Design of a low-fidelity prototype interface for a computerised maintenance management system:

A case study for Thales B.V. Hengelo

Stef Poppe

University of Twente, Enschede

Faculty of Behavioural, Management and Social Sciences (BMS)

Internal supervisors:

dr. Simone Borsci (1st)

prof.dr. Jan Maarten Schraagen (2nd)

External supervisor:

Theo Nijland (R&T Manager, Thales Nederland B.V.)

Contact: s.m.m.poppe@student.utwente.nl

Abstract

Complex systems maintenance management can be supported by using technologies such as computerised maintenance management systems (CMMSs). However, CMMSs face usability issues (Rastegari & Mobin, 2016; Tretten & Karim, 2014). To design a usable system, usability standards should be adopted, which are guidelines that aim to improve product acceptability and user satisfaction. This research focused on designing a usable CMMS prototype interface for a vessel's (ship) complex system as a case study for Thales B.V. Throughout the research, the complex system focus shifted multiple times, due to COVID-19, from general complex systems to TACTICOS: a combat management system from Thales B.V. The structure of the paper is divided into three connected phases:

1) Definition of a framework on issues with complex systems maintenance and CMMSs. A systematic literature review and interviews with stakeholders from a Thales B.V. complex system was performed to identify issues with complex systems maintenance and CMMSs. This resulted in a list of 28 identified issues in total, which were reviewed by participants in a questionnaire. In total, 15 stakeholders participated, from which a list of 5 main issues was identified with complex systems maintenance and CMMSs.
2) *Design of a CMMS prototype: A case study for Thales B.V.* The list of 5 main issues from phase 1, together with a carried-out task analysis on operators and maintainers performing maintenance tasks in complex systems was used to inform the design of a CMMS prototype for a vessel's complex system at Thales B.V.
3) *Usability testing of the CMMS prototype.* Small-scale remote usability testing was performed with 11 stakeholders from Thales B.V. The results indicated a total of 69 feedback remarks, an 'above average' score on the system usability scale (SUS) and a 'below average' net promoter score (NPS). The feedback remarks were considered in a redesign of the prototype.

This research faced three main limitations, from which the most significant one explains that end-users of CMMSs were out of reach. Nonetheless, the phases of method applied can be reused by Thales B.V. in future work when designing a usable CMMS.

Table of contents

1. Introduction **3**

2. Methods **9**

Phase 1. Definition of a framework on issues with complex systems maintenance and CMMSs **11**

2.1 Systematic literature review **11**

2.1.1 Method *11*

2.1.2 Results *13*

2.1.3 Discussion *21*

2.2 Interviews and questionnaire with stakeholders from Thales B.V. **24**

Part 1: Interviews **25**

2.2.1 Method *25*

2.2.2 Results *28*

Part 2: Questionnaire **31**

2.2.3 Method *31*

2.2.4 Results *34*

2.2.5 Discussion *35*

Phase 2. Design of a CMMS prototype: A case study for Thales B.V. **38**

2.3 Interface design principles **40**

2.4 Hierarchical task analysis research **42**

2.4.1 Method *43*

2.4.2 Results *49*

2.4.3 Discussion *53*

2.5 Low-fidelity CMMS prototype explanation **55**

Phase 3. Usability evaluation of the CMMS prototype **58**

2.6 Usability testing **58**

2.6.1 Method *58*

2.6.2 Results *64*

2.6.3 Discussion *71*

3. Overall conclusion **73**

4. Reference list **80**

1. Introduction

Complex technical systems are complex due to their many components and sub-systems which interact and relate to each other. Examples of such systems are: nuclear power plants, transport systems, communication systems, and manufacturing plants. These complex systems deteriorate with age and usage, resulting in the system not being able to perform optimally anymore (Kobbacy & Murthy, 2008). Therefore, maintenance in complex systems is an ever-growing, significant aspect of keeping the systems' life expectancy high. Maintenance in systems can be defined as a "set of activities required to keep physical assets in the desired operating condition or to restore them to this condition" (Pintelon & Parodi-Herz, 2008, p. 22). The concept of maintenance has changed over the past 100 years, from when the technical staff was used to consider maintenance as a series of corrective actions to perform after a system failure, to a set of maintenance strategies in a maintenance management plan to prevent system failures in the future (Kans, 2009; Kobbacy & Murthy, 2008). Maintenance management considers three important levels:

1. The *strategic level* of maintenance is mostly focused on the selection of optimal maintenance strategies (Kobbacy & Murthy, 2008). The most common strategies to consider are (Lee & Wang, 2008):
    a. Corrective maintenance (CM): when a system breaks down, maintenance will be performed.
    b. Preventive maintenance (PM): at fixed modelled time-intervals, parts of the system will be replaced or overhauled. This is usually done either based on the age(-ing) of the system or periodically without considering system deterioration.

2. The *tactical level* of maintenance includes the planning and scheduling thereof, considering the reliability of systems, optimal policy selection, and logistics (Kobbacy & Murthy, 2008).
3. The *operational level* focuses on the execution of maintenance tasks by collecting relevant data and performing analyses (Kobbacy & Murthy, 2008), such as root cause analysis (RCA), which "is a tool designed to help identify not only what and how an event occurred, but also why it happened" (Rooney & van den Heuvel, 2004, p. 45). Collecting data such as RCA will draw results and make recommendations on the future use of these (sub-)systems.

Pintelon and Parodi-Herz (2008) suggest there is a gap between the strategic level on the one hand and the tactical and operational level on the other hand. They found that research is mainly focused on the tactical and operational level while leaving out the 'business-side', or strategic level of maintenance. However, the strategic level should be deemed as an important research topic: Choosing the correct maintenance strategy is an essential task, since this results in less unplanned, corrective interventions by maintainers (Pintelon & Parodi-Herz, 2008). In addition, Kans (2009) found that over the years, strategic maintenance increased in importance and is, in combination with useful tools for maintenance management, beneficial to companies economically.

A corresponding useful tool for maintenance management is eMaintenance solutions, which have been around for approximately 20 years and are continuously in development (Iung, Levrat, Marquez, & Erbe, 2007). eMaintenance has many definitions but can generally be seen as an aid to maintenance (management) by adopting advanced information communication technologies (Iung et al., 2007; Tretten & Karim, 2014). These solutions can support maintenance stakeholders on all levels (strategic, tactical, and operational) in making more efficient and effective maintenance-related decisions in the specific context of their organisation and work (Karim, 2008; Labib, 2004; Tretten & Karim, 2014). eMaintenance technologies are still implemented increasingly by corporations due to the fact that wireless access to the Internet keeps getting easier (Galar & Kumar, 2017). Several advantageous features of eMaintenance solutions to assist maintenance management are: (1) remote maintenance, where system stakeholders can perform actions remotely, (2) cooperative maintenance, in which the information structure makes communication between stakeholders of a system more accessible, (3) immediate maintenance, where a quick response of stakeholders is possible, and (4) predictive maintenance, where companies are supported with "predictive intelligence tools to monitor their assets through Internet wireless communication systems to prevent unexpected breakdowns" (p. 459). From these tools, prognostics data can be gathered and analysed to predict future system behaviours (Galar & Kumar, 2017).

A widely used eMaintenance interface is a computerised maintenance management system (CMMS). CMMSs are computer software packages that aid maintenance processes with planning, managing, optimising, and collecting data for maintenance (Bagida, 2010). Lopes et al. (2016) summarised the main functions of a CMMS to be the management of assets, work orders, preventive maintenance, reports, and inventory control. This means that CMMSs are used by all levels of maintenance management: strategic, tactical, and operational (Labib, 2004). These are all useful functions of a CMMS to support maintenance management, but the system faces some significant problems as well. The leading weakness being usability-related problems, as indicated by several researchers. Labib (2004) and Rastegari and Mobin (2016) found that CMMSs require a lot of input, but do not provide much output in terms of decision support. This means that the system can collect and save a lot of relevant data but cannot analyse this accordingly to give a useful output to the user. In addition, Tretten and Karim (2014) performed a study on usability issues in CMMSs in different industries: pulp and paper, mining, and aerospace, whom all use different CMMS software. Participants from the study were stakeholders from all levels of maintenance management who work daily with the CMMSs that their company employed. They were interviewed on the positive and negative sides of the CMMSs they utilise. The results of the study presented seven issues with the various CMMSs from the different industries: (1) limited access to necessary (technical) documentation, (2) incompatibility with other necessary systems, (3) too much manual input of information, (4) the user interface is not intuitive and difficult to understand, (5) lack of guidance from the system, (6) lack of maintenance decision support, and (7) the CMMS is too complex in its use (Tretten & Karim, 2014). In a previously performed research by Wandt, Tretten, and Karim (2012), the same study was explained, but a more extensive discussion was reported in which the researchers suggest that improvements should be made on context awareness and the interface design of the CMMS. In their research, context awareness is described as “the ability of a system to adapt the operations to the current context without explicit user intervention and thereby respond to changes in the environment in order to make the system behave more relevant to the current situation” (Wandt et al., 2012, p. 3); interface design reflects common usability standards.

Usability standards are guidelines by which to design systems, defined as the “extent to which a system, product or service can be used by specified users to achieve specified goals with *effectiveness*, *efficiency* and *satisfaction* in a specified *context of use*” (International Organization for Standardization (ISO), 2018, 3.1.1 usability section). The four main terms are described in more detail as follows: effectiveness is the “accuracy and completeness with which users achieve specified goals”, efficiency explains the “resources used in relation to the results achieved”, satisfaction denotes the “extent to which the user's physical, cognitive and emotional responses that result from the use of a system, product or service meet the user’s needs and expectations” and context of use defines the "combination of users, goals and tasks, resources, and environment" (ISO, 2018, 3.1.1 usability section). In addition to this, usability is interpreted for maintenance related tasks specifically "in that it enables maintenance tasks to be completed effectively, efficiently and with satisfaction" (ISO, 2018, Introduction: Note 2 section). Complying to the three usability measures, effectiveness, efficiency and satisfaction, and taking into account the specified context of use aims for the design of a user-friendly system, ensures safety, and averts negative outcomes in case of use errors or system breakdowns (Hugo & Gertman, 2016). In addition to that, usable systems are advantageous to companies financially since it enhances the quality of their products and generates "improved product acceptability, increased user satisfaction, and improved product reliability" (Simões-Marques & Nunes, 2012, p. 155). This in turn makes it easier for users to manage the product and provokes higher product demand.

Complex systems maintenance is a set of strategies that enhances systems' operations by means of a computerised system management tool, used by maintenance managers, operators, and maintainers on all levels of maintenance management. Even though usability of systems has become increasingly important, improvements can be made on designing user-friendly maintenance management technologies, such as CMMSs. This research will focus on the design of a usable CMMS prototype interface, as a case study for Thales B.V. The main requirement for this work from Thales B.V. is a workset for CMMSs that can be applied to one of their complex systems. A workset at Thales indicates a role-specific graphical interface design consisting of several components, which is to be used in combination with certain keyboard buttons (T. Nijland, personal communication, March 25, 2021). For this case study, a role-specific CMMS prototype interface for vessels (ships) maintenance was designed, without considering the keyboard buttons. Stakeholders from a Thales B.V. complex system will be asked to participate throughout the different phases of the research. These stakeholders are employed at Thales B.V. and are considered to be experts on complex systems maintenance from a strategic, tactical, and operational point of view, by the company.

The goal of designing a usable CMMS interface will be reached by applying several consecutive phases: First, main issues with complex systems maintenance at all levels and CMMSs will be identified by a systematic literature review and interviews with stakeholders from Thales B.V. The identified issues will be rated on the degree to which the stakeholders find the issues problematic. Following, the final list of key issues, as well as problematic maintenance tasks as identified from a task analysis research, will be resolved in the design of a CMMS prototype interface as a case study for Thales B.V. Lastly, the designed CMMS will be tested with the stakeholders.

2. Methods

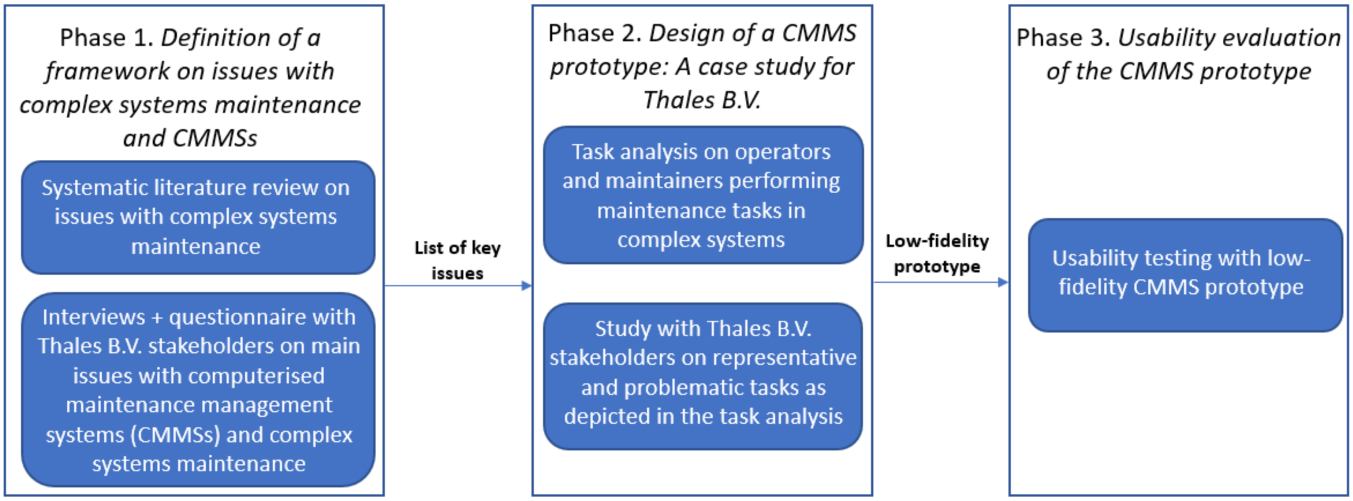
This work was performed in three phases (See Figure 1 for an overview):

1) *Definition of a framework on issues with complex systems maintenance and CMMSs.* The first phase included a systematic literature review to identify and map key issues with complex systems maintenance at all levels. To extend the results of the literature review, stakeholders from a Thales B.V. complex system were involved in interviews and a questionnaire to add to this list of issues and to identify issues with computerised maintenance management systems (CMMSs), resulting in agreement on the key issues.
2) *Design of a CMMS prototype: A case study for Thales B.V.* The stakeholders participating in phase 1 were involved in carrying out a task analysis for maintainers and operators focused on the strategic, tactical, and operational levels of maintenance in complex systems, to understand the main tasks, and the extent to which these tasks are problematic. The results from this, and the results found in phase 1 were used to inform the design of a low-fidelity CMMS prototype interface for a Thales B.V. complex system.
3) *Usability evaluation of the CMMS prototype.* Phase 3 consisted of gaining insights into the usability of the prototype by performing usability tests with stakeholders from Thales B.V.

