

Visualizing AI-Supported Adaptivity in Command and Control Interfaces

A Study on how AI-Supported Adaptivity can Increase
Operator Efficiency and Reduce Workload

Master's thesis in Industrial Design Engineering &
Interaction Design and Technologies

LINUS HALLENBERG
FILIP PALMGREN

DEPARTMENT OF INDUSTRIAL AND MATERIALS SCIENCE

MASTER'S THESIS 2025

Visualizing AI-Supported Adaptivity in Command and Control Interfaces

A Study on how AI-Supported Adaptivity can Increase
Operator Efficiency and Reduce Workload

LINUS HALLENBERG

FILIP PALMGREN



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Industrial and Materials Science
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2025

Visualizing AI-Supported Adaptivity in Command and Control Interfaces

A Study on how AI-Supported Adaptivity can Increase Operator Efficiency and Reduce Workload

LINUS HALLENBERG
FILIP PALMGREN

Supervisor: Lars-Ola Bligård
Examiner: Lars-Ola Bligård

© LINUS HALLENBERG & FILIP PALMGREN, 2025.

Acknowledgments, dedications, and similar personal statements in this thesis, reflect the authors' own views.

Master's Thesis 2025
Department of Industrial and Material Science
Chalmers University of Technology
SE-412 96 Göteborg Sweden
Telephone +46 (0)31-772 1000

Cover:
Illustration of AI interacting with a situation map in a C2 interface.
Cover by Filip Palmgren, © 2025.

Abstract

The user interface of Command and Control (C2) systems allows the radar operators to interact with the system. It enables them to surveil, assess, and detect anomalies in the air, on land, and at sea, thereby gaining situational awareness. The systems are complex, processing vast amounts of information, which is reflected in the user interface. The large volume of information, in combination with the high workload demands, risks causing cognitive overload. This could potentially lead to degraded task performance and fatigue, which, in high-stakes contexts, can have serious consequences.

To address this issue, this thesis, conducted in collaboration with Saab Surveillance, investigates where AI-supported adaptivity can be implemented in C2 systems, and how it should be visualized. The effects of the visualizations are investigated under varying workload to draw conclusions about efficiency and the operator's mental workload.

The project achieved this through an iterative Research through Design approach, where user interviews and research of the systems led to the creation of visual prototypes within three categories of concepts. These were called *Filter*, *Target Prio-List*, and *View*. All of these concepts were evaluated with participants with operational experience. The results of which provided relevant feedback, showing various potential with all concepts, including possible improvements and future research.

An Adaptivity & Workload Test was developed and conducted, investigating operators' efficiency and Situational Awareness under various workloads. These were performed on interfaces with and without an adaptive *Filter*, comparing the results. The results from the test indicated that, with a low workload, performance and Situational Awareness were similar between interfaces with and without a filter; however, participants perceived their control to be higher without the filter. For a high workload, the performance and Situational Awareness were better with the adaptive filter, as well as the participants' feeling of control.

Concluded, the results indicated that an adaptive filter aids the operator under high workload. *Target Prio-List* and *View* were also seen to have potential, however, they need to be developed and tested further, similar to *Filter*, to be properly evaluated.

Keywords: *Command and Control, Artificial Intelligence, Adaptive User Interfaces, Situational Awareness, Mental Workload, Research Through Design.*

Acknowledgments

The making of this master's thesis has been an exciting and rewarding experience. We have many people to thank for their support and guidance in the challenges we have faced along the way. First of all, a special thanks to our Supervisor and Examiner at Chalmers, Lars-Ola Bligård. Your willingness to always make time to help us and give your valuable input has been a huge comfort.

We would also like to extend a huge thanks to the entire Human Factors Integration team at Saab, both for their support in the work itself and especially for making us feel so welcome from the very first day. We are especially grateful to our supervisors at Saab, Evelina Ekholm and Olivia Bång, for all the advice and consultations you have provided throughout the process.

To all the people who have participated in our various tests and interviews, this thesis would not have been possible without you and your contributions, for which we are incredibly grateful.

Finally, thank you to Chalmers University of Technology for these past years and for providing us with a fantastic education. Both of us have made memories and friendships here that will last a lifetime, and we will look back at this time with fondness.

Table of Contents

- 1 Introduction 1
 - 1.1 Aim 2
 - 1.2 Goal 2
 - 1.3 Delimitations 2
 - 1.4 Academic Contribution 2
 - 1.5 Outline of Thesis 2
- 2 AI and The Military Context 5
 - 2.1 Artificial Intelligence (AI) 5
 - 2.2 Saab 7
 - 2.3 Command & Control Systems 7
- 3 Theory 11
 - 3.1 Cognitive Ergonomics 11
 - 3.2 Gestalt Principles 16
 - 3.3 User Interfaces 17
- 4 Methods 23
 - 4.1 User Tests, Analysis & Evaluation Methods 23
 - 4.2 Idea Generation and Concept Creation Methods 25
- 5 General Description of Procedure 27
 - 5.1 Research Questions 27
 - 5.2 Design Process 28
 - 5.3 Preconditions 28
 - 5.4 Overall Project Structure 29
- 6 Phase A: Initial User Study 33
 - 6.1 Objective 33
 - 6.2 Planning & Execution 33
 - 6.3 Result 35
 - 6.4 Summary of Phase 38

7 Phase B: Idea Generation	39
7.1 Objective	39
7.2 Planning & Execution	39
7.3 Result	40
7.4 Scope Delimitation	41
7.5 Summary of Phase	43
8 Phase C: Concept Creation	45
8.1 Objective	45
8.2 Planning & Execution	45
8.3 Result	46
8.4 Summary of Phase	54
9 Phase D: Concept Evaluation.	55
9.1 Objective	55
9.2 Planning & Execution	55
9.3 Result	59
9.4 Summary of Phase	64
10 Phase E: Adaptivity & Workload Testing	65
10.1 Objective	65
10.2 Selecting Concept for Adaptivity & Workload Testing	65
10.3 Planning & Execution	66
10.4 Result	72
10.5 Summary of Phase	79
11 Design Guidelines	81
11.1 Research Question 1	81
11.2 Research Question 2	82
11.3 Transparency and AI Decision Basis	83
12 Discussion	85
12.1 Aim	85
12.2 Goal	85
12.3 Process and Methods	86
12.4 Ethical Considerations	88
12.5 Operator & AI Interaction in C2 Systems	89
12.6 Future Work	89
13 Conclusion	91

References.....	93
Appendices.....	97
Appendix A – Initial Interview Template.....	97
Appendix B – Initial Interview Survey.....	99
Appendix C – Participants’ Task Diagrams.....	100
Appendix D – KJ-representation of Ideas from Phase B.....	101
Appendix E – Concept Evaluation Template.....	104
Appendix F – Prototypes for Filter Evaluation.....	108
Appendix G – Information for Participants in the Adaptivity & Workload Test.....	110
Appendix H – Subsidiary Task Result Template & Self-Assessment Questions.....	114



1

Introduction

With the rapid development of Artificial Intelligence (AI), new opportunities are opening up in many technological fields. An enormous amount of information can be handled fast, in a manner that resembles human intelligence in learning, decision-making and problem-solving. The defence industry also sees potential in AI, with its need to handle a large amount of information and translate it into the right action.

Saab AB is a Swedish defence company with a broad product portfolio. This thesis is written in collaboration with Saab Surveillance, the business area that specializes in radars. These radars are operated by radar operators, who use Command and Control (C2) systems to monitor and analyse what the radar detects in the air, on land and sea, in order to gain situational awareness (SA). This is useful in peacetime as well as in crisis and war. Due to the vast amount of information, C2 systems are often complex, which transcends to the user interface (UI). The context is serious, and sometimes every second counts in the decision-making, with potential severe consequences. Therefore, operators require rigorous training and are experts at what they do.

One important aspect that affects the efficiency of the UI is the amount of information presented in the C2 system. Too much information influences the operators' mental process and causes the risk of exceeding their mental capacity. Such mental overload directly impacts and degrades their performance of the tasks, including the ability to attain SA and multitask. Adaptivity of the interface, based on situations and the operator's mental workload, may be one way to address this issue. This type of adaptivity requires processing a lot of information and could benefit from AI-support. Hence, this thesis will address AI-supported adaptivity of the UI, applied in a C2 context.

1.1 Aim

The project aims to investigate how AI-supported adaptivity can be implemented in Command and Control (C2) systems, including how it should be visualized for varying workload, to increase efficiency and reduce the operator's mental workload.

1.2 Goal

The goal of the project is to develop design proposals in the form of visual prototypes on how AI-supported adaptivity can be implemented in C2 systems, and to test how these affect the mental workload and SA.

1.3 Delimitations

The delimitations of the thesis are the following:

- Due to limited knowledge of AI coding, an actual working AI will not be developed and implemented. Instead, the functions and decisions meant to be performed by the AI will be replicated manually.
- This thesis will only give a brief introduction to AI and suitable types for the application, not decide or propose a specific type of AI model.

1.4 Academic Contribution

This thesis contributes to the existing research by providing design proposals on how AI-supported adaptivity can be implemented and visualized in C2 systems. Furthermore, it addresses the gap in research about how adaptivity in C2 systems affects efficiency and SA with increasing mental workload.

1.5 Outline of Thesis

The report is structured into the following chapters, including descriptions of their content:

Chapter 1: Introduction – Introduces the general topic of the thesis, including the aim, goal, delimitations and its contribution.

Chapter 2: AI and the Military Context – Describes the context of the thesis, including artificial intelligence, and the systems which are in focus.

Chapter 3: Theory – Provides theory and research which is related to the context and utilized during the project.

Chapter 4: Methods – Presents methods which have been implemented in the project.

Chapter 5: General Description of Procedure – Describes the general procedure and structure of the thesis. It introduces the research questions, preconditions, the design process, including an explanation of research through design method, and finally, an overview of the different phases the project is divided into.

Chapter 6: Phase A: Initial User Studies – Presents the initial phase of the project, focusing on the initial user studies.

Chapter 7: Phase B: Idea Generation – Presents the Idea Generation, its results, and the following delimitation of the scope.

Chapter 8: Phase C: Concept Creation – Focuses on the performed Concept Creation and provides explanations of the visual Concepts.

Chapter 9: Phase D: Concept Evaluation – Describes and presents results from the performed evaluation of the created concepts.

Chapter 10: Phase E: Adaptivity & Workload Testing – Presents the final test, focusing on evaluating adaptivity and workload, in relation to performance and situation awareness.

Chapter 11: Design Guidelines – Concludes the evaluations of the created concepts and relates to answering the project's research questions.

Chapter 12: Discussion – Discusses the thesis as a whole, including aspects such as the aim, goal and process of the project, as well as ethical aspects and suggestions for future work.

Chapter 13: Conclusion – Provides the general conclusion of the thesis.



2

AI and The Military Context

The objective of this chapter is to provide an understanding of the context within which this thesis exists. More precisely, the chapter will cover Artificial Intelligence in society and the military, as well as the company Saab AB and Command and Control systems.

2.1 Artificial Intelligence (AI)

AI is a technology which has rapidly developed in recent years and has found a place in the minds and lives of ordinary people, with generative AI models such as ChatGPT by OpenAI (2025). With AI, machines possess the ability to resemble human intelligence, by learning, making decisions and solving problems. Benefits include greater automatization, faster data processing and improved decision-making, to name a few (IBM, 2024).

Through the years, starting from the 1950s, AI has developed through different types and stages. Back in the 1950s, it was referred to as AI when machines performed human intelligence qualities. In the 1980s, machine learning (ML) emerged. ML is AI with the ability to learn from historical data. Several types of algorithms are used in ML. One of the most popular is neural networks (NN), which mimics the structure and function of the human brain. They are built by node layers that are interconnected, and these cooperate to handle and analyse complex data. Hence, ML models can identify complex relationships and patterns in large collections of data.

In more recent times, the 2010s, deep learning (DL) evolved as a subgroup of ML. They use deep neural networks (DNNs), which are NNs with many more hidden layers than typical NNs (NNs usually have one or two, while DNNs have at least three but commonly hundreds). This allows them to simulate the complex function of the human brain even more. DNNs also enable even better data processing and decision-making from big amounts of unstructured and unlabelled data, creating their own interpretation of it (Kosinski, 2024). In other words, DNNs can “be fed with raw input data and automatically discover a representation that is needed for the corresponding learning task.” (Janiesch et al., 2021, pp. 687-688). The latest step in AI is generative AI, which are DL models that create its own content on a descriptive request from the user, e.g., text, images, videos or sound (Kosinski, 2024).

This thesis mainly investigated adaptive UI design and functions to support the operator. No real AI model was used, and it was not investigated exactly what type of models and algorithms to use in the implementation. With this in mind, it was concluded that, in the context of this thesis, DL models likely meet the requirements. They handle large amounts of complex data to make human-like decisions from it, which is what is sought after in a C2 context.

2.1.1 AI - Risks & Ethics

AI provides many benefits, but also presents certain risks. Done right, it automates routine tasks to free up user time, enhances decision-making, reduces human error, is always available and never gets tired. Risks include data security, biases in decision-making, a lack of transparency, and that unsupervised models may start to act unethically or illegally (Kosinski, 2024). To address the last risk, the Artificial Intelligence Act was created by the European Union as the “first-ever comprehensive legal framework on AI worldwide”, which entered into force on August 1st, 2024 (European Commission, n.d.-a). The AI Act is one part of the EU approach to artificial intelligence (European Commission, n.d.-b), aiming to guarantee “fundamental rights, safety and ethical principles” (European Commission, n.d.-a).

The EU has also produced and published ethical guidelines for trustworthy AI (European Commission, 2019). Trustworthy AI is defined as being lawful, ethical and robust. This means adhering to the laws, complying with ethical values, and being technically reliable while also considering the social environment. The guidelines list the following key requirements:

1. Human Agency and Oversight
2. Technical Robustness and Safety
3. Privacy and Data Governance
4. Transparency
5. Diversity, Non-Discrimination, and Fairness
6. Societal and Environmental Well-being
7. Accountability

2.1.2 Explainable AI (XAI)

AI decisions risk being hard to interpret, as to how they were arrived at, due to complex calculation patterns. This kind of AI is referred to as a black box model, where only the inputs and outputs are known to the user, while why and how the outputs were reached are not (Kosinski, 2024). One way to address this problem is to implement explainable artificial intelligence (XAI). XAI is a description of AI models that are transparent in their decision-making, allowing humans to, e.g., understand how a certain result is arrived at. XAI is based on three main methods: prediction accuracy, traceability, and decision understanding. While prediction accuracy and traceability concern the technology itself, decision understanding concerns the human. Suspicion and doubt of AI are rather common. However, in order to achieve optimum performance, users need to understand and trust it. Such increased trust from users is one of the main benefits of XAI, as well as a reduced risk of errors and biases (IBM, 2023).

2.1.3 AI in a Military Context

In military contexts, decision makers are presented with a large amount of information to solve complex problems in non-routine situations. To do this, information has to be collected, analysed and synthesized to understand the situation and to predict the outcomes and consequences of different actions. The large amount of information risk leading to a workload too big for the decision makers, thereby reducing their SA. In a context like this, AI can be very beneficial and will be even more beneficial as it continues to improve (Van Den Bosch & Bronkhorst, 2018).

2.2 Saab

Saab AB is a Swedish defence company with a broad product portfolio. Different branches of the company cover air, land and naval domains. This thesis opportunity was offered by Saab Surveillance, the business area that specializes in radars and other sensors. Saab Surveillance provides various services and solutions, including Airborne Early Warning & Control (AEW&C) and surface radars (Saab AB, n.d.-b). This thesis has investigated the UIs of different C2 systems, which are the systems that bridge the radar operators and the sensors to help the operators gain SA.

2.3 Command & Control Systems

C2 systems are solutions that help commanders plan, coordinate, direct and control forces in missions on land, sea and in the air. The theory of C2 can be traced back 2500 years, to the ideas of Chinese general Sun Tzu. Since then, technology has changed the systems supporting the theory. A modern C2 system processes input data from radio telecommunication and radar signals to gain SA and offer decision support (Saab AB, 2022). To fulfil the goal, Schubert et al. (2018) state that C2 systems must facilitate:

- Collection of data.
- Making sense of this data.
- Identifying needed action.
- Planning needed action.

Due to the large amount of information required to do this, C2 systems are often complex, which transcends to the UI. Therefore, the systems are used by expert users.

The potential of AI in military contexts extends to C2 systems as well, with the large amount of information required to be processed quickly in a human-like manner to gain SA and provide decision support.

In this thesis, different Saab C2 systems were studied, both within AEW&C and surface radars, to gain a broader understanding of C2 systems and how they are used.

2.3.1 Saab's GlobalEye AEW&C

GlobalEye is a multi-domain solution, with a sensor suite for air, sea and land surveillance. The suite includes both active and passive sensors to provide long-range detection and identification of objects. “GlobalEye’s instrumented range is well above 350 NM (650 km)”, and it can be in the air for over eleven hours (Saab AB, n.d.-a). The automatic data and information fusion from on-board sensors and datalinks in the C2 supports the operators in their mission.

Figure 1

GlobalEye AEW&C in coastal environment



Note. Photograph of the GlobalEye AEW&C aircraft. (Saab AB, 2019). © Saab AB.

2.3.2 Saab's Surface Radars

In the surface radar domain, Saab provides several ground-based air surveillance solutions (Giraffe 1X, Giraffe 4A, Giraffe AMB and Arthur). The different solutions provide different strengths, levels of range and flexibility to locate the position of enemy weapons, drones, missiles, etc. (Saab AB, n.d.-c).

The surface radars have a different C2 system and UI than the AEW&C. Compared to the GlobalEye, these systems have fewer operator roles due to their lower number of tasks performed simultaneously.

Figure 2
Giraffe 1X Command & Control (C2)



Note. Photograph of the Giraffe 1X mobile radar system. (Saab AB, 2023). © Saab AB.



3

Theory

This chapter introduces theories and research which have influenced the process and result of this project.

3.1 Cognitive Ergonomics

Cognitive ergonomics is a domain within the scope of ergonomics and human factors. The International Ergonomics Association (IEA) have defined it as follows:

“Cognitive ergonomics is concerned with mental processes, such as perception, memory, reasoning, and motor response, as they affect interactions among humans and other elements of a system. (Relevant topics include mental workload, decision-making, skilled performance, human-computer interaction, human reliability, work stress, and training as these may relate to human-system design.)” (IEA, n.d.).

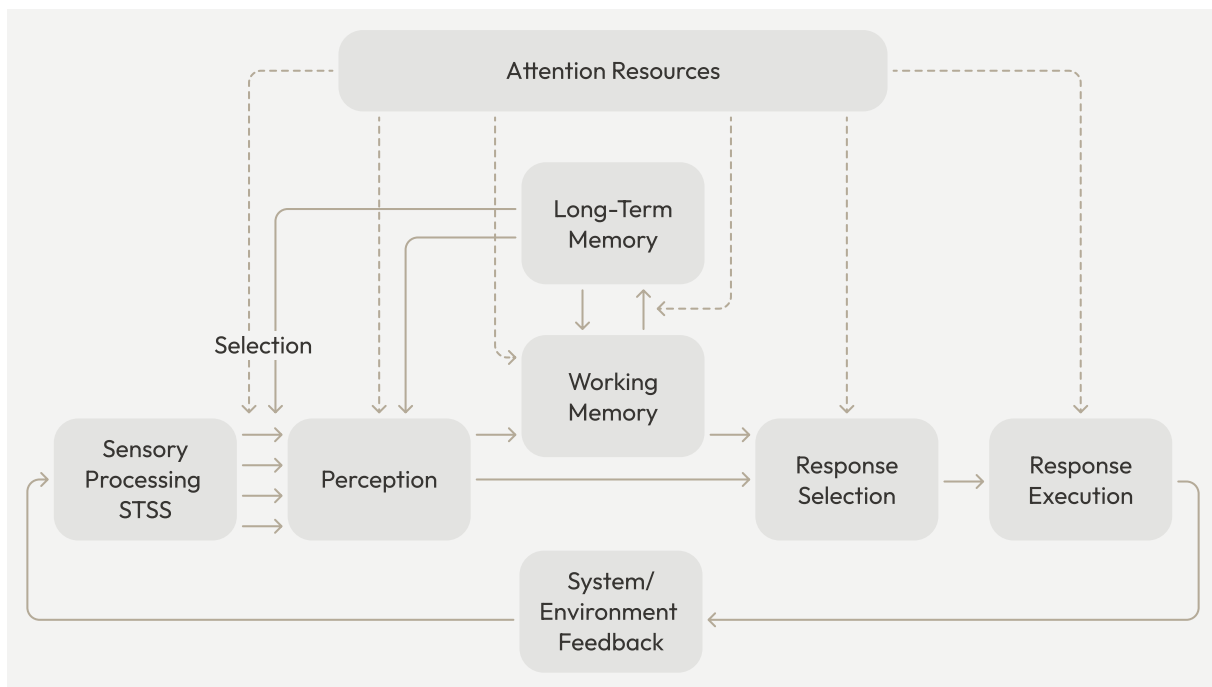
As seen, the area of cognitive ergonomics and mental processes is complex, with a lot of underlying and interlinking areas. With the purpose of this thesis being to increase efficiency and reduce the mental workload of the C2 operators, it is vital to gain an understanding of human cognition and how it is affected.

3.1.1 Information processing

Wickens et al. (2013) present a model of human information processing (Figure 3), which acts as a framework for understanding and analysing the mental processes and information flow when a task is performed.

Figure 3

Model of Human Information Processing Stages



Note. Adapted from Wickens et al. (2013).

The model consists of the following stages:

- **Short term sensory store (STSS)**

This stage involves very short registration and storage of sensory input, such as visual or auditory stimuli. Examples of this include an icon appearing on the screen or an audio alert that goes off.

- **Perception**

Process of interpreting and giving meaning to the sensory input, e.g., organizing into meaningful patterns. Perception is influenced by past knowledge and experiences. To utilize the same example, it could include interpreting the icon as a new target, distinguishing it from other visual clutter based on its movement.

- **Attention**

Attention allocates mental resources to relevant tasks and stimuli. It could be either directed or captured. For instance, the person shifts their attention to monitor the new target, instead of the general overview.

- **Working Memory (WM)**

Temporary storage where active processing of information occurs. For example, keeping information about the new target in mind while comparing it to others to assess potential conflicts.

- **Long-Term Memory (LTM)**

A large storage of knowledge, skills and experiences, previously encoded from the working memory. Can be retrieved and influence, e.g., perception and decision-making. Examples include recalling procedures of how to act when a new target appears.

- **Response Selection**

Selection of an appropriate action, based on the processed information. For example, the user decides to make a classification of the new target.

- **Response Execution**

The selected response is carried out. For example, the user performs actions to make the classification.

- **Feedback**

After the response is executed, feedback is provided. This is processed and evaluated to influence future actions. For example, the system informs the user that the action was performed correctly by changing the classification of the target.

3.1.2 Bottom-up & Top-down Processing

Attentional and mental resources are limited for humans. As mentioned, attention could be either directed or captured. Bottom-up and Top-down are two types of processes which affect attention and perception.

Bottom-up processing is influenced by stimuli and is referred to as data- or stimulus-driven (Salvendy & Karwowski, 2021). Naturally, a vital factor that impacts the ability to perform bottom-up processing and direction of attention is the salience of the stimuli.

Top-down processing, on the other hand, is knowledge-driven, derived from previous experiences (Salvendy & Karwowski, 2021). This type of processing is used, for example, for the expectancy and prediction of information availability, and could thereby affect the allocation of attention accordingly.

3.1.3 Multiple Resource Theory

Wickens (2002) has constructed another model, a Multiple Resource model (Figure 4). It consists of the four dimensions: Stages, Processing Codes, Perceptual Modalities, and Visual Channels. These dimensions all affect time-sharing performance between different tasks. Tasks which compete with each other and demand the same resource interfere with each other more than tasks that demand different resources. The theory and model are applicable as a guide when designing systems where operators are required to multitask.

- **Stages**

This dimension refers to the different stages of information processing, which include perceptual and cognition. For example, based on the work of a radar operator, the operator can detect a new target on the screen (perception) while trying to interpret its meaning (cognition).

- **Processing Codes**

Distinguish between different types of mental processing, including verbal and spatial. Examples of which could include tracking the spatial position of a target on the screen, while also listening to and processing verbal information from another operator.

- **Perceptual Modalities**

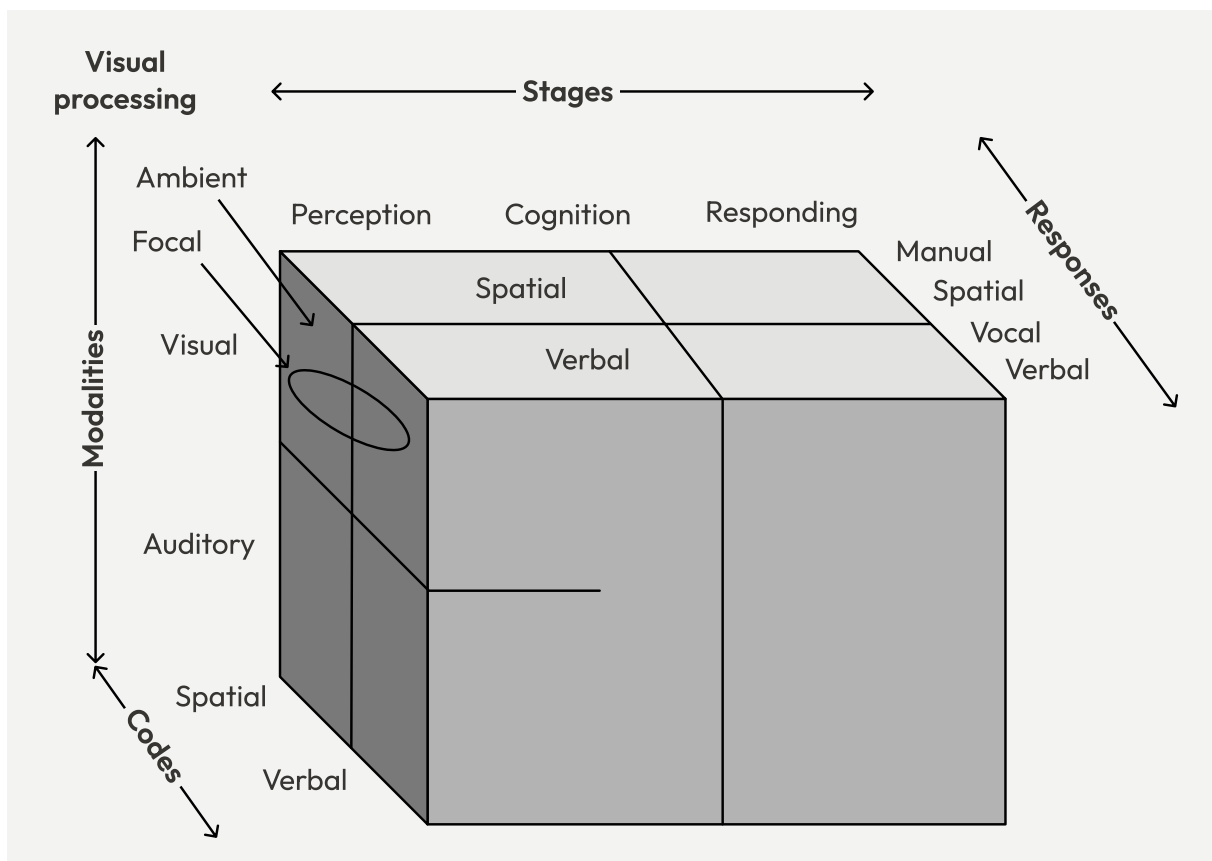
Different senses involved in perception, such as visual or auditory. An example would be monitoring a screen (visual) and listening to the radio (auditory), which is more manageable than, for instance, monitoring two displays simultaneously. Having tasks that use different modalities reduces the competition for the same sensory channel.

