stop feels weird."* in regards to digital controls where the controls could continue even when it in reality should not be possible. One participant was not sure of the safety measures an application would be capable of since the digital controls could imply *"Some problems are hard to notice and alarm."*.

Experience and Learning
Participants were very clear that they often steer from experience, muscle memory, and feeling of the boat. They were still also particular about *"The goal should not be that it takes no learning to use (the application)."* and *"Do not be afraid to demand a little bit more from the user."*. Regarding mental rotation and visuospatial ability they mentioned that it could be hard to train and they have a lot of experience with it through their work. The responses show that individual differences with mental rotation and spatial awareness are difficult to account for. One participant was clear that *"To use a mobile phone application will demand learning."*.

Steer with Sight
From analyzing the interviews it became clear that the participants do not want to look down on the application while steering. Regarding maneuvering, docking, and moments of intense focus one participant said *"There you don't really have the time to sit and look at a screen."*. Having to look down was immediately recognized as an issue *"We use our eyes and our experience, really."* and *"And when you are maneuvering, a chart is not interesting, then you look out. Then it becomes quite distracting. At the same time you want to be able to, you want to have to fiddle as little as possible and still have large instruments, so it is inelegant if you have to switch mode."*. It was also mentioned that they would prefer the phone in an upright position *"If it is in mobile-phone format, then I would probably want it upright, and not just look at it."* and *"I think, if you think landscape, then it feels like then you have to hold it with both hands and look down. But looking down is very unnatural. That is not how you drive a boat."*. From the discussion one could understand that the controls are tools to help with navigation and should not be demanding in attention; one participant said *"Navigation you do by looking out and everything else is an aid to navigation. That's how I think."*.

Tradionality
There was an obvious push back against a digital solution to steering a marine vessel. The idea felt too far from their current practice; one participant said *"I'd rather stand by the levers/joysticks (the fixed stationary controls) than use a touch UI."*.

One participant expressed their skepticism *"I feel like it would be hard to drive in a traditional way in a mobile phone application."* and another comment mentioned *"I have, after all, driven very much with levers and wheels. Then it becomes very natural how one solves, how one drives oneself out of a situation one had not planned. I think it can be difficult with some type of touchscreen solution that is not traditional in the same way."*. At the same time several participants said that they know they themselves are skeptical of change and like the traditional way, that it might be hard to convince them to adapt. One participant said *"I find it really interesting, and I belong to that generation that gets a bit annoyed with modern (...) when it's been too much digital and fiddling, and then there's someone sitting there...It could be your role, but there's someone who doesn't know boats in this case—(and I'm not saying you don't)—but someone who thinks, "This is probably good," (...) But it's coming from my Spotify world, so to speak, it's... So, that way of seeing things is very natural for you, but maybe it's not very boat-like."*

Feedback and Responsiveness
Another major concern the participants had was the issue of feedback and responsiveness. They talked about feeling the feedback from the boat when steering with a traditional joystick and that they want important feedback. One participant said that they *"... combine vibrations and engine sound to know what is going to happen."*, another participant mentioned that *"Some form of haptic feedback would be good if that is possible for the application."*. It was clear that fast response, low latency, to actions within application was important and that they were skeptical of the application being fast enough. One point that was brought up said that *"When there's a delay in the controls, it gets harder to keep track of what's happening around you—it just throws off your situational awareness."*.

Accessibility and Configurability
A matter that was repeatedly brought up concerns that what is natural for some might not be so for others. When discussing visuospatial ability and mental rotation one participant said *"It's very personal, you probably need to talk to a lot of people about that because I was talking about spatial ability. I can turn myself around, so usually do the opposite and I say correctly what is starboard and port and so on. But I know that not everyone really has that ability and instead wants to have, that now it turns. So it might need to have both options depending on what preference you have."*. The subject of color coding and the option of day and night mode was brought up. The matter of different finger sizes and finger dexterity when using the application was brought up *"That's how it is—you might be wearing gloves, so then maybe the display doesn't work or maybe it becomes clumsy or it's too small, because it doesn't work with gloves. You might be able to steer on the display, but the finger might be too thick (on mobile), and so there are so many users."*.

Environmental Circumstances
Environmental circumstances can play a part in which controls are useful and how much. One participant said that *"Exact course is not interesting when you are not out at sea."* and *"Course holder is important but not an absolute value."*. When talking about navigating one participant explained that *"When you steer inside of tight waterways you need to make decisions all the time."* and that *"The seamap is*

not interesting when you are maneuvering (in tight waterways or rivers).". How one steers and what controls are interesting can differ a lot between out at sea, in an archipelago, rivers, or commercial waterways.

Joystick
When discussing the design of the joystick several points and preferences were brought up. For example they did not want that *"A small movement of the joystick should not make the engine rev up a lot."* or that *"Joystick at zero should not act as a virtual anchor."*. It was also discussed that the joystick movement should lead to proportional output in the boat and to think of the joystick movement as a force applied to the boat rather than a specific speed desired.

Digital Advantages
Even though participants were skeptical of the idea of a mobile phone application as a steering device there were positive points brought up. One participant said that this *"Steering with an app should be easier to understand for people that are used to mobile phone games and console controllers."*. Some advantages of a digital solution they mentioned was that it would occupy less space and be more cost-effective. The idea of moving around while steering was something they could appreciate and see the advantage of.

6.1.2 List of Requirements

The research phase aimed to gather foundational insights and domain specific knowledge necessary to compile a list of requirements. By conducting a literature review, multiple interviews, and a TA several key design needs and constraints were identified. Using the MoSCoW method a list of requirements was compiled for each main mode of the application. The MoSCoW method categorizes the findings in the research into four different priority categories:

- Must have – Essential features that are critical for the system to function.
- Should have – Important but not vital; they may be postponed if necessary.
- Could have – Desirable features that can improve the user experience, but are not essential.
- Won't have (this time) – Features explicitly excluded from the current implementation but noted for potential future development.

When using the MoSCoW method in this project it was tweaked to fit into the context of the project. The two categories impacted by this are primarily *Could Have* and *Won't Have*.

The *Must Have*'s of each mode is a feature that needs to be in the application for it to work as intended and for it to be desirable. The *Should Have*'s of each mode are not vital but will help with the usability and desirability. *Could Have*'s are contingent on user feedback and design constraints, such as screen space and functionality. Lastly *Won't Have*'s are features that from feedback and Cstrider directives have been deemed as undesirable and not needed.

At the onset of the project there were some requirements given by Cstrider as can be seen in Table 6.1.

Table 6.1: List of Requirements for each Mode from Cstrider

Modes	**Requirements**
Docking Mode	Joystick Release joystick and ferry drifts
Cruise Mode	Set Power Rudder Angle Throttle/Speed stays at input
Autopilot Mode	Set Course Set Speed Throttle/speed stays at input

The requirements were derived from analyzing the themes and codes in the TA and the controls needed to operate the ferry. Many of the requirements for each mode overlapped with each other, such as: Large Buttons, Effect Meters, Skeumorphism, Disconnect Warning/Feedback, Clear and Succinct labels/text, Moor and Dock buttons. There were some requirements that could not be adequately put into the requirements list. One such requirement has to deal with the want from users to not look down on the application when using the Docking mode. There was no description for that requirement because it would hinge on the success from the other requirements and design choices. Therefore the list of requirements focused on less abstract needs and instead included more features and design styles. The more abstract needs, such as not wanting to look down, were kept in mind and focused on more universally throughout the modes. In Table 6.2 are the requirements for the Docking Mode. The requirements for the other two modes can be found in Appendix D in Tables D.2 (Cruise) and D.3 (Autopilot).

6.2 Design Solution

In this section the final prototypes will be presented and explained. They are high-fidelity prototypes that were developed in Figma and Flutter. The Figma prototype is design focused and has less interactivity while the Flutter prototype had more focus on interactivity. For a better overview of the prototypes they will be presented alongside each other.

6.2.1 Docking Mode

This section will go over the different functions and design decisions in the Docking view of the application.

Top bar and Moor button
At the top of *Docking Mode* there is a button for Moor. Tapping this button makes

Table 6.2: List of Requirements for Docking Mode using MoSCoW method

Must Have	Should Have	Could Have	Won't Have
• Joystick* • Turn function on joystick • Unmoor button • Moor button • Disconnect warning • Joystick snapback function • Disconnect feedback, audio and haptic • Byta mode	• Effect meter • Large buttons • Clear and succinct text/labels • Multitouch function (depends on solution)	• Skeuomorphism • Turning direction indicator • Rudder angle (zeroed or not) • Individual settings and preferences	• Look ahead • Seamap • Status indicators

sure the boat stays moored to the dock. When tapped the Moor button changes color and switches the text to 'Unmoor' in order to show activity and give feedback.