All phases of method end with a short discussion on the work. At the end of all phases, the overall conclusion, limitations, and future work will be discussed.

Figure 1

Structure of the paper



Phase 1. Definition of a framework on issues with complex systems maintenance and CMMSs

2.1 Systematic literature review

A systematic literature review was performed to get a better understanding on the main issues and problems faced with complex systems maintenance at the strategic, tactical, and operational level[1]. The issues were considered in a general form, so they can be applied to any complex system and any level of maintenance management. This literature review resulted in a list of issues commonly found with maintenance at all levels in complex systems.

2.1.1 Method

2.1.1.1 Study design

This systematic literature review was performed on papers published in the past five years that listed or mentioned any issues experienced or found with complex systems maintenance. The reason for having chosen a short timeframe is to get an overview of the most recent issues, not considering out-dated issues. The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) method was applied (Liberati et al., 2009). The full list of inclusion- and exclusion criteria can be found in Table 1. The PRISMA Checklist, with more details regarding the used method for this study can be found in Appendix A[2].

2.1.1.2 Research question

To identify the list of common issues with complex systems maintenance at all levels, the following research question was formulated:

[1] These three levels might be called ‘maintenance at all levels’ or ‘any level of maintenance management’ in future references.
[2] Due to the large number of appendices, these can be requested by contacting the researcher.

- What issues or problems are mentioned and/or experienced with technical complex systems maintenance at the strategic, tactical, and operational levels in literature?

2.1.1.3 Eligibility criteria

For this research, studies published before 2015 were left out to get an insight into the most recent information on issues with complex systems maintenance. In addition, only papers with open access and/or access provided through the University of Twente were used, since others could not be accessed by the researcher. Only English and Dutch papers were read.

Table 1 gives an overview of the detailed inclusion and exclusion criteria for eligible papers. As can be seen, there are only few criteria for inclusion. The reason for this is that the identified issues will be regarded in the most general form, to be applicable to any complex system and level of maintenance management (as long as it does not cover any of the exclusion criteria).

Table 1

Inclusion and exclusion criteria applied to the systematic literature review on complex systems maintenance at all levels

Inclusion criteria	**Exclusion criteria**
Papers that mention issues with technical complex systems maintenance on a strategic, tactical and/or operational level.	Papers based solely on mathematical modelling of optimising maintenance policies and/or decision-making and/or planning/scheduling, without mentioning issues.
Papers published between 2015 and 2020.	Papers that concern complex systems as sociotechnical systems (e.g., hospitals, logistics) instead of technical systems.
	Case studies and/or studies directed towards a specific complex system, without mentioning issues.

	Papers that consider networking- and/or software and/or computing systems related maintenance issues (e.g., cloud computing).

2.1.1.4 Search strategy

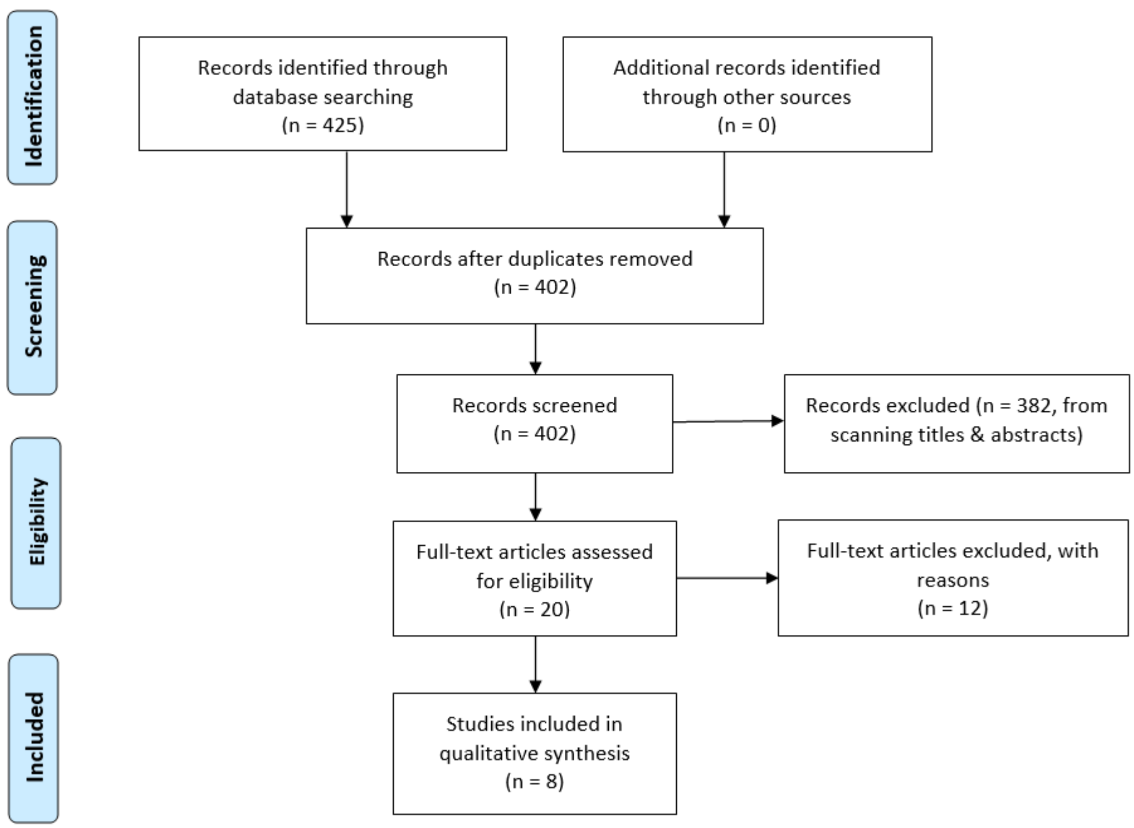
A search string was created after trials with several different search strings. Studies were selected by searching electronic databases. The search was applied to the following databases: IEEE Xplore, SAGE Journals, Web of Science, and the University of Twente library (includes ACM Digital Library, AMS Journals, BioOne, Directory of Open Access Journals, IEEE Publications Database, MEDLINE, PubMed Central, ScienceDirect, SPIE Digital Library, SpringerLink, Staten-Generaal Digitaal: Dutch Parliamentary Papers, Wiley Online Library, and WorldCat.org). The searches in the databases were performed on 24-05-2020. The search string used is the following: "Maintenance AND ("complex system*" OR "complex industrial system*" OR "high-risk system*" OR "high-risk technolog*" OR "decision-support system*" OR "decision-making system*") AND (operator* OR user* OR maintainer*) AND (availability OR reliability OR schedul* OR polic* OR strateg* OR model*) AND (problem* OR issue*) NOT (medic* OR clinic* OR health* OR sustain*)". The detailed search strategy, with search string for the different databases as used in this study can be found in the PRISMA Checklist in Appendix A.

2.1.2 Results

The search resulted in 425 identified papers in total. After deleting duplicates, a total of 402 papers were left. Scanning the titles and abstracts, and considering the inclusion and exclusion criteria, 382 papers were removed from the selection. Papers that mentioned issues related to complex systems maintenance at all levels (except for papers focusing on the exclusion criteria) were retained. After reading full-text papers, 8 were eligible, and used to identify issues with complex systems maintenance from literature. Reasons why other full-text papers were excluded are, amongst others: no mention of issues relating to maintenance, focus on a specific system, focus on specific maintenance terms or concepts, focus not on complex systems, and only providing research on algorithms and/or mathematical modelling of maintenance decision-making processes. See Figure 2 for the complete overview of identifying the papers.

Figure 2

PRISMA flow diagram



Note. Adapted from: "*P*referred *R*eporting *I*tems for *S*ystematic reviews and *M*eta-*A*nalyses: The PRISMA Statement", by D. Moher, A. Liberati, J. Tetzlaff and D. G. Altman, 2009, *PLOS Medicine*, *6*(7), p. 3.

The 8 eligible papers, their research focus, results, and limitations have been summarised in Table 2. The articles focused on a variety of complex systems or tested a tool on a complex system for a case study: 25% of the papers mentioned industrial systems, 12.5% of the papers mentioned production systems, vessels, aircraft systems, and large systems. Two other papers did not mention a specific complex system focus. Six out of eight papers (75%) suggested a tool or system to improve or support maintenance decision-making (37.5%), AR solutions (25%), or troubleshooting (12.5%). The other two papers suggested a 4-steps model to clarify issues related to complex systems maintenance (12.5%), and a framework to establish the maintenance strategy for a specific system (12.5%).

Table 2

Eligible papers of the literature review, their research focus, results, and limitations

Paper	Discussion on the paper
Abele and Weyrich (2017)	*Research focus* This paper proposes a shared decision support system to aid maintainers with fault diagnosis and developers with planning of tests in production systems. *Results* The prototype of the system aids in cooperation between fault diagnosis and test management. *Limitations* There is no explanation on how experts were involved in the design of the prototype, in addition to no explanation on how the prototype was tested (e.g., usability testing, performance measures, interviews, questionnaires). Also, there is limited reporting of the results and participants' experiences with the developed prototype. Therefore, concluding practical applicability of the tool is difficult.
del Amo et al. (2018)	*Research focus* This paper indicates that there is a lack of user-centred design in Augmented Reality (AR) for maintenance processes in industrial settings. They suggest a user-centred design process for AR technologies in the maintenance context, which they applied to two case studies. *Results* The AR maintenance design support tool was validated by performing two test case studies with a random sample

	of conference visitors. *Limitations* The sample size of the random sample of testers was small: 12 participants. Also, the sample was random: people who visited a conference could participate, they had different levels of knowledge on maintenance processes. Nonetheless, they suggested it was a preliminary study that can be replicated with a bigger, pre-selected sample of participants who have experience and knowledge on performing maintenance tasks.
Furda et al. (2015)	*Research focus* This paper describes how making decisions in the maintenance asset management of large systems is challenging for asset managers. A decision support tool to aid maintenance managers with decision-making on selected maintenance policies with limited budgets was developed. A case study with the tool was applied to a power plant system. *Results* The paper suggests and explains a maintenance decision support tool. *Limitations* No results on the case study have been provided, and thus no knowledge is available on the practical applicability of the tool. A developed interface was depicted in the paper, but no explanation on how the interface was designed nor tested has been given.
Gallab, Bouloiz, and Tkiouat (2015)	*Research focus* This paper describes how to overcome risks for operators when performing maintenance activities, by suggesting a tool that detects risks. *Results* The paper suggests and explains a multi-agent decision-support tool. *Limitations* The tool is described as having three parts, two of which have been extensively explained, whereas the last one, the user interface, has not been explained. Therefore, it is unclear how the user interface has been designed and how it could be tested. In addition, practical applications, or results for the usefulness of the tool cannot be established since no testing has been performed.
Lang et al. (2018)	*Research focus* This paper found that, due to the increasing complexity of industrial systems, troubleshooting has become a more complex task for operators. A solution is suggested in the form of an assistance system architecture. *Results* The result of the paper is the proposed system that can aid operators in troubleshooting, in order to decrease system downtime and the need for an experienced technician to perform the troubleshooting. *Limitations*

	The system has been applied to a use case scenario, but not yet tested with actual operators. However, it has been mentioned that this should be performed in future work.
Martinetti, Rajabalinejad, and van Dongen (2017)	*Research focus* This paper explains how AR technologies can aid the maintainer and operator with complex maintenance work to decrease human error and increase safety. *Results* An AR technology was suggested that guides the maintainers through their work by updating operators about the surroundings of their work environment, decreasing the probability of errors and time spent on documentation of tasks, increasing the level of safety, and cutting down time and expenses for maintenance. *Limitations* The suggested tool is described on a highly superficial level: no detailed explanations on how the tool works or how it could aid maintainers have been given, and no tests have been performed with the tool. Therefore, the practical applicability of the tool is difficult to assess.
Martinetti, Schakel, and van Dongen (2018)	*Research focus* This paper suggests a framework to establish a scalable maintenance program using reliability centred maintenance approaches in order to get the highest system reliability in unmanned aircraft systems (UAS). The framework was tested with a UAS manufacturer company in which five operators were invited to participate in the survey. *Results* The results show that the proposed framework was accessible, sustainable, and had high ease of use. However, technical suggestions were made by the operators regarding concepts used in the framework. *Limitations* A literature study on topics considered in UAS mechanical characteristics, and on suitable maintenance approaches to apply was performed, but not explained in terms of its method: merely some studies were referred to, to have an overview of current maintenance concepts. In addition, it is unclear whether a survey, questionnaire, or interview approach was used to test the framework: all words have been used interchangeably. It has not been mentioned or explained how the data from the survey has been analysed, in addition to no quantitative data being available on how many (technical) suggestions were made exactly on which topics. Also, only five operators have been invited to participate, but it is unclear whether they all took part.
Wahid et al. (2018)	*Research focus* This paper aims to clarify the ‘complex naval ship availability issue’. This was done by conducting a literature review and Delphi study with Snowballing Technique, as well as post-survey validation with experts. *Results* A 4-steps availability-oriented model is suggested which increases the understanding of naval ship availability for operators/maintainers/stakeholders and others.