- **Visual Channels**

The two aspects of visual processing, Focal (central vision) and Ambient (peripheral vision). Focal vision includes, e.g., reading small text labels on the screen, and ambient vision could be an alerting light in the periphery.

Figure 4

Four-Dimensional Multiple Resource Model



Note. Adapted from Wickens (2002).

3.1.4 Mental Workload

In the article “Multiple Resources and Performance Prediction,” Wickens mentions the close interlink between resources, as mentioned in multiple resources, and mental workload (2002). He refers to the definition of mental workload, as defined by Moray (1979, as cited in Wickens, 2002), “Mental workload describes the relation between the (quantitative) demand for resources imposed by a task and the ability to supply those resources by the operator.”

Mental workload and cognitive workload are terms which are often used interchangeably. Cognitive workload, however, is more specifically associated with the cognitive, information processing portion of mental workload, implied by Young and Stanton (2002).

When the demand for mental resources is greater than the availability, overloading occurs. Overload results in a degradation of task performance. People can adapt to overload in various ways, either allowing performance to degrade, shifting task execution to a more efficient method, or fully stopping the execution of tasks. If the latter is implemented, it would optimally apply to tasks of lower priority. However, people might as well do it in a less optimal manner for tasks of high importance (Wickens, 2013).

In addition to mental workload, other factors which influence task performance include stress and individual differences between operators. One possible solution mentioned by Wickens (2013), which could remedy negative effects of stress, includes “...reducing the amount of unnecessary information (visual clutter) and increasing its organization...” for tasks where perceptual narrowing and scanning amongst different information occur.

3.1.5 Rasmussen’s SRK Model

Rasmussen (1983) proposed that human behaviour can be separated into three cognitive levels.

- **Skill-based behaviour:** Automatic actions which are based on routine and performed without conscious thought.
- **Rule-based behaviour:** Actions guided by stored rules or procedures, gained from past experience.
- **Knowledge-based behaviour:** Problem-solving in new, unfamiliar, and complex situations without pre-defined rules and know-how.

This SRK-model is valuable for understanding human error and designing systems that support appropriate behaviour levels.

3.1.6 Situational Awareness

SA have previously been mentioned and explained in regard to the terminology’s usage in the military context and C2. However, the term is also prominent within the area of cognitive ergonomics. Endsley (1988, as cited in Wickens et al., 2013) has provided a definition of SA as “the perception of critical elements in the environment, the comprehension of their meaning, and the projection of their status into the future”.

Operators keeping a good level of SA is important for the success and efficiency of task performance (Wickens et al., 2013). Problems with or loss of SA have led to major accidents in the past, and thus, it is extra important in occupations where time and safety are of the essence, such as for radar operators.

The level of SA is impacted by cognitive processes, with two important processes being attention and memory. Attention, as previously mentioned, is a limited resource, and how it is directed or divided affects the ability to perceive critical elements and stimuli that fail to capture attention. Memory, both working and long-term, is relevant for the comprehension and understanding of the situation and the prediction of future events.

Wickens et al. (2013) mention domain experience as a factor which improves the ability to maintain SA. Experts, namely, possess a skill use of storage in long-term memory, described by Ericsson and Kintsch (1995) as long-term working memory (LT-WM). This skill allows for rapid storage and retrieval of relevant information from the long-term memory. Such relevant information includes mental models of the system and states, which are based on their previous experiences.

3.2 Gestalt Principles

The Gestalt principles are a number of rules, widely used in UI and user experience (UX) design, which address the human perception of visual elements. The principles originate from the Gestalt psychology by Wertheimer, which "...emphasizes that the whole of anything is greater than its parts." (Encyclopædia Britannica, 2025).

The following are some core Gestalt principles which are relevant for the purpose of this thesis, and the definitions as stated by Lindwell et al. (2010):

- **Proximity:** "Elements that are closer together are perceived to be more related than elements that are farther apart"
- **Similarity:** "Elements that are similar are perceived to be more related than elements that are dissimilar."
- **Closure:** "A tendency to perceive a set of individual elements as a single, recognizable pattern, rather than multiple, individual elements."
- **Good Continuation:** "Elements arranged in a straight line or a smooth curve are perceived as a group, and are interpreted as being more related than elements not on the line or curve."
- **Figure-Ground Relation:** "Elements are perceived as either figures (objects of focus), or ground (the rest of the perceptual field)."
- **Common fate:** "Elements that move in the same direction are perceived to be more related than elements that move in different directions or are stationary."

3.3 User Interfaces

A UI is the point of interaction between the user and a device or system. The usefulness depends on factors such as the mode of interaction, the user's intention, the level of complexity and the resulting learning curve, the user's background and the technology behind the interface (Miraz et al., 2021). UIs can be both physical and digital. However, this thesis focuses on the digital interfaces of C2 systems.

3.3.1 Usability

Usability is an indication of how easy a UI is to use. The word is also used to refer to design methods during the design process to address ease of use. It is defined by five components (Nielsen, 2012):

- **Learnability:** How easy is it for users to accomplish basic tasks the first time they encounter the design?
- **Efficiency:** Once users have learned the design, how quickly can they perform tasks?
- **Memorability:** When users return to the design after a period of not using it, how easily can they re-establish proficiency?
- **Errors:** How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- **Satisfaction:** How pleasant is it to use the design?

Another important term is utility, which is defined as if the needed features are provided. For a UI to be useful, it has to possess both usability and utility. When actively working with usability in a design process, it takes the shape of usability testing. It is important to test iteratively and to begin early in the design process. Starting testing too late can lead to structural and architectural problems being discovered too late to change (Nielsen, 2012).

Nielsen has developed ten guiding usability heuristics that act as guidelines for making design decisions and analysing the UX of applications. They are applicable on most forms of interactions and complexity levels, from video games to websites. A version aimed towards complex systems (as C2 systems are) looks like the following, with the ten heuristics being the general ones and the comments being additions for complex systems (Kaplan, 2021):

1. **Visibility of System Status:** *The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time.*

An addition in complex systems is that actions requiring longer waiting times are more common. This makes an indication of progress, action and waiting time important so that the user can decide whether to perform another task in the meantime.

- 2. Match Between System and the Real World:** *The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.*

The “real world” can look different in different cultures. It is important to consider the intentional user’s worlds and context.

- 3. User Control and Freedom:** *Users often perform actions by mistake. They need a clearly marked “emergency exit” to leave the unwanted action without having to go through an extended process.*

For users in complex systems, high levels of cognition and time are often invested. Errors may also have bigger consequences. Therefore, the possibility of backtracking or quickly correcting errors is important.

- 4. Consistency and Standards:** *Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions.*

In complex systems, it is common with very domain-specific terms and workflows. It is also beneficial for expert users to be consistent with standards.

- 5. Error Prevention:** *Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.*

This is directly applicable to complex systems, with the addition that the consequences of errors risk being bigger and more serious.

- 6. Recognition Rather than Recall:** *Minimize the user’s memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g., field labels or menu items) should be visible or easily retrievable when needed.*

In complex systems, the amount of information can be overwhelming, while time spent in it is often long. Facilitating recognition saves time and cognitive capacity for the user.

- 7. Flexibility and Efficiency of Use:** *Shortcuts — hidden from novice users — may speed up the interaction for the expert users such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.*

When expert users know a system inside out, they reach an efficiency plateau. By implementing features that speed up the process, accelerators, efficiency can increase again. Shortcuts and personal tailoring options can be such examples.

- 8. Aesthetic and Minimalist design:** *Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.*

In complex systems, all information competes for the users' attention. Getting rid of non-critical information facilitates the process of finding the right information.

- 9. Help Users Recognize, Diagnose, and Recover from Errors:** *Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.*

In complex systems, this heuristic is often not well met. Due to the expert level of the users, they are expected to know enough about the system to know non-explicit error codes. However, that is not realistic and misses an opportunity for learnability, where the expert can learn more from their mistakes.

- 10. Help and Documentation:** *It's best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.*

Complex systems come with a large amount of information about them. However, if possible, the help and documentation should be integrated when needed. Users seldom open large manuals when needing help now.

3.3.2 Adaptive User Interfaces (AUIs)

"Adaptive user interfaces (AUIs) are defined as systems that adapt their displays and available actions to the user's current goals and abilities by monitoring user status, the system state and the current situation" (Lavie & Meyer, 2010, p. 508). Improving user performance and avoiding cognitive overload in complex systems are two of the sought gains. However, if the adaptivity interferes with usability design principles, these gains risk being reduced or disappearing (Lavie & Meyer, 2010). Therefore, the trade-off between adaptivity and usability needs to be considered, to be optimal (Miraz et al., 2021).

Lavie and Meyer (2010) present four factors to consider when implementing adaptivity in a UI. The first, the task, varies according to its characteristics, how challenging it is, and the requirements it puts on the user. The second, the user, varies in e.g., background, personality, cognitive skills, preferences, and mood. Worth adding here is that the adaptivity should be based on user behaviour rather than user feedback (Miraz et al., 2021). The third, the situation, can be routine or non-routine. Finally, the level of adaptivity can vary from complete system-controlled adaptivity to different levels of user-system cooperation in adaptivity (Lavie & Meyer, 2010). In their literature review, Miraz et al. (2021) concluded that the trend is that system-controlled adaptivity is more beneficial than a user-controlled one. This raises questions about Nielsen's usability design principles, especially "User control and freedom" and "Consistency and standards". Interestingly, Gajos et al. (2008) partly answer these questions by showing that the accuracy of the adaption is more important

for user performance than predictability is, while they both increase user satisfaction. They suggest that machine learning algorithms should be considered, due to their potential of being more accurate while risking a decreased predictability.

3.3.3 Intelligent User Interfaces (IUIs)

Intelligent user interfaces (IUIs) utilize AI to improve the usability and user experience in the UI (Brdnik et al., 2022). They adapt their behaviour to users, tasks and context, based on different models (Gonçalves et al., 2019). “Many intelligent interfaces can be described as adaptive interfaces, though not all adaptive interfaces are intelligent.” To be labelled as an IUI, it is not enough to have a UI for an intelligent system, it must be intelligent from the user’s standpoint as well (Brdnik et al., 2022).

In their systematic mapping study of IUIs, Brdnik et al. (2022) found that intelligent agents, algorithms, tools and assistants are the main form AI is used in IUIs. They state that machine learning “offer new ways of newer, faster, and more accurate adaptation, recommendation, help, detection, illustration, and utilization in user interfaces.” These qualities have potential in the context of the thesis scope.

As for AUIs, there is a challenge balancing adaptivity and some usability principles in IUI. AI increases this challenge due to its ability to handle a larger amount of information and to make “black box” decisions. However, as described in the AUIs (see section 3.3.2), Gajos et al. (2008) showed that accuracy was more important for user performance than predictability, despite having a similar effect on user satisfaction. In this case, the performance outweighed the violation of usability design principles. It may therefore not be taken for granted that such a violation is bad, until it is tested. The result from Brdnik et al. (2022) that “81% of studies reported partial or complete improvement in usability or UX” is an incentive to further investigate adaptive IUI in a C2 context.

3.3.4 Intelligent Functions in Command and Control

Command and Control (C2) systems, as stated earlier, are highly complex systems with a high mental workload in a high-stakes context. Zak et al. (2023) performed an experimental study to investigate whether an intelligent filter could reduce the cognitive workload of an Unmanned Aerial Vehicle (UAV) operator. The intelligent filter was applied to the C2 map.

Many C2 tasks are too complex for a single operator and therefore require teams. However, even in small micro-systems operated by just one person, the cognitive workload is high enough to make prioritizing tasks necessary (Zak et al., 2023). The need to divide or sort information is still present on the individual level, despite having already been divided among team members. In the case of this study, it is done by filtering map information. In the current C2 system of this referred study, manual filtering is an option based on set parameters, like showing enemy forces or friendly forces. If shown, potential irrelevant information comes with the layer. If hidden to save cognitive capacity, there is a risk of losing SA. Another problem with this solution is that in the dynamic reality the mission takes place, the information that is irrelevant now and therefore manually hidden, may suddenly become relevant without the operator knowing it. Furthermore, the manual changes between filters

take time from the mission. Therefore, it is important to investigate solutions that reduce cognitive workload while not having a negative impact on SA and the mission performance (Zak et al., 2023).

The AI algorithm in this study makes the decisions about what to filter using a filter protocol based on three factors. The first is the importance of each information item on the map, categorized as important, neutral or distracting. The second one is the probability of areas being within the operator's area of interest. The third one decides how often the filter should be updated. The first two are weighed together. The third one depends on the context and system capacity, and faster update rates, about 20 seconds, seem better than slower ones. The optimal combination of the three depends on mission and environment, combined with the operator's factors, such as workload, SA and performance (Zak et al., 2023).

The study found that the filter reduced the operator's cognitive workload to an extent. It was also shown that a midrange clutter (in terms of amount and complexity of visual information) is the sweet spot, since both high and low informational clutter risk increase cognitive workload and decrease SA. It was also shown to be even more effective when the operator was fatigued in the second stage of the test mission. However, when alert, an operator can handle more information, and the use of a filter has less of an effect. Filtering the map when the operator is alert, or filtering the map too much, was even shown to potentially increase the cognitive workload, since the operator starts to search for more information than the presented one to meet their desired level of information (Zak et al., 2023). The potential of being able to handle more information when alert raises an interesting question: May it be worth to filter harder in routine, non-serious situations to save the operator's capacity for serious situations, instead of filtering harder when the seriousness increases? There, the operator's demand to have more control may be higher. Zak et al. (2023) propose that monitoring current user status and implementing that in the filter protocol of the AI can help balance between keeping SA while reducing cognitive workload, to avoid the dilemma of how to optimize the filter for different situations.

Concluded, the study is a highly relevant test of applied adaptivity in C2 systems, adding to this thesis. Although they do not outright call their filter adaptive, it is an intelligent function that filters and presents information out of user control. And even if the UAV C2 context is not identical to the ones studied in this thesis, the different systems have challenges in common, such as the risk of information overload and the need for SA. Additionally, Zak et al. (2023) propose that in future work, "intelligent filtering can be combined with more sophisticated spatial-temporal visualizations to make the information even more available". These factors make this study highly relevant for the thesis, as more specifically applied information to complement the broader theory presented in this chapter.



4

Methods

This chapter presents methods which have been utilized in one or multiple stages of the project. These include methods related to User Tests, Analysis, Evaluation, Idea generation, and Concept creation.

4.1 User Tests, Analysis & Evaluation Methods

4.1.1 GDPR & Ethics

When conducting tests, interviews, and other activities that collect personal data, it is important to handle it ethically and lawfully. The General Data Protection Regulation (GDPR) is a privacy and security laws constructed by the European Union, which address this. The key regulatory points are (GDPR.eu, n.d.):

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitation
- Integrity and confidentiality
- Accountability

4.1.2 Interviews

Interviews are a qualitative method in which insight is gathered from interviewees. It is possible to gather in-depth information, e.g., regarding their experiences and interpretations. Interviews can be planned, structured, and conducted in various ways in order to best suit the purpose.

One way to categorize interviews is as structured, unstructured or semi-structured. Structured interviews, as the name suggests, are highly structured, with predetermined questions for each interview and little room for flexibility. Unstructured interviews, on the other hand, are more fluid and conversational in nature, allowing the discussion to flow forward. Semi-structured interviews, however, use a mix of both closed- and open-ended questions, which are often followed by some follow-up questions. Conducting semi-structured interviews can be rather time-consuming, but they are beneficial when, e.g., evaluating and examining unknown areas, as well as serving as a good complement to other methods (Adams, 2015).

In the article “How Many Participants for a UX Interview?” Rosala (2021) discuss the appropriate number of participants in studies. Rosala mention the recommendations from UX professionals, such as Nielsen, of five users. However, she presents the argument that this might be too small for exploratory research, stating that the number of participants is determined once saturation is reached.

4.1.3 Applied Cognitive Task Analysis (ACTA)

Applied Cognitive Task Analysis (ACTA) is a user study method useful for extracting information about cognitive demands and skills, as well as representing that information (Militello & Hutton, 1998). The method was developed and is well-suited for studies involving subject-matter experts.

The ACTA consists of three main interview techniques, Task diagram, Knowledge audit, and Simulation interview. Task diagrams provide an overview of the task by having the interviewee divide a task into a specific number of subtasks, as well as identifying which require more or less cognitive skill. The Knowledge audit investigates the levels and aspects of expertise a task or subtask requires. Simulation interview provides insight into the cognitive processes of the expert by presenting them with and simulating a scenario.

4.1.4 KJ Method

The KJ method (or affinity diagramming) is a method to organize, analyse and draw conclusions from complex qualitative data, e.g., from interviews. It was formed by the Japanese anthropologist Jiro Kawakita, hence the name of the method. It consists of four steps: label making, label grouping, chart making and explanation (Lucero, 2015). In label making, the main facts are put on individual notes. They are then discussed and put in categories with similar ones during label grouping. The groups are then placed on a large sheet of paper in relation to each other, based on, for example, dependencies and contradictions, with the relationships illustrated using symbols and signs. During the explanation step, the final chart is first verbally concluded and explained and then written down.

4.1.5 Within-Subject Tests

Within-subject testing occurs when each test participant is tested on all variants, as opposed to between-subject testing, where each participant is only tested on one variant (Price et al., 2017). The main advantage of within-subject testing is that it limits the effect of participant variables, such as IQ, experience, etc. The main disadvantages are different carryover effects.

4.1.6 Carryover Effects

Carryover effects refer to the impact of a participant's experience on a test or question on subsequent tests or questions. Two examples are the learning effect (also known as the practice effect) and the fatigue effect (Price et al., 2017). The practice effect is when performance improves task by task, since the participant learns along the way. The fatigue effect is opposite, since fatigue makes the participant perform worse in later tasks. To counterbalance these effects, it is important to change the order of the tasks for each participant.

4.1.7 N-back Task

The n-back task is a working memory test, with the possibility to vary the cognitive workload (von Janczewski et al., 2021). The subjects are presented with a stimulus sequence, with a pause in between each stimulus. When one stimulus matches the stimulus n steps back in the sequence, the participant acts. How they act depends on the type of stimulus. Several different types of stimuli have been studied, e.g., letters, words, numbers, shapes and tones. It has been shown that the type of stimulus used does not matter significantly. What matters is the factor n . The performance is usually measured in terms of accuracy and reaction time (Gilmour et al., 2019). In their meta-analysis of the n-back test used as a subsidiary task when driving, von Janczewski et al. (2021) state that “the n-back task seems to fulfil the criteria and other requirements that suggest its general suitability as a subsidiary task” (p. 271). This was based on the low level of interaction with the primary task, being easy to measure, and the small amount of equipment and time required.

4.2 Idea Generation and Concept Creation Methods

4.2.1 Prototyping

Prototyping is a central method used in design processes, which embodies and visualizes ideas. Lim et al. (2008) categorize the use of prototypes into four main areas. The first of which is *evaluation and testing*, the second being *understanding of user experience, needs, and values*, the third is *idea generation*, and finally, the fourth is *communication among designers*. It is important to note that these categories are not exclusive; a single prototype can serve different purposes within multiple categories.

Prototypes can vary in level of fidelity, from low-fidelity wireframing and paper models to high-fidelity digital or physical mock-ups. The traditional ways of utilizing prototypes in the past have been for evaluation, such as usability testing. However, as mentioned, it is also serving a valuable purpose as an idea generation and communication tool. Lim et al. (2008) mention how “Prototypes stimulate reflections, and designers use them to frame, refine, and discover possibilities in a design space” (p. 2).

4.2.2 Braindrawing and Brainwriting

Braindrawing and brainwriting are alternative idea generation methods to traditional brainstorming.

When conducting brainwriting, participants write down ideas on paper. After a predetermined amount of time, these papers are passed on to the next participant, acting as inspiration for more ideas. This process continues until the papers have passed between all participants. Litcanu et al. (2015) mention some potential benefits of the process, including elimination of production blocking, reduction of social loafing, and encouragement in the processing of shared ideas.

Braindrawing could be referred to as graphic brainstorming, performed by participants who generate ideas through quick, rough sketches (Wilson, 2013). When performed in a group setting, the method is used similarly to brainwriting, with participants passing sketches between them to draw inspiration from and build on each other's ideas. Utilizing sketches has benefits in being more powerful and often allows for better communication of complex ideas than words.

4.2.3 Crazy 8's

Crazy 8's is a famous method used in Design Sprints (Google, n.d.). The method could be described as an implementation of braindrawing with a specified structure. It is performed in a group of participants, who divide their papers into eight sections, where they are supposed to sketch one idea per section, during a period of eight minutes. The idea is that the method with a limited time frame and a defined number of sketches would provide room for rapid and innovative generation of ideas.

4.2.4 Dot-voting

Dot-voting is a simple and effective tool used for decision-making and prioritizing in group settings (Gibbons, 2019). The process is performed by first gathering the material for which voting should occur, such as sketches and ideas. Thereafter, specify the constraints of the voting, including the number of votes each person receives. Then, the actual voting occurs by placing marks or "dots", giving the method its name. The outcome is calculated, and in case of a tie, a potential narrowing down and revote is conducted.



5

General Description of Procedure

In this chapter, the general thesis procedure and structure are presented to give an overview of the project phases, why they were performed and what they resulted in.

5.1 Research Questions

Two research questions were formed to guide the actions and direction of the thesis. They correspond to the initial brief, however, not covering its entirety due to delimitations made through the project:

- How can AI-supported adaptivity be implemented and visualized in Command and Control (C2) systems?

The first question aims to investigate in which ways C2 systems can be adaptive to add function and help the operator, while also giving concrete suggestions on how this adaptivity can be visualized

- Can AI-supported adaptivity reduce the operator's mental workload while keeping a sufficient level of SA?

The second question aims to investigate if that adaptivity can have an actual positive effect by reducing the operator's mental workload. However, it says itself that reducing information can reduce the mental workload. Hence, simply reducing mental workload is not enough for a total positive effect. To counterbalance this, the thesis also investigates if this can be done while still keeping a sufficient SA.

5.2 Design Process

5.2.1 Research Through Design

This thesis was an explorative research process. Literature review and investigation of Saab's existing C2 systems were complemented by the development of new UI prototypes and testing to address the research questions. The development of UI prototypes, where a new design is used to perform research, was inspired by a method called Research through Design. Zimmerman et al. (2007) present a model of Research through Design to integrate design thinking into human-computer interaction (HCI) research. The model facilitates interaction designers to contribute to research through their ability to solve under-constrained problems. This is done by using HCI research to develop a "product that transforms the world from the current state to a preferred state". This thesis does not implement the actual model itself but follows the general description of the method and uses it as inspiration.

5.2.2 Iterative Design Process

Being an explorative study, the project followed the structure of an iterative design process. Research, development, testing and evaluation were repeated in cycles. This allowed for constant learning and adaptation of the progress and project aim, based on insights from users and other stakeholders. This process aligns with the Research through Design process proposed by Zimmerman et al. (2007), where iteration serves as both a research method and a development strategy.

5.3 Preconditions

The preconditions of the thesis work were the following:

- Restriction to Saab's Surface and AEW based C2 systems.

Due to limited access to other systems, mostly to do with organizational and locational factors, the project was restricted to Surface and AEW&C based C2 systems. For a more comprehensive analysis, the inclusion of even more versions of C2 systems would have been desired.

- Limited access to C2 operators.

There was limited access to C2 operators. Therefore, multiple participants recur throughout the project's various interviews and tests, in order to gain the desired knowledge and experience of expert operators.

- NATO Joint military symbology is utilized.

The visualized adaptive C2 concepts in this project would make use of the NATO joint military symbology (NATO, 2023). Being a standardized and widely used method, the concepts would be able to be implemented on any system. Additionally, replacing them would risk reducing the prototype's fidelity for the tests.

- Mental workload will be compiled into one combined equivalent.

As mentioned, there are many factors which influence mental workload and task performance, including stress and personal differences. This project does not measure these factors individually, such as stress level and fatigue, but instead refers to them collectively as one combined external workload.

- Explorative Research

As mentioned, the research questions are investigated in an explorative way through Research through Design, rather than through conventional research methods. The results from which are possible solutions, requiring further work to be statistically secured.

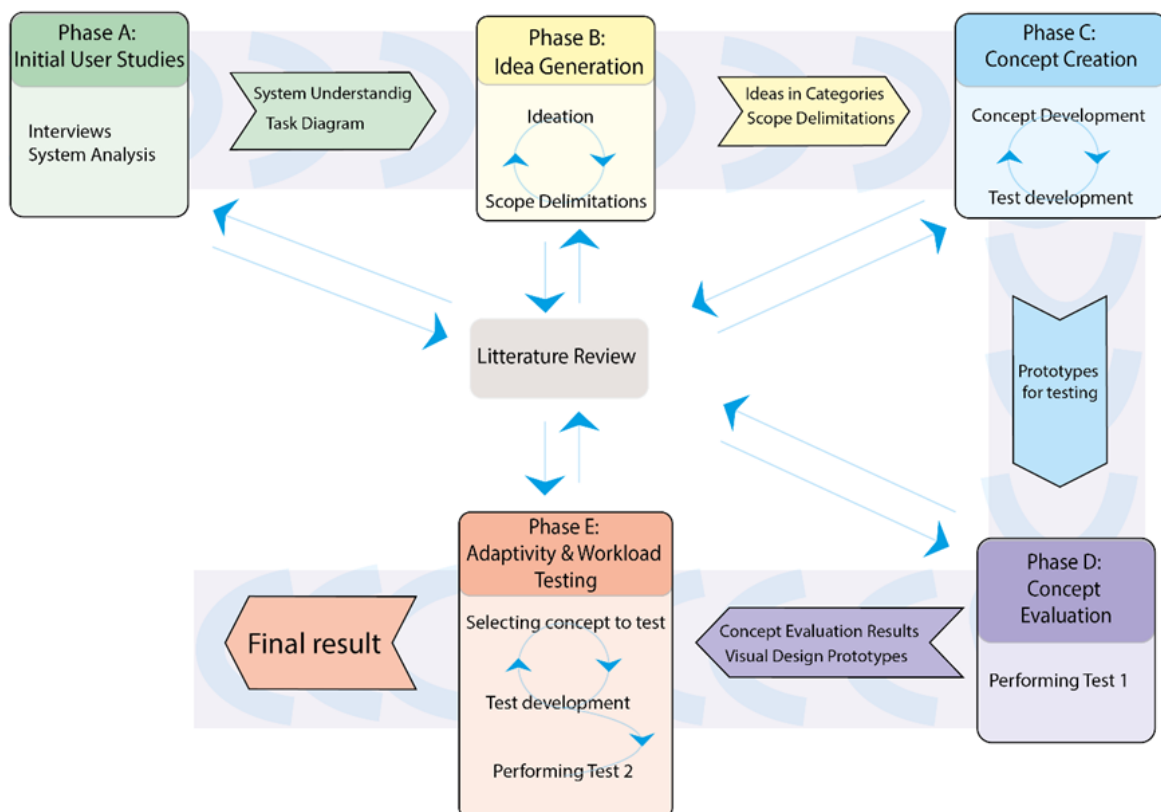
5.4 Overall Project Structure

The thesis consisted of five project phases:

- Initial User Studies
- Idea Generation
- Concept Creation
- Concept Evaluation
- Adaptivity & Workload Testing

Figure 5

Visualization of the Process of the Project



5.4.1 Phase A: Initial User Studies

In this phase, user and C2 system studies were conducted, complemented by a review of existing research. The objective of this phase was to lay a foundation of knowledge for the thesis, specifically regarding C2 systems and their operators. The results are presented in section 6.3.