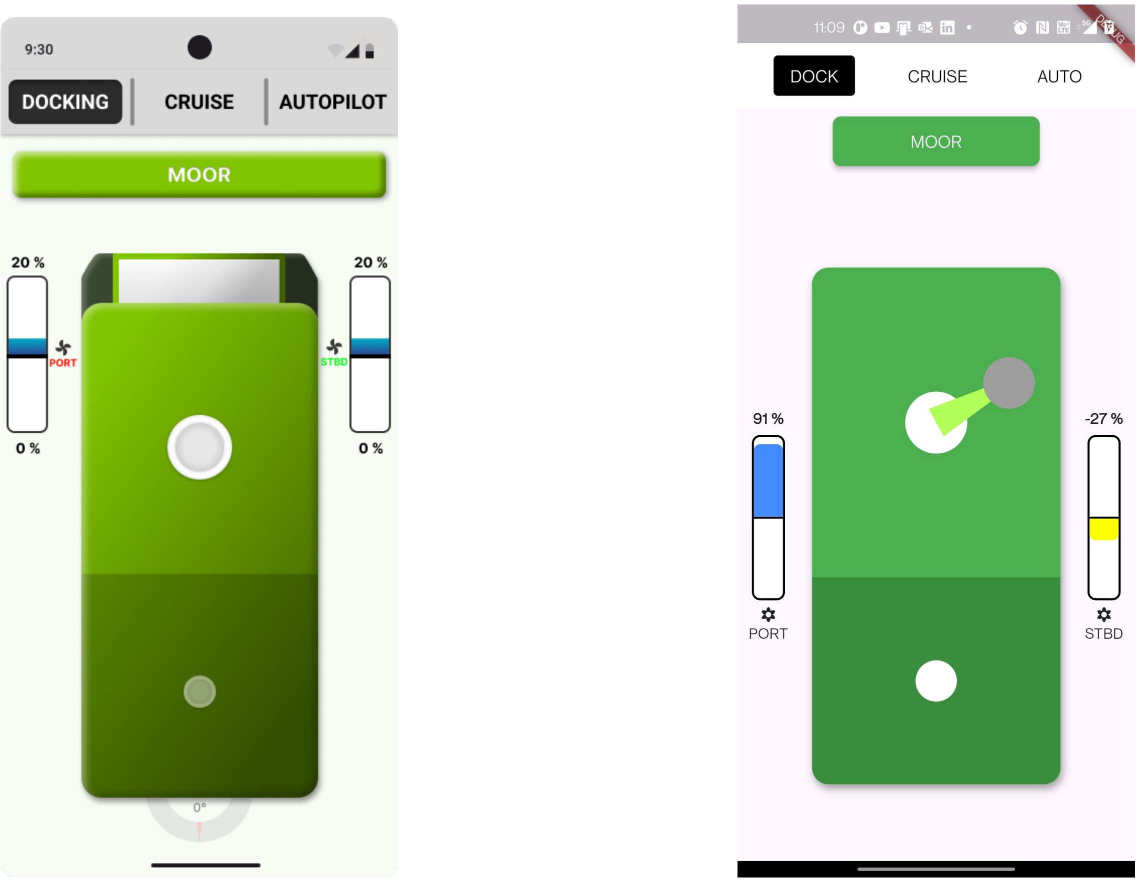
The top bar is a navigation bar between the different modes: Docking, Cruise, and Autopilot. It shows which mode is activated with a gray background. The Moor button and the top bar are the same in every mode.

Joysticks
In docking mode there are two joysticks. The larger joystick is used to maneuver the boat. The direction in which the joystick is moved is the direction toward which the boat will turn. Moving the joystick towards the upper right corner, as seen in Figure 6.1b, will primarily engage the port engine in a forward thrust, resulting in a starboard turn. The opposite is true if you move the joystick towards the upper left corner. Moving the larger joystick towards the lower right or left corners results in the boat engaging a reverse thrust towards port or starboard, as can be seen in Figure 5.4 in section 5.4.1.

The lower, smaller joystick is for rotating the boat around different axis. Moving the joystick in one direction will rotate the boat for the corresponding axis on the boat. This is again visualized with vectors and movements in Figure 5.4 in section 5.4.1.

In the Figma design (Figure 6.1a) the joysticks are made to look more pliable. Increasing the pliableness of the joysticks is to make them pop out more and look more interactive. This is done by adding both inner shadows and drop shadows on the joystick. The joystick that is activated is the one whose surrounding area is lit up. To switch between joysticks the user double taps the inactive joystick. The necessary action of double tapping to switch between joysticks is to increase the safety, by decreasing the risk of accidentally switching between joysticks. There is a small vibration of the phone when the switch is made.

Boat Representation
The area upon which the joysticks are placed is the boat representation. This is a



(a) Figma Docking mode

(b) Flutter Docking mode

Figure 6.1: Docking Mode shown as in Figma and Flutter side by side.

top-down view of the Cstrider ferry designed in Figma. This representation aims to lessen the degree of mental rotation and cognitive workload placed on the ship officer when steering in Docking Mode. The representation removes one degree of possible confusion by clearly showing the joystick movement, the force placed upon the boat, in relation to the boat.

Effect Meters

On the sides of the boat representation there are two effect meters in the shape vertical bars. These meters will show what percentage output each engine has and in which direction. Positive output, forward thrust, is colored blue and negative output, reverse thrust, is colored yellow, as seen in Figure 6.1b. It was important to avoid typical color connections for positive and negative, such as red and green, while still separating them for effective feedback and visibility. For each effect meter there is a propeller icon and corresponding colored text for either *Port* or *Starboard*. Placing the effect meters higher up by the boat representation, as seen in figure 6.1a, increases the visibility of the meters when moving the joystick with one's fingers.

Rudder Indicator

At the bottom of the boat representation there is a rudder indicator (See Figure 6.1a that only shows what angle the rudder is at currently. When steering in docking mode the rudder should be at zero, if it is then the indicator will be at low opacity, if it is not then it will be at higher opacity and show the angle degree. The rudder indicator is there to inform the ship officer of the rudder angle so that they keep the angle in mind when steering, if it is not at zero.

6.2.2 Cruise Mode

This section will go over the different functions and design decisions in the Cruise view of the application (See Figure 6.2). There are two different functions in cruise mode: Navigation and Effect. In the navigation function there are two different features: COG and RUD. COG stands for Course Over Ground and RUD stands for Rudder angle. In the effect function there are two different features: PWR which stands for Power and SPD which stands for Speed. The user will be able to choose which feature of Navigation and Effect that they feel most comfortable using.

(a) Figma Cruise mode

(b) Flutter Cruise mode

Figure 6.2: Cruise Mode shown as in Figma and Flutter side by side.

COG

The COG feature allows the user to set a heading for the boat to follow. There are three different degrees shown in the COG feature: Current Heading, Target Heading, and COG. Current Heading is the direction in which the bow of the boat is pointing. Target Heading is the desired direction, the target, that the boat bow should be pointing in. COG is the direction the boat currently travels with environmental factors included, such as wind and current, it is the deviation of the boats travel direction to the current heading.

The yellow box on the middle of the arc is the current heading (See Figure 6.2a). The current heading is always centered on the arc so that one can quickly see what their current heading is and easily see the number progress towards the target. The part of the arc that is colored yellow shows the progress to the desired target heading. The yellow color is used throughout the different features, and is always a color signifying the remaining progress to the desired target.

The black box with the number 210° (See Figure 6.2a) is the target heading which can be tweaked by either dragging the arc or by tapping the increment/decrement buttons on the sides. When dragging or changing the value there is a very small vibration for each step, this is the same for all values that can be changed. There are two different increment buttons, one arrow means lowering or raising the value by one, and two arrows is instead lowering or raising the value by ten. To confirm the target heading the user have to press the blue button *SET HEADING*. This button serves as a confirmation before changing the navigation, this is there to prevent possible accidental changes. The gray arrow below the arc indicates the target heading and is meant to visualize and help understand the distance between the current heading and the target.

In the upper right corner of the navigation function there is a black box showing the COG (See Figure 6.2a). The COG is not as important in most scenarios and will mainly serve as a check to see how much deviation there is and if it has to be corrected for.

(a) Figma Rudder Angle Feature

(b) Flutter Rudder Angle Feature

Figure 6.3: Rudder Angle feature shown as in Figma and Flutter side by side.

RUD

The Rudder Angle feature lets the user change the angle of the rudder towards port or starboard upwards of 40 degrees (See Figure 6.3). The angle can be changed by either dragging the pointer to the desired degree or by tapping anywhere on the arc to immediately move the pointer to that spot. The port side is red, and the starboard side is green. If the pointer is on the port side then the degree is negative, as can be seen in the black box in the top center, the degree is positive if the pointer is on the starboard side. The yellow line shows at what degree it is currently and the yellow arc shows the remaining distance to the target.

(a) Power Feature Single

(b) Power Feature Double

Figure 6.4: Power Feature on the Effect function.

PWR

The PWR feature lets the user individually control the port and starboard power output (see Figure 6.4a). Two vertical bars marked with 0 at the bottom and 100 at the top show the percentage of power granted to each motor. The red handle can be dragged up or down to continuously adjust the power levels, or the buttons can be pressed to move the power up or down in increments of 1% or 10%. As the power is adjusted, a floating numeric bubble tracks alongside the handle so that the exact power is always known. The 'Double' button can be tapped to toggle to a version of the dial with a single wider bar and a single set of buttons that controls the power of both port and starboard in lockstep (see Figure 6.4b). The SOG (Speed Over Ground) meter on the lower face of the dial shows the actual speed of the boat relative to the Earth, as opposed to its speed through water, taking into account factors such as wind resistance, currents, and tides.

SPD

The SPD feature lets the user choose the boats speed in knots. It features a three-quarter-circle arc that's shaded in a green-yellow-red gradient to give a visual cue of low, medium and high speeds. Along the arc there are major and minor ticks, the major arcs are labeled with whole knot values. A user can either touch the red pointer to drag it smoothly around the arc or tap anywhere on the arc to quickly move the pointer to that spot. When the dial has been moved, a yellow progress arc appears that shows the difference between the current- and desired speed as the vessel speeds up or slows down (see Figure 6.5). The SOG meter functions as in the PWR feature.

Figure 6.5: Speed Feature on the Effect function.

Effect Meters
The effect meters in Cruise Mode (See Figure 6.2a) serve the same purpose as in Docking Mode except that they cannot be negative. Because of a limited amount of space, the meters are instead placed into boxes. With the vertical bars not appropriate the effect meters are instead only showing the percentage output and colored abbreviations for port and starboard.