Limitations

A literature review on Downtime Influence Factors (DIFs) was performed, but no explanation on the method of the literature review was given; merely the resulting list of DIFs was provided. In addition, no detailed information on the Delphi study was given as it was unclear which questions were asked to the participants in the questionnaire. Also, it was unclear who the experts participating in the Delphi study exactly were: no demographic information was provided.

From reading the 8 eligible papers, 32 issues in total were identified. These 32 identified issues were merged, based on similarity as perceived by the researcher, in order to reduce the number of items in the list and create a clear overview, resulting in a total of 14 main issues. Some issues were rewritten to imply the meaning behind the issue or show familiarity of terms to make the explanation of the issue clear. Table 3 provides an overview of the merged list of issues; the number of times the issue was mentioned in the different papers; the complex system(s) focus of the issues; and a description of the identified issue as taken from the paper(s). Considering the most commonplace issues mentioned, three different papers (37.5%) discussed issues related to documentation of work and/or communication; four of the papers (50%) explained the duration of maintenance work and/or money-related maintenance issues (50%). The most frequent issue, as mentioned by five different papers (62.5%) indicated issues regarding the complexity of maintenance work.

Table 3

List of merged issues found; the number of times the issue was mentioned; the complex system(s) focus of the paper(s); and a description of the issue from the paper(s)

Issues	Count	Complex system(s) focus	Description from the paper(s)
1. No documentation of important data, such as communication between people/teams	4	Production systems Unmanned aircraft systems	1. Results of the communication between maintainers and developers are not documented well (Abele & Weyrich, 2017). 2. Results of the communication are only kept between the people who communicated, and not shared with relevant others (Abele & Weyrich, 2017). 3. Information and knowledge management (Wahid et al., 2018). 4. The absence of and insufficient quality of maintenance documentation (Martinetti et al., 2018).
2. The amount of alarms/notifications	1	Industrial systems	5. The amount of alarms/notifications in case of a system error/issue (Lang et al., 2018).
3. The duration of looking for instructions	1		6. The duration of looking for instructions (Martinetti et al., 2017).
4. Availability of systems	2	Industrial settings Naval ships	7. The availability of systems should be maximised (del Amo et al., 2018). 8. Operational availability (Wahid et al., 2018).
5. Duration of maintenance tasks	4	Industrial settings Industrial systems	9. Duration of maintenance tasks (del Amo et al., 2018). 10. Duration of maintenance tasks (Gallab et al., 2015). 11. Repair time in case of failure too long (Lang et al., 2018). 12. The duration of performing maintenance tasks (Martinetti et al., 2017).
6. Inefficient method of sharing knowledge	2	Production systems	13. Developers and maintainers have no efficient method to share knowledge regarding problems in systems (Abele & Weyrich,

between developers and maintainers			2017). 14. Telephone/email communication is inefficient, and interferes with work processes (Abele & Weyrich, 2017).
7. Difficulties in making optimal maintenance decisions, due to the amount of policies available	1	Large systems, with case study on power plants	15. Difficulties in making optimal maintenance decisions, due to the amount of policies available (Furda et al., 2015).
8. Difficulties in making optimal maintenance decisions, due to trade-offs to be made between costs and availability, and reliability	1	Large systems, with case study on power plants	16. Difficulties in making optimal maintenance decisions, due to trade-offs to be made between costs and availability, and reliability (Furda et al., 2015).
9. Difficulties in making optimal maintenance decisions, due to conflicting decision criteria	1	Large systems, with case study on power plants	17. Difficulties in making optimal maintenance decisions, due to conflicting decision criteria (Furda et al., 2015).
10. Insufficient training of operators/maintainers	3	Industrial settings Industrial systems Naval ships	18. Need for skills and knowledge of maintainers to perform tasks (del Amo et al., 2018). 19. The operator cannot always fix a problem themselves (Lang et al., 2018). 20. Knowledge management; training, knowledge, skills of maintainers (Wahid et al., 2018).
11. Budgeting	5	Industrial settings Large systems, with case study on power plants Gas turbines Naval ships	21. The costs of maintenance should be minimised (del Amo et al., 2018). 22. Maintenance budget constraints (Furda et al., 2015). 23. Timely and expensive process of training maintainers (Martinetti et al., 2017). 24. Costs of maintenance activities (Martinetti et al., 2017). 25. Distribution of maintenance budget, lack of cashflow, and increase in resource prices (Wahid et al., 2018).

12. Complexity of maintenance tasks	5	Industrial settings Industrial systems Naval ships	26. Maintenance processes are getting more complex (del Amo et al., 2018). 27. The complexity of maintenance tasks (Gallab et al., 2015). 28. Complexity of systems (Lang et al., 2018). 29. The complexity of maintenance tasks (Martinetti et al., 2017). 30. Complexity of systems (Wahid et al., 2018).
13. Priority setting of maintenance tasks	1		31. Difficulties with prioritising tasks (Gallab et al., 2015).
14. Availability of Original Equipment Manufacturer expert support	1	Naval ships	32. The availability of the Original Equipment Manufacturer expert support (Wahid et al., 2018).

2.1.3 Discussion

The results of the literature review suggest that there is a limited amount of literature regarding issues with complex systems maintenance at all levels. From the 8 eligible papers identified by the literature review, 14 common issues with complex systems maintenance were found. The most frequent issues mentioned in the literature appear to be related to documentation, the duration of tasks, budgeting, and the complexity of maintenance work.

Regarding the budgeting issue, Lundgren, Skoogh, and Bokrantz (2018) indicate it is difficult for corporations to quantify the benefits of investing in maintenance. The authors found a gap between current maintenance models and their practical applicability, due to the number of models available, as well as unclarity regarding concepts. This means that maintenance, in terms of budgeting, might be reviewed by businesses as an issue, since no evident research has been performed that presents economic profits. On the other hand, Pintelon and Parodi-Herz (2008) suggest that "…the economic implications of maintenance action are comprehended, a direct impact on the maintenance policies is expected" (p. 30). Considering the complexity of maintenance work, this is an observable issue since maintenance systems have become more complex over time: starting off as a simple construction, having turned into a complex entity with many sub-components (Kobbacy & Murthy, 2008). In addition, Kobbacy and Murthy (2008) suggest that maintenance includes different aspects: management, technology, logistics, and operations, making it increasingly complicated to understand. This also results in predicaments for maintenance managers, who must figure out a balance between technical predictions and business promises. The increasing complexity of the systems is also related to the duration of maintenance work: because of the complexity, more interruptions take place, meaning the system must make more stops. This, in combination with the complex design of the system, compels complex repairs which take longer to regulate (Lundgren et al., 2018). Looking at the issue with manuals, years back Chaparro and Groff (2002) found that maintenance manuals in aviation are acceptable for conveying technical information, but lack usability. They also found that research into the satisfaction levels with manuals at companies was mainly focused on solving errors in terms of instructions or depictions, while not considering user-friendly operability of the manual. Some causes of non-usable documentation of manuals are the speed at which maintenance work is performed, which in return influences the financial impact (Chaparro & Groff, 2002). These are all problems mentioned with paper-based manuals. A solution by Jorgensen (1994) has been, and is continuously advancing over the years, which might decrease the issues with paper-based manuals: The Interactive Electronic Technical Manual (IETM) (Jorgensen, 1994). Research showed that complex work, such as troubleshooting, is significantly improved when using the IETM (Jorgensen, 1994). More recently, Su, Liang, Wang, Wang, and Pecht (2019) found that while performing troubleshooting, searching through various interactive manuals is time-consuming and decreases the efficiency with which the work can be performed. They suggested a solution in which the maintainer gets supported by the system, which was tested by a case study. The results show that the maintainers solely have to make a few clicks to get all the relevant information for his troubleshooting work, contributing to faster and more efficient manual searches.

The 14 found issues from the literature review will be expanded by conducting interviews with stakeholders of a Thales B.V. complex system. This will be explained in more detail in the next section (See Section 2.2).

2.2 Interviews and questionnaire with stakeholders from Thales B.V.

Following the literature review, interviews and a questionnaire were performed with stakeholders of TACTICOS working at Thales B.V. TACTICOS is a combat management system that Thales B.V. developed and continuously advances (Thales, n.d.). TACTICOS is a complex system, which consists of many (sub-)components. The stakeholders of TACTICOS are employees at Thales B.V., and have varying roles, knowledge, and experience with designing, creating, testing, and selling a complex system such as TACTICOS. The goal of the interviews and questionnaire was to identify key issues with computerised maintenance management systems (CMMSs) in a complex system (such as TACTICOS), as perceived from the stakeholders' point of view. End-users of a CMMS could not be reached for this part of the study. Therefore, and because Thales B.V. desired to use their employees for the research, the stakeholders were approached. In addition, Thales B.V. deems their stakeholders as experts on maintenance from a strategic, tactical, and operational point of view, and useful in drawing opinions on CMMSs.

Two phases of data collection were applied: (1) interview questions via email with follow-up phone interviews, and (2) an online questionnaire. The results from the interviews and systematic literature review were used to design the online questionnaire from which the extent of agreement on issues with complex systems maintenance and CMMSs amongst the stakeholders was determined, and the key issues were identified.

Part 1: Interviews

The interview phase consisted of a qualitative investigation to identify perceived issues with CMMSs in complex systems from stakeholders' point of view. This was done by conducting interviews with stakeholders working at Thales B.V.

2.2.1 Method

2.2.1.1 Participants

Participants of the study consisted of stakeholders of TACTICOS, working at Thales B.V. The stakeholders have different roles, such as: product managers, human factors specialists, trainers, and maintainer application designers. Throughout the whole research, these roles have been clustered into four roles, to be able to keep a better overview: engineer, manager, human factors specialist/UX designer, and instructor. Purposive sampling was used: 18 stakeholders were approached via email to ask if they were interested in joining the first part of the study. See Appendix B for the email to contact the stakeholders.

In total, 14 stakeholders (*M*ExperienceLevel[3] = 56.9, *SD*ExperienceLevel = 26.4) participated, taking the responses by email and phone calls combined. Most participants were male (71.4%), and most participants work as instructors at Thales B.V. (42.9%). Also, many participants indicated to have over 20 years of experience working with complex systems (42.9%), whereas few participants indicated 0-5 years of experience (14.3%). The concept of 'experience working with complex systems' indicates experience at any level, such as: designing, creating, testing, and/or

[3] Question in questionnaire: 'How experienced would you rate yourself working with maintenance (any kind, e.g., software, hardware, clients, etc.) in complex systems?' 'Working with maintenance in complex systems' is defined as having experience with maintenance at the strategic, tactical and/or operational level in complex systems.

selling a complex system. The age of the stakeholders was not asked, as this was not deemed as relevant information for this study. Rather, participants were asked to rate their experience level with complex systems maintenance at all levels.

An overview of the participants and their demographics can be found in Appendix C. This appendix gives an overview of the participants taking part in all phases of the research (interviews, questionnaire, task analysis, usability testing). In this appendix, participants are indicated by their work role at Thales B.V. A green block means a participant joined in a certain phase of the study.

2.2.1.2 Materials

The following materials were used in the interview phase:

- Open-ended personalised interview questionnaires regarding CMMSs, complex systems, and maintenance, as created in a Word document (See Appendix B). The stakeholders were asked to not focus on a specific system from Thales B.V. (such as TACTICOS) but to generalise their answers to any complex system that perceives issues with CMMSs. The reason for this is that the researcher did not have access to documents specified at TACTICOS, due to COVID-19, and therefore the research was generalised to any complex system. The questionnaires were personalised due to the varying roles of the stakeholders with and knowledge regarding maintenance and complex systems. All interview questionnaires contained at least questions regarding the stakeholders' opinion on current CMMSs, important aspects of CMMSs and suggestions of improvements in CMMSs. Questions asked are for instance: 'What do you think could be improved in maintenance systems/environments?' and 'What is your opinion of current maintenance environments of complex systems (design, ease of use, efficiency of use, understandability etc.)?'.

- Atlas.ti was used for qualitative analysis of the interview responses (v. 8.4.25, ATLAS.ti Scientific Software Development GmbH, 2020).

2.2.1.3 Procedure

The study was approved by the Ethics Committee of the Faculty of Behavioural, Management and Social sciences (BMS) of the University of Twente, with IRB approval code: BCE 200495. A first email was sent to the stakeholders with the personalised interview questionnaires attached in a Word document. After a few weeks, a second email was sent, asking the stakeholders for a phone interview to ask questions about the email responses, CMMSs, and their perceived issues. The phone interviews took between 30 minutes and 1 hour. The interviews were unstructured and due to the exploratory nature of the interviews, other introduced topics deemed relevant by participants were discussed. This resulted in participants also discussing issues with complex systems maintenance at all levels (as similarly done in the literature review). All communication with the stakeholders was done in Dutch. The answers to the questions were used as preliminary information-gathering on issues with CMMSs in complex systems, and, due to suggestions made by the stakeholders, complex systems maintenance at all levels from the stakeholders' point of view.

2.2.1.4 Data analysis

During the phone interviews, notes were taken by the researcher: no recording was made due to the exploratory nature of the interviews. Information from the interviews was analysed qualitatively using Atlas.ti, applying the method of thematic analysis (Braun & Clarke, 2006). The specific method chosen for this part of the study is theoretical, meaning that an already known theme and/or research question was discussed; explicit, meaning that the actual issues 'behind' what has been said was recognised; and realist, which “reports experiences, meanings and the reality of participants” (Braun & Clarke, 2006, p. 81). The answers from the interviews were uploaded into Atlas.ti and read through to get a general overview of the comments. After that, initial codes, identifying issues with CMMSs and complex systems maintenance at all levels were established. The initial codes were merged, based on similarity, and listed as issues.