Interviews were conducted with five expert users with previous operational experience. Both surface and AEW&C C2 systems were represented. The gained information was structured and analysed through the KJ method and compiled in a task diagram.

The phase provided insight into various Saab C2 systems and how they are utilized, as well as methods and preferences from the expert users. Literature reviews gave further theoretical insight into human factors, AI and adaptive user interfaces.

5.4.2 Phase B: Idea Generation

The objective of the second phase was to put the knowledge gained in the first phase to use and turn it into ideas for implementing adaptivity and AI in C2 systems. The results are presented in section 7.3.

In this phase, several ideation methods were performed, and ideas emerged about adaptivity in C2 systems. Braindrawing and brainwriting were used, followed by the KJ-method to sort the ideas into categories. The categories from the KJ analysis were cross-referenced with the task diagram from Phase A to make the first major scope delimitation of the project. This was a suitable time to make delimitations. Earlier would have risked reducing the potential to understand the big picture in the interviews and the wide span of the ideas during ideation. Waiting longer would have affected the possibility of reaching an interesting conclusion in the end.

The phase resulted in numerous ideas, sorted into categories. The big delimitation of the project scope, made at the end of the phase, influenced the continuation of the project.

5.4.3 Phase C: Concept Creation

The objective of this phase was to develop and prototype concepts, which would be evaluated with expert users in the next phase, Phase D. While these are presented as two subsequent phases, in reality, the concept creation and test planning were performed in parallel and highly influenced each other. The results are presented in section 8.3.

The Concept Creation began by discarding the ideas outside the new delimitations, conducting another KJ analysis on the remaining ideas and voting on the most relevant ideas within the most relevant categories. Following this, the ideas were prototyped in Adobe XD (Adobe Inc., 2025b) to be evaluated with expert users in the upcoming phase.

5.4.4 Phase D: Concept Evaluation

The objective of this phase was to evaluate the concepts created in the previous phase, Phase C. This included planning, conducting and analysing the results from the evaluation. The results are presented in section 9.3.

The evaluation consisted of some quantitative tests, but the majority were semi-structured interviews regarding the concepts, with the prototypes as discussion material. The result was analysed mainly through another KJ analysis.

This phase resulted in qualitative data on the evaluated visual concepts, which contribute to answering the first research question about how AI-supported adaptivity can be visualized in C2 systems.

5.4.5 Phase E: Adaptivity & Workload Testing

The objective of the fifth and final phase was to conduct more specific and quantitative evaluations, investigating one of the created concepts in relation to adaptivity and mental workload, in order to answer the second research question. The results are presented in section 10.4.

The phase included selection and refinement of the concept implemented in the test based on the previous Concept Evaluations, while iteratively developing the second test. A significant part of the phase's focus was on developing the test itself, ensuring it was balanced and quantifiable to be able to draw conclusions from it. The resulting test consisted of one main, C2-related task and one subsidiary task that increased the mental workload. Six operators participated in the test, and the resulting data were analysed. After which, final conclusions were drawn.

The outcome of this phase was data on how one of the adaptivity concepts affected mental workload and SA, to answer the second research question. It also contributed to making recommendations for further research and future work.



6

Phase A: Initial User Study

This chapter presents the initial phase of the project, in which an initial user study was performed. The planning, execution, and results of the study are presented, leading into the phase's contribution to the project.

6.1 Objective

The objective of this study was to gather knowledge and gain an understanding of the different systems and their usage, including and specifically focusing on the operator's role and interaction. The knowledge and findings from this study would lay a foundation for the future design process to build upon.

6.2 Planning & Execution

6.2.1 Participants & Systems

This thesis aims to investigate and develop concepts and suggestions of AI implementations for adaptive C2 systems. Since the scope of the thesis is unspecific to a particular version or alteration of C2 system, this would need to be reflected in the implementation of the user studies. Thus, the inclusion of participants from different domains and operational experiences of different systems was desired.

The recruitment of participants was, as previously mentioned, affected by the availability of experienced operators as well as the access to different C2 systems. The initial sample selection consisted of five participants, corresponding to the recommendations made by Nielsen, with the possibility of conducting more if sufficient saturation was not reached.

The participants were all Saab employees who possessed expert knowledge of the respective C2 system, as well as having some sort of past military operational experience involving C2 systems. Two of the participants were focused on surface-based C2 systems, while the other three focused on AEW&C systems. The participants were distributed as seen in Table 1:

Table 1*Information and Distribution of Participants Interviewed in Phase A*

Participant:	C2-system:	Self-estimated level of operational experience: (No-Low-Medium-High)	Self-estimated level of C2-system experience: (No-Low-Medium-High)
1	Surface	Medium	High
2	Surface	High	High
3	AEW	High	High
4	AEW	High	High
5	AEW	High	High

6.2.2 Design of Study

The design of the study consisted of two main phases, a semi-structured introductory interview, followed by a walkthrough of the system. The study was facilitated by one interviewer and documented by a note-taker, both of whom were able to intervene and stray from the original interview template, for example, to provide further clarification or additional details. In addition to note-taking, a voice recorder was used as a means for documentation and support. The duration of each study was planned to be approximately 1.5 hours.

The design of the study took inspiration from existing methods aimed at investigating user experts. One such example is Applied Cognitive Task Analysis (ACTA), which is useful for extracting information about cognitive demands and skills, as well as representing that information.

In this study, the task diagram technique from the ACTA method was adapted with some adjustments to better suit the purpose. The participants were asked to select and describe two tasks/assignments of different levels of seriousness, one “every day” and one “serious”, and break them up into 3–6 subtasks. The ambition was to gain an understanding of different tasks and the operator’s procedure on an overarching level, without getting hung up on specific details.

These tasks and subtasks laid a foundation for and led into the following walkthrough in the system, where the participants would act as an operator and perform the tasks and subtasks they described. This procedure is similar to the simulation interview technique from ACTA, providing insight into the cognitive processes of experts as well as a deeper understanding of the system and its functions.

Before each study, the participants were provided with a consent form explaining the purpose of both the project at large and the specific interview, the type of data that would be collected, as well as its intended use. The participants were informed that they were able to revoke their consent or end their participation at any time during the study. This procedure made sure that the General Data Protection Regulation (GDPR) was followed.

The interview phase of the study also included some other general questions, as well as a short survey, mainly aimed at mapping the user's experience and background. The template for the interview and the survey can be seen in Appendices A and B.

6.2.3 Analysis of Study

In order to analyse the results from the study, two main methods were used. Firstly, the task diagrams from each participant were summarized, structured and visualized together in a table to get a clear overview. This allowed for easy comparisons between the different participants. The comparison aimed to identify commonalities in tasks and procedures for operators, regardless of which alteration or domain of C2 was used. These results were presented in an illustration described as a domain independent task diagram, from here on also called just "task diagram".

Secondly, the information from the entirety of the studies was analysed using the KJ method. Findings were written down on sticky notes, placed on a whiteboard, and sorted into categories and themes, such as problem areas.

6.3 Result

The task diagram interview technique, as well as the following walkthrough, provided deep insight into the systems and tasks. The summarized task diagrams for all participants can be found in Appendix C. These include the titles of each subtask and additional information on procedures, functions and steps that the participants mentioned or performed.

The task diagrams of ordinary and advanced tasks were chosen to be combined into one common for each participant. This is due to the majority of the procedure being the same or similar, with only slight adjustments or additional steps. The difference between ordinary and advanced tasks, mentioned in some instance by all the participants, was due to external factors. Most often, the type of target, the arrival of new targets, etc. One participant described the difference between tasks as, "advanced tasks are ordinary tasks that escalate and become more serious."

An overview and comparison of the different participants' task diagrams identified the following subtasks as main events in an operator's work process, independent of the specific task and C2 system.

- **Mission Brief:**

Mission Brief is the first stage of the operator's work process, where they, as the name suggests, receive a brief of the mission from higher command. The brief includes various levels of relevant information, such as the target type or area being searched for. This step highly influences the operator's continued process.

- **Positioning:**

The positioning involves transportation to and the startup of the system in the correct geographical position. It is based on the mission brief and influenced by external factors such as geographical and weather conditions.

- **Sensor & Radar settings:**

Radar mode and settings are selected and adjusted by the operator according to the type of target being searched for. Including selection of sensor direction(s) and area(s).

- **Presentation settings:**

The operator makes presentational settings in the interface, based on their tasks as well as personal preferences.

- **Surveillance & Assessment:**

The operator perceives and assesses the information provided by the system and gains SA. Includes monitoring activity of targets through movements, sensor and link sources, etc., as well as monitoring status of the system itself. The operator tries to distinguish deviations from the “normal”/nominal state. This subtask requires active work from the user, but is also assisted by systems aids, e.g., visual and audible alerts.

- **Communication:**

The operator can communicate with and receive information from other operators, other units, and higher command. It includes, e.g., information about events and targets. All received and communicated information is well documented.

- **Action:**

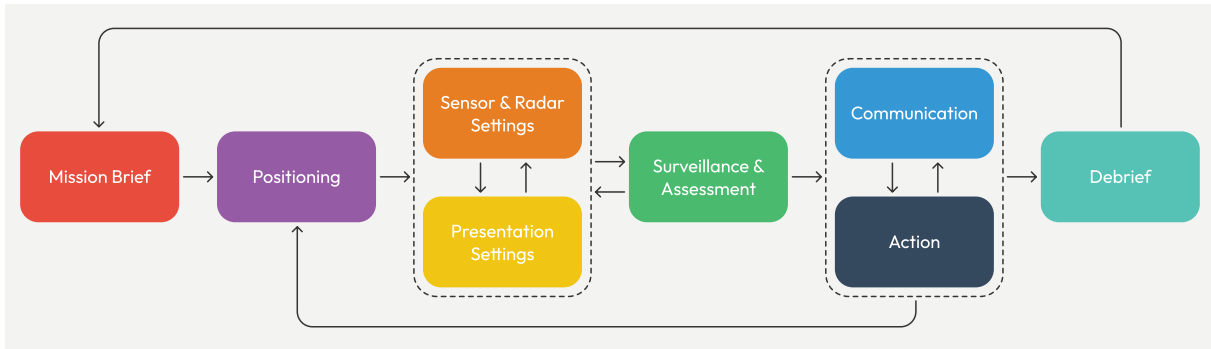
Communication can be received, for which the operator should perform some new actions. For example, changes to the mission, requiring adjustments involving previous steps.

- **Debrief:**

The final step includes analysis of and reporting events from the mission. This information can be a part of the intelligence received in the briefing for a following mission.

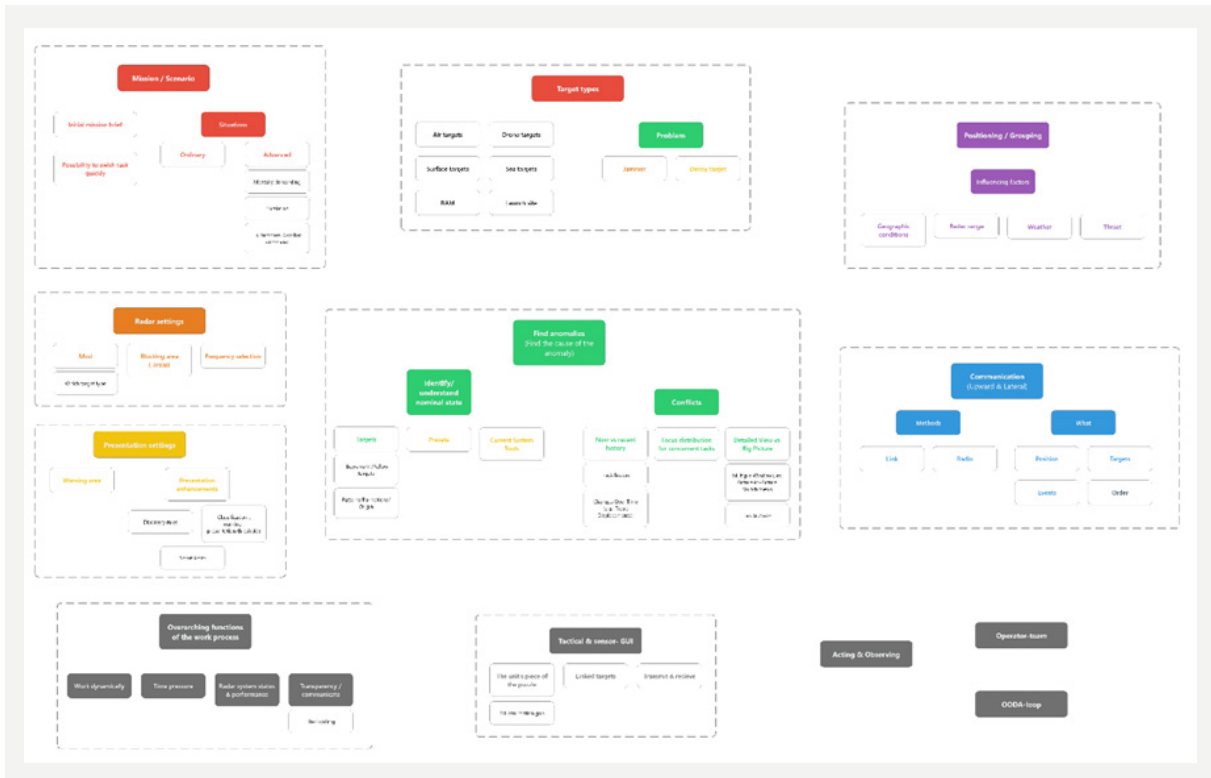
The internal relationship between the steps and the contents of them is illustrated in the domain independent task diagram seen in Figure 6. As seen, the process is not a strictly linear one. The work process is highly dynamic, with the operator constantly processing and acting on new information and feedback provided by the system, making iterations and adjustments, such as changing different settings.

Figure 6
Domain Independent Task Diagram



Analysis made by constructing the KJ diagram gave many insights in mapping and grouping all the participants' responses. The diagram includes a range of information, such as the types of targets that exist, and how and to what sorts of information is transmitted. A digital representation of the diagram, summarizing the most important or prominent groupings of information, is presented in Figure 7.

Figure 7
Digital Representation and Summation of the KJ-diagram



Note. The KJ-diagram is colour coded to the corresponding step in the domain independent task diagram.

One important finding, mentioned by all participants in some capacity, had to do with the perceived purpose of their work. This was formulated as “finding anomalies” (and their cause). This could be identified and described as the main task goal. In order to reach this goal, the participants mention the importance of identifying and understanding the “normal”/ nominal state. This perception of a nominal state are mental models that the expert users have formed from their previous experiences. With these, they are able to utilize the long-term working memory and make rule-based decisions.

6.4 Summary of Phase

The initial user study proved very useful in providing knowledge of the system and tasks, as well as insight into the operator’s role, work and cognitive process. The domain independent task diagram and affinity diagram complement each other well in representing the findings from the study and acted as a support in the following phases.

Sensor & Radar settings, Presentation settings, and Surveillance & Assessment are identified as the most relevant steps in the work process for the scope of this thesis. It is predominantly in these stages that the operator spends their time and effort. Surveilling, reacting and acting on much and constant information, splitting their attention and cognitive capacity between multiple stimuli and tasks. The potential of AI supporting the user is thereby deemed greatest during these stages.



7

Phase B: Idea Generation

This chapter presents the idea generation phase, describing how it was performed and the resulting ideas in terms of their general categories. The phase led into a subsequent delimitation of the project scope, the result of which is also presented.

7.1 Objective

The objective of this phase was to begin exploring the solution space and generate ideas of how AI-supported adaptivity can be implemented in C2 systems. These would be built upon in the upcoming concept generation and evaluation.

7.2 Planning & Execution

7.2.1 Ideation

To perform the ideation, papers, different coloured pens, post-it notes and a timer were placed on a table. Initially, the ideation was broad, due to the size of the solution space, which encompassed mainly but not exclusively the areas of Sensor & Radar settings, Presentation Settings and Surveillance & Assessment (see Figure 6). As time passed, the ideation would be funnelled down to more specific prompts and solution areas.

Hence, a prompt-free round of combined mindmapping and brainwriting was performed first, followed by explanation and discussion of the result. From this, the ideation prompts were set to scenario based. Three levels of seriousness were set: “Normal tasks”, “Non-urgent incidents” and “Urgent serious incidents”. From these, the ideation went towards solution-oriented prompts: “UI information filtering”, “What AI has the potential to do” and “Task-supporting functions”. The two final prompts were “What is important to see?” and “How should the system be interacted with?”. After this, this iteration of ideation was deemed to be saturated.

7.2.2 Idea Sorting

Following the ideation, the KJ method was used to sort the ideas into categories, to get an overview of the ideas and understand where in the operator's work process they were applied. The result was transferred to an Adobe XD file to save physical space and ensure easy accessibility and adjustability.

7.3 Result

The performed ideation thoroughly explored the solution space and resulted in multiple ideas of varying levels. Through the use of the KJ-method, these ideas were sorted, and a total of five main categories of ideas emerged. The five main idea categories were:

- **Generally applicable**

The generally applicable ideas were ideas that could be implemented in many different situations and functions within the system. They included ideas about, for instance, different AI automation levels, visual transparency of AI actions and means of interaction between the operator and the system.

- **Filtered Situational Display**

These ideas are the ones that address different ways of reducing, filtering or sorting the visual information presented in the map view of the UI. The two main idea areas here regard different ways to filter or enhance prioritized targets and different ways of presenting information in split windows.

- **Adaptive UI support**

This category of ideas focuses on how AI can guide the operator to information or actions of interest, such as alerts or incidents. The ideas vary in how forcing the support was, from colour changes to the AI moving the screen to the intended interaction, out of operator control.

- **Mission-, setup-, and process support**

As the name suggests, this idea category contains the ideas about how AI can support operators in understanding the mission, optimize settings, and in the process of identifying targets.

- **Other ideas**

The rest of the ideas were categorized as other ideas. Here, ideas about more automated and leader-oriented AI were placed, as well as ideas about how to keep the operator in-the-loop in such solutions.

7.4 Scope Delimitation

While the scope and ideation were deliberately kept open, to facilitate creativity and unsuspected discoveries, it was now time to make delimitations and narrow down the scope. This decision was made due to several reasons.

The first reason was that the following steps of the process were concept creation and concept evaluation. Conducting a concept creation that incorporates every good idea would have been overwhelming, both in terms of selection and prototyping. Furthermore, a larger number of concepts would have significantly increased the time required for the evaluation phase.

The second reason was to find a way that was relevant to the use cases and addressed gaps in previous research. Some of the ideas, for example, have already been investigated in a previous thesis at Saab (Bung Tidblom & Jansson, 2021). Hence, the delimitations were made in agreement with the supervisors.

The third reason was testability. Some elements of both the initial scope and the ideas presented several challenges in forming tests to obtain insightful results. Such challenges included comparability and the ability to represent more complex ideas in prototypes. This reason was the most challenging to handle and balance. On the one hand, the delimitations should not be decided with testability in mind, since this risk reduces the innovation before it has even started. On the other hand, if test results that can be analysed and compared are desired, it makes no sense to develop a concept that cannot be tested and compared with the available resources of this project. Worth mentioning, this conflict was also present during the following phases as well (C: Concept Creation and D: Concept Testing). There, it has a more direct effect on the concept creation, while in this phase, it affected the scope limitation more generally.

7.4.1 Delimitations of Scope

Here are the delimitations presented with individual motivations.

Assessed Workload and Human Parameters

In the initial thesis brief, several factors were suggested to assess the workload, for instance, pulse, eye movement, reaction time, task complexity, and the number of tasks. A delimitation was set here: the human parameters to base the adaptive presentation on will not be designed for or tested. The area requires proper investigation to reach useful results. Therefore, for the purpose of this thesis, an assumption is made that AI can make a judgment on how to merge different input parameters to a total workload level. The total workload level will be simulated by adding additional cognitive tasks with increasing difficulty, which will be performed simultaneously with the main task in the Adaptivity & Workload Test.

The Radar Surveillance Operator Role

Since this thesis investigates adaptivity in C2 systems for both AEW&C and surface radars, different operator roles have emerged. However, to keep the research independent of the type of C2 system, the operator role focused on from here on was the radar surveillance operator role, which is present in both areas. Solutions for other operator roles will not be further investigated.

General Visual Presentation

Another delimitation is that the visual concepts will be general ways to filter, sort, and prioritize information, rather than specific functions that support the operator. This increases the number of C2 systems, situations, and functions for which the result is relevant. Furthermore, it allows for a more general testing process, where the fidelity of existing systems can be lowered, and the biases of the expert users regarding functions and strategies will play a smaller role. The testing will assess how the operators take in and process C2-generalized information when the mental workload increases, and how adaptivity affects that performance.

Surveillance & Assessment rather than optimizing settings

In the previous phase, a delimitation was made to focus on Sensor & Radar settings, Presentation settings and Surveillance & Assessment in the domain independent task diagram. Now, after further deepening of knowledge, the project will focus on the Surveillance & Assessment part, as this is where most of the SA is gained. Today, Surveillance & Assessment is closely related to Sensor & Radar settings and Presentation settings, since it is an ongoing iterative process between the three, where the latter two are support functions to optimize the SA gained in Surveillance & Assessment. From here on out, we make the assumption that Sensor & Radar settings and Presentation settings are optimized to display accurate information for Surveillance & Assessment. How they are optimized is not technically specified. Concluded, the information in Surveillance & Assessment will not be negatively affected by errors in the settings. Furthermore, this delimitation results in the settings not being included in the tasks of the operator and thereby not regarded as tasks in the concepts and testing. This delimitation is primarily made due to three factors that emerged from the interviews. Namely, that Surveillance & Assessment is a crucial, time-consuming, and challenging task. Improving its workflow has the potential to significantly improve efficiency and reduce the workload on the operator. Hence, this is where the following work will be focused.

Current System Controls

The means of interaction will be kept the same as they are now: mouse, keyboard and headset. This will make testing easier and the results easier to implement. There is potential for investigating alternative means of interaction, but it will not be explored further in this thesis. This aligns with the decision not to test human parameters, since such tools will not be included.

The AI makes the right decisions

The ideas focus on what AI-supported adaptivity can be and how it can be visualized. As mentioned during the initial delimitations (see section 1.3), an actual AI would not be developed and implemented. Instead, the decisions supposed to be made by the AI would be replicated manually.

To develop and test the different ideas, an assumption is made that these replicated decisions, supposed to be made by the AI, are correct. If the AI decisions have to be critically evaluated, there are too many variables to draw conclusions about the adaptivity and how it is visualized. Therefore, the AI decisions are assumed to be correct.

7.5 Summary of Phase

This phase was initiated with an idea generation, resulting in multiple ideas, which were in turn sorted into the five categories, *Generally applicable*, *Filtered Situational Display*, *Adaptive UI support*, *Mission-, setup- and process support*, and *Other ideas*.

Due to multiple factors, including the wide range of ideas, the phase led to a delimitation of the project scope. The result of which is presented, and the following conclusion can be drawn.

The scope is the SA in a cognitive demanding, information overloaded routine surveillance scenario within a general system C2. The concepts and testing of concepts will assess how information and understanding are gained through visual adaptive sorting, filtering and prioritizing, when the workload increases.



8

Phase C: Concept Creation

As previously mentioned, the third and fourth phases of the project, involving Concept Creation and the subsequent Concept Evaluation, were partially performed in parallel and highly influenced each other. This chapter focuses on the Concept Creation.

8.1 Objective

The objective with this part was to develop the ideas presented in the idea generation into more refined concepts. These concepts would contribute to answering, mainly the first of the two research questions, through subsequent evaluations.

- How can AI-supported adaptivity be implemented and visualized in Command and Control (C2) systems?
- Can AI-supported adaptivity reduce the operator's mental workload while keeping a sufficient level of SA?

8.2 Planning & Execution

8.2.1 Concept Selection

In order to develop the ideas presented in the idea generation into more refined concepts, an initial idea elimination was necessary. This was first done by eliminating the idea themes that no longer fit into the updated, delimited scope of the project (presented in section 7.4.1 Delimitations of Scope).

Following this, the elimination was conducted on a more detailed level. The physical sketches and wireframes created during the Idea Generation phase (for concepts that had not yet been eliminated) were cut up and sorted once more in a new KJ diagram. The main argument for this method was to more thoroughly and organized go through, and thus avoid the risk of missing any potential solutions. Prioritization and selection of the ideas were performed

through dot-voting, based on judged relevance and feasibility from the knowledge gained so far. This method not only helped with the selection, but it also contributed to and transitioned naturally into the upcoming refinement and creation of the concepts.

8.2.2 Concept Creation

After the selection, the next step was to start the refinement and creation of individual concepts within the four selected categories. Therefore, an additional brainstorming session was performed. The method selected was based on the Crazy-8 method, with some modifications, including a shorter time duration and specific prompts to keep in mind for each iteration. Prompts include, for example, “How could the layout of [concept category] look?” and “Could some [concept category] ideas be combined?”. The intention was to increase the efficiency of the process and aimed at adding realism and feasibility to the concepts, contrary to the initial iterations of the idea generation, which had an exploratory focus.

The next step included an iterative process of prototyping the ideas with an increasing level of fidelity. This was achieved using graphical design tools, including Adobe XD (Adobe Inc., 2025b) and Illustrator (Adobe Inc., 2025a), with the first containing functionality for creating interactive prototypes. This was desirable for imitating the interactive and adaptive properties of AI-supported adaptivity in a C2 system. However, Adobe XD comes with some technical limitations in interactivity, resulting in some compromises in the coming tests. More about this later, in the situations it affected choices and delimitations.