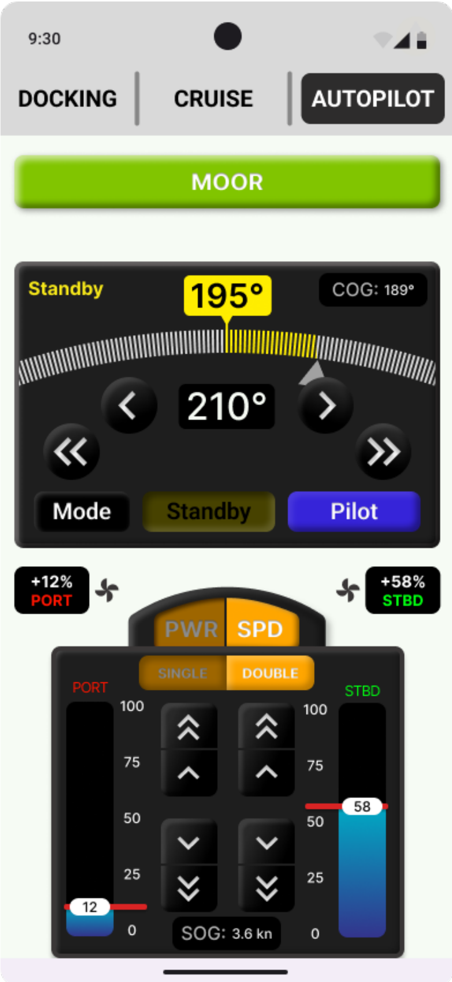
6.2.3 Autopilot Mode

This section will go over the different functions and design decisions in the Autopilot view of the application. The effect function is the same in Cruise Mode and Autopilot and will therefore not be presented.

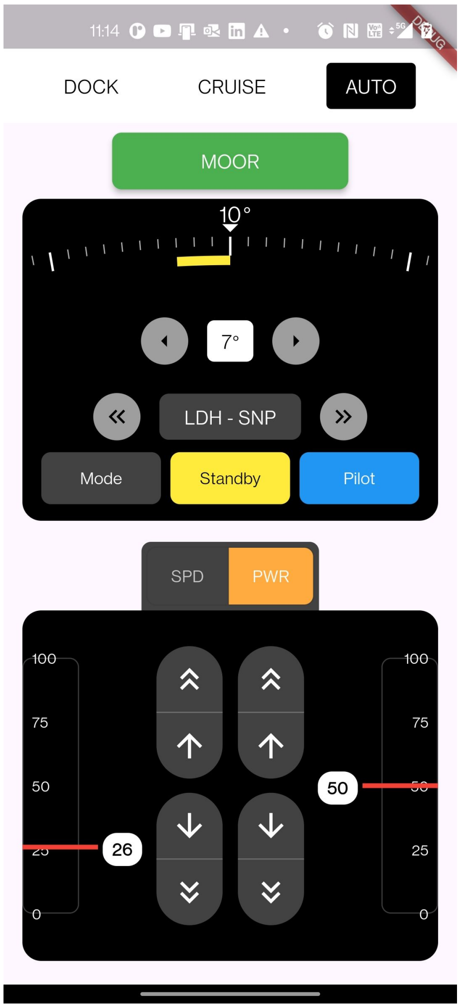
Autopilot Function
The autopilot function shares many similarities with the COG feature. Therefore, only the unique aspects will be presented here.

There are three buttons on the bottom of the autopilot function: *Mode*, *Standby*, and *Pilot* (See Figure 6.6a). When tapping Pilot it is the same as when tapping the Set Heading button, it confirms the route/course. The difference between the two buttons is that Pilot can navigate a plotted route. Tapping Standby puts the Autopilot on standby, meaning it is disengaged. When in standby it is shown in the upper left corner and when tapped it gives immediate visual feedback by changing to a darker hue for a short moment. When tapping Mode it changes the view to a list display of available routes (See Figure 6.7a). In this list display the user can choose a route or a free route. The route start and end destination abbreviations can be seen, as well as when the route was received. When a route is active, meaning that the boat is currently navigating it, there is a green indicator next to the route description. When a route is active it shows up on the standard autopilot function view right below the target heading (See Figure 6.7b). The black box containing the selected route switches every ten seconds between showing the route and showing

Distance To Waypoint (DTW).

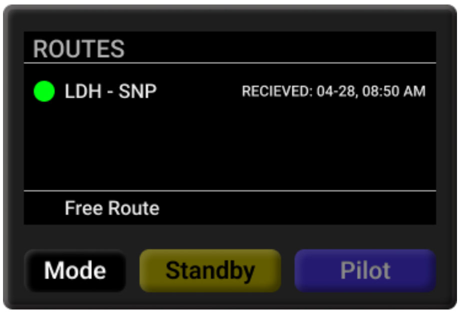


(a) Figma Autopilot Mode

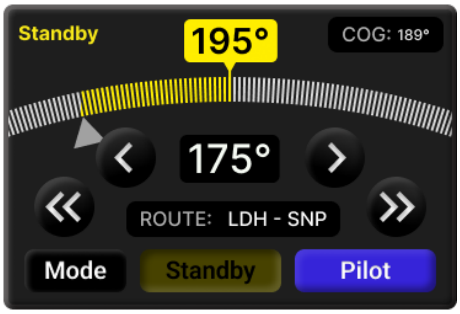


(b) Flutter Autopilot Mode

Figure 6.6: Autopilot Mode shown as in Figma and Flutter side by side.



(a) Route Selection View



(b) Route selected View

Figure 6.7: Route selection and Route On views in the Autopilot function.

Disconnect Screen

This is a view that is universal across the application and shows up if the application loses connection (see Figure 6.8). There is a large red warning indicator and text that directs the user to switch to the nearest alternative steering. There is a rotating loading icon along with a *Trying to reconnect...* text.

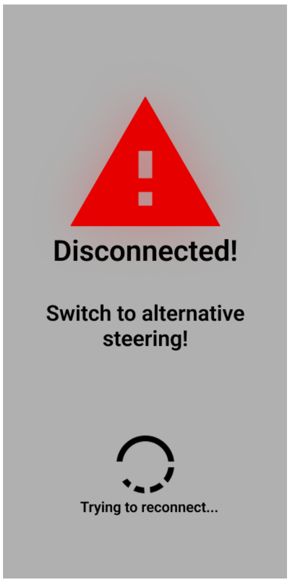


Figure 6.8: Disconnect screen.

6.3 Evaluation Results

6.3.1 NASA-TLX

In this subsection the ratings from the NASA-TLX will be presented. Table 6.3 shows each participant's raw NASA-TLX ratings across the six workload sub-scales. The complete data, with ratings, tallies, weights and the calculated overall workload can be found in Appendix C. The overall workload scores will not be considered due to the low sample size (n = 3). Such a small sample size leads to a very low statistical power and a high sensitivity to outliers. Instead, the factor sub-scales will be inspected in order to identify consistent pain points in the UI. These findings can serve as a basis for a future, more comprehensive, evaluation of the UI when it has been properly integrated into its intended use environment.

Table 6.3: NASA-TLX Ratings

Workload Factor	**P1**	**P2**	**P3**
Mental Demand	25	80	25
Physical Demand	15	50	5
Temporal Demand	20	50	10
Performance	25	80	20
Effort	20	25	10
Frustration	20	35	5

It should be noted that some of the participants (all native Swedish speakers) expressed difficulty during the assessment. Either because of the wordings used to describe the six factors, or difficulty in understanding the pairwise comparison section of the assessment.

Across the three expert mariners, Mental Demand and Performance exhibited the highest average ratings, while Effort and Frustration were the lowest.

Two participants (P1 and P3) consistently reported light to moderate demands on every subscale (all ratings ≤ 25), whereas P2 was a clear outlier, scoring high on Mental Demand (80), Physical Demand (50), Temporal Demand (50), Performance (80), and Frustration (35).

In sum, Mental Demand emerged as the single largest contributor to workload among participants, closely followed by Performance, while Effort and Frustration had the smallest perceived impact.

6.3.2 SUS

Table 6.4: System Usability Scale Ratings

Item	**P1**	**P2**	**P3**
I think that I would like to use this system frequently	2	4	3
I found the system unnecessarily complex	1	2	2
I thought the system was easy to use	5	3	5
I think that I would need the support of a technical person to be able to use this system	1	2	3
I found the various functions in this system were well integrated	4	4	5
I thought there was too much inconsistency in this system	2	2	1
I would imagine that most people would learn to use this system very quickly	4	4	5
I found the system very cumbersome to use	1	2	1
I felt very confident using the system	4	4	4
I needed to learn a lot of things before I could get going with this system	1	1	2
Calculated score	82.5	75	82.5

The SUS scores were calculated by converting each item response (odd-numbered items: score - 1; even-numbered items: 5 - score), summing the adjusted scores, and multiplying by 2.5. As shown in Table 6.4, Participants 1 and 3 each achieved a score of 82.5, while Participant 2 scored 75.0. These scores would correspond to two A grades and one B grade on the Sauro-Lewis CGS. The aggregated score ($M = 80$, $SD \approx 4.33$) would correspond to an A- grade, indicating that each expert mariner judged the phone-based autopilot UI to be highly usable under these test conditions.

6.3.3 Interview

Below the results from the interviews will be presented in a summary format for each category of questions. See Appendix E for the list of interview questions in Swedish.

6.3.3.1 Docking Mode

Is it clear that the joystick is a representation of the boat?
The responses show that the representation of the boat is not immediately clear and that a clearer distinction is needed between stern and bow.

Could use a more standardized representation of a boat that is more clear of what the stern and bow is. Would perhaps lessen the degree of connection to the Cstrider brand and boat.

Do you understand how to steer the boat with the joysticks? Is it easy to switch between joysticks by double tapping?
It is reasonably clear and easy. The respondents show an appreciation for the safety measure that double tapping would create, lessens the risk of accidentally activating the wrong joystick mode.

There was also a clear response that navigating with the joysticks is not immediately apparent how to do, but the larger main joystick was thought of as very easy to learn and adapt to while the smaller rotation joystick was more contingent on the circumstances and might require more learning how to use and in what circumstances.

Is it clear which joystick is activated?
Responses show that it is clear what area is activated, but that it could be even more clear.

Is it logical/intuitive how the rotation works?
The steering is intuitive and easy to learn, but it is not immediately transferrable from regular on board joysticks typically found in ferry's or boats. The effect meters are helpful to understand the rotation and direction of the boat.

Are the effect meters easy to understand and read? Do they help with understanding the vector force of the boat?
Responses show a clear appreciation for the effect meters. They mention how they will not be looking down most of the time but that the effect meters are helpful feedback.