2.2.2 Results

In total, 18 issues were found by identifying similarity in comments from the participants (See Table 4). Most issues were related to a missing or insufficient aspect with CMMSs such as: a missing workset (5x), a missing global software package (1x), insufficient fault reporting (1x), or no spare parts overview (2x). The most frequently reported issues were ‘no available maintenance workset’, named 5 times by the participants, and ‘no general overview of the (sub-)systems’, mentioned 6 times by the participants. Table 4 presents an overview of the list of issues found with CMMSs in complex systems and complex systems maintenance at all levels from the interviews, the number of times the issues were mentioned, and example quotes from participants to clarify the identified meaning of the issues.

Table 4

Merged issues found with computerised maintenance management systems, and complex systems maintenance at the strategic, tactical, and operational levels from interviews with stakeholders from Thales B.V.; the number of times the issue was mentioned; and example quotes from participants

Issue	**Count**	**Example quote(s)**
1. No available maintenance workset	5	'The missing of an unambiguous maintenance workset caught my attention…' (Participant 3). 'Activities such as the development of a maintenance workset have not been started until now' (Participant 4).
2. No general overview of the whole system + its subsystems	6	'Overview: what is the status of the system or subsystem? … can be improved' (Participant 3). 'It should mainly give a good overview of the complete system … so you can see the status. You should also be able to click through to lower levels of the system' (Participant 5).
3. Missing of global software package	1	'A general global maintenance software package for the complete system is missing, now they are all just loose pieces put together' (Participant 12).
4. Insufficient fault/error reporting	1	'Monitoring and reporting … should be improved' (Participant 4).
5. Insufficient fault/error monitoring	4	'Fault monitoring … is not supported' (Participant 4). 'A maintenance monitoring system is still very limited' (Participant 9).
6. Inefficient and slow fault/error diagnosis	2	'Fault detection is extremely difficult because of the lack of maintenance tools' (Participant 4).
7. No understanding of the impact of a system fault/error	4	'Overview: what are the consequences of an error?… can be improved' (Participant 3). 'It should be clear what the impact of a system failure is' (Participant 10).
8. No logging or documentation of knowledge or occurrences	3	'Very often specific knowledge on the complex system is required to perform good maintenance, which is 'in the heads' of people' (Participant 5). 'It (the system) should make a backlog of everything that has happened' (Participant 12).

9. Manual: too complex (no ease in use, too technical for skills maintainers have)	3	'The sometimes-ambiguous description of procedures in the … manual does not always help maintainers that show to have limited reading skills in English' (Participant 4).
10. Manual: root cause/fault finding unclear	3	'The complexity of the system requires that he (the maintainer) has good manuals …. For pinpointing the root cause of errors…' (Participant 1). '…These procedures are far from complete and do not cover the full system. The part that is lacking is fault finding on system level' (Participant 4).
11. Too many or unclear system alerts	1	'… Instead of a storm of related or unclear errors' (Participant 1).
12. More automation of systems	2	'Many more things can be automated, if the system can do it, let the system do it' (Participant 12).
13. Time management in terms of fault finding	1	'Fault detection is sometimes slow…' (Participant 6).
14. Time management in terms of communication between people/teams	1	'Maintainers walk around the ship a lot; it takes a lot of time and it makes communication between parties difficult' (Participant 2).
15. The downtime of a system	1	'Reduce downtime to a minimum' (Participant 14).
16. Level of difficulty of maintenance tasks	4	'How can we make the maintainer's daily work easier?' (Participant 3). 'The level of schooling of maintainers is much lower than the level of operators' (Participant 7).
17. No clear overview of spare parts	2	'Spare parts: is a certain spare part still available? … can be improved' (Participant 3).
18. No clear calendar for preventive maintenance tasks	1	'Planning system: can a preventative maintenance task be planned easily? … can be improved' (Participant 3).

Part 2: Questionnaire

Building upon the interview phase, in which 18 key issues were identified with computerised maintenance management systems (CMMSs) and complex systems maintenance at all levels, and the literature review in which 14 issues were identified with the latter mentioned, the goal of phase 2 was to determine to what extent the stakeholders from Thales B.V. agreed on the main identified issues with CMMSs and complex systems maintenance at all levels by a questionnaire.

2.2.3 Method

2.2.3.1 Participants

A total of 25 stakeholders, including participants of the interview phase, were invited by email to participate in the questionnaire. In total, 18 responses to the questionnaire were recorded. Of those, three responses were deleted due to having questionnaire progress of only 0%, 4%, and 8% thus no relevant data were collected from these responses. Of the 15 responses left, 12 filled out the complete questionnaire. The other three responses filled out 38%, 58%, and 58%, respectively, therefore these were used in the analysis because relevant data was still collected. Appendix C shows a complete overview of the participants. Most participants (*M*ExperienceLevel = 60.8, *SD*ExperienceLevel = 27.5) that filled out the questionnaire were male (80%) and managers by profession (40%). In addition, most participants have over 20 years of experience working at Thales (46.7%), and over 20 years of experience working with complex systems (66.7%).

2.2.3.2 Materials

A questionnaire regarding the 14 issues identified from the literature review (See Table 3) and 18 problems identified by the interviews (See Table 4) was created in Qualtrics (Qualtrics, 2021) (See

Appendix D). The issues reported in the questionnaire were 28 in total, as 4 issues were merged because of commonality or similarity in the two lists, as interpreted by the researcher (See Appendix E for the merged list of issues). Some issues were rewritten to indicate the meaning behind what has been said or to make the content clear to participants of the questionnaire.

The questionnaire started with an introduction to the thesis and its goal, the way information was gathered in the interview phase, and how this questionnaire could help quantify the already gathered data. It was explained that the survey would take no more than 15 minutes, and that data would be processed anonymously and saved securely. If participants agreed, they could continue by clicking on ‘I agree to the abovementioned informed consent’. After continuing, the questionnaire started with presenting the list of 28 items, which can be identified as the merged list of issues with CMMSs and complex systems maintenance at all levels. Presented with this list, participants were asked ‘On a scale from 1 (strongly disagree) to 7 (strongly agree), how much do you agree the following is a problem when performing tasks with maintenance in complex systems?’. Participants were also able to choose the option ‘I don’t know’. To be able to define problem, a general definition has been considered first: “a *problem* is considered to be a task, a situation, or person which is difficult to deal with or control due to complexity and intransparency” (Seel, 2012, p. 2690). In this research, the focus of ‘problem’ or ‘problematic’ is on tasks or situations difficult to deal with or control, due to the complexity of the system. After the 28 items were presented, five demographic questions were asked: ‘What is your gender?’ (options: male, female, other), ‘What is your role within Thales?’ (open text block), ‘How many years of experience do you have working at Thales?’ (options: 0-5 years, 6-10 years, 11-15 years, 16-20 years, >20 years), ‘How many years of experience do you have working with complex systems?’ (options: 0-5 years, 6-10 years, 11-15 years, 16-20 years, >20 years), and ‘How experienced would you rate yourself working with maintenance (any kind, e.g., software, hardware, clients, etc.) in complex systems?' (rating scale from 0 - 100). At the end of the questionnaire, participants were asked if they had any further comments, which they could indicate in a text box. They were thanked for their participation and told that any further questions or suggestions could be sent to the researcher via email.

2.2.3.3 Procedure

Participants were contacted by email and asked if they wanted to participate in the questionnaire. After two weeks, a reminder email was sent to generate more responses. Participants filled out the questionnaire online on Qualtrics (Qualtrics, 2021) on their own device.

2.2.3.4 Data analysis

Data was exported from Qualtrics (Qualtrics, 2021) to an Excel file. The Excel file was adapted so it could be imported into R (*R-3.6.2 for Windows (32/64 bit)*, 2019). The full code used for the analysis in R can be found in Appendix F. First, in Excel, certain participants were deleted from the analysis based on not having filled out the complete questionnaire. Next, in R, frequency analysis was performed to outline the demographics of the participants. Following, a consensus analysis was applied to analyse the degree to which participants agreed on the 28 items from the questionnaire to be problematic. Median scores of the answers to the 28 items were calculated to identify answers above and below which 50% of the answers fell. This was done by summarising the data in R. Furthermore, interquartile ranges (IQRs) were used to outline the spread of the data, and to determine the level of (dis)agreement between the stakeholders on the items. The IQR is calculated as follows: Q3 - Q1. Looking at studies that have applied consensus analysis; they set their range on agreement/disagreement themselves (See Hassan & Barnett, 2002; Huijben et al., 2019; Polisena et al., 2018). Since there are not many participants in this study, the definition of (dis)agreement is set tight, so a clear division can be noticed. Therefore, the level of agreement in this study is determined as follows: (1) strong agreement: median > 5 and IQR < 2; (2) strong disagreement: median < 3 and IQR < 2; (3) moderate agreement: median < 3 and IQR < 3.

2.2.4 Results

The median and IQR of the 28 items of the questionnaire were calculated, from which the level of agreement amongst the stakeholders can be assessed. Relevant results in this study are items on which stakeholders strongly agree to the issue being problematic when performing maintenance tasks in complex systems. This means that items with a median > 5 and an IQR < 2, imply strong agreement and will be considered in the results. A total of five items were deemed as problematic by the participants (See Table 5). These items are related to fault reporting, the impact of a system error, the manual, spare parts, and priority setting of tasks. The resulting scores of all items can be found in Appendix G.

Table 5

Median scores and interquartile ranges of the most problematic[1] items from the questionnaire

Items	Median	IQR*
1. Insufficient fault/error reporting	6	1
2. No understanding of the impact of a system fault/error	6	1
3. Manual: root cause/fault finding unclear	6	1
4. No clear overview of spare parts	6	1
5. Priority setting of maintenance tasks	6	1

Note. [1]Tasks or situations difficult to deal with or control, due to the complexity of the system, *Interquartile range

2.2.5 Discussion

From the interviews, a list of 18 issues with computerised maintenance management systems (CMMSs) and complex systems maintenance at all levels was generated. The two most frequently mentioned issues were that there is no maintenance workset (x5), and the missing of a general overview of the (sub-)systems (x6). The first issue can be explained by the fact that the stakeholders from Thales B.V. do not seem to be aware of the concept of a CMMS, even though this is a similar notion to what they call ‘maintenance workset’. Next to that, the clear overview of systems is something that the stakeholders relate to the ‘maintenance workset’, and name as a requirement for the workset to contain. Thus, if such a workset or CMMS has been designed and introduced, including a clear system overview, the stakeholders might not perceive this to be an issue anymore.

The results from the questionnaire indicated a list of five main issues as agreed upon by the stakeholders related to fault reporting, the impact of an error, the manual, spare parts, and priority setting of tasks. These issues will be discussed in more detail in the next paragraph. An interesting finding is that the most frequently mentioned issues in the interviews, ‘no maintenance workset’ and ‘no general overview of the (sub-)systems’, as mentioned 5 and 6 times respectively, are not apparent as main issues that were identified from the questionnaire, even though there was a great overlap in participants from the interviews and questionnaire. The reason for this can be that the stakeholders did not agree on these issues being as problematic as all other issues that had to be rated, thus having indicated a ‘lower’ Likert score. In addition, the questionnaire contained a list of merged issues from the literature review and interviews: Participants might have noticed issues in the merged list that they had not considered yet and deemed these as more problematic than issues they had mentioned in the interviews.

The issue regarding no clear overview of spare parts is, as mentioned before, related to the fact that there is currently no CMMS adopted. This means that the stakeholders might not have a perception of what such an overview resembles. Nonetheless, this will be a suggestion in the design of the CMMS prototype in the next phase of the research. Looking at the issue of insufficient fault/error reporting, this can be done inadequately by operators and maintainers, or even by an automated system. When the reporting is performed by operators or maintainers, this process can be improved in the CMMS prototype by making the course of actions to be taken understandable through usable interface design. When fault reporting is done by an automated system, Tagliente, Lospinuso, and DiRosa (2017) mention that, in vehicles such as vessels, ‘smart line replaceable units’ are capable of reporting faults by using the built-in test function. However, this results in substantial amounts of data, which results from burdensome work to find the relevant output, which again leads to lower operational availability. They suggest a vehicle health management system that establishes only the relevant data (Tagliente et al., 2017). Considering the issue of technical manuals, this has been discussed before: Paper-based manuals were found to lack usability (Chaparro & Groff, 2002), but are advancing due to the continuous development of Interactive Electronic Technical Manuals (Jorgensen, 1994). In addition, IETMs were found to be time-consuming to examine, but a solution has been suggested to make the inquiries more efficient (Su et al., 2019). The last main issues to be discussed are related to the understanding of a system error and priority setting of tasks. This can be examined considering reliability centred maintenance (RCM). Reliability centred maintenance (RCM) is an often applied and essential approach to constructing a maintenance plan (Rausand & Vatn, 2008). RCM is defined by the International Electrotechnical Commission (IEC) (1999) as: “systematic approach for identifying effective and efficient preventive maintenance tasks for items in accordance with a specific set of procedures and for establishing intervals between maintenance tasks”. Rausand and Vatn (2008) divided the RCM analysis into 12 steps. In step 4, the functional failure analysis is carried out, where the criticality and severity of system failures are ranked and determined. The severity ranking can have four outcome classes: safety of personnel, environmental impact, production availability and economic losses. These are several types of impacts and consequences a system failure can have. The severity ranking is calculated in combination with criticality ranking, which is done based on the frequency of the functional failure of a system. Both frequency and criticality are ranked on low impact, medium impact, or high impact (Peeters, Basten, & Tinga, 2018; Rausand & Vatn, 2008). Items that are scored on medium or high should be constructed to further analysis. Following, from step 6 on, a failure mode, effects and criticality analysis (FMECA) is performed from which the foremost goal is "the early identification of all catastrophic and critical failure possibilities so they can be eliminated or minimized through design correction at the earliest possible time" (Department of Defense, 1980, p. iii). The results of the FMECA lead to a risk priority number (RPN) that can range from 1 (best) to 1000 (worst) (Braaksma, 2012; FMEA-FMECA, 2006). From the RPN, maintenance actions that need to be taken can be prioritised. Thus, as can be seen, the issues relating to impact and prioritisation can be improved and understood when performing an RCM, which takes maintenance management through the analyses step by step.