8.3 Result

8.3.1 Concept Selection

The initial idea elimination, based on the updated delimitations, directly eliminated a large amount of the previous ideas (as visualized in Figure 8).

In the following, more detailed KJ-method iteration, four categories of ideas were identified and selected for continuous development. Compared to the first KJ sorting of five idea categories presented in the Idea Generation phase and described in section 7.3, these categories are more refined and homogeneous, and serve as the foundation for the rest of the project. They are: *Filter*, *Target Prio-List*, *View*, and *Transparency*.

Filter:

The ideas belonging to the filter category involve alterations of how adaptivity could reduce the mental workload of operators by visually prioritize targets deemed important or otherwise notable by the AI. The ideas include visual changes to the targets or map, e.g. changes in size, opacity, or colour. Such increased salience of the stimuli would aid bottom-up processing and draw the operator’s attention to the right target.

Target Prio-list:

This category contains ideas of how a list of targets (target-list) could be made adaptive, for instance by prioritizing and sorting targets automatically with changed target priority.

8.3.2 Concept Creation

After iterations and adjustments during the process, a few concepts belonging to each concept category of *Filter*, *View*, and *Target Prio-list* were decided to move forward with based on judged relevance, feasibility, novelty, and the spanning width of ideas. To exemplify the “spanning width”, if two similar concepts were both deemed relevant, feasible, and new, one could be replaced by a third concept deemed “worse”, but that would add broader total insight during the coming testing and evaluation. Regarding *Transparency*, as mentioned earlier, these ideas are meant to be applied to other concepts and adaptivity in general. Therefore, *Transparency* was not conceptualized further from here on its own but instead implemented when needed in the other three concept categories.

At this stage, a delimitation was made for the visual prototypes. From the NATO Joint military symbology, only “Unknown Air Target” and “Unknown Surface Target” were used in the concept prototypes. This delimitation has to do with the subsequent evaluation and will be further elaborated upon in that section.

Filter

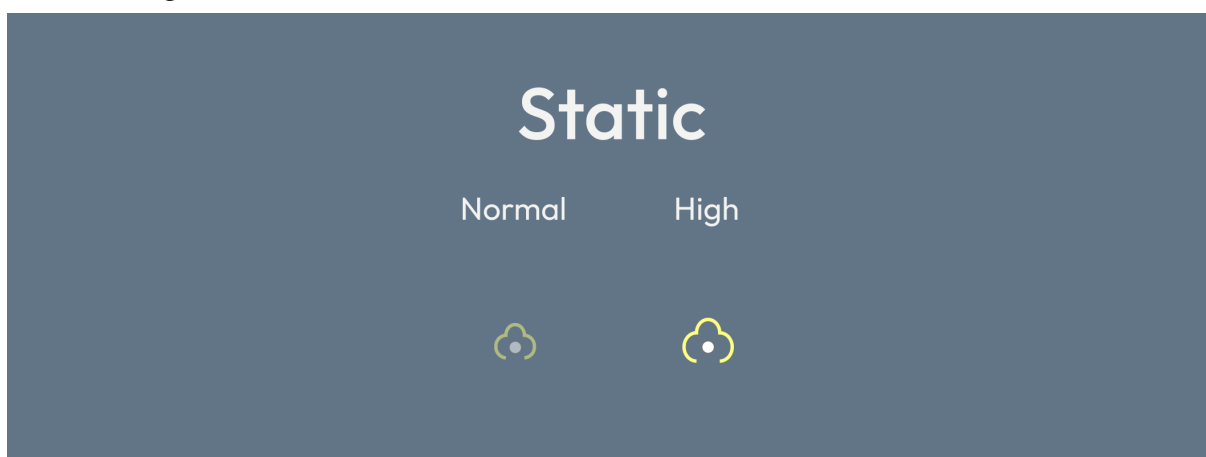
Four different concepts within the filter category were decided upon. These are referred to as: *Static*, *Dynamic*, *Map Change* and *Tunnel Vision*.

Static

The first concept, named *Static*, utilizes a combination of opacity and size to convey the difference between prioritized and normal targets. Prioritized targets are larger, and normal targets are regular sized with reduced opacity, to capture and focus the operator’s attention on the ones prioritized by the AI. An illustration of *Static* can be seen in Figure 9 below, on NATO-symbols of unknown air target, however, the concept can be applied on any alteration of NATO-symbol.

Figure 9

Static Concept

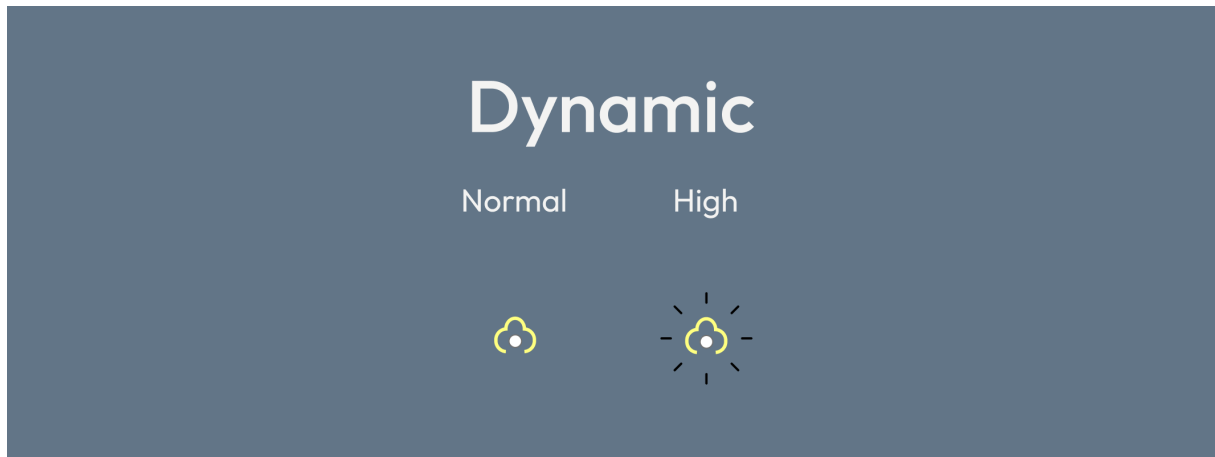


Dynamic

The Dynamic concept, similar to Static, differentiates between normal and targets decided as high priority by the AI. Contrary to Static, Dynamic does not affect the normal targets, it only adds an indication on the prioritized ones. This indication is illustrated in Figure 10 and revolves continuously around the symbol in the digital interface, thus capturing the operator's attention through movement.

Figure 10

Dynamic Concept

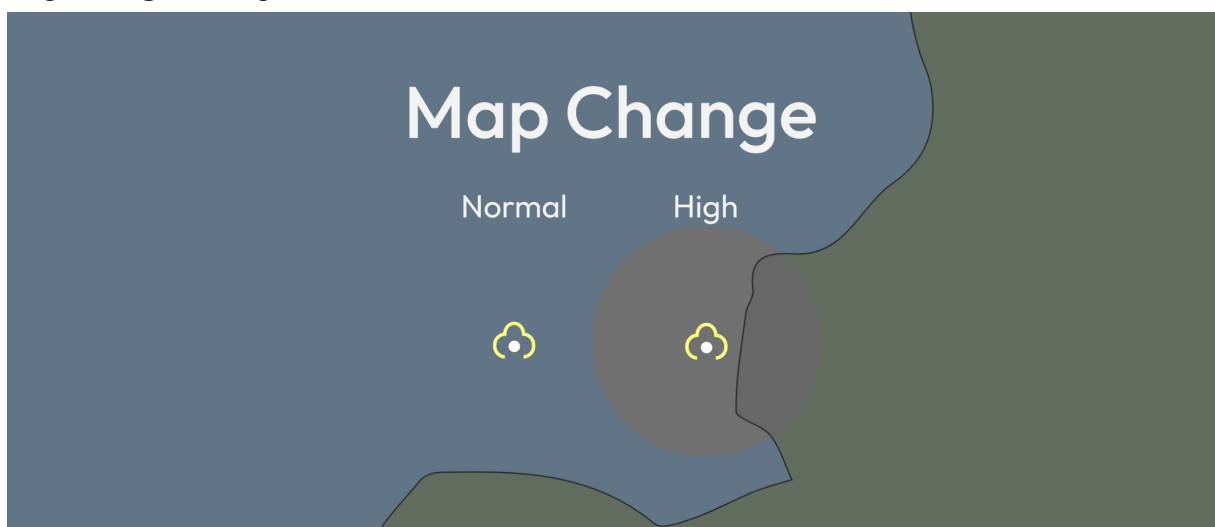


Map Change

Map Change, as the name suggests, does not alter the symbol of the target, but the surrounding area on the map. The map, as demonstrated in Figure 11, is altered in hue around the high priority targets, intended to make them stand out more from the surroundings through visual hierarchy and colour contrast.

Figure 11

Map Change Concept



Tunnel Vision

The last *Filter* concept is named Tunnel Vision, a nod to its function of intentionally replicating both the tunnel vision metaphor, where one focuses on only one thing at a time and misses what is happening elsewhere, as well as the physical phenomenon of tunnel vision, where one loses their peripheral vision. This replication is performed by blurring everything except those targets which are prioritized by the AI and an area surrounding them. The intention is that the contrast between blurred and sharp sight will capture the operator's attention.

Figure 12

Tunnel Vision Concept



Target Prio-list

Four concepts of *Target Prio-list* were created, with different variations on how to indicate the priority of targets in a list format, based on live AI decisions. In all concepts, the target in the list and in the map are connected when the operator clicks on them. If the operator wants to know the location of the highest priority from the list, they click on it, and the target will be marked on the map. In the prototypes, it is indicated by a black ring around the target, which is not a final solution. It works the same way in reverse: if the operator finds one target in the map interesting and clicks it, the target gets outlined in the list. The variations of these concepts are named: *Sorting*, *Size*, *Opacity*, and *Bar*.

Sorting

Sorting is, as the name suggests, a prioritization by sorting the targets in order of priority, with the highest at the top and then in descending order. When priority changes, the list adjusts accordingly, switching the order of the targets.

Size

This concept prioritizes the targets by changing the size of their corresponding card in the list. Larger cards are more prioritized targets than the smaller ones. When AI makes changes of priority, the cards adapt accordingly in size.

Opacity

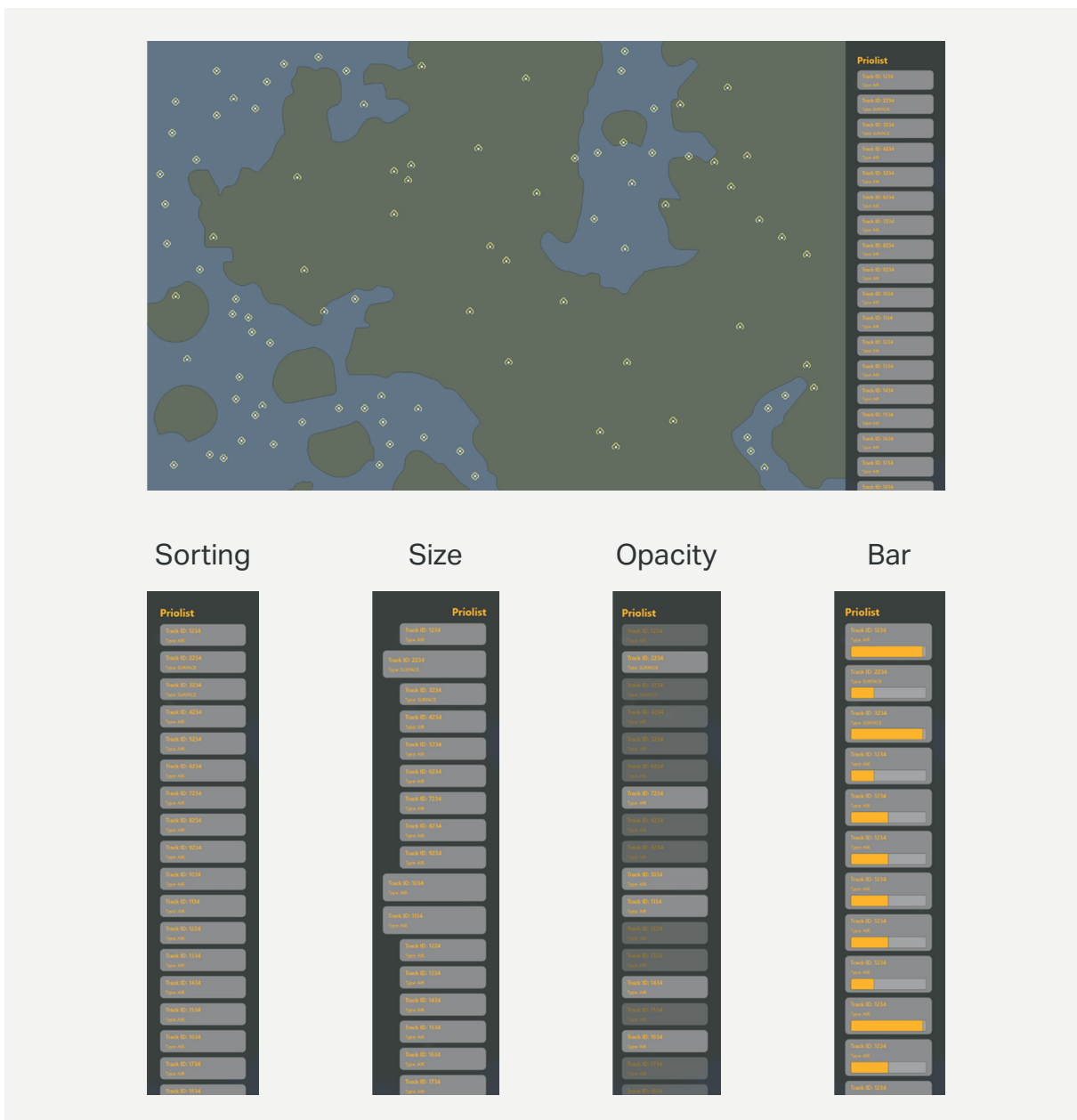
In this concept, the prioritization is visualized by reducing the opacity of the cards of less prioritized targets.

Bar

The bar concept conveys the prioritization of targets through a linear indication or “bar” in each corresponding card. This provides more detailed information on the priority status of each target. In Figure 13, the linear indication is represented in an orange colour, filling the bar horizontally from left to right depending on the target’s priority status.

Figure 13

The four Target Prio-list Concepts



Note. Sorting is visualized twice, first to illustrate the placement of all concepts in the interface, related to the map. Then, all four concepts are shown.

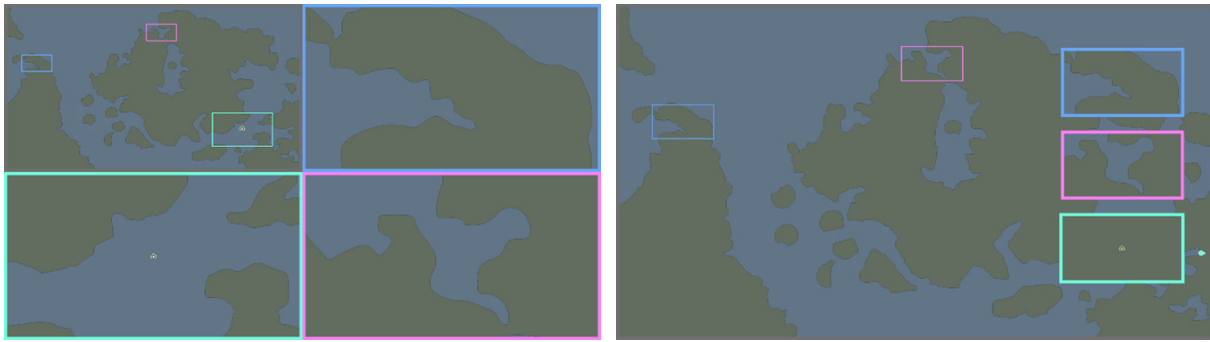
View

As mentioned, the *View* category addresses both the possibilities of how different windows can contain adaptive content and how the windows themselves can be adaptive. The main idea with *View* is to address the conflict between seeing the big picture and details simultaneously. From the initial interviews, it was found that this can be a challenge in current systems.

Two main methods of splitting into different views were prototyped: split-screen (left) and picture-in-picture (right), as shown in Figure 14. All concepts include a total overview at all times. In the split-screen version, the total overview is the upper-left quarter. In the picture-in-picture, the total overview is the large background window. To facilitate recognition and understanding of scale and relative placement in both versions, the windows and their representation in the overview are colour-coded.

Figure 14

View Concept: Split-Screen (left) and Picture-In-Picture (right)



Numerous versions of how to utilize these different views with added functionality and adaptivity were visualized for each view method. The functionalities include:

- **Overview**

As mentioned earlier, a constant overview of the full operational area is always present.

- **Target following**

A prioritized target is locked in on and followed in its own dedicated, zoomed-in window (turquoise in Figure 15). The target and its surrounding area covered by the view are visible in the overview (top-left in Figure 15). If the target exits the zoom level of the overview, an indication of its position appears on the edge of the overview, while the operator can still see the target in the detail window (turquoise in Figure 16).

- **Area of interest**

A window can also be dedicated to a specific area or location, which is deemed important by the AI. This is represented by pink in Figure 16 to help the operator keep focus on the narrow strait leading to a bay.

Figure 15

Illustration of View Concept - 1

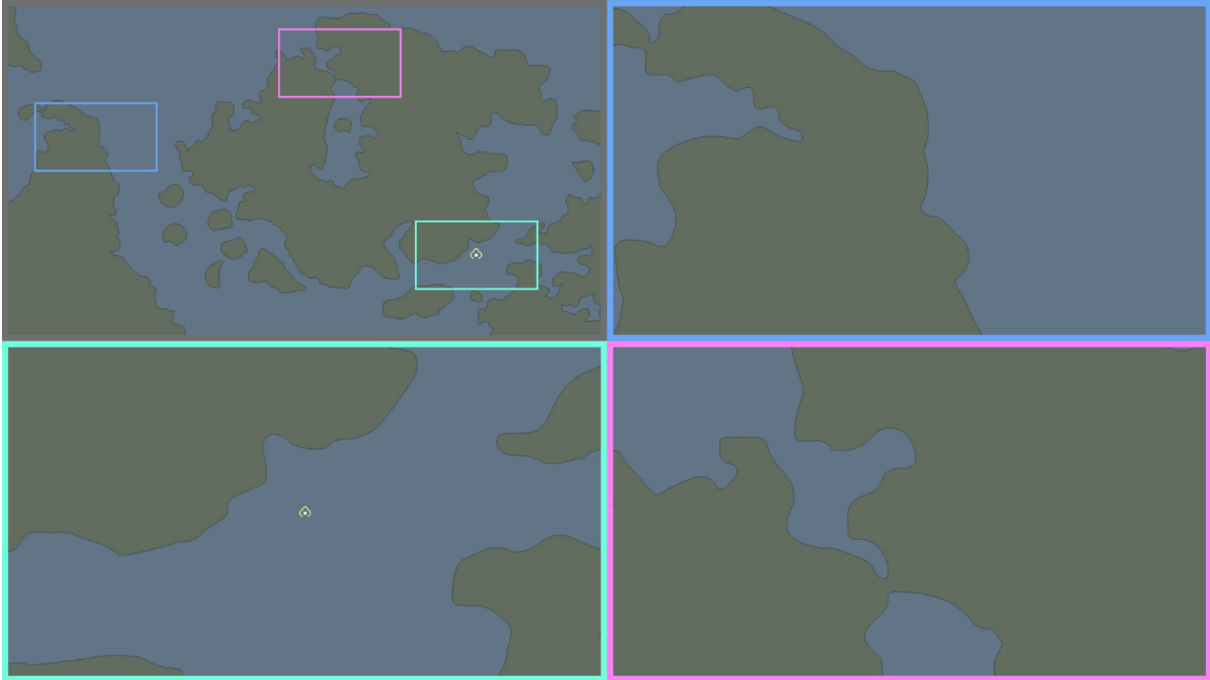
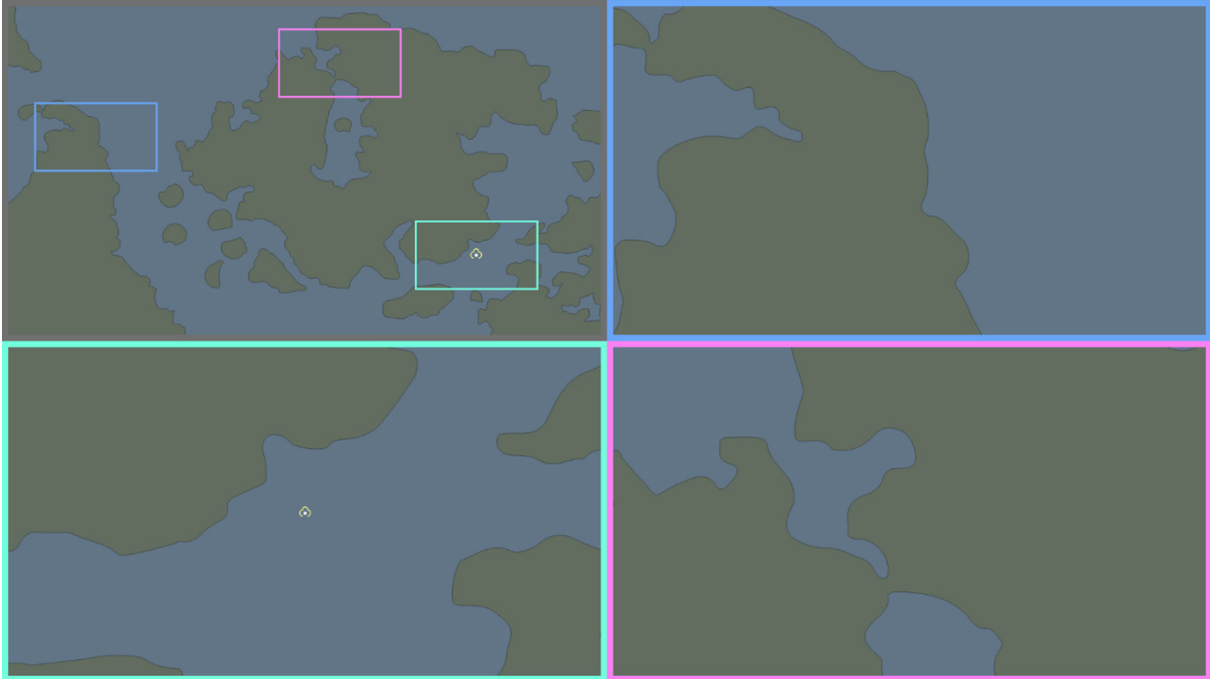


Figure 16

Illustration of View Concept - 2

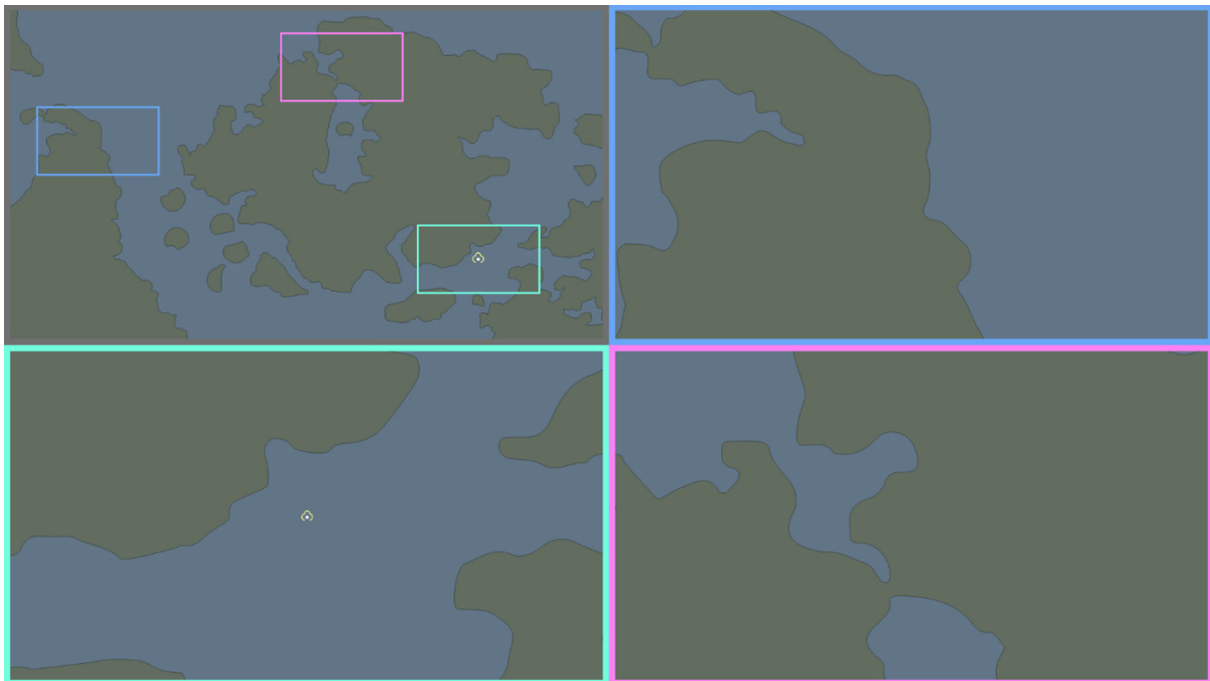


- **Additional functions and information**

Another functionality relevant to have in the adaptive windows (when the AI decides that it is beneficial) is functions and information used by and interacted with by the operator in different situations. In this version of the concept, additional functions or information are represented by a *Target Prio-list*, sorting targets in descending order of priority. Concluded, the overview is located at the top left, detailed views of the target (turquoise) and area of interest (pink) are displayed below, and the interactive function, which adds more information, is located to the right.

Figure 17

Combination of View Concept & Target Prio-list



8.4 Summary of Phase

This phase resulted in the selection and prototyping of concepts belonging to the categories, *Filter*, *Target Prio-list* and *View*. A fourth category, *Transparency*, was also discussed; however, it was decided not to continue with it on its own, since the idea is not stand-alone but could be implemented in other categories.

The different categories and their respective individual concepts were to be studied in the following evaluation, contributing to answering the research questions.



9

Phase D: Concept Evaluation

This chapter concerns the part of the project in which the concepts created in the previous phase were evaluated.

9.1 Objective

The evaluation was planned and performed in order to evaluate the potential of the selected concept categories and their respective solutions. As mentioned in Phase C, this would contribute to answering predominantly the first of the two research questions.

9.2 Planning & Execution

The structure of the study was divided into three parts, one for each category, in the order *Filter*, *Target Prio-list* and *View*. The participants in the evaluation were the same experts who were interviewed during the study performed in the first phase of the project (see section 6.2.1). In the part of the study regarding the *Filter* category, these were complemented by human factor experts in order to increase the dataset and reliability.