Do you understand the purpose of the rudder angle indicator at the bottom of the boat representation?
The rudder angle at the bottom of the boat representation is helpful and helps in understanding the structure of the boat, meaning that the boat has a centered rudder

6.3.3.2 Cruise Mode

Is the *Set Heading* button necessary? Would you prefer that the heading changes immediately, to not have a confirmation button?
The responses were very divided for whether they want a comfirmation button or not. One respondent was very clear that they would not want it since the extra action of tapping confirm is unnecessary when the adjustment of direction wont overstep

the point chosen. One was fine with it being there but also thought it might be a personal preference. The last thought it was good with a confirmation button, because of the forgiving nature of the function - and that it allows one to set a direction before actually needing to switch from the current one.

Do you feel that the color choice for the buttons is correct, are they clear?
Responses were clear that the colors were fine but would prefer the current button indication to be the highlighted one while the other was grayed out, this is the current design in the Flutter version. One respondent would like the text or text-outline to be highlighted in a contrasting color to make it pop out more. They mentioned accessibility where some combination of red and green could be difficult to differentiate from each other and that the text on the orange switch feature buttons should be black and clearer.

6.3.3.3 Autopilot Mode

Which direction should be in the middle of the arc? Current heading or the target heading?
Responses show a clear preference for the current heading to be the central one on the arc. Two respondents would like a clearer indication of which degree is which, a HDG near the current one and a Target or set for the one you set. There was also a desire for the COG degree to be differentiated more from the goal degree and for it to be put on the arc as well.

Do you understand the purpose of the button mode? Is the word *Mode* appropriate?
If the sole purpose of the button is to see and pick a route then the button could be called route but otherwise the responses were fine with the button being called mode.

6.3.3.4 General Questions

These are general questions that concern the design overall. Some of the general questions were answered in the other categories, therefore the questions below include only the ones that have not been touched upon already.

Is there any wording or abbreviations that do not make sense or that could be clearer?
The main modes of Docking, Cruise and Autopilot could be revisited for a name change, since the cruise mode feature COG shares some functionality with the autopilot function. Unless only Docking mode can dock then the name could be changed to something more descriptive. They were aware of the future possible implementations where autopilot might be able to dock and therefore thought that the names should be more mode descriptive and appropriate. It was no confusion or preference for other abbreviations in the application.

Is there any function that is too small, hard to read or press?
Two respondents thought that the PWR feature could be a bit difficult to get precise values, but were also aware that precise values are not very interesting in the circumstances of the boat navigation. One or a couple of values of does not mean much they said. One respondent mentioned that the buttons with one increment would most likely be unnecessary and it could be changed to a value of five.

Does the navigation through the app feel smooth and easy to understand?
They all said that the navigation within the application felt natural and would be easy to get used to. One respondent mentioned a desire for optional navigation with sliding fingers across the screen to switch between main modes and a button for going back to previous location.

Are the different buttons and functions reachable, if they are not, is there any other way you would prefer it?
Only one respondent felt completely at ease reaching all the functions with only one hand, the others thought that they would use both hands for some features and also that it would be a more comfortable method. This does not mean using both hands to touch the screen but it means one hand holding the phone while the other hand navigates the app.

One respondent mentioned a desire and preference for having the phone be stationary, more like a joystick. They said it could be disorientating to turn around with the phone and then steer.

They were all in agreement that the phone was the limiting factor in reaching some functions and that the application itself was fine in that regard.

7

Discussion

Implementing a mobile phone application to steer and navigate a small passenger ferry has presented both opportunities and challenges. This chapter will discuss the process of which the final solution was created and reflect on the strengths and weaknesses of the final solution.

Besides discussing the project process and final design the chapter will also delve into ethical considerations of the project, limitations of the project, and future work.

By reflecting on both the strengths and constraints of the project, this chapter seeks to situate the study within the broader domain of maritime human-computer interaction and contribute to the understanding of user-centered design for safety-critical applications.

7.1 Reflection on Process

Throughout the project, multiple design approaches and methods were utilized to guide the development of final solution. While no single framework was followed rigidly, the design process primarily followed a combination of participatory design (see section 4.3) and RTD (see section 3.1.2). The iterative nature of RTD with emphasis on collaboration and user involvement ensures the design is user-centered. This is a crucial aspect when designing for a safety-critical and context-specific domain, such as a ferry operator with passenger responsibility.

The combination of participatory design and RTD helped to explore the design of a system intended for a professional and unfamiliar user group. Involving users, even in a limited capacity, proved valuable in uncovering needs and challenges that may not have been immediately obvious at first. This user involvement helped shape both functional requirements and interaction patterns, and highlighted the importance of simplicity, clarity, and reliability in the interface. RTD supported a reflective cycle of prototyping and evaluation, which allowed design directions to evolve based on stakeholder feedback. This method encouraged an engagement with the design space, where each iteration could reveal further insights, for example related to button placement, legibility, or visuals. The flexibility of this approach was especially beneficial in a project where domain knowledge was initially limited and needed to be developed alongside the design itself.

What became evident during the process was that designing for a specialized context such as ferry operation required different priorities than traditional UI design. Rather than focusing on broad accessibility or aesthetic differentiation, the design needed to emphasize robustness, simplicity, and situation awareness. The integration of participatory and prototype-based methods helped to realize specific needs, despite the constraints of limited access to users and real-world testing environments.

7.1.1 Evaluation Phase

During the Evaluation Phase, a mixed method approach was implemented, combining standardized workload (NASA-TLX) and usability (SUS) assessments and semi-structured interviews. The NASA-TLX provided insights into perceived workload and highlighted variations among participants, particularly regarding cognitive demands. The SUS scores contributed an overall assessment of the system's perceived usability and consistency. The interviews complemented these quantitative measures by identifying specific areas of confusion, ambiguity, or discomfort in the user interface. And to get feedback on specific UI elements.

Conducting evaluations with experienced mariners in a controlled, off-vessel environment allowed the identification of interface issues without interference from environmental factors, such as sea state. However, this approach has limitations, notably reduced ecological validity. The absence of realistic maritime conditions may have led to underestimations in the physical, temporal, and frustration workload factors. To accurately assess the usability and workload implications of the UI, future evaluation should involve users actively controlling the vessel in real maritime conditions. This evaluation would ideally be conducted onboard the intended vessel, allowing participants to interact with the interface while receiving immediate and realistic feedback from the boat's movements, environmental conditions, and operational demands.

A clear limitation of the evaluation was the small number of participants ($n = 3$). Because of this, the results can only be viewed as preliminary trends rather than conclusive findings. Such a small sample size makes it difficult to perform reliable statistical analyses, and the results can be significantly influenced by individual differences or outliers. Future evaluations would benefit from including more participants to improve the reliability and usefulness of the results. Participants also reported issues with understanding the NASA-TLX instructions, particularly the terminology and the pairwise comparison component, suggesting a need for clearer instructions or adapted phrasing in future evaluations.

Taken together the three evaluations paint a nuanced picture. The SUS scores place the usability of the interface somewhere between good and excellent, and participants universally gave the UI low complexity scores. At the same time NASA-TLX subscales seem to show that Mental Demand and to a lesser extent Performance are the most prominent contributors to workload, especially for one outlier participant. This could indicate that users feel the UI is easy to navigate, but still invest significant mental effort to map phone-based controls onto real-world vessel behavior.

Two of the three experts reported only light to moderate workload in all TLX factors, matching their high SUS ratings and suggesting that, under calm bench-top conditions, the design functions well.

When looking at the individual SUS scores, the participants seemed to agree that the various functions of the UI were well integrated, that the system would be quick to learn and that they felt confident while using the system. They also disagreed with the statements that the system was cumbersome to use and that a lot had to be learned before they could 'get going' with the system. These scores suggest that users found the layout logical and accessible and that they perceived the UI to be easy to learn. Participants 2 and 3 explicitly mentioned during their interviews that even a brief additional familiarization period (around an hour) would significantly increase their level of comfort with the application, further highlighting the importance of realistic onboard training.

7.2 Reflection on Design

This section will cover the different modes found in the application and the motivations behind the design. This application has a very narrow target group where experience and preferences within that group can vary widely. The solution is not meant to be generalizable, but to fit the context of Cstrider ferry operations. This means balancing the needs and wants of the users to the requirements given by the company.

7.2.1 Design Decisions

The design of the application was informed by the research done at the beginning of the project and the feedback received from end-users. The design wanted to achieve a simple and intuitive UI that would minimize potential dangers of information overload, mental rotation, lack of situation awareness, and complacency. A positive outcome would be an application that, at the minimum, would work as an alternative to the stationary controls when needed.

Compared to traditional stationary control systems, which are larger, the application is limited by the size of a phone. The design aimed to solve this by retaining only essential functions and not straying far from familiar system visuals. This would hopefully keep the application simple and intuitive, which boosts usability, and reduces mental effort (Banga & Weinhold, 2014). The different functions such as COG, RUD, PWR, and SPD tried to resemble physical controls as much as possible. The intention was to make use of the mariners' existing procedural and semantic knowledge that they have built using traditional stationary controls (Poldrack & Packard, 2003; Reder, 1996). It can be concluded from the evaluation that using skeumorphism did indeed help the users understanding what the functions did and how to work them, in line with research by Banga and Weinhold (2014). It can be seen from the evaluation, see section 6.3.3, that some of the labels and text need more legibility. This could be solved by using stronger contrasting colors (Huang, 2008; Wickens et al., 2013), such as on the orange buttons using black text and having a clearer distinction of which button is pressed. The application had to make a trade-off between readability and design in the case of labels and abbreviations (Wickens et al., 2013, p. 168).

In all three modes there is very little use of icons, this was a conscious decision enforced by the available space and the possibly unclear connection between features and icons. The icon used is solely the icon representing the propellors by the effect meters. This icon works because of the limited and consistent use and the concreteness of it (Isherwood et al., 2007). Label design instead went with abbreviations and text to increase the cognitive clarity, keep a consistent design language and reduce potential mental workload.