In the next phase of the research, a CMMS prototype interface will be designed, in which the resulting five main issues will be resolved in the design.

Phase 2. Design of a CMMS prototype: A case study for Thales B.V.

In this phase of the research, a case study was performed for Thales B.V., in which a low-fidelity computerised maintenance management system (CMMS) prototype interface for performing maintenance tasks in vessels (ships) complex systems was designed. In this study, the complex system focus of the prototype is TACTICOS. However, due to COVID-19, the researcher was unable to visit Thales B.V. to get specific information and documents on the complex system, which resulted in some parts of the system having been interpreted as a general complex system. The requirement from Thales B.V. for the prototype is what they call a 'maintenance workset', which indicates a role-specific graphical interface design consisting of several components, which is to be used in combination with certain keyboard buttons (T. Nijland, personal communication, March 25, 2021). For this research, only the graphical interface will be considered without the keyboard buttons, and 'maintenance workset' has been translated to a CMMS.

While designing the prototype, multiple aspects have been taken into account, and applied:

1. *Requirements* regarding colour scheme and interface components from Thales B.V. These have been identified from a former, similar work done on a maintenance workset for Thales B.V. (de Noord, 2019).
2. The *results* from phase 1 of the research regarding the *5 main issues* with CMMSs and complex systems maintenance at all levels. Issues that can be resolved in the design of the prototype have been considered. The issue regarding insufficient fault reporting has been resolved by giving the operator and maintainer an easy-to-follow issue report screen that they can fill out. The issue regarding the manual has been solved by integrating an eManual/IETM in the system, which makes sure

maintainers can inquire the troubleshooting procedure faster. Following, the issue of no clear overview of spare parts has been solved by the supervisor having an overview of the current spare parts, and the amount thereof. Lastly, priority setting of maintenance tasks has been solved by given the supervisor insight into the urgency the operator assigned to the issue; based on that, it can be easier to prioritise tasks. The issue regarding no understanding of system impact has not been integrated or solved in the prototype.

3. Some of the most commonly used *interface design guidelines and principles* (Nielsen, 1994b), and *MIL-STD-1472G* (Department of Defense, 2012). The guidelines and military standards considered in the prototype will be discussed shortly below (See Section 2.3).
4. *A performed hierarchical task analysis* for operators and maintainers. The task analysis was carried out by identifying the representativeness of, and main problematic tasks for operators and maintainers performing maintenance tasks in complex systems. The method and results of this part will be explained in more detail below (See Section 2.4).

2.3 Interface design principles

Design principles or guidelines aim to prevent system users to make slips (unintended errors), or mistakes (intended errors) and to recover easily from these errors made (Johnson, 2021). Over the years, many lists of design principles and guidelines have been established by several researchers (See for example: Johnson, 2021; Norman, 1983a, 1983b; Stone, Jarrett, Woodroffe, & Minocha, 2005). For the design of the CMMS prototype in this study, it has been chosen to take the 10 usability heuristics by Nielsen (1994b) into consideration while designing the prototype. The reason for this is that the research done by Nielsen (1994a) refined the list of heuristics based on a factor analysis of 249 usability problems. The results of his research indicated a list of 10 heuristics to be considered in system interface design: (1) visibility of system status, (2) match between system and the real world, (3) user control and freedom, (4) consistency and standards, (5) error prevention, (6) recognition rather than recall, (7) flexibility and efficiency of use, (8) aesthetic and minimalist design, (9) helping users recognize, diagnose, and recover from errors, and (10) help and documentation (Nielsen, 1994b). Some examples of how certain guidelines have been used are: a calendar widget has been integrated into the prototype, which adheres to the 'match between system and the real world', since users are most likely familiar with a similar type of calendar widget, such as Google Calendar. Error prevention has been applied by asking an extra verification question before operators and maintainers can send in an issue report. Aesthetic and minimalist design has been included by presenting an overview screen which has minimal text and components, to not make users get distracted by information overflow.

In addition to general interface design guidelines, principles of the MIL-STD-1472G have been taken into consideration in the prototype design. This military standard aims "to present human engineering design criteria, principles, and practices to be applied in the design of systems, equipment, and facilities so as to: a. Achieve required performance by operator, control, and maintenance personnel, b. Achieve required manpower readiness for system performance, c. Achieve required reliability of personnel-equipment combinations, and d. Foster design standardization within and among systems" (Department of Defense, 2012, p. ii). The reason for also using principles from this standard is that the system interface considered in this case study aims to be suited during (military) vessels missions for both operators and maintainers. The principles applied are regarding colour scheme, which indicates that cool colours (blue, green) should be used to "display information used infrequently and to convey status of background information" (Department of Defense, 2012, p. 91), whereas warm colours (red, orange) should be used to "convey action or the requirement for a response" (p. 91). In detail this means the following colours should be applied in the following cases (p. 91):

- Red: OFF, failure/error, stop, critical condition
- Green: ON, acceptable, ready/proceed
- Yellow: warning/impending danger, abnormal state
- Orange: alert/hazard, the value between red and yellow

In this prototype, the colour principles have been applied to, for example, the system overview component of the prototype by giving systems with critical failures a red colour, systems with mild failures an orange colour, and functional systems a green colour. In addition, the colour blue has been used in the prototype to display informative messages.

The application of the heuristics and principles can be found in Appendix H in which the prototype interface is explained.

2.4 Hierarchical task analysis research

To aid in the design of the CMMS prototype interface, a hierarchical task analysis on maintenance tasks performed by operators and maintainers working with complex systems was performed, to understand the flow through which the stakeholders that participated would operate the system interface designed for the case study. The method of hierarchical task analysis by Annett (2003) aims to "identify actual or possible sources of performance failure and to propose suitable remedies, which may include modifying the task design and/or providing appropriate training" (p. 18). In this research, the task analysis intends to identify the representativeness of the depicted tasks, and to identify problematic tasks for operators and maintainers, so as to resolve these problematic tasks by making the system aid the user in the CMMS prototype interface. As mentioned before, in this research, the definition of 'problematic' is tasks or situations difficult to deal with or control, due to the complexity of the system.

In this part of the study, the stakeholders had to identify themselves with operators and maintainers to evaluate which tasks are commonly performed by them. Even though this is unusual Human Factors practice, since actual operators and maintainers are experts in the tasks they perform, Thales B.V. desired to use their own employees throughout the research. In this part of the study the same group of stakeholders participated as in the interviews and questionnaire in phase 1 of the research. First, exploratory interviews were conducted to identify and confirm the tasks, followed by a questionnaire in which representativeness of tasks and problematic tasks were determined and identified, respectively. The aim of the operator task analysis is for the operator to be able to report an issue; the aim of the maintainer task analysis is for the maintainer to be able to solve an issue, both by using the CMMS interface. The reason for these goals is that these are the main functions the users should be able to perform when using the system (as indicated in the former work by de Noord (2019)).

2.4.1 Method

2.4.1.1 Participants

The same participants as in phase 1 (interviews and questionnaire) of the research took part. To summarise the interview participants: 14 stakeholders (*M*ExperienceLevel = 56.9, *SD*ExperienceLevel = 26.4) participated. Most participants were male (71.4%), and work as instructors at Thales B.V. (42.9%). Almost half of the participants indicated to have over 20 years of experience working with complex systems (42.9%), whereas few participants only had 0-5 years of experience (14.3%). To summarise the participants involved in the questionnaire: 15 stakeholders participated. Most participants (*M*ExperienceLevel = 60.8, *SD*ExperienceLevel = 27.5) that filled out the questionnaire were male (80%) and managers by profession (40%). In addition, most participants have over 20 years of experience working at Thales, and experience working with complex systems (46.7% and 66.7%, respectively). See Appendix C for an overview of the participants that joined.

2.4.1.2 Materials

The following materials were used while carrying out the hierarchical task analysis research:

- Mindomo for the visualisation of the task analyses[4] of the operator and the maintainer. This program is used for, amongst other services, creating mind maps (Mindomo, n.d.). Therefore, it was also a useful program to make the depictions of the task analyses.

[4] The plural 'task analyses' is used, since two task analyses have been carried out for the maintainer and operator, separately.

- An email with questions sent to the stakeholders. These were sent before conducting the phone interviews so the stakeholders could prepare. The email included the first draft task analyses, an explanation of the topic of the thesis, an explanation on how to read the task analyses, six questions about the task analyses for feedback, and a suggestion for a phone interview (See Appendix I).
- Atlas.ti was used for qualitative analysis of the interview responses (v. 8.4.25, ATLAS.ti Scientific Software Development GmbH, 2020).
- Qualtrics (Qualtrics, 2021) was used to create a questionnaire which contained eight questions concerning the task analyses, and five demographic questions (See Appendix J for questionnaire). The questionnaire aimed to assess to what extent the stakeholders agree on the representativeness of tasks, and to identify problematic tasks. The questionnaire included the same introduction and informed consent as previously described in the questionnaire in phase 1 of the research. Following, the operator task analysis was depicted and explained with text. Question 1 (Q1) listed all depicted tasks and asked, ‘On a scale from 1 (strongly disagree) to 7 (strongly agree), how much do you agree with the current process being representative of the tasks an operator performs when reporting an issue (as depicted above)?’. Participants were also able to choose the option ‘I don’t know’. Next, it was asked whether the participants missed any important tasks, and if so, they could report which ones in a text box (Q2). Following, participants were asked to group the operator tasks as problematic or not problematic, and to rank those tasks in the different groups from most (un)problematic to least (un)problematic (Q3). In addition, it was asked why participants found the three highest problematic ranked tasks problematic (Q4). The next section of the questionnaire was the same, however, this time for the depicted maintainer task analysis, with the starting question (Q5): ‘On a scale from 1 (strongly disagree) to 7

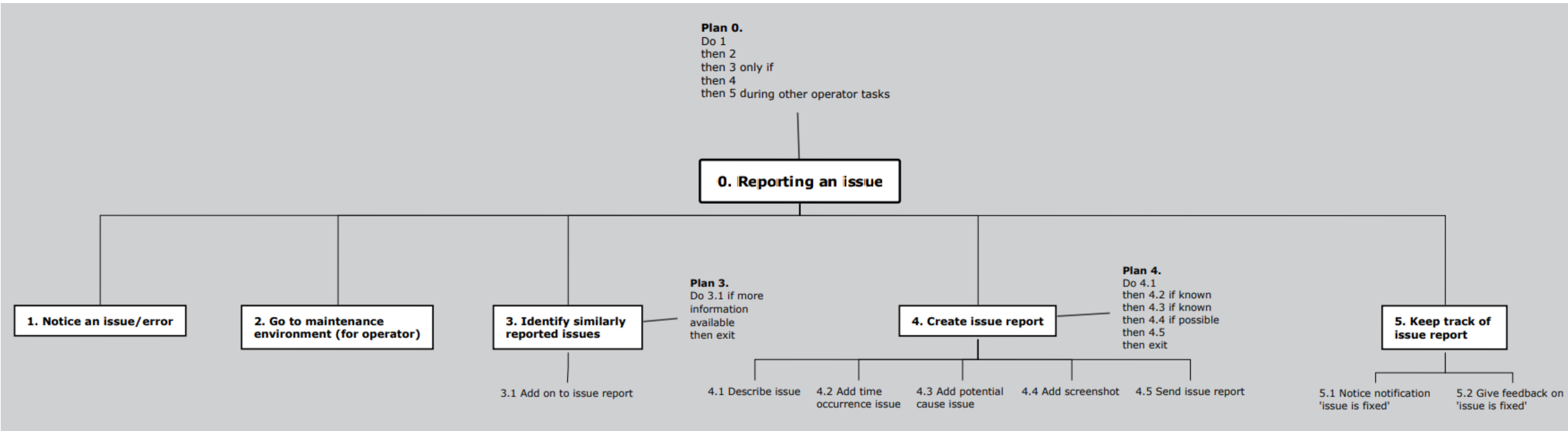
(strongly agree), how much do you agree with the current process being representative of the tasks a maintainer performs when performing general maintenance tasks (as depicted above)?'. Following, it was asked whether any important tasks were missing (Q6); the grouping and ranking of (un)problematic tasks (Q7); and why tasks were deemed problematic (Q8). After the task analyses questions, demographic questions were asked similar to the five questions from the questionnaire in phase 1 of the research regarding gender, role within Thales, years of experience working at Thales, years of experience working with complex systems, and rating of experience level of working with maintenance in complex systems. At the end of the questionnaire, participants were asked if they had any further comments, which they could indicate in a text box. They were thanked for their participation and told that any further questions or suggestions could be sent to the researcher via email.

2.4.1.3 Procedure

A first draft task analysis for operators and maintainers separately was carried out by the researcher (See Appendix K). The drafts were based on a similar case study done by de Noord (2019) in which he created a workflow for operators and maintainers who perform maintenance tasks on TACTICOS. The first drafts were used in the interviews, to ask for feedback from the participants. After applying the feedback from the interviews, a second draft was visualised and used in the questionnaire (See Figures 3, 4 and 5). The questionnaire was distributed amongst the stakeholders, who filled this out on their own device.