Similarly to the previous interviews, the participants were informed firsthand about the project at large as well as the purpose of the particular study, including their rights and what data would be collected. These were followed by a similar survey of questions that mapped their experience and background (Appendix B), differentiating only with added questions regarding their eyesight, in order to determine if they had any problems that could potentially affect the test results. The actual interview questions can be seen in Appendix E.

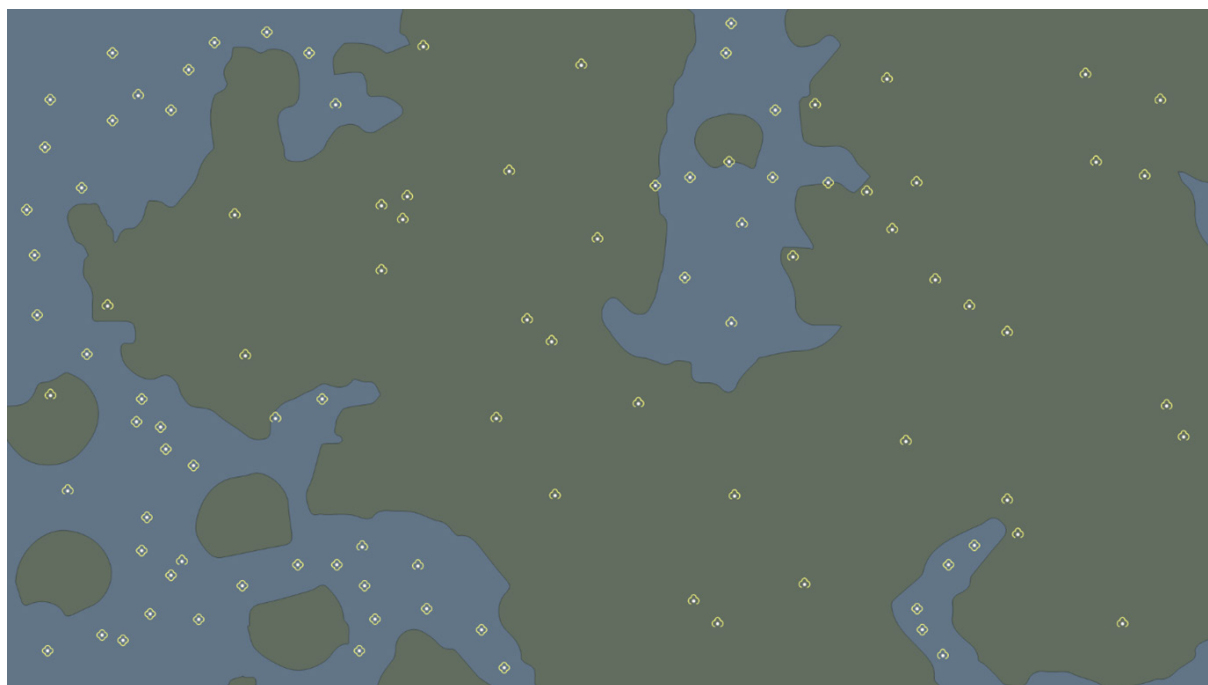
For each category, digital prototypes of the respective solutions (see section 8.3.2) were presented to the participants. During the creation of the digital prototypes, a few important realizations impacting the evaluation were made, followed by decisions and solutions to address these realizations. First, the situation map for each prototype was standardized. The situation map should not change regarding colours, location, etc. between the different concepts, since this could affect their perception and the results of them. The same conclusion applies to the scenario of targets presented, which was thereby also standardized. In addition,

the map should depict a location which is unfamiliar to the operators in evaluations made to investigate the performance of the concepts. Reasons being to avoid any preconceptions and biases originating from their expert knowledge and previous experiences.

As mentioned during the concept creation, another decision was made during the prototyping regarding the target types. These would only include Unknown targets of Air and Surface. This decision was based on two main factors: the decision only to use two types was to reduce the number of variables, and the choice of “Unknown” in particular was made to avoid operator bias. For example, choosing “Hostile” could cause the operator to see it earlier than other identities, as that is how they are trained. Additionally, the realism of the prototypes in the eyes of an operator could be reduced if the “AI” in the tests does not prioritize the hostile targets. The chosen map and targets based on these decisions can be seen in Figure 18.

Figure 18

Situation Map Used for Evaluations



The following sections present how each concept category evaluation was performed.

9.2.1 Filter

The *Filter* part of the study investigated how easy it was to spot the prioritized targets in the respective solutions: Static, Dynamic, Map Change, and Tunnel Vision. This was performed by exposing the participants to four situation maps, each with a different filter method applied. In these maps, ten targets were prioritized from a total of 100. The maps are located in Appendix F. The participants were tasked with identifying these prioritized targets, which they indicated by clicking on them. They were not informed that the number was ten, so they would not know by the number when they were done. Due to technical limitations in Adobe XD, it was only possible to make three of the four concepts interactive when clicked on. In these three, their targets were crossed out after being clicked on, helping the participant see what they had already clicked on. The exception was “Dynamic”. To be able to make the dynamic movement work, it was not possible to make it crossed out when clicked on, since

this stopped the moving animation loop. To counterbalance this, the participant was reminded that this would be the case just before that particular test started. When they perceived themselves as finished with the task, they would verbally express so.

The participants were timed manually using a stopwatch, measuring both the time until their perceived completion and the actual completion of the task. Additional collected data were the number of correctly identified targets, along with the number of errors. *Filter* were the only concepts that were measured in a quantitative way in the first round of testing, while also collecting qualitative data. *Target Prio-list* and *View* were only assessed in a qualitative manner. The reason for this was that the four *Filter* concepts were easier to test performance-wise in this phase, and they were also more directly comparable to each other.

In order to identify and minimize carryover effects between the four tests, the order in which the participants performed each filter method was mixed, as seen in Table 2. As mentioned, in order to aid with this and increase the reliability of the qualitative data, this test was performed with additional participants, all of whom were human factor (HF) employees. All evaluations were within-subject.

At the end of this section of the evaluation, follow-up questions were asked in order to extract the participants’ opinions regarding both the use of *Filter* at large and the input of each individual solution.

Table 2
Filter Test Order

Participant		Order			
Experts	1	A	B	C	D
	2	B	D	A	C
	3	C	A	D	B
	4	D	C	B	A
Human Factor	5	A	D	B	C
	6	B	A	C	D
	7	C	B	D	A
	8	D	C	A	B

Note. A = Static, B = Dynamic, C = Map Change, D = Tunnel Vision.

9.2.2 Target Prio-list

The part of the study regarding *Target Prio-list*, which differentiates it from *Filter*, does not evaluate the performance of the different solutions in the form of collecting any data points. Instead, interactive prototypes of each solution, that is, Sorting, Size, Opacity, and Bar, are presented for and interacted with by the participant, acting as a basis for interview and discussion. All prototypes include an animation of re-prioritization of different targets. Questions were asked after each interaction to help probe the information, including, e.g., how the different ways of prioritizing are experienced, if added features like this would affect the work process of identifying targets, etc.

9.2.3 View

In the evaluation of *View*, similar to *Target Prio-list*, the different visualization methods previously mentioned, including the various functionalities, were presented to the participants in digital prototypes, which acted as a basis for interviews and discussion. The aim was to evaluate:

- Which type of screen dividing is preferred, split-screen or picture-in-picture?
- Are the different functions, Overview, Target following, Area of interest, and Additional functions and information, which are meant to be shown in windows, relevant and useful functions?
- Do the operators understand the way the adaptivity *within* each window, that is, how the function is adaptive (e.g., if a target of interest moves outside the Overview), is visualized?
- How many of these adaptive changes can happen at the same time in different windows, and still be comprehensible and not confuse the operator? For example, if one window shows when a target moves outside the overview, while another shows an area of interest being zoomed in, and a third window changes from one interesting target to another, higher prioritized.
- Can the position of the windows be adaptive as well? For example, can the view change from split-screen to picture-in-picture as the AI determines that the Overview is more important than the details in the windows and vice versa? Or can the windows themselves change order based on the prioritizing of them?
- Can the window content and the windows themselves be adaptive at the same time?

The questions asked were of the same nature as during the evaluation of the *Target Prio-list*. They were, for example, “Do you understand what is happening here?”, “What do you think about it?” and “How do you feel about that AI taking this decision to change your view?”

Finally, after all three evaluations were performed, the operators were asked what they base their decisions on when identifying and evaluating targets. This was asked to gather information on what an AI could base its decisions on, to be both correct and be trusted by the operators.

9.2.4 Analysis

To compile and analyse the qualitative results from the tests and evaluations, the KJ-method were used once again, for each concept category. This was done by writing down the results, answers, and other inputs from the operators onto notes, then sorting them into categories such as “General opinion”, “Potential improvement”, and “Feeling towards adaptivity”.

The quantitative results from *Filter* were compiled in a Microsoft Excel sheet for easy overview and analysis of the times (actual and perceived), and the number of clicked targets (correct and error).

9.3 Result

9.3.1 Filter Evaluation

The results from the *Filter* will be presented first from the timed tests, and then from the interpretation of the opinions combined with observations about each concept.

Timed Tests

Generally, the total hit rate was high for all concepts, with 95% being the lowest total scoring hit rate on a *Filter* concept. However, all errors were made by one participant here. The highest total hit rate for a concept was 98.75%. The difference was not big enough among the eight participants to draw conclusions. Dynamic, the only test where a clicked target did not get crossed out, had a disadvantage on one hand, since the participants had to remember which of the ten targets had already been clicked on. However, during the tests, the concepts where the targets did get crossed out sometimes had a delay in the appearance of the cross. Furthermore, if not clicked on correctly, the cross did not appear. This resulted in slight delays at times. Hence, it is hard to say if the disadvantage of keeping Dynamic targets in the head outweighed the advantage of being clicked on faster. Additionally, it could be seen that the time difference between being done and feeling done was, with some exceptions, only a few seconds. This indicates that the participants have a good overview of all prioritized targets. Keep in mind that they were not informed that it was ten, so they did not count to and knew they were done just by reaching ten targets. In conclusion, the quantitative results from the *Filter* evaluation differed only slightly. No conclusions about which concept is best are drawn from the times. The small differences may as well depend on other factors, such as the learning effect, delays in the crossing out, a slight error margin in the measured time, etc.

Qualitative Evaluations

Since the quantitative results did not provide much insight, the qualitative results were more important for the overall outcome. The following presents these results for each concept.

Static

It was concluded that Static did not have a big enough difference between prioritized and ordinary targets. However, this concept is scalable according to preference, so the difference can be altered. It did not disturb or have an annoying effect on the participants. The colour of ordinary targets changes differently depending on the background, due to the opacity. This was not mentioned as a problem.

Dynamic

The opinions differed on whether this concept made prioritized targets easy or hard to discover. Many agreed that it can be too cognitively demanding over time or with an increasing number of prioritized targets. The spinning movement may also visually interfere when the targets move on the map, as they do in reality, but not in the test. Dynamic as an idea, in the sense of making prioritized targets draw attention through movement, was not dismissed. It was this particular version of movement that was. Pulsating was suggested as an alternative. Also, e.g., pulsating can be combined with other *Filter* concepts to provide different levels of prioritizing.

Map Change

This concept was polarizing. Some thought it was pleasant to look at, while others thought it drew too much attention. It was challenging to distinguish closely located, prioritized targets. Also, due to the Map Change being a black and white filter applied on different colours of the map (blue and green), the colour of the Map Change differed slightly. Therefore, participants thought the two resulting variations had different meanings. Furthermore, some kind of optical illusion caused the participants to perceive the grey indications as red or purple.

Tunnel Vision

For this concept, some participants expressed concerns about removing too much SA. They mentioned that the blur removes too much detail from the targets, which will become an even bigger problem in reality when more variations of NATO standards symbols are used. Additionally, it can be demanding to look at blurred images for extended periods. However, the prioritized targets are relatively easy to see.

Furthermore, some general insights were gained. To begin with, there is a gain in having different levels of prioritizing through visual differences, to see the whole picture and the prioritized at the same time and know where to act first. Another insight was that some participants used tactical scanning patterns. Finally, in reality, when targets move and the information level increases, the situations are different. Therefore, these results have to be tested and iterated with added complexity.

9.3.2 Target Prio-list

Today, there is a list function that provides information about the targets. There is no intelligent prioritization in it. How much it is used varies between operators. It is used to gain more information about the targets, sort by specific properties, or as a target encyclopedia, but it is not where the work is primarily performed. Three out of four operators thought that there is potential in using an intelligent *Target Prio-list* to identify targets. The results of each concept evaluation are presented in the following sections.

Sorting

All four participants thought that the prioritization should be made from top to bottom in descending order, which was said for the other three concepts as well. However, they thought that a clearer indication of the re-prioritization need to be implemented for it to work. When it just moves, it is hard to know what has happened. A suggestion for improvement was the

addition of small arrows and numbers to indicate the position changes made in the list, similar to adaptive scoreboards, along with an auditory indication of the change. Additionally, they thought the changes has to happen with the right speed. Not so fast that it gets missed with the blink of an eye, but not so slow that it takes time from work. Furthermore, some form of indication for the total relative importance was wished for, as in the concept Bar. Otherwise, 100% importance or urgency could appear as important as 50% importance or urgency, if it is located at the top of the list. All these wishes and suggestions correlate to the fourth concept area, “Transparency”. As stated earlier, “Transparency” was not conceptualized but meant to be applied to the other three. Here, it is clear that transparency is very important to avoid confusing the operators.

Furthermore, the participants wished that targets that are currently clicked on, which are under investigation, should stay clicked on when a re-prioritization occurs, in order not to be confusing.

Size

As mentioned, it was not popular that the following three concepts did not prioritize top-down. However, Size was experienced to be clear in its binary form, prioritized or normal. This contradicts the wish for prioritization relative to a total max described under Sorting. Questions arose about what happens if a target that is not in the list suddenly gets prioritized. Would the operator have to scroll to see it? Additionally, the size difference removed more from the map than the other concepts.

Opacity

This concept was unclear compared to the other three. The operators thought it was not easy to know if the targets deemed as prioritized in the list were actually prioritized, or if they had newly appeared on the radar. Also, the normal or unprioritized targets, with the lower opacity, were confused with not being active any more. This could be because interactive elements with lower opacity in interfaces generally indicate that the element is not meant to be interacted with. Concluded, the visual contrast was clear but confusing.

Bar

As stated under Sorting, the function of having this indication of prioritization relative to the total and not only other current targets was liked. An opinion that came forward during this evaluation was that it is hard to see a threshold value for when a target becomes “important”, as the binary Opacity and Size provide.

General Conclusions about Target Prio-list

For all concepts, it became evident that the prioritization of the target in the map has to be made clearer, as in the *Filter* concepts. One operator suggested that it could be interesting to always have the top five most important targets automatically prioritized in the map as well. It could be seen that operators wanted to have both a threshold value, important or not (like the binary Opacity and Size), and an indication relative to the total importance (like Bar). When one was present, the other was missed. A solution could be to combine them, to have all important targets, for instance, bigger, and to have the bar on them.

9.3.3 View

In the systems today, an operator can present multiple windows simultaneously. However, the windows are static and manually placed, and neither their content nor their placement adapts intelligently. Earlier, a list of questions to answer was presented (see section 9.2.3), which were aimed to be answered in the evaluation. The evaluation results addressing those questions are presented next:

- Which type of screen dividing is preferred, split-screen or picture-in-picture?

Split-screen was preferred to see the details, while picture-in-picture is easier to have the overview. However, for picture-in-picture, the details could be so small that they were hard to see. A general opinion was that the windows are positive as long as they add more information than they remove. In both, the colour coding was deemed to be positive and intuitive.

- Are the different functions, Overview, Target following, Area of interest, and Additional functions and information, which are meant to be shown in windows, relevant and useful functions?

In general, potential was seen in all four functions. Target following and Area of interest were the most interesting. However, when testing animations of a re-prioritization in the shape of a window selecting another more important target or area, it proved to be potentially confusing. Additional functions or information, which were represented by the *Target Prio-list*, did not give more insight than testing it on its own. It was indicated that it gave too much information at the same time. The case may be different if it was more interactive tools than adaptive.

- Do the operators understand the way the adaptivity *within* each window, that is, how the function is adaptive (e.g., if a target of interest moves outside the Overview), is visualized?

In general, yes. Three out of four understood the Target following, even when it disappeared outside the overview. The case was the same with the Area of interest. However, as mentioned above, if the windows changed target or area, it could be confusing. Additionally, a window with *Target Prio-list* could be confusing with too much information.

- How many of these adaptive changes can happen at the same time in different windows, and still be comprehensible and not confuse the operator? For example, if one window shows when a target moves outside the overview, while another shows an area of interest being zoomed in, and a third window changes from one interesting target to another, higher prioritized.

Here, it was clear that a change within one window was understandable. Within two was confusing, while three changes simultaneously were incomprehensible. Then all three changes were lost.

- Can the position of the windows be adaptive as well? For example, can the view change from split-screen to picture-in-picture as the AI determines that the Overview is more important than the details in the windows and vice versa? Or can the windows themselves change order based on the prioritizing of them?

It can, if it happens one thing at a time, as shown in the previous question. However, its use was questioned by the operators.

- Can the window content and the windows themselves be adaptive at the same time?

With the knowledge from simultaneous adaptivity, the short answer, without further research, is no.

9.3.4 Operators' Opinions on AI

A recurring theme for all operators participating in the studies of the thesis (though only rarely for the Human Factors employees who contributed to the *Filter* evaluation only) was scepticism towards AI. Even if an explicit precondition of all evaluation was that the AI bases its decisions on the right factors and comes with a correct prioritization, the participants frequently stated that they needed to know what the prioritization was based on. In *View*, it reaches another level, since in this concept, the AI not only decide what the operator should see, but also how. Some doubted that AI could perhaps decide what they wanted to see now, but not how the operator forms their strategy to foresee the future.

Generally, the participants want some way to override the AI if needed, or be able to turn it off or have it as a support that does not decide the visualizations.

To be able to increase the trust for the next iteration of tests and draw conclusions about the required input for the AI later, the participants were asked about what they base their identification decisions on. The following is the compiled answer, in no particular order:

- Mission type
- Target type (surface/air)
- Geographical context
- Personal responsibility area
- Signal of identification
- Target behaviour
- Altitude
- Velocity
- Direction
- Target number
- Distance
- Radar target area

9.4 Summary of Phase

This phase has evaluated the categories of concepts created during the previous concept creation phase. The results provided various relevant insights and opinions about both the concept categories as a whole, as well as their respective individual concepts. There exists potential with multiple concepts, and information regarding possible improvements is drawn.

One important aspect that participants with previous operational experience expressed was regarding the control of the AI. They all expressed restriction or doubt that AI could make the right decision, which indicates a need for transparency and explainability from the AI. Some of the information which the users would like the AI to take into consideration when making their decisions is listed in section 9.3.4.

10

Phase E: Adaptivity & Workload Testing

This chapter presents the final phase of the project, in which Adaptivity & Workload testing were performed. It presents the selection of and refinements made to the tested concept, how the tests were performed, and provides their results.

10.1 Objective

From the tests in the previous phase D, insight and understanding were gained regarding the solution areas, *Filter*, *Target Prio-List*, and *View*, to help answer the first research question:

- How can AI-supported adaptivity be implemented and visualized in Command and Control (C2) systems?

Even if the participant was timed during the Filter test, the focus was on the qualitative data. The objective of the following Phase E was to answer the second research question through a quantitative test. The question was:

- Can AI-supported adaptivity reduce the operator's mental workload while keeping a sufficient level of SA?

10.2 Selecting Concept for Adaptivity & Workload Testing

The development of the tests and the concept to be tested was iterative, as the two interlocked. Testability was a significant factor in the choice of the solution area to represent adaptivity in the second test round. The solution area deemed to be most suitable was *Filter*. As shown in the previous test round, *Filter* was compatible with quantifiable tests, which was preferred to answer the remaining research question.

The chosen variant of *Filter* was an improved variant of Static. From the result of the previous concept evaluation, the time did not differ enough to serve as a decision basis. Therefore, the

choice was made based on opinions and reflections about the four alternatives. Static had too small a difference between prioritized and normal targets. Dynamic was not popular in its demonstrated form with a surrounding revolving indication, but movement was not bad per se. Pulsation was mentioned as an alternative. Map Change did not get the intended effect, since no participant realized that it was connected to the map. It was interpreted as connected to the target. Tunnel Vision took away too much of the Situational Display, which was deemed to affect the SA too much in a negative direction.

Initially, the conclusion was to use a variant of Static, where the size and opacity differences were increased, combined with a pulsating movement on the prioritized targets. However, when testing the pulsating target in the prototype, two conclusions were drawn. It could be overwhelming when many targets would be prioritized, as well as during longer time periods. Additionally, it would require extensive prototyping, since the program is limited when it comes to cyclic movements. Therefore, only the improved version of Static, with increased size and opacity difference, was used in the second test round.

10.3 Planning & Execution

As stated in the objectives of this phase, the goal of the second test round was to investigate whether adaptivity can reduce the operator's mental workload while keeping sufficient SA. To test this, the formation of the actual test was crucial to get relevant results. Therefore, the choice of which solution areas to test also depended on the testability. To get the sought-after insight, the test had to include:

- Different levels of mental workload.
- Different levels of adaptivity.
- A measurable task for the test participant.
- A way to judge how SA is affected.

One way to balance these four aspects and keep down the number of variables was to divide the test into two parallel parts. These parts were a main and a subsidiary task, where the purpose of the main task was to perform a C2-related task, while the purpose of the subsidiary task was to alter the mental workload.

10.3.1 Main Task

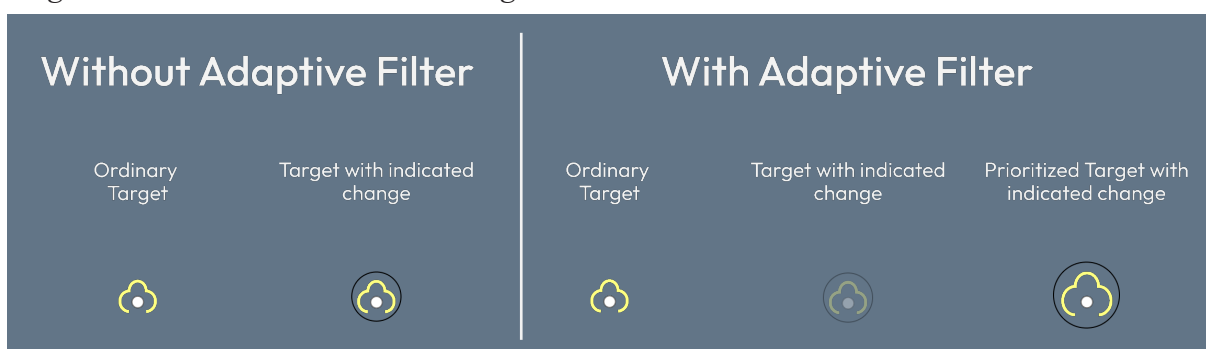
The main task is comparable to "Whac-A-Mole", but in a C2 context. The concept prototypes used for the main task were created in Adobe XD, the same program used for the first test round concepts. Each task round was one minute long, and this minute is divided into six ten-second segments. During each of these segments, five random targets out of 150 were marked with a black ring, as an indication that something about that target is interesting. The black ring was chosen to avoid interference with operators' associations, for instance, a visual identity change from unknown to hostile. These five indicated targets should, similar to the Filter evaluation in the first test round, be clicked on by the participant. In the next ten-second segment, the rings appeared around five other random targets, and this continued for one minute. To make sure the participant understood when the next ten seconds began, the screen flashed. In the end, the participant had a final score out of a maximum of 30 for each

task round. Due to technical limitations in Adobe XD, it was not possible to make the targets crossed out after being clicked, so the participants had to know for themselves which of the five they had clicked on.

This was done in two variations: one without and one with adaptivity. How the targets looked is shown in Figure 19 for both variations. Without the adaptivity, all targets in the map were shown equally, except for the five marked with black rings. With the adaptivity, three out of the five marked targets were highlighted through the filter function. The other two marked targets and their black rings were filtered down with lower opacity, like the rest of the filtered targets.

Figure 19

Target With & Without Indicated Change



Note. The Figure is depicted with symbols of unknow air target. It is applied similarly to unknown surface targets in the test.

The reason that only three of the five interesting targets were highlighted by the filter was to see if the adaptive filter could keep sufficient SA while reducing mental workload, despite also filtering away interesting information. To let the filter show all five interesting targets and compare it to the same situation without a filter would have a too predictable and weak hypothesis. Instead, the test aimed to address the conflict with all filters: is it beneficial even if it also filters out relevant information? In a realistic C2 context, this would probably be the case. The amount of relevant information risk being too much for the cognitive capacity of the operator; could the total result therefore improve even if it filters too much? Important to add here is that filtered information is not removed information, but just lowered in opacity.

In conclusion, the main task is the primary measurable task for the operator and encompasses the variable of adaptivity (with and without an adaptive filter). Also, by filtering out “too many” of the interesting targets, the effect on SA is measured.