The navigation of the application focused on safety and accessibility. By placing the different modes in a top bar instead of in a bottom bar the reach is further but accidentally switching modes is less likely. The consistent layout of the functions, features, and buttons is meant to decrease cognitive switching costs and boost user performance (Wickens et al., 2013, p. 136). By placing the functions and buttons in the center of the application the hope was that it would suit both right handed and left handed users. Looking at the SUS scores from each participant it can be interpreted that the experience of using and navigating the application was comfortable and smooth.

Docking Mode
The feasibility of the application being practical and accepted as an alternative steering system hinged mostly on the docking mode design. There are big differences between a traditional stationary joystick or steering system and a wireless on boat control system. Therefore the docking mode had to be a comfortable bridge for the end-users between the traditional controls and the new controls. The focus for this mode was always on making it a smooth and intuitive experience.

The joystick is a representation of the Cstrider ferry from a top-down view. The motivation behind this is to create a clearer connection between the ship officer and the actual ferry and to maintain a clear connection to the Cstrider brand. By having the joystick be a representation of the boat, the goal was to remove one layer of confusion. When the joystick moves on top of the boat, the direction and movement of the boat should be more easily discerned. As opposed to a more traditional circular joystick, as seen in figure 5.2 in section 5.3.1, the boat representation helps to reduce the mental rotation when, for example looking at the stern and maneuvering. It could be interpreted when using a traditional joystick that one is steering from an egocentric frame centered on the user. When creating the boat representation it was meant to hint, to the user that the steering is based on the egocentric frame of the boat, not the user. From the evaluation, it can be seen that the boat representation was not immediately clear. Having a more standard boat shape as the representation could possibly be more effective in minimizing the mental rotation, but at the cost of losing the direct Cstrider brand connection. A strong point in the design, as seen from the evaluation, was the decision to include the effect meters on the side. The main benefit of the effect meters is that it supports situation awareness by providing immediate information about boat travel direction, rotation, and movement (Endsley, 1993). Although these indicators are primarily beneficial during initial learning, they help users quickly perceive, comprehend, and project the vessel's current and future state (Vidulich & Tsang, 2012). When users are comfortable with steering on the application, the effect meters would most likely only be looked at when the boat is not moving as expected.

As presented in sections 2.2 and 2.3 the biggest contributing factor for maritime accidents is the human factor, information overload and situation awareness. Therefore it was imperative that the docking mode would not require too much attention. The decision of having two joysticks, one to steer and one to only rotate, was made due to the need of being able to perform fine maneuvering. While switching between two joysticks slightly increases workload it's a tradeoff that enhances situation awareness and precision for fine maneuvers, helping operators get a more accurate understanding of the orientation and movements of the vessel (Vidulich & Tsang, 2012; Wickens, 1980). It could on the other hand be a drawback of having to perform an extra action to switch between joysticks when maneuvering, it could impact the smoothness of the steering.

There were discussions about including sensors in the boat connected to the application. The sensors would be placed at the stern and bow and flip the steering of the joysticks, depending on which direction the phone was pointed. This was discarded because of the potential cost of implementing such sensors and the dangers of flipping steering suddenly. One variant of this was to have sensors at the stern and bow but instead of flipping steering controls it always has the representation pointing at the bow of the actual ferry. This was also discarded because of the continual mental rotations that the user would have to perform in order to align the joystick movements correctly with the boat. Another issue of that variant was that the boat representation would not accurately portray the joystick's directional force output on the boat. By requiring manual cognitive engagement rather than automated orientation adjustments, the design maintains active user monitoring and reduces the risk of complacency and inattentiveness (Endsley & Kiris, 1995; Kaber et al., 1999).

Cruise Mode and Autopilot Mode

Both cruise mode and autopilot mode are modes that do not require active engagement. The user can set a course and speed to follow and then put the application down until they need to change the values. Therefore these modes had to be simple and efficient for the users to interact smoothly. In the interviews, it became evident that the preferred method of steering differed between participants, therefore both the navigation and effect functions have two different control methods, COG/RUD and SPD/PWR, respectively. To support users' working memory, the design incorporates redundancy by providing information through multiple sensory modalities, aligning with Wickens' multiple resource theory (Wickens, 2008). By visually representing Speed over Ground (SOG) through redundant features (e.g., text, symbols, visual indicators) the system effectively engages both the visuo-spatial sketch pad and phonological loop components of working memory. This dual encoding helps reduce cognitive load and enhance the user's ability to quickly interpret and remember critical information (Wickens et al., 2013). This principle was applied in the design by numerically displaying the speed over ground (SOG) for both effect features, while also using visual elements to indicate their status, such as a speed pointer in the SPD feature and effect bars in the PWR feature. Another version of this is also applied with the effect meters in docking mode, which aids in understanding the directional force output. The COG and PWR features were designed to allow multimodal interaction, either dragging the arc/effect bar or tapping the buttons to increase or decrease the value. The motivation behind designing both a slider and buttons is that the slider focuses on quick adjustments, and the buttons focuses on precise control.

The yellow progress bars on the SPD, RUD, COG, and autopilot features indicators enhance situation awareness by allowing users to perceive both the vessel's current state and its trajectory towards the target state (Vidulich & Tsang, 2012). By making the vessel's adjustment process transparent, the progress bars help users maintain an accurate mental model of system behavior, which is particularly important when the autopilot is engaged. In this way, visual feedback keeps the user in the loop, supporting continuous awareness and reducing the risk of complacency during automated control (Endsley & Kiris, 1995).

The orange colored buttons adhere to Cstrider brand style and signify switching between two features. The idea behind the Figma design was to utilize skeumorphism, therefore the orange buttons are differentiated between a pressed/active state and a non-pressed/inactive state by a change of hue and realistic shadowing. The darker hue was meant to imply the active state, but according to the evaluation most participants confused the two states. They would have preferred the flutter design version. This points to that there are aspects where skeumorphism is not the best choice of design style, alternatively designing such that the active button is highlighted by a RGB color - as seen in some traditional control systems.

Although the low-engagement nature of Cruise and Autopilot reduces workload, it can also encourage automation complacency (Endsley & Kiris, 1995; Kaber et al., 1999). According to de Waard et al. (1999), automated steering failures are often overlooked, leading to delayed or absent manual interventions. One effective way to reduce over trust and automation complacency is to expose users to automation failures during training, before they begin using the system in real scenarios. Research shows that encountering a failure early in their interaction with automation encourages users to adopt a more skeptical approach, making them more likely to monitor the system and verify automated actions. This method has been shown to be significantly more effective than simply warning users about the possibility of failures (Bahner et al., 2008).

7.3 Ethical Considerations

The use case of this application is in a work environment where the user has responsibility of others and company property. For these reasons it was very important for the application to be deemed safe and easy to use. If the application could be used with the same level of comfort and freedom as the stationary controls then that would be considered a success. There were limitations of what could be designed for possible scenarios of the application losing connection. The disconnect screen shows what would appear on the phone screen in case of disconnection, and the design focused on clarity, immediate course of action, and current status feedback. The design is very simple to enhance the existing text and warning symbol, the use of the red color would signify a problem and pops out on the gray background. The exact action of what the boat and application would do in such scenarios was outside of the project scope and designer abilities. That course of action and functionality lies with Cstrider and the discussion from the interviews regarding safety could serve as relevant data for Cstrider.

It is important to evaluate the application in relation to current practices because it has to uphold a certain standard if the users would ever want to use it. But at the same time, the application in this project, had to be evaluated in a vacuum with no similar products to compare it to and no real work environment. This does not offer a fair chance for the participants to evaluate the application and does not paint the whole picture for Cstrider.

It was very important that participants were aware of the project outline and scope while also what they would be contributing to. If the application did not live up to their expectations, in some level at least, they might feel their contribution has been neglected. That is also why the consent form was important for the participants to read and understand.

Going into this project with little to no experience within the maritime domain meant that it was even more important to listen and respect the users. That is also one reason why the interviews were conducted in a semi-structured manner, because of a lack of pre-existing knowledge from the designers the discussion needed to be open. But it can also be argued for that no pre-existing knowledge minimizes bias and could open up the design space further.

7.4 Limitations

While this study provides valuable insights into the design and implementation of a mobile application for ferry navigation and control, several limitations should be acknowledged.

First, the number of participants involved in both the research interviews and the subsequent evaluation was limited. A larger and more diverse sample could have provided a broader range of perspectives and more robust feedback, potentially enhancing the design and reliability of the findings.

Second, testing was not conducted on the actual ferry, which would have offered a more realistic and contextually relevant environment. Evaluating the application under real operational conditions—particularly with the app actively connected to the ferry's control systems—could have revealed important usability and performance factors that were not observable in the simulated settings used during this study.

Another significant limitation was the research team's limited prior knowledge of the maritime domain. This unfamiliarity sometimes made it challenging to fully assess the relevance or significance of certain interview responses or functional requirements.

Domain-specific nuances may have been underappreciated or overlooked during the design and analysis phases.

Lastly, the topic of mobile-based remote control for small ferries is still relatively underexplored in academic literature. As a result, there was a lack of directly applicable prior research to inform the design process or theoretical framing. Much of the work had to be developed from foundational principles or adjacent fields, which may have constrained the depth of the theoretical grounding.

These limitations highlight opportunities for future research and iterative refinement, particularly through expanded testing, deeper collaboration with domain experts, and broader engagement with the target user community.

7.5 Future Work

The first step of future work involves implementing the suggested changes and feedback identified during the evaluation phase. This includes improving the clarity of certain buttons and labels, optimizing the layout for more effective use of space, and refining visual elements such as the boat representation. Additional minor enhancements, such as adding vibration feedback for function confirmation and integrating a course-over-ground (COG) degree indicator into the arc visualization, could further improve usability and situational awareness.

Following these refinements, the next stage would be to conduct further user testing in more realistic and context-appropriate environments. Specifically, testing the application while connected to the ferry's actual control systems and in operational conditions would yield more meaningful data regarding system performance, mental workload, situational awareness, and user acceptance. Given the novel context of operating a ferry via a mobile phone interface, it would also be beneficial to introduce a learning period for participants. This would allow for the collection of data reflecting both initial user impressions and post-familiarization experiences.