Figure 3

Operator task analysis



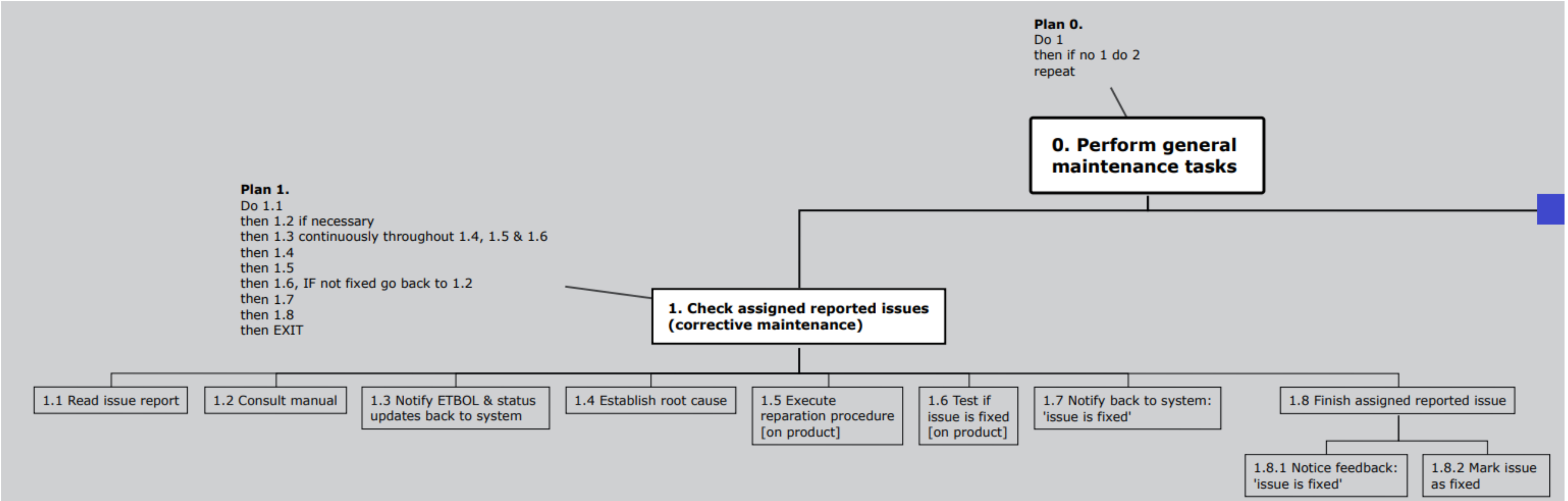


Figure 4 (top)

Maintainer task analysis (part 1): Figures 4 and 5 are a connected picture but have been cut in two in order to be able to read the text on the picture. The blue coloured blocks indicate where the figures should be attached.

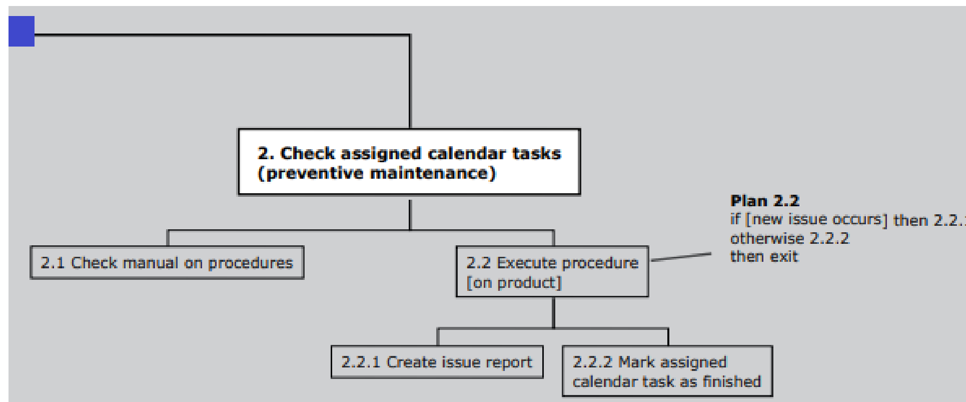


Figure 5 (bottom)

Maintainer task analysis (part 2)

2.4.1.4 Data analysis

Data was exported from Qualtrics (Qualtrics, 2021) to an Excel file. The Excel file was adapted so it could be imported into R (*R-3.6.2 for Windows (32/64 bit)*, 2019). The full code used for the analysis can be found in Appendix F. First, in Excel, certain participants were deleted from the analysis based on not having filled out the complete questionnaire. Next, in R, frequency analysis was performed to outline the demographics of the participants. Following, similar to the analysis used in the questionnaire in phase 1 of the research, consensus analysis was applied to analyse questions 1 (agreement on the representativeness of tasks of operators) and 5 (agreement on the representativeness of tasks of maintainers) (See Appendix J). To summarise the analysis: median scores of the items were calculated by summarising the data in R and interquartile ranges (IQRs) were used to determine the level of (dis)agreement between the participants on the items. The level of agreement in this study is determined as follows: (1) strong agreement: median > 5 and IQR < 2; (2) strong disagreement: median < 3 and IQR < 2; (3) moderate agreement: median < 3 and IQR < 3.

Next, the answers to questions 3 and 7 were analysed (ranking and grouping of (un)problematic operator and maintainer tasks, See Appendix). Relevant data taken from these questions are the tasks that have been identified by the participants as problematic, therefore, only the data on these tasks was analysed. The number of times a task was grouped as problematic was counted in Excel for the maintainer and operator separately. Next, in R, mean ranks per task grouped as problematic were calculated, to identify the three highest ranked problematic tasks for operators and maintainers. A mistake was noted in the questionnaire while analysing the data. In question 7 (grouping and ranking of maintainer tasks), the two items ‘None of the tasks is problematic’ and ‘I do not have enough knowledge on this topic to answer this question’ were not taken out, which should have been the case. Therefore, for every participant, these items were taken out of the grouping data. In addition, the rank these items were placed on was deleted, and all other items were moved up by one place in the rank. This way, the ranking remained the same, and the two irrelevant items were not taken into account.

Lastly, to analyse questions 2 and 4 regarding missing tasks, and questions 6 and 8 regarding an explanation on why tasks are problematic, Atlas.ti was used (See Appendix J for questions). Thematic analysis was applied to identify similarities in responses on problematic tasks and identified missing tasks for operators and maintainers (Braun & Clarke, 2006). The specific method chosen for this part is the same as previously applied.

2.4.2 Results

First, the extent of agreement on the measures for tasks will be discussed, followed by the commentary questions.

2.4.2.1 Level of agreement on the representativeness of tasks and problematic tasks for operators and maintainers

The median and IQR of the items of question 1 and 5 (representativeness of depicted operator and maintainer tasks, See Appendix J) were calculated, from which the extent of agreement between participants on the depicted tasks as being representative could be assessed. Relevant results in this study are those items on which stakeholders strongly or moderately disagree with being representative since these identified tasks might be left out of the prototype interface. This means that the items with a median < 3 and IQR < 2 or median < 3 and IQR < 3, will be considered in the results. The results show that only task 3 for the operator: ‘Identify similarly reported issues’, has a high level of disagreement on being representative (See Table 6[1]). This task explains the following: when operators report an issue, they first identify whether there are any issues similar to the issue they want to report, so as to not overwhelm maintainers with many similarly reported issues.

Question 3 and 7 of the questionnaire (ranking and grouping of (un)problematic tasks, See Appendix J) were analysed by counting the number of times a task was grouped as problematic: the three tasks with the highest count are indicated in Table 6[2]. In addition, question 3 and 7 were analysed by calculating the mean rank per task that has been grouped as problematic, from which the three highest-ranked problematic tasks for operators and maintainers were identified (See Table 6[3]). In total, five tasks were identified for the operator as problematic and four tasks for the maintainer. The first identified problematic task for the operator is noticing an issue, which means the operator is working behind his console and notices one of the systems important to his work has an error. Following, the operator can go to the maintenance environment, and from there report the error, which is the second identified problematic task. The third task has been explained before, which is identifying similarly reported issues. Also, describing the issue is deemed problematic, which is a form in which the operator explains what the error is, and possibly why this is a problem to him. The last identified problematic task is adding a potential issue cause, which indicates that the operator might have experience with similar issues, and therefore recognise what the solution to the error can be. For the maintainer, one of the four identified problematic tasks include checking the assigned reported corrective maintenance (CM) issues, which is where the maintainer sees which CM tasks he has to perform. Following, notifying estimated time back online and other status updates back to the system, which is where the maintainer gives updates on the repair of the issue to stakeholders, so everyone has an overview. Establishing a root cause was also deemed problematic, which is where the maintainer determines what the root cause of the system error is.

The last identified problematic task is testing whether the issue is fixed, which is where the maintainer performs tests to see if the system error is solved. The complete results of all items can be found in Appendix L.

Table 6

Median scores, interquartile ranges, counts and mean ranks of representativeness and identified problematic tasks for operators and maintainers. The tasks are numbered according to the depiction in the hierarchical task analysis

Operator tasks	Median*	IQR**	Count***	Mean rank****
1. Task 1: Notice an issue/error[3]	6	1	3	0.684
2. Task 2: Go to maintenance environment (for operator)[2]	5	1	9	2.263
3. Task 3: Identify similarly reported issues[1,2,3]	2.5	1.75	14	1.543
4. Task 4.1: Describe issue[3]	6	4	4	1.649
5. Task 4.3: Add potential cause issue[2]	3	3	10	2.544
Maintainer tasks	Median*	IQR**	Count****	Mean rank***
1. Task 1: Check assigned reported issues (corrective maintenance)[3]	6	1.5	4	1.000
2. Task 1.3: Notify Estimated Time Back Online & status updates back to system[2,3]	4.5	3.25	8	1.000
3. Task 1.4: Establish root cause[2,3]	5.5	3	11	1.000
4. Task 1.6: Test if issue is fixed [on product][2]	6	1	5	1.546

Note. *Median score on agreement of representativeness, **Interquartile range score on agreement of representativeness, ***Count in grouped as problematic, ****Mean rank in problematic task list, [1]agreement on unrepresentative task, [2]highest count problematic task, [3]highest rank problematic task.

2.4.2.2 Comments

To answer questions 2 and 6 regarding missing tasks (See questions 2 and 6, Appendix J), and questions 4 and 8 asking about an explanation on why the task is deemed problematic (See questions 4 and 8, Appendix J), Atlas.ti was used for the analysis. Several categories of comments were identified per question: two categories of answers were established for the questions on missing tasks for the operator and maintainer, namely: 1) missing tasks, and 2) general suggestions (Q2, Q6). Three categories were identified per explanation why the task was problematic. For the operator: 1) non-existent tasks, 2) identifying similar issues, and 3) no knowledge of potential error cause (Q4). For the maintainer: 1) non-existent tasks, 2) difficulty of the tasks, and 3) no aid of the system (Q8). Appendix L gives an overview of the identified categories per question; a short explanation of the definition of the categories; and summarised comments as given by managers, engineers, instructors, and human factor specialists/UX designers.

Considering question 2, a total of three different tasks were suggested to be missing in the operator task analysis, for example: Task 5 should have included ‘estimated time back online’ (ETBOL). In addition, a total of six different suggestions were given: A suggestion given twice is aiding the operator with the issue reporting process by using a smart Q&A system or automating parts of the process. For question 6, a total of six missing tasks were recognised for the maintainer task analysis, for example the fact that the maintainer should ask for approval to reset or switch off a system. A total of three suggestions were given, such as: A root cause analysis cannot always be done before repairing the system. Considering questions 4, a total of three different non-existent tasks were identified for the operator. It was mentioned twice that there is currently no issue report process in place, as well as no existing maintenance environment. In addition, for an operator to identify similar issues is suggested to be too time-consuming, difficult, misleading, or doubtful, as mentioned four times by the participants. Three different comments were made on the category that the operator has no knowledge of a potential error cause, in which it was mentioned twice that the operator has no understanding or skills to do so. For question 9, a total of three tasks were suggested to be non-existent currently for the maintainer, namely: performing a root cause analysis, the issue report process, and the ability to determine ETBOL if the root cause is unclear. Lastly, a total of five tasks were suggested to be difficult to perform for the maintainer in which it was mentioned twice the task to perform a root cause analysis, as well as the maintainer not having enough skills and knowledge in general.

2.4.3 Discussion

The analysis of the tasks was used to assess the level of agreement amongst stakeholders on the representativeness of the depicted tasks for operators and maintainers, and to identify problematic tasks. The results show there is only one task in the operator task analysis that is not representative, namely: ‘identify similarly reported issues’. This means, according to the stakeholders participating in this research, this task is not usually performed by operators. Therefore, it will be considered to leave the task out of the task flow in the prototype. However, the participants are no experts on the performance of the tasks, nor end-users of the system, and a final conclusion on leaving out tasks based on the results found is not a reliable choice. Nonetheless, as indicated before, Thales B.V. deems the stakeholders to be experts on knowledge and experience with complex systems maintenance. Therefore, and because the prototype is the company’s case study, the task can be left out, but this is regretful Human Factors practice. In terms of problematic tasks, five were identified for the operator, and four for the maintainer. For the operator, it seems problematic to be involved in the issue reporting process. For the maintainer, the tasks related to ‘estimated time back online’ and root cause analysis are most problematic, which agrees with the commentary feedback given by the participants. The problematic tasks identified in the task analysis research have been considered in the design of the CMMS prototype. For example, task 4.1: ‘Describe issue’, was solved by giving the operator the ability to describe the issue they are experiencing in a report issue screen. In addition, task 1: ‘Check assigned reported issues (corrective maintenance)’, was solved by providing the maintainer with a list of tasks he can work on, from which he can check the assigned issues. Appendix H gives an overview of all tasks that have been solved in the interface.