Figure 20

Example of a Situation Map Without Adaptivity, as used in Main Task

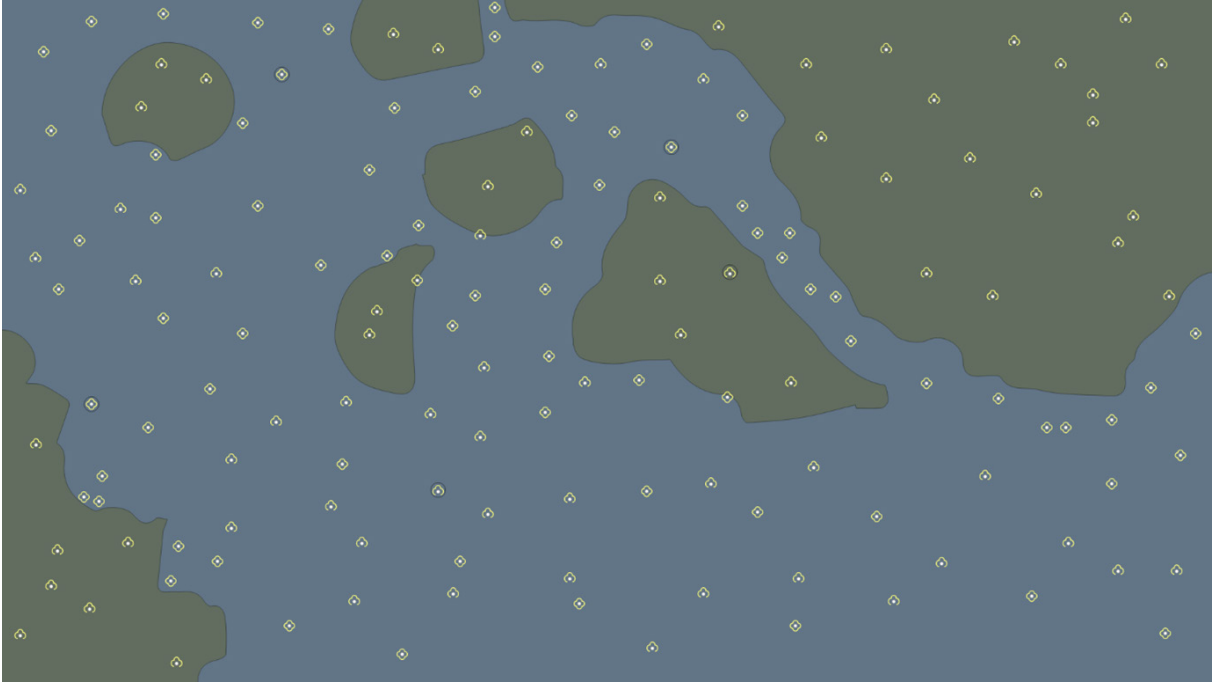


Figure 21

Example of a Situation Map With Adaptivity, as used in Main Task

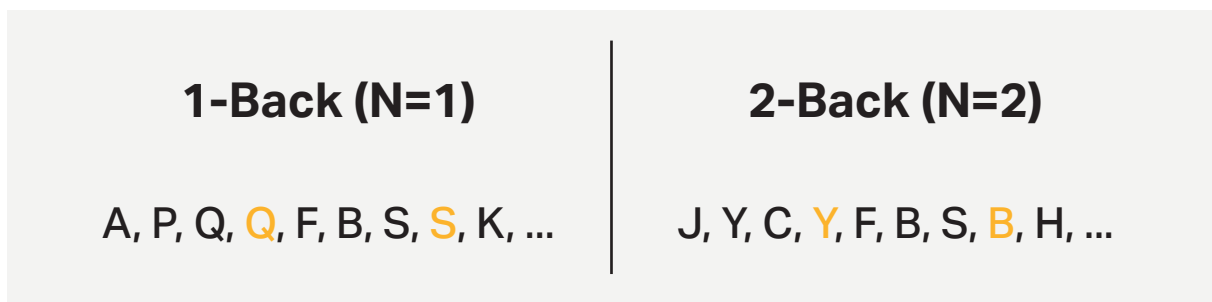


10.3.2 Subsidiary Task

The subsidiary task addressed the need for different levels of mental workload and thereby completed the previously mentioned list of required qualities of the second test round. It simulated an accumulated workload that is a combination of factors such as task complexity, stress level, and fatigue. An n-back task was used for this purpose. To keep a low level of interaction with the visual primary task, it used a different perceptual modality, being auditory, thus not competing with the same sensory channel. The sequence consisted of 30 letters and matched the length of each task round, which was one minute. Three levels of cognitive workload were applied on both adaptivity levels: no n-back task as a baseline, $n=1$ and $n=2$. Additionally, to establish a baseline for n-back and to help the participant understand it before it was combined with the main task, one round each of only 1-back and 2-back was performed. Figure 22 provides a visual explanation of the 1-back and 2-back tasks.

Figure 22

Examples of N-Back Sequences



Note. Yellow letters indicate where the matches occur.

To form the N-back, ChatGPT (OpenAI, 2025) was used to get an initial random sequence of 30 letters. Then, it was manually manipulated to add six matches in a 1-back and 2-back. From these originals, two more on each level were created by moving each letter one step forward in the alphabet. These three sequences were given back to ChatGPT, with the prompt to make them into one-minute-long sound files, where the letters are read every other second in an even pace. The result was three versions of 1-back and three versions of 2-back, altered enough not to be recognized by matching letters, but not so much that the sequences differ in difficulty. The sequences are presented in Appendix H.

Even if the focus lay on the score of the main task, the score of the subsidiary task was kept as well. Keeping it helped to analyse the results and gain an understanding of whether a participant, for instance, loses control over one of them while scoring well on the other. However, since the scores in the two tasks may not weigh the same (e.g., one point in the main task may be easier to get than one point in the subsidiary task), and the maximum score differs, the different task scores were kept separate.

10.3.3 Total Task Round Variants

In total, there were eight task round variants of the test, listed below. NA stands for No Adaptivity and AD stands for Adaptivity. 0, 1, and 2 represent the mental workload levels, with no workload (n=0), n=1, and n=2.

- Only 1-back
- Only 2-back
- NA-0
- NA-1
- NA-2
- AD-0
- AD-1
- AD-2

As in the concept evaluation of *Filter*, this test was within-subject to get as many results as possible on each task variant, while avoiding the effect of participant variables. To counterbalance the learning effect, half of the test participants started with no adaptivity, while the other half started with adaptivity. The workload was gradually increased in different zigzag patterns between NA and AD to even out the learning effect between them. The different zigzag patterns were mirrored, so the participant performed them in pairs. Table 3 depicts the different orders in the test variants.

Table 3
Performed Test Order for Each Participant

Participant	Order							
	Only n-back		0-back		1-back		2-back	
	<i>1st</i>	<i>2nd</i>	<i>3rd</i>	<i>4th</i>	<i>5th</i>	<i>6th</i>	<i>7th</i>	<i>8th</i>
1	1-back	2-back	NA-0	AD-0	NA-1	AD-1	NA-2	AD-2
2	1-back	2-back	AD-0	NA-0	AD-1	NA-1	AD-2	NA-2
3	1-back	2-back	NA-0	AD-0	AD-1	NA-1	AD-2	NA-2
4	1-back	2-back	AD-0	NA-0	NA-1	AD-1	NA-2	AD-2
5	1-back	2-back	NA-0	AD-0	AD-1	NA-1	NA-2	AD-2
6	1-back	2-back	AD-0	NA-0	NA-1	AD-1	AD-2	NA-2

10.3.4 Self-Assessment Questions

After each round, the participants were asked two self-assessing questions with the aim of quantifying their perceived performance and control. The questions were answered on a scale of 1-10, and were as follows:

- How do you judge your performance, only based on the result?
- How do you judge your control over the situation?

These two questions were chosen to investigate how their self-judgment correlates with the actual result and whether the feeling of control differs from this. For example, can they feel like they had no control, yet still think they performed well, or vice versa?

10.3.5 Participants

The study was planned for six participants to perform the tests. All were Saab employees who possess expert knowledge of respective C2 system, as well as having some sort of military operational experience in C2 systems. They represented both air and ground radar C2 systems.

Due to unforeseen circumstances, one of the planned participants had to cancel their participation at the last minute. In order to keep to the timetable and perform the tests as planned, a Human Factor engineer, familiar with the C2 system, replaced this participant.

Table 4

Participants of the Test and the Domain of C2 They are Experienced With

Participant	C2-system
1	Surface
2	AEW
3	AEW
4	AEW
5	Human Factor
6	Surface & AEW

10.3.6 Conducting the Tests

To perform the test, an external 24-inch desktop screen was placed on a table in a closed, quiet room. The screen was used to resemble the size and position of a system monitor. The test participant used a conventional, external computer mouse for the same reason.

The two test leaders started the tests by giving instructions on how the tests would be performed (see Appendix G). One recurring doubt from the operators during the first round of testing was the AIs' capability. Despite being initially and repeatedly informed during the tests that the AI was assumed to be correct, the operators had often questioned what the AI based

its decisions on. Therefore, as part of the briefing for the second round of testing, the operators were informed about what the AI made the decisions on (based on the answers on the question “What information do you search for when evaluating a target?” from the previous phase) to avoid this doubt in the second round of tests. This information was a simplification, and it had not yet been decided exactly what the AI should base its prioritization on at this point.

During the tests, the test leaders kept track of the scores. One focused on the main task, and the other on the subsidiary task. The screen was recorded to document for potential deeper analysis later on, or to identify any potential errors in the scorekeeping. The sound was recorded together with the screen to hear what was said and how the subsidiary task was performed. After each round, the two self-assessment questions mentioned previously were asked.

10.3.7 Analysing the Test Results

When all tests were completed, the results from the main task, the subsidiary task and all assessment questions were compiled separately, both within categories and between participants. They were then analysed in different combinations until a clear visualization of the results could be seen.

10.4 Result

From all six test participants, patterns in the results could be seen in the raw data. They are presented in this part.

10.4.1 Main Task Result

In Table 5, the highest scored between the non-adaptive and adaptive filter versions of the main task are presented, for the different n-back workloads.

Table 5

Main Task Result

Participant	0-back (No n-back)		1-back		2-back	
	NA-0	AD-0	NA-1	AD-1	NA-2	AD-2
1 (Older)		Yellow	Yellow		Yellow	
2 (Older)		Yellow		Yellow		Yellow
3 (Older)		Yellow		Yellow		Yellow
4 (Older)		Yellow		Yellow		Yellow
5 (Younger)	Yellow		Yellow		Yellow	
6 (Younger)	Yellow		Yellow		Yellow	

Note. Yellow rectangle indicates which version the participant performed better on. If both are yellow, the participant scored a tie.

As shown in the table, without n-back, four out of six participants scored higher with adaptivity. With 1-back, three scored better without it, and three with it. With 2-back, four scored better with adaptivity, one without adaptivity, and one scored a tie. In total, adaptivity was scored higher on. The two youngest scored better without adaptivity in the tests without n-back and with 1-back, indicating that age may play a role in how adaptivity supports the operator. It could also be that chance played a role. This should be investigated in further research.

The benefit of the adaptive filter was greatest when the workload rose to 2-back. A pattern is seen for adaptivity to first be better, then worse and then significantly better than without adaptivity for increasing workload. This indicates that the test needs to be fine-tuned for future iterations and performed with more participants to determine if the patterns are correct or a coincidence. To reduce the big jump in learning effect after the absolute first round for each participant, the participants could have been handed a warm-up round to familiarize themselves and learn tactics.

The prioritized targets in the adaptivity rounds were almost scored full on by all participants. The ones who missed only missed one, and the missed one was in the first round of adaptivity with no workload, when the tactics were still figured out. In the rounds with added workload, no prioritized target was missed at all, by any of the participants. This result shows that even with a high mental workload, the targets deemed most important by the AI will be registered by the operator. The way this prioritization is visualized, facilitates bottom-up processing through captured attention from the ambient vision, while the filtered away and non-adaptive targets need deliberate directed attention from the participant. However, the non-prioritized targets with a black ring in rounds with adaptivity were harder to find than black ringed targets without adaptivity.

10.4.2 Subsidiary Task Results

In Table 6, the different n-back results are presented, including both the round with only n-back tests and the rounds with 1-back and 2-back tests combined with the main task.

Table 6

Subsidiary Task Result

Participant	Only n-back		1-back		2-back	
	1-back	2-back	NA-1	AD-1	NA-2	AD-2
1 (Older)	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow
2 (Older)	Yellow	Yellow	White	Yellow	Yellow	Yellow
3 (Older)	Yellow	White	Yellow	Yellow	White	Yellow
4 (Older)	Yellow	White	Yellow	Yellow	Yellow	White
5 (Younger)	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow
6 (Younger)	Yellow	White	Yellow	White	Yellow	White

Note. Yellow rectangle indicates which version the participant performed better on. If both are yellow, the participant scored a tie.

Unsurprisingly, each participant scored better on “Only 1-back” than “Only 2-back”, or a tie (full score). For the 1-back and 2-back tests that were performed simultaneously with the main task, it can be seen that the results of the n-back correlate with the results between non-adaptive and adaptive rounds in Table 5. The n-back score is better or tied for each winner of NA vs. AD in Table 5, with only one exception. On the 2-back workload level, Participant 4 scored better with adaptivity on the main task but better without adaptivity on the side task. Concluded, this result indicates that the main task and subsidiary task scores correlate and do not contradict the total score of the two, simplifying the following analysis of the results. However, to reduce the number of ties on the 1-back and 2-back workload levels to strengthen this correlating result further, the maximum score on the n-back tasks should have been higher than six.

10.4.3 Self-Assessments

Here, the result from the self-assessments is presented and connected to the actual performances.

Only Subsidiary Task

For the rounds with only n-back tests, one clear result was visible for all the participants. As seen in Table 7, for all participants, the assessed percental decrease in performance between “Only 1-back” and “Only 2-back” was worse than the actual percental decrease in performance. Furthermore, the percental decrease in assessed control was even bigger or the same as the decrease in assessed performance. This shows that the negative feeling about the increased mental workload did not correlate proportionately with the actual result.

Hence, from a user satisfaction perspective, just measuring the performance when the mental workload increases is not enough. This negative feeling, which is over-proportioned compared to the actual result and performance, is also important to address. However, it is important that the assessed result and control never exceed the actual performance. This could have severe consequences if an operator became overconfident and considered themselves falsely complete with a mission.

Table 7

Percentual Decrease of Performance & Control

Participant	Actual decrease of performance from “Only 1-back” to “Only 2-back” in %. (Max score was 6)	Self-assessed decrease of performance from “Only 1-back” to “Only 2-back” in %. (Max score was 10)	Self-assessed decrease of control from “Only 1-back” to “Only 2-back” in %. (Max score was 10)
1 (Older)	0	20	20
2 (Older)	0	20	20
3 (Older)	17	40	60
4 (Older)	17	22	44
5 (Younger)	0	70	80
6 (Younger)	17	44	44

Only Main Task

From the rounds with only the main task, presented in Table 8, it was shown that the self-assessed performance on non-adaptivity versus adaptivity aligns with the actual performance, for all participants except participant 4. That is, five out of six participants assessed their best performance the same as their actual best performance between NA-0 and AD-0. But for the self-assessed control, all six thought that their control was higher or the same without adaptivity than with it. Concluded, for the rounds with only the main task, four out of six performed better with the adaptive filter, and three out of those four thought that they did so. But all six thought that the control was higher without the adaptive filter. This contradiction is not present for participants 5 and 6. Their actual performance, assessed performance, and control align fully.

Table 8

Result of Main Task Without N-Back

Participant	Actual performance		Self-assessed performance		Self-assessed control	
	1-back	2-back	NA-1	AD-1	NA-2	AD-2
1 (Older)		Yellow		Yellow	Yellow	
2 (Older)		Yellow		Yellow	Yellow	
3 (Older)		Yellow		Yellow	Yellow	
4 (Older)		Yellow	Yellow		Yellow	
5 (Younger)	Yellow		Yellow		Yellow	
6 (Younger)	Yellow		Yellow		Yellow	

Note. Yellow rectangle indicates which of non-adaptive and adaptive the participant actually performed better on, assessed the best performance, and assessed the highest control. If both are yellow, it was tied.

Main Task and 1-Back

The results are presented in Table 9. When adding the 1-back test to the main task, things changed from without performing a side task. Which version the participants actually performed better on only changed for participant one. However, the patterns in self-assessed performance and control are not as clear and aligned, indicating that this is a turning point for the assessments.

Table 9

Result of Main Task With 1-Back

Participant	Actual performance		Self-assessed performance		Self-assessed control	
	1-back	2-back	NA-1	AD-1	NA-2	AD-2
1 (Older)	Yellow		Yellow		Yellow	
2 (Older)		Yellow		Yellow		Yellow
3 (Older)		Yellow	Yellow		Yellow	
4 (Older)		Yellow		Yellow		Yellow
5 (Younger)	Yellow		Yellow		Yellow	
6 (Younger)	Yellow			Yellow		Yellow

Note. Yellow rectangle indicates which of non-adaptive and adaptive the participant actually performed better on, assessed the best performance, and assessed the highest control. If both are yellow, it was tied.

Main Task and 2-Back

The results are presented in Table 10. When the workload was increased to 2-back, things really started to change. The actual performance leaned clearly in favour towards the adaptive filter. Also, their assessed performance starts to align with this. All participants judged their performance to be tied or better with adaptivity. For the assessed control, there is a clear advantage towards the adaptive filter.

Table 10

Result of Main Task With 2-Back

Participant	Actual performance		Self-assessed performance		Self-assessed control	
	1-back	2-back	NA-1	AD-1	NA-2	AD-2
1 (Older)	Yellow		Yellow		Yellow	
2 (Older)		Yellow	Yellow			Yellow
3 (Older)		Yellow		Yellow		Yellow
4 (Older)		Yellow		Yellow		Yellow
5 (Younger)		Yellow	Yellow			Yellow
6 (Younger)	Yellow		Yellow			Yellow

Note. Yellow rectangle indicates which of non-adaptive and adaptive the participant actually performed better on, assessed the best performance, and assessed the highest control. If both are yellow, it was tied.

10.4.4 Concluded Result – Adaptivity Testing

With no workload, the participants performed in favour of adaptivity, and five out of six assessed their performance in line with this. However, they indicated feelings that they had more control without adaptivity.

With a 1-back workload, an equal amount performed better with and without adaptivity, and their assessments started to shift. When the workload rose to 2-back, the performance was in favour of adaptivity. The assessed performance started to shift in that direction, with many ties between adaptive and non-adaptive assessed performance. The biggest difference, however, was the assessed control. This shows that the adaptive filter gives both improved results and a better feeling of control during high workloads, which is important for user satisfaction. If AI-supported, high transparency and low error margin in the filter decisions can increase this user satisfaction. It is important to remember that the assessed control in this test comes with the precondition that the AI does not make wrong decisions. If it did, the filter would probably not increase the feeling of control with a higher mental workload.

The results are clearest for high workload, where an adaptive filter is beneficial for both performance and user satisfaction compared to having no adaptive filter. It is also clear that the assessed control is higher without an adaptive filter when there is no added mental workload.

10.5 Summary of Phase

In this phase, a concept to test was selected iteratively with the test development. The test aimed to answer the second research question and evaluate adaptivity combined with different workloads and the effect on SA. The test was developed for the “Static” *Filter* and performed with six participants. The results showed that the adaptive filter has a positive effect on both performance and experienced control, compared to not having an adaptive filter.

11

Design Guidelines

In this chapter, the results from Phases D and E will be compiled and concluded in a shorter version. If and how the results answer the research questions will be discussed.

11.1 Research Question 1

The first research question was as follows:

- How can AI-supported adaptivity be implemented and visualized in Command and Control (C2) systems?

Three concept areas were developed, prototyped, and evaluated as solutions on how AI-supported adaptivity can be implemented in C2 systems. They were *Filter*, *Target Prio-list* and *View*.

11.1.1 Filter

Adaptive filtering of prioritized targets is one way to visualize adaptivity. It emphasizes important targets and makes them easier to find through captured attention by increasing the salience of the stimuli. This aids the operator during the Surveillance & Assessment. One efficient way to do it is the concept “Static”, which emphasizes without completely removing information. Prioritized targets have full opacity and are larger than ordinary ones, which are reduced in opacity. The prioritization is based on factors affecting the identification work, e.g., mission, geographical context, identification signal, etc., and changes with the new information.

11.1.2 Target Prio-list

A second way to implement and visualize AI-supported adaptivity in C2 systems is to place the target list on the right side of the screen, make it intelligent and adaptive by prioritizing targets in the list from the top down in descending order of importance. The prioritization is made based on the same factors as for *Filter*. This supports the Surveillance & Assessment work in the way that the operator always knows that the highest prioritized target (as deemed by the AI) is located at the top of this list. When clicking on it, it gets emphasized in the

map, so the connection between AI prioritization and geographical context is clear. This list is adaptive in the way that the re-prioritization is made at a suitable frequency. This frequency has not been decided and requires further investigation. Concluded, this change helps the identification work for the operators, by constantly providing them with a suggested prioritization to evaluate on the map.

11.1.3 View

A third way to implement and visualize AI-supported adaptivity in C2 systems is to present information in different windows, to support the operator in perceiving and assessing both the overview and details simultaneously. The results show that dividing the screen into four sections, where the top left section is an overview, where operators can see the entire map, is a promising way to do it. The overview window is meant to be interacted with. In the other three windows, different adaptive functions are chosen by the AI and presented. The functions can be, for instance, to show a target of interest (regardless of whether it leaves the overview), show an enlarged version of an area of interest, or provide other support functions or information deemed to be relevant at the time. If connected to the map, the windows are colour-coded to the map section they depict. However, making the position of the windows themselves adaptive out of user control risks confusing the user. So does showing too many changes at the same time, for example, if one window changes the target of interest, another changes the area of interest, and the third provides a new support function. One change of window information at a time is preferred. If more than one happens simultaneously, it can result in the operator not understanding any of them. Then, the AI acts counterproductively, causing an overload by providing too much information. A solution is to perform the changes in sequence.

11.2 Research Question 2

The second research question was as follows:

- Can AI-supported adaptivity reduce the operator's mental workload while keeping a sufficient level of SA?

This question was aimed to be answered through the Adaptivity & Workload test, where *Filter* was tested in a more quantitative way. For a high mental workload (2-back), the result indicated that adaptivity helps the operator score better than without adaptivity. It was also shown that with a high mental workload, the operators felt that they had more control with an adaptive filter than without. Even if a higher score is not directly translated to a lower mental workload, it can be said that adaptivity, in the form of a filter, helps the operator perform better and feel more in control under high mental workload. Additionally, the score shows that the SA, in the form of the number of found targets, was generally higher with a filter, despite the fact that the two "hidden" targets in the adaptive rounds were harder to find than targets in the non-filter rounds. This indicates that, especially under high workload, an adaptive filter is beneficial for SA despite also hiding relevant information. Two factors that may affect this are the duration of time the work is performed and the operator's age. Since fatigue affects task performance, adaptive filters probably improve the operator's performance more over time. The filter seemed to help younger participants less than the older ones, which need to be investigated further. Including other factors, such as e.g., gender, which this test did not take into consideration, would also be beneficial in future investigations.

One important result here is that, if possible, information that is less important should not be filtered away completely. By toning it down instead of removing, it was still possible for the operators to find more than the three prioritized in the adaptive rounds, if they had the time and capacity. If only the prioritized targets remained on the map, the SA would be very limited. Now, the operator can choose to go further than the captured attention and direct their attention if there is time and mental capacity for it.

11.3 Transparency and AI Decision Basis

Correlating to both research questions, AI transparency is crucial for AI-supported adaptivity to be beneficial. This is twofold. The first type of transparency is from a human factors perspective. Visual clues about adaptive changes have to be given to the operator, so that they can perceive and understand both that something is happening and what is happening. For example, when a prioritization is updated in *Filter*, *Target Prio-list* or *View*, it has to be indicated when it happens and what happens. In the Adaptivity & Workload test, this was visualized with the screen blinking between the ten-second intervals. Without this indication, the participants would risk missing when they should look for new targets. Preferably, this transparency would also show what had happened after the change, so that changes can be interpreted also after the actual change is performed. For example, if the *Target Prio-list* blink when the list order changes, the operator would not know what changes were made. But if, for example, arrows with a number indicated how many steps up or down a target took, they would understand what had happened.

The other type of important AI transparency is how it is defined, for example, by the European Union, that AI outputs should be transparent about how the decisions were arrived at. This type helps the operator to understand and trust the AI decisions, which was a recurring concern through initial interviews and the Concept Evaluation. The suggested decision basis for the AI is listed below:

- Mission type
- Target type (surface/air)
- Geographical context
- Personal responsibility area
- Signal of identification
- Target behaviour
- Altitude
- Velocity
- Direction
- Target number
- Distance
- Radar target area

Furthermore, the update frequency for the adaptivity should be an input for the AI, as it is in the model used in the experiment by Zak et al. (2023), presented in section 3.3.4.

12

Discussion

In this chapter, the thesis is discussed as a whole. The aim, goal, strengths and weaknesses in the process and execution, method choices, suggestions for future work, and ethical aspects are all included.

12.1 Aim

The project aimed to investigate how AI-supported adaptivity can be implemented in Command and Control systems, including how it should be visualized for varying workload, to increase efficiency and reduce the operator's mental workload.

The aim is deemed to have been met fairly well. Through research and literature studies, knowledge has been gathered on areas such as AI, adaptivity, human cognition, and other relevant topics. This knowledge has supported the explorative Research through Design process, in which the implementation of AI-supported adaptivity in C2 systems has been investigated, and visual concepts have been created. However, the project has not investigated different visualizations for varying workload. Instead, only one adaptive visualization and its performance under varying workload have been studied, in comparison to a non-adaptive version.

By measuring the operator's performance with the increased workload, the project has managed to gain insight into how efficiency and mental workload are affected by the suggested adaptive visualization.

12.2 Goal

The goal of the project was to develop design proposals in the form of visual prototypes on how AI-supported adaptivity can be implemented in C2 systems, and to test how these affect the mental workload and SA.

The project delivered the said design proposals, however, not with a functioning AI supporting it. It was tested how SA was affected. As mentioned under Aim, the mental workload was not measured in itself. However, measuring the operator's performance with increased workload gave insight into how it was affected.

12.3 Process and Methods

In this part, the chosen process and execution are discussed, including whether the right decisions were taken to answer the research questions.

12.3.1 General Procedure

The project was, as described, an explorative study utilizing the research through design method. By conducting a design process, including, for example, user and literature studies, concept development, and evaluations, the initial aim and goal of the project were, as mentioned, largely met and delivered.

As with many design processes, one could, in hindsight, reflect on how time could have been spent more efficiently. It is easy to think that time spent on ideas and thoughts not directly visible in the final result could have been used better. However, these ideas and thoughts are part of the nature of iterative design processes and also contribute to the final result. In an explorative project, the final result is not possible to decide from the start. It has to be tested and experimented towards. Therefore, all ideas, thoughts and results are part of the final result, since the final result is also what did not work, and how it had to be done instead. Hence, even though it is easy to reflect like this, the time was spent efficiently during the project, since all the work contributed to progress in the iterative process.

The overall structure and planning of the project worked well. Multiple factors, some of which have already been discussed, including limited access to experienced operators, limitations with the prototyping software, etc., contributed to a restriction in time towards the end. This mainly influenced the report writing and the final phase (see section 12.3.5 for more details). The following sections include discussions in more detail regarding each individual phase of the project.

12.3.2 Phase A – Initial User Research

During this phase, information was gathered in order to gain an understanding of the context of the thesis. The methods and decisions taken here were conventional, with user interviews and literature studies. However, the main improvement here would be to interview a larger variety of C2 operators, and have a bigger variety in ages, genders, etc. The positive aspect was that operators of different C2 systems were interviewed to understand C2 more generally. Only interviewing AEW&C or Surface radar operators would not provide a general understanding, since they showed to have some substantial differences. Additionally, it would improve the result to interview one more operator of surface-based C2 systems, since only two of the five participants were surface operators. Interviewing operators with active or more recent experience would also strengthen the result. More information could be gathered in this phase, e.g., through system observation; however, it was intentionally left out not to have biases entering the next phase.

12.3.3 Phase B – Idea Generation

The purpose of actively avoiding biases when entering ideation proved beneficial. Enough ideas about existing functions and solutions were present as it was. The ideation was conducted and iterated until a sufficient saturation was reached. Then, the ideas had to be investigated and eliminated for new ideas and concepts to form.