Prior to conducting such tests, it would be valuable to undertake a more comprehensive study investigating ship officers' attitudes and disposition toward wireless control systems. This research could provide important insights into psychological, cultural, and operational barriers to adoption, and help determine whether a mobile phone is perceived as a credible and trustworthy interface for critical vessel operations.

In parallel with these efforts, exploring alternative interface concepts could also be a fruitful direction. For example, designing a dedicated handheld control unit with physical buttons and haptic feedback may offer a more tactile and familiar experience for operators used to conventional control systems. Such an approach could bridge the gap between modern digital interaction and the need for robust, low-distraction interfaces in high-risk environments.

Looking further ahead, one potential avenue of exploration is the extension of the control interface to enable remote operation from shore. This would involve a deeper investigation into the technical, functional, and safety-related requirements necessary to support remote ferry operation. Evaluating whether a mobile phone remains an appropriate platform in this expanded context, or whether alternative devices or interfaces would be better suited, would be essential.

Finally, comparing the final design against existing commercial or experimental solutions—particularly those used in professional marine contexts—could help validate the effectiveness of the application and highlight areas for continued improvement. Such comparative studies would ideally take place in real working environments to assess not just usability, but long-term reliability, acceptance, and operational efficiency.

8 Conclusion

The aim of this project was to design a mobile phone application for the company Cstrider to safely steer a small passenger ferry through understanding the needs and potential pain-points of the end users, ship officers. This project was done in collaboration with Chalmers University of Technology and Cstrider. Through incorporating the design approaches RTD and participatory design the goal was to understand the needs of ship officers and Cstrider by following the research question:

> **How can a wireless mobile phone application be designed and implemented to ensure safe and efficient navigation and control of a small transport ferry?**

It was found that some of the critical requirements gathered from ship officers for the successful implementation of the proposed solution was steering with sight, simplicity, intuitive controls, traditionality, and safety. Focusing on these requirements the final application utilized skeumorphism to resemble traditional controls, included only the essential tools, followed a simple navigation and layout, and little necessary attentional interaction. User evaluations indicated that the application was generally perceived as intuitive and easy to use. However, participants expressed hesitation about replacing traditional, stationary controls with a wireless mobile interface, especially in a safety-critical context. This can reflect the challenges of introducing new interaction paradigms in fields where reliability and control fidelity are paramount.

Given the novelty of wireless ferry control and the limited body of existing research in this area, the outcomes of this project contribute both a concrete design prototype and a set of insights that can inform future development. The application itself can serve as a design annotation for further iteration, while the findings highlight important directions for future work—such as real-world testing, deeper studies on user acceptance, and exploration of alternative interaction models. This research can facilitate further investigation into how mobile and wireless interfaces can be safely and effectively integrated into maritime navigation systems.

Bibliography

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors*, *37*(1), 85–104. https://doi.org/10.1518/001872095779049462

Alsaawi, A. (2014). A critical review of qualitative interviews. *European Journal of Business and Social Sciences*, *3*(4), 149–156. https://doi.org/10.2139/ssrn.2819536

Amadeo, R. (2018). Google starts a push for cross-platform app development with flutter sdk. https://arstechnica.com/gadgets/2018/02/google-starts-a-push-for-cross-platform-app-development-with-flutter-sdk/

Baddeley, A. D. (1986). *Working memory*. Clarendon Press.

Baddeley, A. D. (1995). Working memory. In M. S. Gazzaniga et al. (Eds.), *The cognitive neurosciences* (pp. 755–784). MIT Press.

Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198528012.001.0001

Bahner, J. E., Hüper, A.-D., & Manzey, D. (2008). Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies*, *66*(9), 688–699. https://doi.org/10.1016/j.ijhcs.2008.06.001

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Baker, C. A., & Grether, W. F. (1954). *Visual presentation of information (vol. 54)*. Wright Air Development Center, Air Research and Development Command, United States Air Force.

Baker, S. C., Rogers, R. D., Owen, A. M., Frith, C. D., Dolan, R. J., Frackowiak, R. S., & Robbins, T. W. (1996). Neural systems engaged by planning: A pet study of the tower of london task. *Neuropsychologia*, *34*(6), 515–526. https://doi.org/10.1016/0028-3932(95)00133-6

Banga, C., & Weinhold, J. (2014). *Essential mobile interaction design*. Addison-Wesley.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Barnett, B. J. (1990). Aiding type and format compatibility for decision aid interface design. *Proceedings of the Human Factors Society Annual Meeting*, *34*(19), 1552–1556. https://doi.org/10.1177/154193129003401928

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bhardwaj, U., Teixeira, A. P., & Guedes Soares, C. (2022). Casualty analysis methodology and taxonomy for fpso accident analysis. *Reliability Engineering & System Safety*, *218B*. https://doi.org/10.1016/j.ress.2021.108169

Blomberg, J., Giacomi, J., Mosher, A., & Swenton-Wall, P. (1993). Ethnographic field methods and their relation to design. In D. Schuler & A. Namioka (Eds.), *Participatory design: Principles and practices* (pp. 123–155). CRC Press. https://doi.org/10.1201/9780203744338

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

Braver, T. S., Cohen, J. D., Nystrom, L. E., Jonides, J., Smith, E. E., & Noll, D. C. (1997). A parametric study of prefrontal cortex involvement in human working memory. *NeuroImage*, *5*(1), 49–62. https://doi.org/10.1006/nimg.1996.0247

Brooke, J. (1996). Sus: A “quick and dirty” usability scale. In P. Jordan, B. Thomas, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

Buchanan, R. (1992). Wicked problems in design thinking. *Design Issues*, *8*(2), 5–21.

Camacho, M. J., Steiner, B. A., & Berson, B. L. (1990). Icons vs. alphanumerics in pilot-vehicle interfaces. *Proceedings of the Human Factors Society Annual Meeting*, *34*(1), 11–15. https://doi.org/10.1177/154193129003400104

Caridis, P. A. (2001). *CASMET CASUALTY ANALYSIS METHODOLOGY FOR MARITIME OPERATIONS final report for publication october 1999*.

Casey, S. (1988). *Set phasers on stun*. Aegean Press.

Charness, N. (1989). Age and expertise: Responding to talland’s challenge. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 437–456). Cambridge University Press. https://doi.org/10.1017/CBO9780511759390

Charness, N. (1995). Expert performance and situation awareness. In D. J. Garland & M. R. Endsley (Eds.), *Experimental analysis and measurement of situation awareness* (pp. 35–42). Embry-Riddle Aeronautical University Press.

Charness, N., & Tuffiash, M. (2008). The role of expertise research and human factors in capturing, explaining, and producing superior performance. *Human Factors*, *50*(3), 427–432. https://doi.org/10.1518/001872008X312206

Chen, Q., Qi, X., Pang, S., Meng, X., Hu, H., & Liu, Z. (2015). Analysis of current status and causes of maritime accidents. *EMLE-15 Proceedings*, 279–283. https://doi.org/10.2991/emle-15.2015.64

Cheng, H.-H. (2023, October). *Trends in maritime autonomous surface ships (mass) technology*. IOP Publishing. https://doi.org/10.1088/978-0-7503-3151-7

Clarke, V., & Braun, V. (2016). Thematic analysis. *The Journal of Positive Psychology*, *12*(3), 297–298. https://doi.org/10.1080/17439760.2016.1262613

de Waard, D., van der Hulst, M., Hoedemaeker, M., & Brookhuis, K. A. (1999). Driver behavior in an emergency situation in the automated highway system. *Transportation Human Factors*, *1*(1), 67–82. https://doi.org/10.1207/sthf0101_7 de la Campa Portela, R. (2005). Maritime casualties analysis as a tool to improve research about human factors on maritime environment. *J. Marit. Res.*, *2*, 3–18.

D'Esposito, M., Zarahn, E., & Aguirre, G. K. (1999). Event-related functional mri: Implications for cognitive psychology. *Psychological Bulletin*, *125*(1), 155–164. https://doi.org/10.1037/0033-2909.125.1.155

Druckman, D., & Bjork, R. A. (Eds.). (1991). *In the mind's eye: Enhancing human performance*. National Academy Press. https://doi.org/10.17226/1580

Durso, F. T., & Alexander, A. L. (2010). Managing workload, performance, and situation awareness in aviation systems. In E. Salas & D. Maurino (Eds.), *Human factors in aviation (2nd ed.)* (pp. 217–247). Academic Press. https://doi.org/10.1016/B978-0-12-374518-7.00008-0

Endsley, M. (1993). Situation awareness and workload: Flip sides of the same coin. In R. S. Jensen & D. Neumeister (Eds.), *Proceedings of the seventh international symposium on aviation psychology* (pp. 906–911).

Endsley, M. R. (1990). A methodology for the objective measurement of pilot situation awareness. *AGARD, Situation Awareness in Aerospace Operations (AGARD-CP-478)*, 1-1–1-9.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, *37*(2), 381–394. https://doi.org/10.1518/001872095779064555

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245. https://doi.org/10.1037/0033-295X.102.2.211

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General*, *119*(1), 63–76. https://doi.org/10.1037/0096-3445.119.1.63

Galster, S., & Parasuraman, R. (2001). Evaluation of countermeasures for performance decrements due to automated-related complacency in ifr-rated general aviation pilots. *Proceedings of the International Symposium on Aviation Psychology*, 245–249.

Gaver, W. (2012). What should we expect from research through design? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, 937–946. https://doi.org/10.1145/2207676.2208538

Glaser, R. (1987). Thoughts on expertise. In C. Schooler & K. W. Schaie (Eds.), *Cognitive functioning and social structure over the life course* (pp. 81–94). Ablex Publishing.