Some of the identified tasks from this research can be compared to a study performed by Schaafstal, Schraagen, and van Berlo (2000). In this study, four sub-tasks to perform troubleshooting for technicians were suggested: 1) formulate problem description, 2) generate causes, 3) test, and 4) repair and evaluate (p. 79). When comparing these tasks to the identified tasks in the task analysis of the maintainers, overlap can be found: sub-task 2, generate causes, can find similarities with ‘establish root cause’. However, as also mentioned by participants of the study, root cause analysis is a difficult task to perform for maintainers, so it can be suggested to change the task to ‘establish cause’, which is a different practice from performing root cause analysis. Task 3 and 4 as suggested by Schaafstal et al. (2000), have substituted places in the task analysis carried out in this study. In this study, it is first suggested to perform a reparation procedure, followed by testing. Although it can be suggested that, while performing reparation procedures, in between testing might be necessary to identify whether parts of the system are functional again. As can be seen, even though comparisons can be found between the identified tasks in case of performing maintenance work, future research can focus on creating a task depiction for maintainers and operators working with complex systems maintenance that concerns commonly performed tasks. This depiction can be useful in recognising problematic tasks or procedures, and improving the task flow, as is the goal of hierarchical task analysis (Annett, 2003).

2.5 Low-fidelity CMMS prototype explanation

After the 2 phases of data collection, the researcher designed a computerised maintenance management system (CMMS) prototype in Adobe XD, focused to a great extent on TACTICOS (v. XD31, Adobe, 2020). The prototype has been divided into three parts: the maintenance workset (general overview), the operator part, and the maintainer part (See Figures 6, 7, 8). The maintenance workset can be used by someone taking on a supervisory role and wanting to keep an overview of incoming/current/solved issues, the system, the maintainers, maintenance planning, and inventory of spare parts; the operator part by operators that need to report an issue; and the maintainer part by maintainers performing maintenance tasks. As can be seen, the identified tasks of the maintainer in the task analysis research have now been divided into two roles: the supervisor and the maintainer. The reason for this is that we found that too many tasks were not performed by the maintainer but by a 'higher-level' maintainer that keeps an overview. Thus, the tasks that identify with a supervisory role will be selected for the supervisor. The interfaces for the supervisor and operator are specified for a static device such as a computer screen, while the maintainer interface is designed for a dynamic device such as a tablet. The reason for this is that maintainers on vessels usually change location a lot, and this way they can keep the information with them at all times. Figures 6, 7 and 8 give an overview of the initial interfaces for the CMMS for the maintenance workset, operator part and maintainer part, respectively. The complete CMMS prototype interface design explanation can be found in Appendix H. This appendix gives an overview of the components, notes on which identified issues from phase 1 have been resolved, and notes on which identified problematic tasks from phase 2 have been resolved. The interface is also made interactive: the link to the interactive screens can be found in the link in Appendix H. Some parts of the interface are blurred due to including classified data from Thales B.V.

Figure 6

Maintenance workset (general overview)

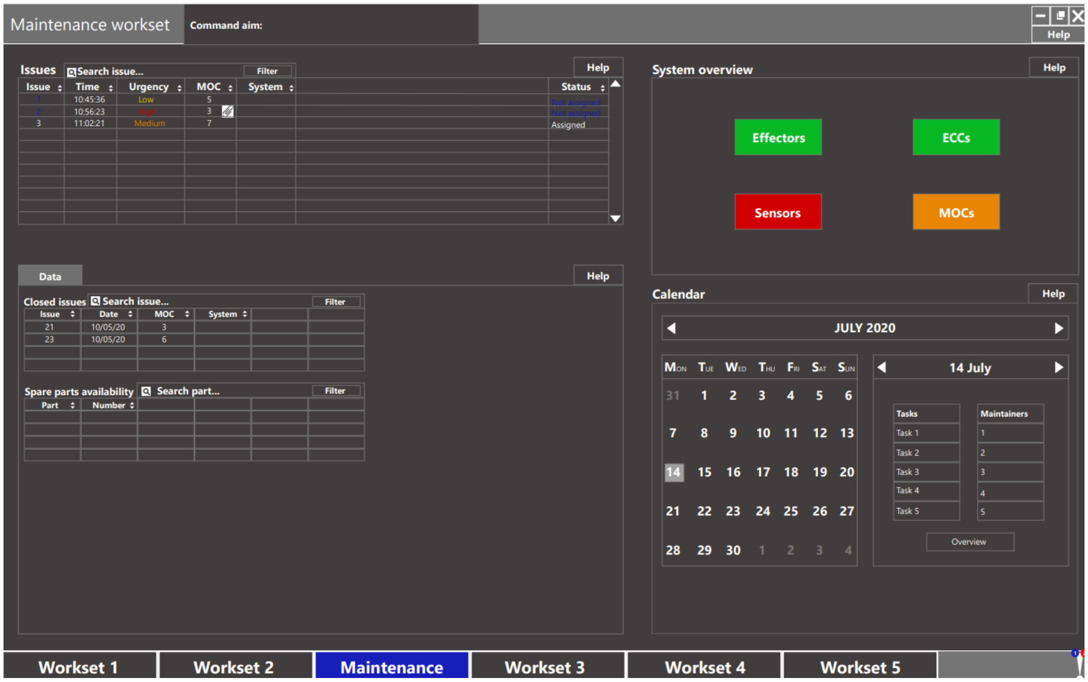


Figure 7

Operator part

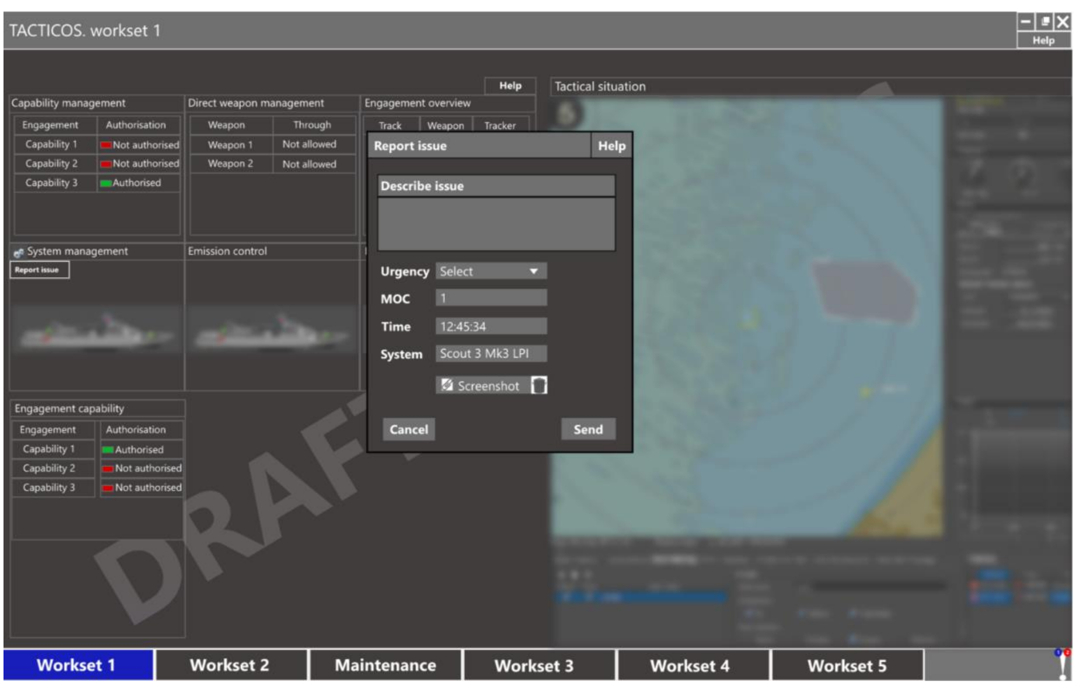
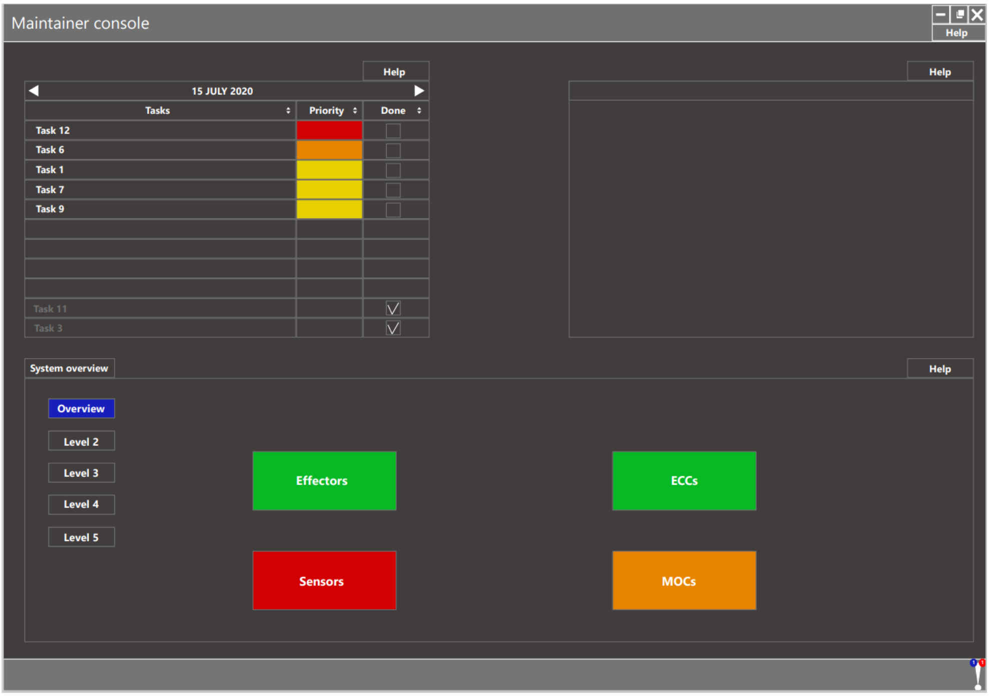


Figure 8

Maintainer part



Phase 3. Usability evaluation of the CMMS prototype

2.6 Usability testing

The designed low-fidelity CMMS prototype interface was tested using small-scale, remote usability testing. All three parts of the interface: the maintenance workset (general overview), the operator part and the maintainer part were explored. Quantitative performance measures, as well as qualitative feedback on the design were recorded. The interface of the prototype was made interactive, so participants could click on certain buttons in the interface, which would lead them to different screens (See Appendix H for the link). Due to COVID-19, usability testing was performed online, using Google Meets, screen sharing, and screen recording.

2.6.1 Method

2.6.1.1 Participants

An email was sent to 23 stakeholders working at Thales B.V., who were invited to participate in the usability testing. All stakeholders that received an invite to the usability testing had already been contacted before in other phases of the study. A total of 11 participants joined the usability tests (*M*ExperienceLevel = 58.6, *SD*ExperienceLevel = 30.5). The participants that joined were mostly male (81.8%), and over half of them had over 20 years of experience working with complex systems (54.5%). Most participants were either engineers (36.4%) or instructors (36.4%). Appendix C shows the participants that joined the usability tests.

2.6.1.2 Materials

For the usability tests the following materials, measures and scales were used:

- *Google Meets*: to call participants remotely for the usability testing.

- *Feedback on the interface components*: A combination of retrospective and concurrent think-aloud was applied. The former means participants perform a task, and afterwards share their thoughts about this task (van den Haak, De Jong, & Jan Schellens, 2003). This will make sure the participant is able to focus on the task at hand only, thus it will not interfere with their cognitive workload. The latter means that the participant can also share their thoughts during the performance of the task. Retrospective and concurrent think-aloud were applied in combination due to the expectation that participants would not only focus on the task at hand. This expectation came from experiences with talking to this set of participants in former parts of the study. Part of the feedback on the interface was also three questions asked by the researcher to get a general idea of participants' preference on certain components. These questions were:

  - Maintenance workset: in the system overview component, would you rather want to see a functional block overview (schematic depiction), or a picture of the vessel?

  - Operator part: do you think that this is the correct place for the 'report issue' button?

  - Operator part: do you think that the extra verification question is necessary when sending/cancelling the issue report?

- *Performance measures on tasks*: Scenarios were depicted and tasks were performed by the participants (See Table 7). For all tasks, the following performance measures were used:

- *Success rate*: task completed yes/no.
- *Error rates:*
    - Critical errors: the participant is not able to finish the task at all (usability.gov, n.d.).
    - Non-critical errors: the participant is able to finish the task in a less efficient manner (usability.gov, n.d.).

Table 7

Performed tasks usability testing per component and per part of the interface

Interface part + components	**Tasks**
Maintenance workset (general overview)	
Issue report screen	1. You want to get a better overview of ‘Issue 2’, you want to assign this issue to ‘maintainer 5’, and after some time (which you don’t have to take now) you want to give feedback that the spare parts to be used are available.
System overview	2. You want to get a better overview of the ECC’s.
Dynamic screen	3. You are looking for ‘closed issues’, because you want to get a better overview of ‘issue 21’.
Calendar	4. Today is not the 14th of July anymore, but the 15th. You want to switch to this day and get a detailed overview of the tasks to be done, and the available maintainers.
Notifications	5. You want to see an overview of the notifications.
Operator part	
Button report issue and Report issue screen	1. Imagine being an operator working with this complex system, and you notice your map is not working correctly. So, you want to report an issue, using the system.
Button report issue and Report issue screen	2. Imagine being an operator working with this complex system, and you notice your map is not working correctly. So you want to report an issue, using the system. However, when filling out the report, the map is working fine again, so you cancel the issue report.
Maintainer part	

Tasks overview and Troubleshooting form	1. You want to start working on task 12. You want to get a better overview of the task, you start reading the issue report, and want to set ETBOL* to a certain time. After that you want to check the manual, and you go to part 3 of the troubleshooting form. You realise that the incorrect attributes have disappeared, so you are done performing the task.
Tasks overview and Troubleshooting form	2. You want to start working on task 12. You want to get a better overview of the task, you start reading the issue report, and want to set ETBOL to a certain time. After that you want to check the manual, and you go to part 3 of the troubleshooting form. You realise that the incorrect attributes have NOT disappeared, so you go back to part 1 of the troubleshooting form.