The delimitation of the scope of the project, which followed the Idea Generation, was necessary in order to achieve a final result on a desired level, within the available time and demands of the thesis. As previously mentioned, the scope had been deliberately kept wide before this, in order to explore the solution space and not be too restricted in which path the project would follow. Why the particular delimitations were made is motivated (see section 7.4). However, just because the project followed down this particular path, it does not mean that the other alternative routes and possible solution areas which were excluded do not have potential.

12.3.4 Phase C & D – Concept Creation & Concept Evaluation

In Phase C, when developing concepts, the decision not to make every concept “perfect” had both ups and downsides. To take *Target Prio-list* as an example, it deliberately only focused on one function in each of the four concepts. The upside was that time was saved on developing for lower fidelity and getting more isolated answers regarding the focused function. On the downside, this resulted in the operators, during the Concept Evaluation in Phase D, sometimes focusing on the most annoying lack of function. In the *Target Prio-list* case, it was that three of the four concepts did not sort from top down in descending level of priority. Later, during the analysis, this resulted in all concepts having more negative feedback than positive. This could make the analysis require more afterthought and interpretation, since no concept scored obviously best of the four. The result of the Concept Evaluation is an analysis of how to combine and improve the ideas. In this type of analysis, the KJ-method showed to be very suitable. However, despite confidence in this method, there was no quantitative data to show. This could make it hard to motivate stakeholders as to why a certain decision was made.

Furthermore, the discussion about participant diversity held during the Initial User Interviews (see section 12.3.2) also applies to the evaluations in Phase D.

12.3.5 Phase E – Adaptivity & Workload Testing

During Phase E, determining which concept would be tested was heavily dependent on how it was possible to test it. Due to the available time and technical limitations of the prototyping software, the test had to be formed in a narrow and niche way. During these circumstances, the test had the potential to give insightful results. However, the Adaptivity & Workload Test only tested one *Filter* concept, “Static”. Therefore, the total project results for *Filter* were more rigid than those for *Target Prio-list* and *View*. It would have been interesting to find a way to test these two in a quantitative way as well, and maybe compare *Filter* to *Target Prio-list*, to see if the latter adds value and improves the operators’ work process. Or if they could be combined.

Regarding the results from the Adaptivity & Workload Test, they were not as insightful with clear patterns over all the workloads, as hoped. It was evident that the adaptive filter helped both the result and the assessed control when the workload level was 2-back. Additionally, it was indicated that adaptivity seemed to help older participants more. A thought beforehand was that age and experience would perhaps affect the total score, but not the advantage or disadvantage between non-adaptive and adaptive. This could, however, be coincidental or have other explanations than age. This, combined with the unclear pattern observed without n-back and with 1-back, between adaptivity and non-adaptivity, can also be a sign that more data needs to be gathered. This was challenging due to the limited number of available test participants with an operational background. So was accounting for gender and age as factors.

As previously mentioned, one operator had to cancel their participation in the test at the last minute and had to be replaced (participant 5). The replacement was a young Human Factors engineer. Since the test was purely cognitive and did not include any assessment, this lack of operational experience was probably not a disadvantage. On the contrary, the lack of operational experience could have been to an advantage, since no pre-knowledge or search patterns affected the ability to find a suitable tactic for the tests.

Furthermore, the data could have been increased within the tests as well. The number of marked targets per ten seconds could be increased from five to, for example, ten, and made slightly more visible to compensate. If so, the possibility of getting them marked as clicked is more important, since keeping the information about which have been interacted with or not takes up mental resources and limits the possibility to multitask. The n-back task could have been constructed with more than six matches to make differences in score more statistically certain. Furthermore, it could have been investigated if there was a way to make a weighted combination of the total test score between the main and subsidiary tasks. That would have helped the analysis of the total performance.

12.4 Ethical Considerations

This thesis topic contributes to the defence industry and products used in peace, crisis and war. Hence, ethical considerations are of great importance when it comes to the implementation of AI in this context. The human must remain accountable and responsible, where transparency and explainability of AI reasoning and decisions are key. The user must also be able to override the AI if it is believed to be wrong.

The focus of this thesis is on enhancing usability and efficiency for the operator while reducing the mental workload. This makes the work environment more ergonomic and pleasant, which can be considered a positive effect.

Furthermore, there are personal data aspects to consider if the adaptive AI would include and consider human factors like stress level, eye movement, pulse, etc., of the operator in its decision basis. Such personal data must be handled securely and ethically, with operators informed about the usage and storage, and their consent obtained. This aligns with GDPR principles, contributing, for example, to their aim of lawfulness, transparency, integrity, and accountability.

12.5 Operator & AI Interaction in C2 Systems

The usage of AI to support the users has been evaluated and discussed in this project through the visualized adaptive concepts. However, some general insights into the implementation of AI have also been gathered.

The most important of these insights might arguably be transparency. Firstly, in regard to being transparent about *what* the AI does or has done, which affects the user. For instance, informing that it has visually altered the information presented to the user on the display, such as making visual priorities in the *Filter* concepts. Secondly, transparency in regard to *why* the AI has done it, meaning how the AI works and what it bases its decisions on. Both aspects should be addressed and taken into consideration when designing a system with AI and operator interaction, especially one where safety and reliability are of great concern.

This thesis has investigated and presented some data that the operators would like the AI to include in its decision-making. This could act as guidelines for the development of the functioning AI model. Additional information, such as human factors (e.g., measured workload), could also be of value, which is supported by the discussion made by Zak et al. (2023), as presented in section 3.3.4.

12.6 Future Work

For future studies, there are a number of things to investigate further, with a few already being mentioned. To begin with, *Target Prio-list* and *View* need further investigation on different levels. The *Target Prio-list* has more set frames and, therefore, has come further in its development. It should be prototyped and tested in a similar quantitative way as *Filter*, and also be compared and potentially combined with *Filter* to see if it is a supporting adaptive function that adds value. *View*, however, does not have as set frames and is more complex. As of now, some conclusions have been drawn about potential gains and discarded *View* solutions. However, so far, the conclusions are that there could be potential benefits with the concept. These must be tested with more interactive prototypes in a scenario-based test. *View* is also the solution that is applicable in most situations, since its main purpose is to be able to see and access both details and the big picture at the same time. Each application must be tested separately.

As for *Filter*, the Adaptivity & Workload Test should be modified to be more balanced and conducted with a larger number of operators with a bigger variety to see if clearer patterns can emerge for no n-back and 1-back. If it does work, the fidelity and realism have to be increased step by step, with more distracting information such as areas, movement, and different kinds of targets, etc.

To evaluate the use of the three concept categories, *Filter*, *Target Prio-list*, and *View*, visual transparency of the adaptivity must be implemented on all. Otherwise, the use may be discarded with the belief that the user cannot understand what is happening. Additionally, a method for being transparent with the decision basis should be developed.

Finally, to assess if it actually works, an AI model has to be developed to see if the right prioritization can be made in a simulated scenario, to then apply the solutions to this model and validate the use of AI-supported adaptivity in a simulated scenario.

13

Conclusion

In this chapter, the main conclusions from the project are presented. They include both results, processes and future work.

- Three ways of implementing AI-supported adaptivity in C2 systems were developed. *Filter*, *Target Prio-list* and *View*. *Filter* adaptively filters targets, which are prioritized by the AI. *Target Prio-list* sorts the most important targets in an adaptive list, and *View* divides the screen into four windows, adaptively showing the presently important details and their relation to the overview.
- The results from the conducted test indicate that for a high mental workload, the operator's situational awareness and feeling of control were better with an adaptive filter than without, despite the filter "hiding" two of five important targets. For no added workload and medium workload, the results were not as clear. However, it was clear that with no added workload, participants assessed their control higher without adaptivity, which was not the case for their actual performance.
- It was concluded from qualitative data from user tests that *Target Prio-list* and *View* have the potential to help the operator in his or her work process.
- All three ways of implementing AI-supported adaptivity need further development and investigation, where *Filter* has come the furthest, followed by *Target Prio-list* and then *View*.
- The process and methods were beneficial in answering the research questions. However, the result would be strengthened and broadened if a wider variety of operators were available and chosen, in the form of operator role, gender, and age. For the Adaptivity & Workload Test, more tests should have been performed to get clearer results.
- Finally, AI transparency is deemed to be of utter importance, both for the operators to understand what is happening and to be able to trust the AI and the decisions it makes.

References

- Adams, W. (2015). Conducting semi-structured interviews. In J. Sharp, J. Peters, & K. Howard (Eds.), *The researcher's toolkit: The complete guide to practitioner research* (Chapter 19). Wiley. <https://doi.org/10.1002/9781119171386.ch19>
- Adobe Inc. (2025a). *Adobe Illustrator* (Version 29.5.1) [Computer software]. <https://www.adobe.com/products/illustrator.html>
- Adobe Inc. (2025b). *Adobe XD* (Version 57.1.12.2) [Computer software]. <https://www.adobe.com/products/xd.html>
- Brdnik, S., Heričko, T., & Šumak, B. (2022). Intelligent user interfaces and their evaluation: A systematic mapping study. *Sensors*, 22(15), 5830. <https://doi.org/10.3390/s22155830>
- Bung Tidblom, J., & Jansson, E. (2021). *Artificial Intelligence to Improve Situation Awareness for AEW&C Operators: A research through design approach to create guidelines and visual concepts* (Master's thesis, Chalmers University of Technology). <https://hdl.handle.net/20.500.12380/303948>
- Encyclopædia Britannica. (2025, April 12). Gestalt psychology. *Encyclopædia Britannica*. <https://www.britannica.com/science/Gestalt-psychology>
- 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>
- European Commission. (2019, April 8). *Ethics guidelines for trustworthy AI*. Shaping Europe's Digital Future. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- European Commission. (n.d.-a). *AI Act*. Shaping Europe's Digital Future. Retrieved January 22, 2025, from <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- European Commission. (n.d.-b). *European approach to artificial intelligence*. Shaping Europe's Digital Future. Retrieved January 22, 2025, from <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>
- Gajos, K. Z., Everitt, K., Tan, D. S., Czerwinski, M., & Weld, D. S. (2008, April). Predictability and accuracy in adaptive user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)* (pp. 1271–1274). Association for Computing Machinery. <https://doi.org/10.1145/1357054.1357252>
- GDPR.eu. (n.d). *What is GDPR, the EU's new data protection law?* <https://gdpr.eu/what-is-gdpr/>
- Gibbons, S. (2019, July 7). *Dot voting: A simple decision-making and prioritizing technique in UX*. Nielsen Norman Group. <https://www.nngroup.com/articles/dot-voting/>

- Gilmour, G., Porcelli, S., Bertaina-Anglade, V., Arce, E., Dukart, J., Hayen, A., Lobo, A., Lopez-Anton, R., Merlo Pich, E., Pemberton, D. J., Havenith, M. N., Glennon, J. C., Harel, B. T., Dawson, G., Marston, H., Kozak, R., & Serretti, A. (2019). Relating constructs of attention and working memory to social withdrawal in Alzheimer's disease and schizophrenia: Issues regarding paradigm selection. *Neuroscience & Biobehavioral Reviews*, *97*, 47–69. <https://doi.org/10.1016/j.neubiorev.2018.09.025>
- Gonçalves, T. G., Kolski, C., de Oliveira, K. M., Travassos, G. H., & Strugeon, E. G. L. (2019, December). A systematic literature review on intelligent user interfaces: Preliminary results. In *Adjunct Proceedings of the 31st Conference on l'Interaction Homme-Machine* (pp. 1-8). <https://doi.org/10.1145/3366551.337034>
- Google. (n.d.). *Crazy 8's*. Design Sprint Kit. <https://designsprintkit.withgoogle.com/methodology/phase3-sketch/crazy-8s>
- IBM. (2023, March 29). *What is explainable AI?*. <https://www.ibm.com/think/topics/explainable-ai>
- IBM. (2024, August 9). *Artificial intelligence (AI)*. <https://www.ibm.com/think/topics/artificial-intelligence>
- IEA. (n.d.). *What Is Ergonomics (HFE)?*. Retrieved May 14, 2025, from <https://iea.cc/about/what-is-ergonomics/>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Kaplan, K. (2021, August 15). *10 usability heuristics applied to complex applications*. Nielsen Norman Group. <https://www.nngroup.com/articles/usability-heuristics-complex-applications/>
- Kosinski, M. (2024, October 29). *What is black box AI?* IBM. <https://www.ibm.com/think/topics/black-box-ai>
- Lavie, T., & Meyer, J. (2010). Benefits and costs of adaptive user interfaces. *International Journal of Human-Computer Studies*, *68*(8), 508-524. <https://doi.org/10.1016/j.ijhcs.2010.01.004>
- Lim, Y.-K., Stolterman, E., & Tenenberg, J. (2008). The anatomy of prototypes: Prototypes as filters, prototypes as manifestations of design ideas. *ACM Transactions on Computer-Human Interaction*, *15*(2), Article 7. <https://doi.org/10.1145/1375761.1375762>
- Lindwell, W., Holden, K., & Butler, J. (2010). *Universal principles of design: 125 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design* (Rev. & updated 2nd ed.). Quarto Publishing Group USA.
- Litcanu, M., Prostean, O., Oros, C., & Mnerie, A. V. (2015). Brain-writing vs. brainstorming: Case study for power engineering education. *Procedia – Social and Behavioral Sciences*, *191*, 387-390. <https://doi.org/10.1016/j.sbspro.2015.04.452>

- Lucero, A. (2015). Using affinity diagrams to evaluate interactive prototypes. In J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, & M. Winckler (Eds.), *Human-Computer Interaction – INTERACT 2015* (Lecture Notes in Computer Science, Vol. 9297, pp. 231–248). Springer, Cham. https://doi.org/10.1007/978-3-319-22668-2_19
- Militello, L. G., & Hutton, R. J. B. (1998). Applied cognitive task analysis (ACTA): A practitioner’s toolkit for understanding cognitive task demands. *Ergonomics*, *41*(11), 1618–1641. <https://doi.org/10.1080/001401398186108>
- Miraz, M. H., Ali, M., & Excell, P. S. (2021). Adaptive user interfaces and universal usability through plasticity of user interface design. *Computer Science Review*, *40*, 100363. <https://doi.org/10.1016/j.cosrev.2021.100363>
- NATO. (2023). *APP-6E: Joint military symbology*. NATO Standardization Office. <https://nso.nato.int/nso/nsdd/main/standards/ap-details/3169/EN>
- Nielsen, J. (2012, January 3). *Usability 101: Introduction to Usability*. Nielsen Norman Group. <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>
- OpenAI. (2025). *ChatGPT 4o* [Large language model]. <https://openai.com>
- Price, P. C., Jhangiani, R. S., Chiang, I. A., Leighton, D. C., & Cuttler, C. (2017). *Research methods in psychology* (3rd American ed.). Washington State University. <https://opentext.wsu.edu/carriecuttler/>
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(3), 257–266. <https://doi.org/10.1109/TSMC.1983.6313160>
- Rosala, M. (2021, October 31). *How many participants for a UX interview?*. Nielsen Norman Group. <https://www.nngroup.com/articles/how-many-participants-ux-interview/>
- Saab AB. (2019). *GlobalEye AEW&C in coastal environment* [Photograph]. Saab Brand Portal. <https://brand.saab.com/>
- Saab AB. (2022, December 2). *The art of command & control*. <https://www.saab.com/newsroom/stories/2019/march/the-art-of-command--control>
- Saab AB. (2023). *Giraffe IX deployment set* [Photograph]. Saab Brand Portal. <https://brand.saab.com/>
- Saab AB. (n.d.-a). *GlobalEye AEW&C*. Retrieved May 14, 2025, from <https://www.saab.com/products/globaleye>
- Saab AB. (n.d.-b). *Organisation*. Retrieved January 24, 2025, from <https://www.saab.com/about/company-in-brief/organisation>
- Saab AB. (n.d.-c). *Surface radar*. Retrieved May 14, 2025, from <https://www.saab.com/products/land/surface-radar>

- Salvendy, G., & Karwowski, W. (2021). *Handbook of Human Factors and Ergonomics* (5th ed.). John Wiley & Sons.
- Schubert, J., Brynielsson, J., Nilsson, M., & Svenmarck, P. (2018, November). *Artificial intelligence for decision support in command and control systems*. In *Proceedings of the 23rd International Command and Control Research & Technology Symposium: Multi-Domain C2* (Vol. 2, pp. 18–33).
https://www.researchgate.net/publication/330638139_Artificial_Intelligence_for_Decision_Support_in_Command_and_Control_Systems
- Van Den Bosch, K., & Bronkhorst, A. (2018). *Human-AI cooperation to Benefit Military Decision Making*. NATO Science and Technology Organization.
<https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-IST-160/MP-IST-160-S3-1.pdf>
- von Janczewski, N., Wittmann, J., Engeln, A., Baumann, M., & Krauß, L. (2021). A meta-analysis of the n-back task while driving and its effects on cognitive workload. *Transportation research part F: traffic psychology and behaviour*, 76, 269-285.
<https://doi.org/10.1016/j.trf.2020.11.014>
- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Science*, 3(2), 159-177. <https://doi.org/10.1080/14639220210123806>
- 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>
- Wilson, C. (2013). Braindrawing. In C. Wilson (Ed.), *Brainstorming and beyond* (pp. 63–75). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-407157-5.00003-8>
- Young, M. S., & Stanton, N. A. (2002). *Attention and automation: New perspectives on mental underload and performance*. *Theoretical Issues in Ergonomic Science*, 3(2), 178-194. <https://doi.org/10.1080/14639220210123789>
- Zak, Y., Parmet, Y., & Oron-Gilad, T. (2023). Facilitating the work of unmanned aerial vehicle operators using artificial intelligence: an intelligent filter for command-and-control maps to reduce cognitive workload. *Human Factors*, 65(7), 1345-1360.
<https://doi.org/10.1177/00187208221081968>
- Zimmerman, J., Forlizzi, J., & Evenson, S. (2007, April). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 493–502). Association for Computing Machinery. <https://doi.org/10.1145/1240624.1240704>

Appendices

Appendix A – Initial Interview Template

Interview questions (in Swedish)

Please note that your participation is voluntary and you can choose to cancel your participation at any time.

Introduktion:

Vi gör ett exjobb på Chalmers och för Saab, där vi undersöker hur AI kan integreras för ett anpassningsbart C2, samt hur det ska visualiseras. Därför behöver vi en övergripande kunskap om scenarion, hur det används, samt vilken information som är viktig.

Intervjun kommer att ha fokus på två saker:

1. Vi kommer börja med att fråga på en övergripande nivå för C2 som koncept, hur det ämnas användas och hur scenarion ser ut i verkligheten, både vardagliga och allvarliga.
2. Till sist vill vi gärna att du visar funktionerna i C2

Intervjuperson:

1. **Vad är din roll nu?**
2. **Vad har du för erfarenhet?**
3. **Vad har du för koppling till systemet? (Vilket specifikt system?)**
4. **Har du arbetat med något/några andra system? (Vilket/vilka?)**

Del 1: Övergripande nivå (Brasklapp – Gå inte in på krångel och buggar)

1. **Hur ser ett typiskt uppdrag ut?**
 - Hur lång tid är detta? (Total tid, kontinuerligt, raster?)
2. **Vilka arbetsroller finns i användandet av systemet?**
 - Om praktisk erfarenhet, vilken roll/flera?
3. **Beskriv ett/några typiska vardagliga scenarion, om det finns?**
 - Välj ett. Hur agerar man då? Bryt ner i 3-6 steg.
 - Vilka steg är mest kognitivt krävande? Med 'kognitivt krävande' menas ex. bedömning, utvärdering och problemlösning.

4. Beskriv ett/några möjliga allvarigare scenarion

- Välj två. Hur agerar man då? Bryt ner i 3-6 steg.

- Vilka steg är mest kognitivt krävande? Med 'kognitivt krävande' menas ex. bedömning, utvärdering och problemlösning.

5. Vilka är de mest användbara funktionerna? Vilka är dina favoriter? (Fråga vid behov)

6. Beskriv ett/några scenarion som leder till hög arbetsbelastning. (Fråga vid behov)

Del 2: Demonstration

1. Kan du visa oss/utföra de tre tidigare nämnda scenariona i C2 som du beskrev i steg?

- **OBS!** Vi utvärderar ej funktionernas buggar/problem. Intresserade av deras syfte och funktion taktiskt.

2. Job Smarts: "När du utför denna uppgiften, finns det smarta sätt att jobba för att uppnå mer med mindre arbete?"

Följande frågor ställs i mån av tid:

3. Self Monitoring: "Finns det någon gång där du insett att du behöver ändra ditt arbetssätt för att få gjort uppgiften? Orsaker kan vara stress, trötthet, yttre faktorer m.m."

4. Big Picture: "Kan du ge mig ett exempel på vad som är viktigt för helheten för den här uppgiften? Vad är de viktigaste delarna av uppgiften?"

5. Equipment Difficulties: "Har det varit gånger då programvaran pekat i en riktning men ditt eget omdöme sa något annat? D.v.s. du behövt lita på din erfarenhet för att inte bli vilseledd av programvaran"

Appendix B – Initial Interview Survey

Survey

Please note that your participation is voluntary and you can choose to cancel your participation at any time. All data fields are optional to fill in.

1. **Gender (circle one):** **Male** **Female** **Other/prefer not to say**

2. **Year of birth:** _____

3. **Do you have system (C2) related military service (circle one):**

Yes No Prefer not to say

4. **Rate your level of operator experience (circle one):**

No operator experience

Low level of operator experience

Medium level of operator experience

High level of operator experience

5. **Rate your level of experience with the C2 interface (circle one):**

No experience

Low level of experience

Medium level of experience

High level of experience

Appendix C – Participants' Task Diagrams

Figure C1

Participants' Task Diagrams



Note. The individual task diagrams of the participants of the first user study (in Swedish). Colour-coded to the corresponding steps in the created Generic Task Diagram (see Figure 6).

Appendix D – KJ-representation of Ideas from Phase B

Figure D1

Generally applicable (Övergripande)

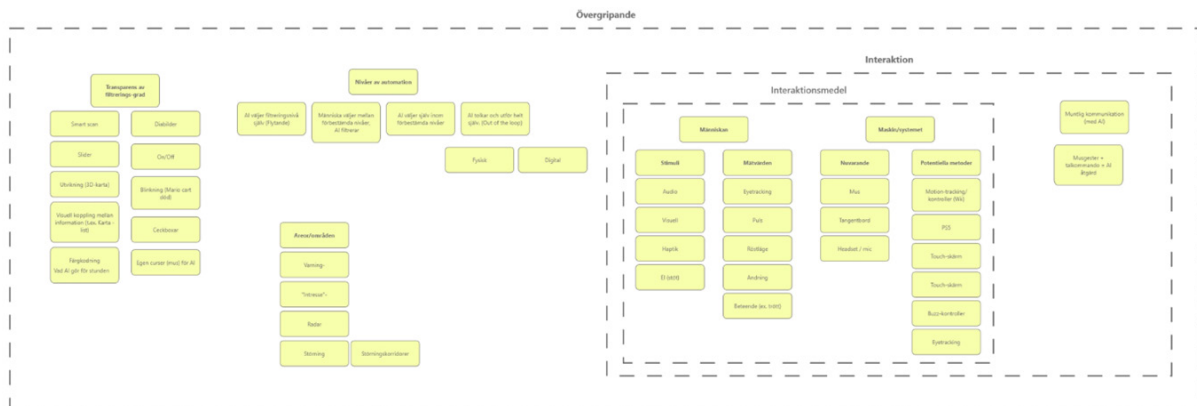


Figure D2

Filtered Situational Display (Avskalad lägesbild)

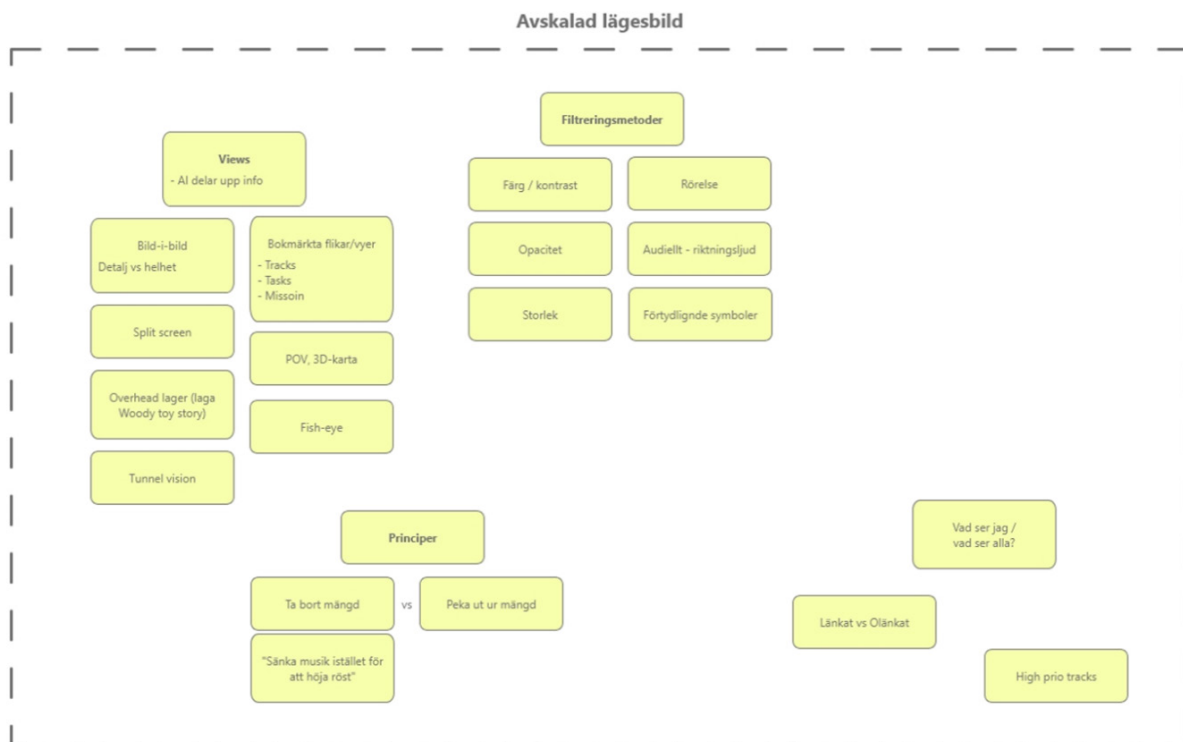


Figure D3

Adaptive UI support (Adaptivt GUI-stöd)