Grech, M. R., Horberry, T., & Smith, A. (2002). Human error in maritime operations: Analyses of accident reports using the leximancer tool. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*(19), 1718–1721. https://doi.org/10.1177/154193120204601906

Haier, R. J., Siegel, B., Tang, C., Abel, L., & Buchsbaum, M. (1992). Intelligence and changes in regional cerebral glucose metabolic rate following learning. *Intelligence*, *16*(3), 415–426. https://doi.org/10.1016/0160-2896(92)90018-M

Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908. https://doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology* (pp. 139–183, Vol. 52). https://doi.org/10.1016/S0166-4115(08)62386-9

Hetherington, C., Flin, R., & Mearns, K. (2006). Safety in shipping: The human element. *Journal of Safety Research*, *37*(4), 401–411. https://doi.org/10.1016/j.jsr.2006.04.007

Hockey, G. R. J., Gillard, A. W. K., & Coles, M. G. H. (Eds.). (1986). *Energetics and human information processing*. M. Nijhoff. https://doi.org/10.1007/978-94-009-4448-0

Horn, J. L., & Masunaga, H. (2000). New directions for research into aging and intelligence: The development of expertise. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 125–159). Oxford University Press. https://doi.org/10.1093/oso/9780198524380.001.0001

Huang, K. (2008). Effects of computer icons and figure/background area ratios and color combinations on visual search performance on an lcd monitor. *Displays*, *29*, 237–242. https://doi.org/10.1016/j.displa.2007.08.005

Interaction Design Foundation - IxDF. (2018a, March). What are high-fidelity prototypes?

Interaction Design Foundation - IxDF. (2018b, March). What are low-fidelity prototypes?

Interaction Design Foundation - IxDF. (2023, March). What is participatory design?

International Maritime Organization. Maritime Safety Committee. (2019). Guidelines for the standardization of user interface design for navigation equipment (msc.1/circ.1609) [Accessed: (no date provided)].

Isherwood, S. J., McDougall, S. J. P., & Curry, M. B. (2007). Icon identification in context: The changing role of icon characteristics with user experience. *Human Factors*, *49*(3), 465–476. https://doi.org/10.1518/001872007X200102

Jorgensen, A. H., Garde, A. H., Laursen, B., & Jensen, B. R. (1999). Applying the concept of workload to it work. In L. Straker & C. Pollock (Eds.), *Virtual proceedings of cyberg 1999: The second international virtual conference on ergonomics*.

Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science*, *4*(1), 56–88. https://doi.org/10.1080/14639220210159735

Kaber, D. B., Onal, E., & Endsley, M. R. (1999). Level of automation effects on telerobot performance and human operator situation awareness and subjective workload. In M. W. Scerbo & M. Mouloua (Eds.), *Automation technology and human performance: Current research and trends* (pp. 165–170). Erlbaum.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kostev, R. S. (2023). Challenges and problems of the moscow method application in erp system implementation. *2023 International Scientific Conference on Computer Science (COMSCI)*, 1–4.

Lewis, J. R. (2018). The system usability scale: Past, present, and future. *International Journal of Human-Computer Interaction*, *34*(7), 577–590. https://doi.org/10.1080/10447318.2018.1455307

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *149*, 5–55.

McDougall, S. J. P., de Bruijn, O., & Curry, M. B. (2000). Exploring the effects of icon characteristics on user performance: The role of icon concreteness, complexity, and distinctiveness. *Journal of Experimental Psychology: Applied*, *6*(4), 291–306. https://doi.org/10.1037/1076-898X.6.4.291

McDougall, S. J. P., Reppa, I., Smith, G., & Playfoot, D. (2009). Beyond emoticons: Combining affect and cognition in icon design. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (pp. 71–80). Springer Berlin Heidelberg.

McKie, R. (n.d.). Maritime autonomous surface ships (mass) and sar [Accessed: (no date provided)]. *International Maritime Rescue Federation*.

McVay, J., & Kane, M. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35*, 196–204. https://doi.org/10.1037/a0014104

Metzger, U., & Parasuraman, R. (2005). Automation in future air traffic management: Effects of decision aid reliability on controller performance and mental workload. *Human Factors*, *47*(1), 35–49. https://doi.org/10.1518/0018720053653802

Moroney, W. F., Biers, D. W., & Eggemeier, F. T. (1995). Some measurement and methodological considerations in the application of subjective workload measurement techniques. *The International Journal of Aviation Psychology*, *5*(1), 87–106. https://doi.org/10.1207/s15327108ijap0501_6

Neyedli, H. F., Hollands, J. G., & Jamieson, G. A. (2011). Beyond identity: Incorporating system reliability information into an automated combat identification system. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *53*(4), 338–355. https://doi.org/10.1177/0018720811413767

Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, *88*(1), 1–15. https://doi.org/10.1037/0033-295X.88.1.1

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*(1), 44–64. https://doi.org/10.1016/0010-0285(75)90004-3

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors*, *29*(6), 695–706.

Parasuraman, R., & Manzey, D. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, *52*(3), 381–410. https://doi.org/10.1177/0018720810376055

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced "complacency". *International Journal of Aviation Psychology*, *3*(1), 1–23. https://doi.org/10.1207/s15327108ijap0301_1

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160. https://doi.org/10.1518/155534308X284417

Parks, R. W., Loewenstein, D. A., Dodrill, K. L., Barker, W. W., Yoshii, F., Chang, J. Y., Emran, A., Apicella, A., Sheramata, W. A., & Duara, R. (1988).

Cerebral metabolic effects of a verbal fluency test: A pet scan study. *Journal of Clinical and Experimental Neuropsychology*, *10*(5), 565–575. https://doi.org/10.1080/01688638808402795

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 245–251. https://doi.org/10.1016/S0028-3932(02)00157-4

Previc, F. H. (1998). The neuropsychology of 3-d space. *Psychological Bulletin*, *124*(2), 123–164. https://doi.org/10.1037/0033-2909.124.2.123

ProductPlan. (n.d.). Moscow prioritization.

Reder, L. M. (Ed.). (1996). *Implicit memory and metacognition*. Psychology Press. https://doi.org/10.4324/9781315806136

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements. *Psychological Science*, *12*(1), 48–55. https://doi.org/10.1111/1467-9280.00309

Robertson, T., & Simonsen, J. (2012). Participatory design: An introduction. In J. Simonsen & T. Robertson (Eds.), *Routledge international handbook of participatory design* (pp. 1–18). Routledge. http://www.routledge.com/books/details/9780415694407/

Ryan, F., Coughlan, M. J., & Cronin, P. (2009). Interviewing in qualitative research: The one-to-one interview. *International journal of therapy and rehabilitation*, *16*, 309–314. https://doi.org/10.12968/ijtr.2009.16.6.42433

Sauer, J., Wastell, D. G., & Schmeink, C. (2009). Designing for the home: A comparative study of support aids for central heating systems. *Applied Ergonomics*, *40*(2), 165–174. https://doi.org/10.1016/j.apergo.2008.03.002

Sauro, J. (2011). *A practical guide to the system usability scale (sus): Background, benchmarks & best practices*. Measuring Usability LLC.

Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: Practical statistics for user research (1st ed.)* Morgan Kaufmann.

Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research (2nd ed.)* Morgan Kaufmann.

Schröder, S., & Ziefle, M. (2008). Effects of icon concreteness and complexity on semantic transparency: Younger vs. older users. In K. Miesenberger, J. Klaus, W. Zagler, & A. Karshmer (Eds.), *Computers helping people with special needs. icchp 2008. lecture notes in computer science, vol 5105* (90–97 (approx.)). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-70540-6_12

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.

Stratmann, T., & Boll, S. (2016). Demon hunt - the role of endsley's demons of situation awareness in maritime accidents. In L. N. in Computer Science (Ed.), *Distributed, ambient, and pervasive interactions: Dapi 2016* (pp. 203–212). Springer International Publishing. https://doi.org/10.1007/978-3-319-44902-9_13

Taylor, J. L., O'Hara, R., Mumenthaler, M. S., Rosen, A. C., & Yesavage, J. A. (2005). Cognitive ability, expertise, and age differences in following air-traffic control instructions. *Psychology and Aging*, *20*(1), 117–133. https://doi.org/10.1037/0882-7974.20.1.117

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*(4940), 301–306. https://doi.org/10.1126/science.2296719

Vidal-Balea, A., Fraga-Lamas, P., & Fernandez-Carames, T. M. (2024). Advancing nasa-tlx: Automatic user interaction analysis for workload evaluation in xr scenarios. *2024 IEEE Gaming, Entertainment, and Media Conference (GEM)*, 1–6. https://doi.org/10.1109/GEM61861.2024.10585425

Vidulich, M. A. (2003). Mental workload and situation awareness: Essential concepts for aviation psychology practice. In P. S. Tsang & M. A. Vidulich (Eds.), *Principles and practice of aviation psychology* (pp. 115–146). Erlbaum.

Vidulich, M. A., & Tsang, P. S. (2012). Mental workload and situation awareness. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics (4th ed.)* (pp. 243–273). John Wiley & Sons, Incorporated. https://doi.org/10.1002/9781118131350.ch8

Vu, V. D., Lützhöft, M., & Emad, G. R. (2019). Frequency of use – the first step toward human-centred interfaces for marine navigation systems. *Journal of Navigation*, *72*(5), 1089–1107. https://doi.org/10.1017/S0373463319000183

Wang, L., Jamieson, G. A., & Hollands, J. G. (2009). Trust and reliance on an automated combat identification system. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *51*(3), 281–291. https://doi.org/10.1177/0018720809338842

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance viii* (pp. 239–257). Psychology Press.

Wickens, C. D. (2001). Workload and situation awareness. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload, and fatigue* (pp. 443–450). CRC Press. https://doi.org/10.1201/b12791

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449–455. https://doi.org/10.1518/001872008X288394

Wickens, C. D., & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomics Science*, *8*(3), 201–212. https://doi.org/10.1080/14639220500370105

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2013). *Engineering psychology and human performance (4th ed.)* Psychology Press. https://doi.org/10.4324/9781315665177

Wickens, C. D., Self, B. P., Andre, T. S., Reynolds, T. J., & Small, R. L. (2007). Unusual attitude recoveries with a spatial disorientation icon. *The International Journal of Aviation Psychology*, *17*(2), 153–165. https://doi.org/10.1080/10508410701328821

Wiener, E. L. (1981). Complacency: Is the term useful for air safety? *Proceedings of the 26th Corporate Aviation Safety Seminar*, 116–125.

Wong, E. (2023). What are wicked problems and how might we solve them?

Zimmerman, J., Stolterman, E., & Forlizzi, J. (2010). An analysis and critique of research through design: Towards a formalization of a research approach.

Proceedings of the 8th ACM Conference on Designing Interactive Systems (DIS '10), 310–319. https://doi.org/10.1145/1858171.1858228

A

Research Phase Interviews

A.1 Swedish version

Allmän förståelse för styrning och navigation

1. Kan du beskriva din erfarenhet av att styra eller navigera mindre fartyg?
2. Vilka är de viktigaste faktorerna du tar hänsyn till för att styra en liten färja säkert och effektivt?
3. Hur brukar du ta emot och bearbeta navigationsinformation vid styrning av en färja?

Uppfattning om mobila handhållna styrningsenheter

4. Vad är dina första tankar om att använda en mobiltelefon som styrningsenhet för en liten färja?
5. Vilka potentiella fördelar eller utmaningar ser du med att använda en handhållen enhet för styrning?

Kognitiva och fysiska aspekter av handhållen styrning

6. Hur tror du att hålla och orientera telefonen (stående eller liggande läge) kan påverka din förmåga att fokusera på styrningen?
7. Vad skulle vara det bästa sättet att säkerställa situationsmedvetenhet när man använder en mobil enhet för navigation och kontroll?
8. Hur tror du att styrning via en pekskärm jämförs med traditionella fysiska kontroller när det gäller användarvänlighet och mental arbetsbelastning?

Nödvändiga funktioner och systemimplementation

9. Vilka funktioner anser du är nödvändiga för att ett mobilt handhållet styrsystem ska vara säkert och tillförlitligt?
10. Hur bör viktig information (t.ex. hastighet, kurs, hinder) presenteras för att möjliggöra snabba och enkla beslut?
11. Vilken typ av feedback (visuell, ljud, haptisk) skulle hjälpa dig att känna kontroll när du styr med en handhållen enhet?

Säkerhet och redundans

12. Vilka risker ser du med att använda en mobilapp för styrning, och hur skulle dessa kunna minimeras?
13. Om en handhållen styrningsenhet tappar anslutning eller slutar fungera, vilka reservsystem skulle du förvänta dig att ha som backup?

Användarförväntningar och preferenser

14. Hur mycket manuell kontroll tycker du att användaren bör ha jämfört med automatiskt stöd vid användning av en handhållen digital styrningsenhet?
15. Vad skulle få dig att lita på en mobil enhet som ett tillförlitligt styrningssystem för en liten färja?

A.2 English Version

General Understanding of Steering & Navigation

1. Can you describe your experience with steering or navigating small vessels?
2. What are the key factors you consider when steering a small ferry safely and efficiently?
3. How do you typically receive and process navigational information when steering a ferry?

Perceptions of Mobile Handheld Steering Devices

4. What are your initial thoughts on using a mobile phone as a steering device for a small ferry?
5. What potential advantages or challenges do you see in using a handheld device for steering?

Cognitive and Physical Aspects of Handheld Steering

6. How do you think holding and orienting a phone (portrait vs. landscape) could impact your ability to focus on steering?
7. What would be the best way to ensure situational awareness when using a mobile device for navigation and control?
8. How do you think steering with a touchscreen compares to traditional physical controls in terms of ease of use and mental workload?

Essential Functions & System Implementation

9. What functions do you think are necessary for a mobile handheld steering system to be safe and reliable?
10. How should critical information (e.g., speed, heading, obstacles) be displayed to ensure quick and easy decision-making?

11. What kind of feedback (visual, audio, haptic) would help you feel in control when steering with a handheld device?

Safety & Redundancy Considerations

12. What risks do you foresee in using a mobile app for steering, and how could they be minimized?
13. If a handheld steering device were to lose connection or malfunction, what backup systems would you expect to be in place?

User Expectations & Preferences

14. How much manual control do you think a user should have versus automated assistance when using a handheld digital steering device?
15. What would make you trust a mobile device as a reliable steering system for a small ferry?

B

System Usability Scale Questionnaire

Participant ______________________ Date __________

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

C

NASA-TLX Results

Table C.1: Participant 1

Workload Factor	**Rating**	**Tally**	**Weight**
Mental Demand	25	4	$\sim$.27
Physical Demand	15	1	$\sim$.067
Temporal Demand	20	2	$\sim$.13
Performance	25	5	$\sim$.33
Effort	20	3	.20
Frustration	20	0	0
Overall $\approx$ 22.67			

Table C.2: Participant 2

Workload Factor	**Rating**	**Tally**	**Weight**
Mental Demand	80	4	$\sim$.27
Physical Demand	50	3	.20
Temporal Demand	50	3	.20
Performance	80	1	$\sim$.067
Effort	25	2	$\sim$.13
Frustration	35	2	$\sim$.13
Overall $\approx$ 54.67			

Table C.3: Participant 3

Workload Factor	**Rating**	**Tally**	**Weight**
Mental Demand	25	4	∼.27
Physical Demand	5	0	0
Temporal Demand	10	3	.20
Performance	20	5	∼.33
Effort	10	2	∼.13
Frustration	5	1	∼.067
Overall = 17			

D

MoSCoW Lists of Requirements

Must Have	Should Have	Could Have	Won't Have
• Joystick* • Turn function on joystick • Unmoor button • Moor button • Disconnect warning • Joystick snapback function • Disconnect feedback, audio and haptic • Byta mode	• Effect meter • Large buttons • Clear and succinct text/labels • Multitouch function (depends on solution)	• Skeuomorphism • Turning direction indicator • Rudder angle (zeroed or not) • Individual settings and preferences	• Look ahead • Seamap • Status indicators

Table D.1: List of Requirements for Docking Mode using MoSCoW method

Must Have	Should Have	Could Have	Won't Have
• Disconnect warning • Rudder angle • Dock button • Moor button • set Course • Byta mode • set Speed • Disconnect feedback, audio and haptic	• Effect meter • Large buttons • Speed over ground meter • Clear and succinct text/labels • Skeuomorphism • Progress to desired goal	• Compass • see Course	• Look ahead • Status indicators • Seamap

Table D.2: List of Requirements for Cruise Mode using MoSCoW method

Must Have	Should Have	Could Have	Won't Have
• Dock button • Moor button • Adjust course • set Course • set Speed • Disconnect warning • Byta mode	• Clear and succinct text/labels • Large buttons • Speed over ground meter • Skeuomorphism • Progress to desired goal • Map with route	• Effect meter • Compass • Preset courses • Distance to waypoint	• Look ahead • Status indicators • Seamap • Plotting

Table D.3: List of Requirements for Autopilot Mode using MoSCoW method

E

Evaluation Phase Interviews

E.1 Swedish Version

Docking Mode

1. Är det tydligt att det är båten som representeras?
2. Är det lätt att byta mellan joysticks genom att dubbelklicka?
3. Är det tydligt vilket område/vilken joystick som är aktiverad?
4. Är det logiskt/intuitivt hur rotationen fungerar? Hade man velat att det hållet man drar åt är det hållet båten roterar?
5. Är effektmätarna lätta att läsa av? Hjälper dem en att förstå vektorkraften av båten?
6. Förstår man rodervinkeln i botten av båten? Är den för synlig eller osynlig?

Cruise Mode

1. Känns det nödvändigt eller bra att ha en knapp för att 'Set Heading', eller borde den ändra heading samtidigt som man justerar, alltså inte en bekräftelse på ens beslut?
2. Är färgvalet för knapparna tydligt, känns det rätt att de är de färgerna?

Autopilot Mode

1. Vilken riktning ska vara i mitten, den man satt eller den man har?
2. När man klickar på knappen 'Mode', förstår ni var ni hamnar och vad man kan göra där? Är valet av ordet 'Mode' rimligt eller borde det vara något annat?

General Questions

1. Är texten och siffrorna som är på knapparna eller på bakgrunden tydliga?
2. Är det något ordval eller förkortning som inte passar in där den är eller som kan bli tydligare?
3. Saknas det någon funktion? Eller är det kanske för mycket funktioner?

4. Känns färgvalet bra på de olika funktionerna, är det tydligt vad de olika färgerna symboliserar för något?
5. Är det någon funktion som är för liten, svår att läsa av eller trycka på?
6. Känns navigationen smidig och lättförståelig?
7. Når man de olika knapparna, om det är för långt för fingret att nå, hur hade du kanske velat ha det annorlunda?
8. Fungerar det för både vänster och höger hand att använda appen?

E.2 English Version

Docking Mode

1. Is it clear that the object being represented is the boat?
2. Is it easy to switch between joysticks by double-clicking?
3. Is it clear which area/joystick is currently active?
4. Is the rotation function logical/intuitive? Would you prefer the boat to rotate in the direction you drag?
5. Are the power meters easy to read? Do they help you understand the vector force of the boat?
6. Is the rudder angle at the bottom of the boat understandable? Is it too visible or not visible enough?

Cruise Mode

1. Does it feel necessary or helpful to have a ‘Set Heading’ button, or should the heading change directly as you adjust it—without requiring confirmation?
2. Are the button color choices clear, and do they feel appropriate?

Autopilot Mode

1. Which direction should be shown in the center — the one you’ve set or your current heading?
2. When you click the ‘Mode’ button, do you understand where you end up and what you can do there? Is the word ‘Mode’ appropriate, or should it be something else?

General Questions

1. Are the texts and numbers on the buttons or background easy to read?
2. Is there any wording or abbreviation that doesn’t quite fit or could be made clearer?
3. Is any function missing? Or are there perhaps too many features?

4. Do the color choices across different functions feel appropriate? Is it clear what each color symbolizes?
5. Is any feature too small, difficult to read, or hard to tap?
6. Does the navigation feel smooth and easy to understand?
7. Can you reach the various buttons easily, or are some too far for your finger? If so, how would you prefer it to be?
8. Does the app work equally well for both left- and right-handed use?