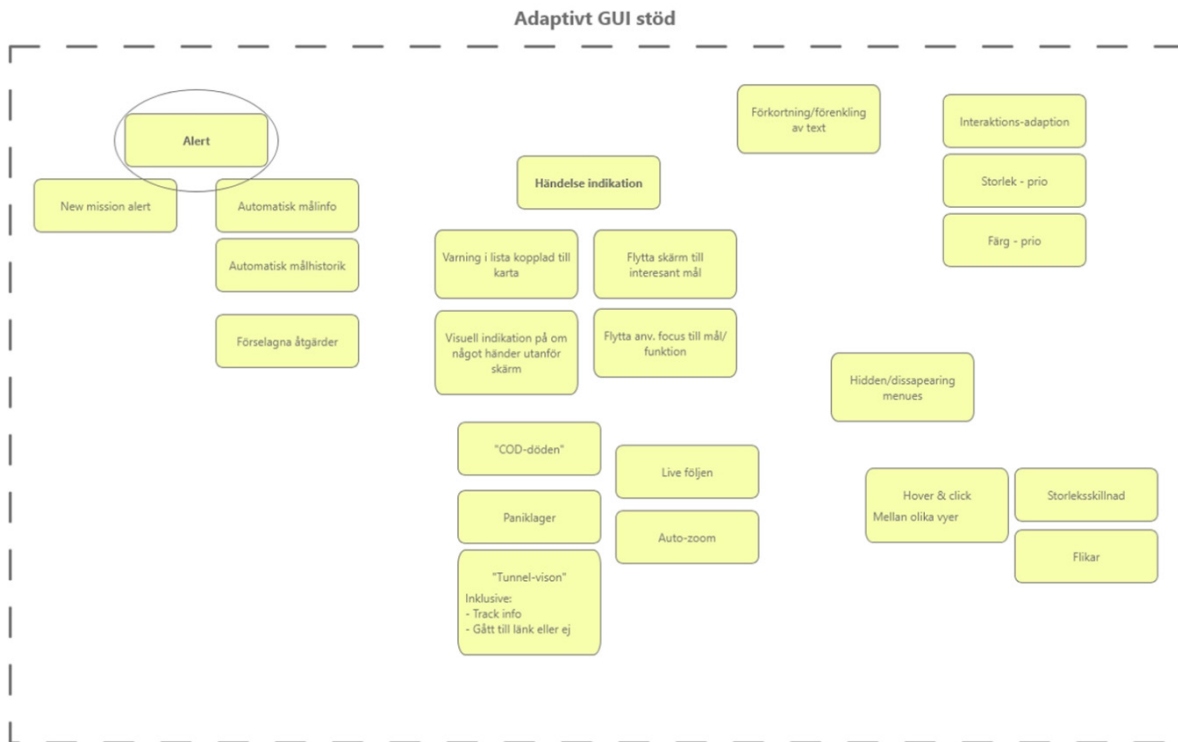


Figure D4

Mission-, setup-, and process support (Uppdrags-, Uppstarts-, Processtöd)

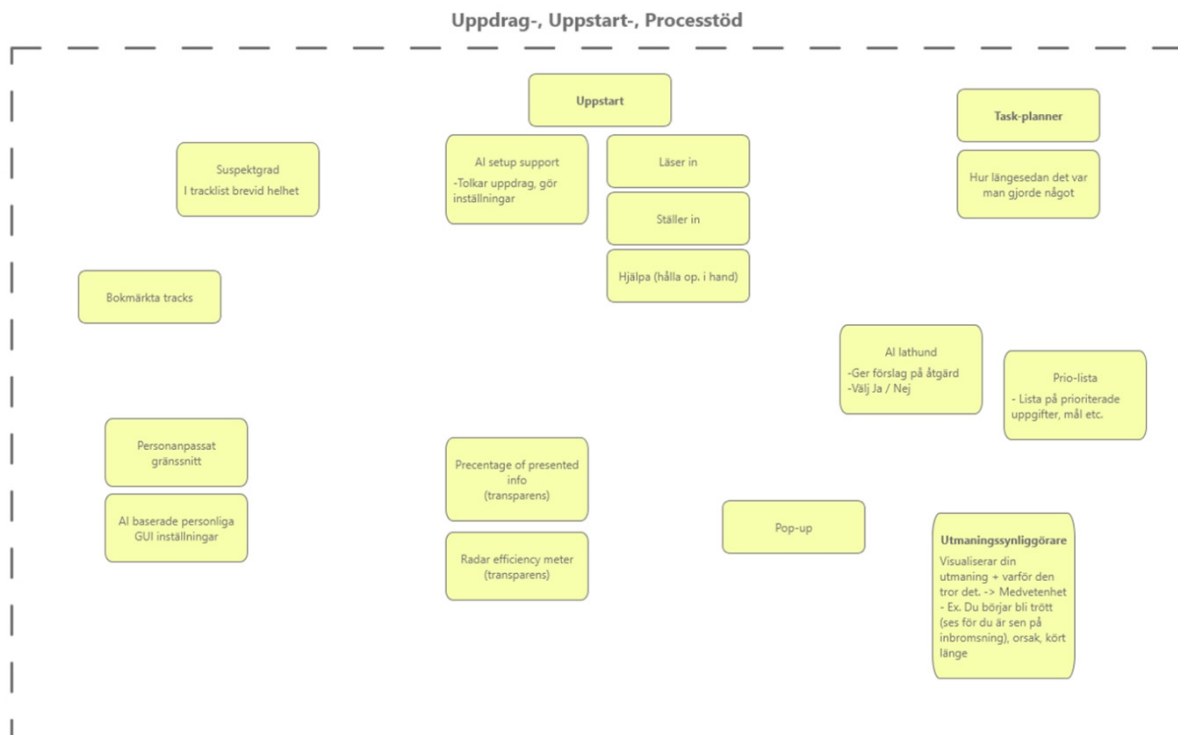
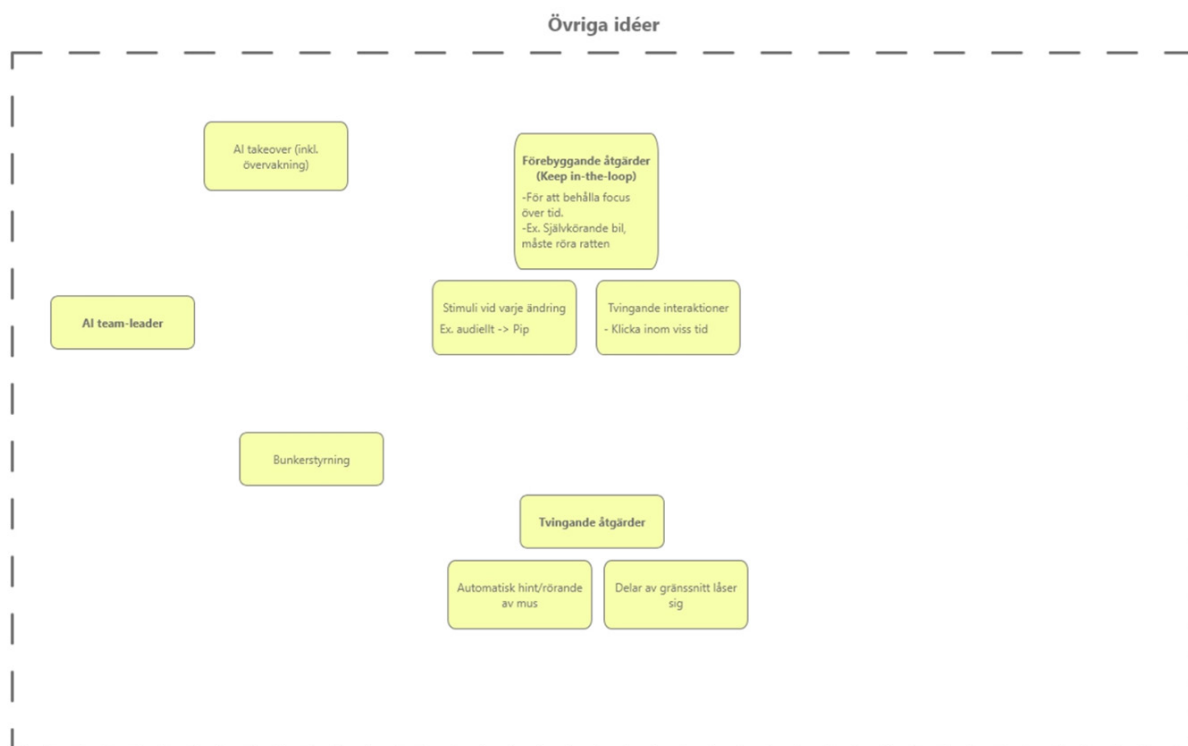


Figure D5

Other ideas (Övriga idéer)



Appendix E – Concept Evaluation Template

Intro

För att ge en liten repetition, så handlar vårt exjobb om att undersöka hur vi kan göra C2 adaptivt med hjälp av AI. Vi har tagit fram 3 st lösningsområden, varav ni kommer att få göra tester på det första, medan de två andra kommer att vara diskussionsunderlag.

Bra att veta:

- Det är idéer som är menade att vara allmänt applicerbara, det är inte specifikt för ett system
- Idéerna är prioriterade över realismen i form av antal tracks, information osv.
- Vi utgår från att AI:n är bra nog att göra rätt bedömningar för vilka saker som är viktigt att se

Filter

Till att börja med kommer vi att visa lösningsområdet vi kallar Filter. Det handlar om hur AI:n lyfter fram vissa tracks som är extra viktiga att se. Vi kommer att visa 4 varianter, där du ska klicka på de tracks som är visuellt framhävda och vi tar tid på hur lång tid det tar för dig att bli klar. Du kommer att få värma upp på en demo först.

A. Static

<i>Time to completion:</i>	
<i>Time to feeling done:</i>	
<i>Antal rätt:</i>	
<i>Antal fel:</i>	

B. Dynamic

<i>Time to completion:</i>	
<i>Time to feeling done:</i>	
<i>Antal rätt:</i>	
<i>Antal fel:</i>	

D. Map Change

<i>Time to completion:</i>	
<i>Time to feeling done:</i>	
<i>Antal rätt:</i>	
<i>Antal fel:</i>	

E. Tunnel Vision

<i>Time to completion:</i>	
<i>Time to feeling done:</i>	
<i>Antal rätt:</i>	
<i>Antal fel:</i>	

Efter alla:

- Vad tyckte du om var och en?
- Bäst/sämst?

Prio list

Nu går vi in på det andra lösningsområdet, som vi kallar Prio list. Nu är koncepten diskussionsunderlag. Detta är en funktion där Prio List ligger till höger, och med hjälp av AI prioriteras målen efter vilka som ska tittas på först. Vi har gjort det på 4 olika sätt. Modellerna är inte fullt interaktiva, men om man klickar på de tre högst prioriterade lyser rätt track upp.

A. Sortera

Hur upplevde du denna?

B. Storleksordning

Hur upplevde du denna?

C. Opacitet

Hur upplevde du denna?

D. Stapel

Hur upplevde du denna?

Efter alla:

- *Känns detta som en funktion som gör nytta i identifieringsarbetet?*
- *Vilken gillade du mest/minst och varför?*
- *Har du egna idéer?*

Views

Nu går vi in på det tredje lösningsområdet som vi kallar Views. Dessa koncept är också diskussionsunderlag. Den här koncepten behandlar olika sätt som AI:n delar upp information visuellt i olika fönster.

Adding window

- *Förstår du vad som händer?*
- *Vad tycker du om det?*
- *Vad känner du om att systemet bestämmer detta?*

Outside screen

- *Förstår du vad som händer?*
- *Vad tycker du om det?*

Outside screen + change tracks + zoom area

- *Förstår du vad som händer?*
(Om ej: Förklara för dem efteråt vad som händer.)
- *Vad tycker du om varje enskild funktion?*
- *Vad tycker du om att det sker samtidigt?*

View change

- *Förstår du vad som händer?*
- *Vilken view föredrar du?*
- *Vad tycker du om att systemet bestämmer?*

Outside screen + view change vs outside screen -> view change

- *Förstår du vad som händer?*
- **Visa andra*. Vilken föredrar du?*

(Om vi behöver: Everything everywhere all at once)

- *Förstår du vad som händer?*
- *Vad tycker du?*

View + tracks list

Nu visar vi en kombo.

- *Förstår du vad som händer?*
- *Vad tycker du om det?*

Övergripande följdfråga:

- *Vilken typ av info vill du ha för att identifiera ett tracks?*

Appendix F – Prototypes for Filter Evaluation

Figure F1

Filter Test of Dynamic

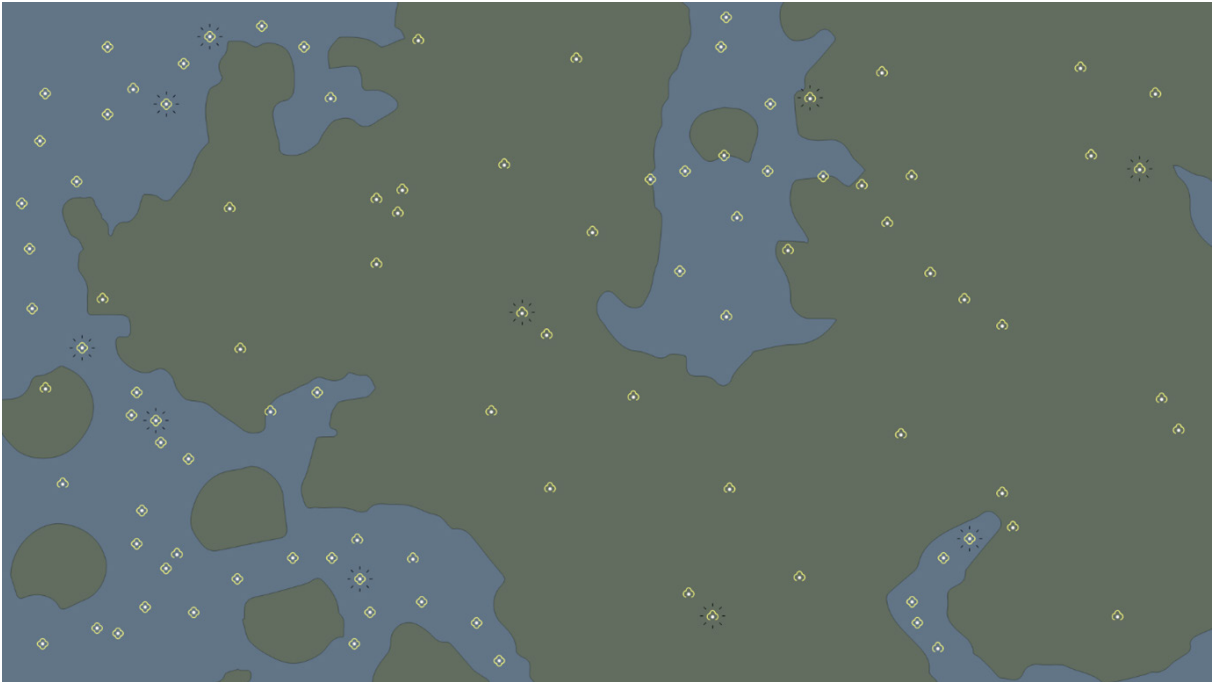


Figure F2

Filter Test of Map Change

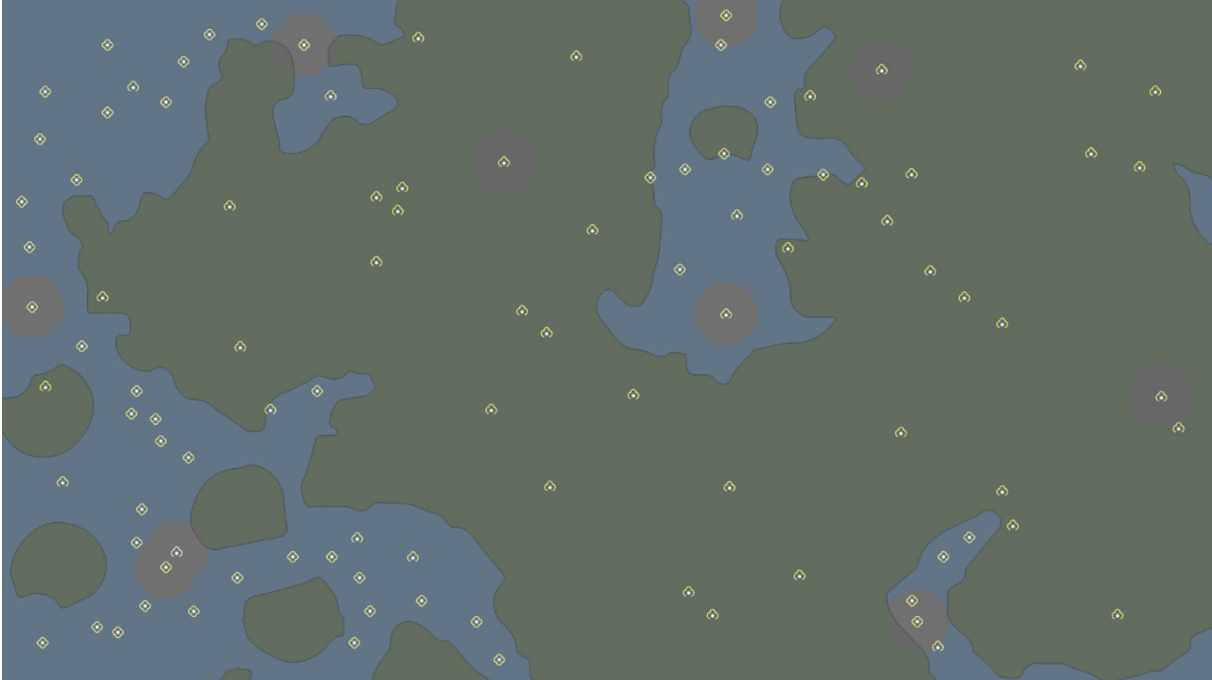


Figure F3

Filter Test of Static

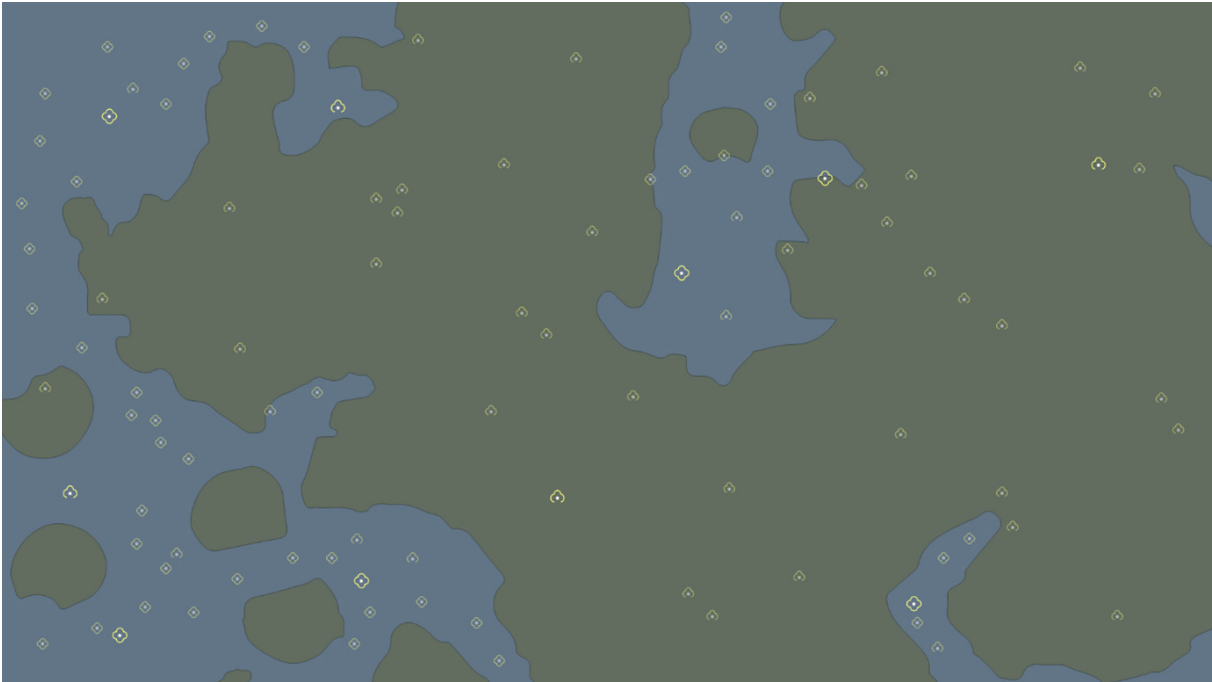
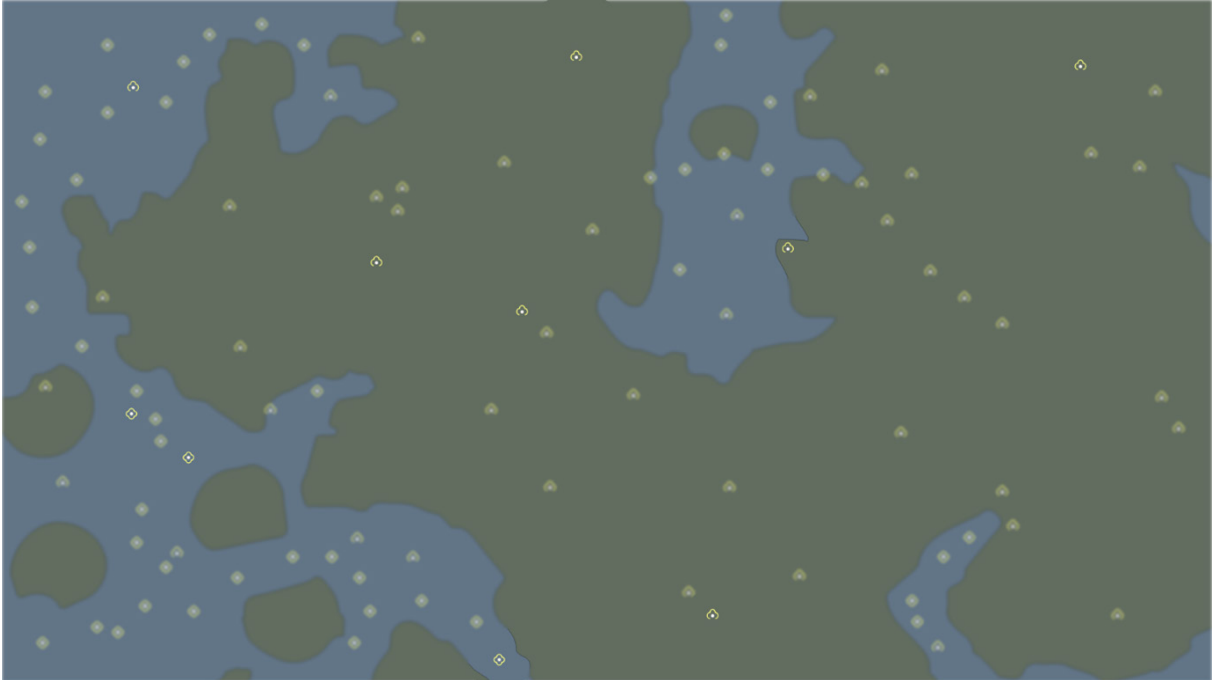


Figure F4

Filter Test of Tunnel Vision



Appendix G – Information for Participants in the Adaptivity & Workload Test

Figure G1

Slide 1 - Introduction

Välkommen till Sluttestet

- Vårt exjobb handlar om att undersöka hur AI-stöttad adaptivitet kan implementeras och visualiseras i C2.
- Hittills har vi genomfört intervjuer, idegenererat, utvecklat koncept och utvärderat dem.
- Nu ska vi testa hur ett av koncepten, ett adaptivt filter, påverkar operatörens arbetsbelastningen och situational awareness.

Figure G2

Slide 2 - Introduction

- Testet kommer bestå av två parallella uppgifter, en huvuduppgift och en sidouppgift.
- Huvuduppgiften består av 6 omgångar som kan liknas med bonka bäver, 1 minut långa var.
- Sidouppgiften är en uppgift som kallas N-back test. De är också en minut långa och görs parallellt med huvuduppgiften, svårare och svårare.
- Nu till testen:

Figure G3

Slide 3 - Main Task

Huvuduppgift – ”Bonka bäver”

- Två varianter av bonka bäver med tracks
- Utan och med adaptivt AI-filter.
- Varje omgång, dvs 1 minut, är uppdelad i 6 st 10-sekundersintervall
- Under varje intervall, har 5 tracks en svart ring runt sig som representerar en förändring
- Du ska klicka på så många du hittar i varje intervall. De kommer inte visa att du klickat på dem, så håll koll på vilka du klickat.
- Intervallbytena är tydliga, skärmen ”byter blad”. Då får 5 nya tracks en svart ring runt sig.
- Vi räknar hur många du hittar. Precision räknas också.

Figure G4

Slide 4 - Main Task Examples

Huvuduppgift

Utan adaptivt filter



Med adaptivt AI-filter (3 framlyfta, 2 i bakgrund)



Figure G5

Slide 5 - Side Task

Sidoupppgift – N-back test

- Ett test för kognitiv arbetsbelastning
- Du kommer höra en sekvens bokstäver.
- N innebär antalet steg bakåt i sekvensen en match ska ske.
- När det sker en match, säger du till.
- Du kommer att göra huvuduppgiften, både med och utan filter, tillsammans med en stegrad svårighet på N-back.
- Utan, N=1 och N=2

Figure G6

Slide 6 - Side Task Examples

Sidoupppgift

N=1, dvs 1-back test

A, P, Q, **Q**, F, B, S, **S**, K

N=2, dvs 2-back test

J, Y, C, **Y**, F, B, S, **B**, H

Du ska få en testminut på
båda för att värma upp!

Figure G7

Slide 7 - AI Decision Basis

AI:n tar sitt beslut vad som ska lyftas fram i det adaptiva filtret på en sammanvägning av bland annat:

- Uppgift
- Måltyp
- Geografisk kontext
- Identifieringssignal
- Fart
- Höjd
- Position
- Beteende
- Avstånd
- Radarmålyta

Appendix H – Subsidiary Task Result Template & Self-Assessment Questions

Participant number ___

1-Back 0

N, F, I, I, D, M, B, P, Q, Q, O, R, T, T, W, Y, A, A, V, E, E, H, J, K, K, U, S, B, C, N

1-Back 1

L, D, G, G, B, K, Z, N, O, O, M, P, R, R, U, W, Y, Y, T, C, C, F, H, I, I, S, Q, Z, A, L

1-BACK 2

M, E, H, H, C, L, A, O, P, P, N, Q, S, S, V, X, Z, Z, U, D, D, G, I, J, J, T, R, A, B, M

2-Back 0

A, S, H, N, H, Q, Z, Q, E, D, G, D, J, Y, G, Y, O, P, P, T, P, N, J, C, V, C, Y, B, A, W

2-BACK 1

Y, Q, C, L, C, O, X, O, C, B, Y, B, H, W, E, W, M, N, N, R, N, L, H, A, T, A, W, Z, Y, U

2-BACK 2

Z, R, D, M, D, P, Y, P, D, C, F, C, I, X, F, X, N, O, O, S, O, M, I, B, U, B, X, A, Z, V

Test:

Hur bedömer du din prestation, endast i form av resultatet?

1 2 3 4 5 6 7 8 9 10

Hur bedömer du din kontroll över situationen?

1 2 3 4 5 6 7 8 9 10



CHALMERS
UNIVERSITY OF TECHNOLOGY