Note. *Estimated time back online

- A Qualtrics survey containing scales that measured subjective experiences of the participants (Qualtrics, 2021) (See Appendix M):
  - Subjective satisfaction levels of the system's usability using the *System Usability Scale* (SUS) (Brooke, 1996). This scale consists of 10 items, as measured on a Likert scale from 1 (strongly agree) to 5 (strongly disagree). The calculation of the scores is as follows: first, the scores to items 1, 3, 5, 7, 9 have to be adapted to 'scale position – 1', followed by the scores to the items 2, 4, 6, 8, 10 which have to be adapted to '5 – scale position' (Brooke, 1996). Next, the sum of the item scores is calculated and multiplied by 2.5 to get the end-score. A SUS-score of 68 can be seen as average (Sauro, 2011a, 2011b). The scores can range from 0 – 100. The SUS is a valid measure and gives reliable results for small sample sizes (Sauro, 2011a).
  - *Net Promoter Score* (NPS) (Reichheld, 2003). This is a single question that asks if the participant would recommend product X to a friend or colleague. The question is measured on a scale from 0 (not at all likely) to 10 (extremely likely). The scores can be divided into three categories of responses: 'Detractors' (0-6), 'Passively

satisfied' (7-8), and 'Promoters' (9-10). The outcome is calculated by subtracting the percentage of Detractors from the percentage of Promotors (Gitlin, n.d.). The resulting score ranges from -100 to 100, and can be compared to other or similar industries' benchmarks (Barnum, 2020). The outcome of this question relates to the customer's loyalty towards the company, (Fisher & Kordupleski, 2019), in which a score of +35 indicates average (Gitlin, n.d.).

- *Demographic questions:* similar to the five questions asked in the questionnaire in phase 1 and 2 of the research regarding gender, role within Thales, years of experience working at Thales, years of experience working with complex systems, and rating of the experience level of working with maintenance in complex systems.

2.6.1.3 Procedure

Participants that signed up for one of the available timeslots were sent an email with a link to a Google Meets call with the researcher. At the beginning of the usability test, the participant was given a Qualtrics link to the questionnaire (Qualtrics, 2021), which contained the explanation of the study (in English), and informed consent. When agreeing to the informed consent, the participant had the chance to ask any remaining questions. Following, the researcher asked for permission to record the meeting, so the researcher could focus on the procedure of the usability test and analyse the recording later. In addition, the participant was asked to share their screen, so the researcher could see exactly what was happening, and follow the participant's movements and clicks in the interactive interface. The next screen on the Qualtrics questionnaire indicated a link to the CMMS prototype, which started with a white screen. The participants had to click once more on the white screen to see the 'maintenance workset (general overview)' interface, which is the screen that all usability tests started with. Next, the participant was given two tasks in the maintenance workset, one task in the operator part and one task in the maintainer part: thus, every participant got four tasks in total. The exact script and division of tasks amongst participants can be found in Appendix N. After each task per part of the prototype interface, general feedback could be given on the different components in the interface. So, for example, when the participant finished the two tasks on the maintenance workset screen, the researcher would ask feedback on each component of that screen (issue reports, system overview, dynamic part, calendar, notifications). After all tasks on the prototype were finished, and participants were done giving feedback, screen sharing was stopped, the recording was stopped, and participants could continue filling in the last part of the Qualtrics questionnaire consisting of the SUS, the NPS and the demographic questions. When the usability test was finished, participants were thanked for their participation, and told they would be kept updated on the progress of the thesis. All usability testing was performed in Dutch and took between 45 minutes and 1 hour. During some usability tests, due to time constraints or extensive feedback on the components, certain tasks could not be performed, or feedback could not be asked for on every component of the prototype. Therefore, not all components have an equal amount of feedback or measures on the performed tasks.

2.6.1.4 Data analysis

Feedback on the interface was categorised by the name of the components of the interface. The feedback on these components was qualitatively analysed in three drawn up codes: issues, comments, and suggestions. In this case, issues mean any problems participants faced while using the interface, comments mean general notes participants made about the interface, and suggestions indicate recommendations from participants on how to improve the interface. The feedback was analysed using Microsoft Word, by highlighting parts of the text with colours per code. Atlas.ti was not used again, since there appeared to have been issues with the sharing of personal, and/or sensitive data in Atlas.ti (ATLAS.ti Scientific Software Development GmbH, 2020), as indicated by the University of Twente's software distributors (Notebook Service Center, n.d.). The analysis used the method of thematic analysis by Braun and Clarke (2006), as applied before. The specific method chosen for this part is the same as previously applied, except that this time a combination of the semantic and explicit approach has been used, which indicates that both exact words and meaning behind what participants said will be considered.

The success and error rates were counted and noted down while re-watching the recordings of the usability tests. The System Usability Scale was calculated as explained before, and the average of all participants was taken. The Net Promoter Score was calculated by taking the % of Detractors, and subtracting them from the % of Promotors to model the likelihood of use of the system (Gitlin, n.d.).

2.6.2 Results

First, the qualitative feedback given by participants on the interface of the prototype will be discussed, followed by the performance measures and scales (SUS, NPS).

2.6.2.1 Feedback on the design of the prototype

Appendix O shows all transcripts of the usability tests. The transcripts have been coded on the feedback given, divided by issues, comments, and suggestions. A total number of 69 feedback remarks were collected: 30 on the maintenance workset, 16 on the operator part, and 23 on the maintainer part. In total, 10 remarks were given regarding issues, 22 regarding comments and 37 suggestions on how to improve the prototype interface. Thus, most remarks were made on the maintenance workset part (30), and predominantly suggestive remarks were given (37). Table 8 provides an overview of the number of issues, comments and suggestions given per component of the interface, as well as example feedback.

Considering the feedback given in the ‘maintenance workset part’, it was noted by 8 out of the 11 participants that the colours red and blue were not visible on the grey background. The component with most suggestions was the ‘issue report screen’ (8x): it was mentioned twice that the flow of filling in data was unclear for the supervisor and should be improved. Other suggestions given were: use dropdown menus, use different colours, provide a manual, show a progression status. For the ‘system overview’ component, 6 out of 11 participants had the same comment: they all expected to be able to click through to more detailed levels of the system overview, whereas this was not possible in this interface. In the maintenance workset, the question was posed for the system overview component whether participants rather wanted to see this as a functional block overview (schematic depiction), or a picture of the vessel. There was division in opinion, as four participants wanted to see both, two participants only wanted to see the picture of the vessel, and another participant indicated that he would prefer for the operator to be able to see the picture of the vessel, and for the maintainer to be able to see the schematic depiction. A total of 9 remarks were made on the calendar component: A comment was made regarding the calendar being too big, which connects to a suggestion given twice, to make the calendar component smaller or a minimalised screen.

In the ‘operator part’ of the prototype, there was a considerable division of opinion on the placement of the ‘report issue’ button, as well as on the extra verification screen before sending in an issue. Five participants suggested that the button was placed correctly, whereas another five indicated the button to be in the incorrect place: two suggestions were given on where else to place the button, namely at the top/bottom of the screen or in the ‘status overview’ of the system. In terms of the verification step, three participants mentioned no verification was necessary, whereas another five indicated it was deemed necessary. A frequently (5x) suggested solution to having no verification was the ability to change or resolve the issue report when the issue changed or was solved. One other suggestion was to use a wizard to lead the operator through the issue report and/or system-aid such as a pop-up screen.

The ‘maintainer part’ of the interface was the newest one, therefore this was most difficult to understand for the participants. In this part, participants mentioned four times that they expected more guidance from the system by suggesting for example a wizard, or for the system to automatically fill in parts of the troubleshooting form. This troubleshooting form seemed to cause most predicaments for the participants, as they had no experience with or knowledge on such a form. Nonetheless, four participants indicated to be happy with the fact there is an online or e-manual available.

Appendix P shows a complete overview of all (merged) comments given.

Table 8

Overview of the number of issues, comments, and suggestions given per component of the interface, with example feedback

Components	**Issues**	**Count**	**Comments**	**Count**	**Suggestions**	**Count**
Maintenance workset (general overview)						
Issue report screen	Red and blue colours not visible.	2	The MOC is also a system.	3	Show task-progression of the maintainer.	8
System overview		0	Expected to be able to click on the blocks.	2	Show status of system parts by using fly-overs or pop-ups.	1

Dynamic screen		0		0	Add date and user actions.	2
Calendar	Unclear whether arrow moves a month or a year.	2	Calendar is too big.	2	Show a weekly/monthly overview.	5
Notifications	Too small to notice.	1		0	Show notifications window continuously.	2
Operator part						
Button report issue		0		0	Place button at the top/bottom of the screen.	2
Report issue screen		1	The MOC indicates the user.	2	Flowchart/wizard that leads you through the issue report.	8
Notifications	ETBOL* is unclear.	1	Expected differentiation between notices.	1	Ability to 'also report this issue' in notices.	1
Maintainer part						
Tasks overview		0	Expected more detailed information in the tasks overview.	2	Add: calendar (when should the task be done), subsystem, serial number & 'working on this task' button.	2
Troubleshooting form	Too time-consuming to fill out.	3	Happy with E-manual.	8	Use a wizard/fault-tree/help-desk idea.	4
System overview		0	Expected the system overview to go 'with' maintainer's actions.	2	Minimise system overview screen.	2

Note. *Estimated time back online

2.6.2.2 Success and error rates

Table 9 gives an overview of the success rates, error rates (critical and non-critical), and the flow participants used when an error was made, per task that was analysed. The flow, in this case means the pathway followed by the participant during the interaction with the interface. In addition, due to the two tasks in both the operator and maintainer part being very similar, these were analysed together and made into one task per interface part each. Even though, occasionally, during the performance of tasks participants gave a lot of feedback or did not understand components of the prototype, the outcomes of the success and error rates represent an initial overview of the difficulty and level of understanding of the different tasks and components.

As can be seen in Table 9, tasks 1, 2 and 3 in the maintenance workset were feasible to perform, since 3/5, 4/4 and 5/5 participants completed the tasks, respectively. Specifically, tasks 2 and 3 were achievable since no errors were made. In task 1, a critical error was made four times by participants clicking different buttons, or other components in the interface that the researcher had not expected. Still, 3/5 participants were able to finish the task. Task 5 regarding notifications was deemed as the most difficult task, since participants indicated to not know where to search in the interface when looking for notifications. This can also be observed by the number of critical errors made on this component, which is six in total. Connecting this to comments made by participants: The notifications button might have been too small to notice, and therefore difficult to detect. Task 4 was also not completed by many participants (2/5). It was noticed that the second part of the task was difficult to perform, in which participants had to find the ‘overview’ button but did not know where to locate this. In the operator part of the interface, it can be seen that half of the participants finished the task, and five different flows were used by participants. The reason for this is that participants had difficulty finding the ‘report issue’ button, which can also be observed by the number of feedback remarks given (x5) of the button being in the wrong place, as formerly described. In the maintainer part of the interface, tasks were mostly finished, however, the task was often explained in great detail before participants were able to understand and finish it (7/9). The most difficult part was the troubleshooting form, which participants indicated to not understand or be unfamiliar with.

Table 9

Success rate, error rates, and flow used per task

Task	**Success rate**	**Critical errors**	**Non-critical errors**	**Flow used**
Maintenance workset (general overview)				
1. You want to get a better overview of ‘Issue 2’, you want to assign this issue to ‘maintainer 5’, and after some time (which you don’t have to take now) you want to give feedback that the spare parts to be used are available.	3/5	4	2	System overview only, dynamic screen (history of issues), or did not click on details of issue
2. You want to get a better overview of other levels of the system.	5/5	0	0	-
3. You are looking for ‘closed issues’, because you want to get a better overview of ‘issue 21’.	4/4	0	0	-
4. Today is not the 14th of July anymore, but the 15th. You want to switch to this day and get a detailed overview of the tasks to be done, and the available maintainers.	2/5	2	0	System overview, issue report
5. You want to see an overview of the notifications.	1/5	6	1	System overview, issue report, other worksets
Operator part				
You want to report an issue (send/cancel the	5/10	2	4	Notifications, vessel pictures, help

issue).				button, top or bottom of screen
Maintainer part				
You want to start working on task 12. You want to get a better overview of the task, you start reading the issue report, and want to set ETBOL* to a certain time. After that you want to check the manual, and you go to part 3 of the troubleshooting form. You realise that the incorrect attributes have (NOT) disappeared, so you go back to part 1 of the troubleshooting form.	7/9	2	0	System overview

Note. *Estimated time back online

2.6.2.4 System Usability Scale & Net Promoter Score

The resulting score of the system usability scale (SUS) is 80.7 (See Table 10). Much research has been done on the interpretation of SUS scores; therefore, multiple meanings can be given to the scores. The score can be seen as ‘Above average’ (Sauro, 2011b), as well as ‘Good’, or to be graded with a ‘B’, following the American grading system (Bangor, Kortum, & Miller, 2009). Thus, it can be said participants were generally satisfied with the system.

The result of the Net Promoter Score (NPS) is 0 (See Table 10). In this study, two participants scored the system with a 9 (promoters), two other participants scored the system with a 5 (detractors), and the remaining seven participants scored the system with a 7 or 8. This means the percentage of the two detractors get subtracted from the percentage of the two promotors, and thus the outcomes of the percentages get cancelled out. This indicates that there are no critics, but also no advocates for the system. Global benchmark numbers indicate that a score of +35 is average, and scores of 0 and below are ranked in the lower quartile of performers (Gitlin, n.d.).

Table 10

Scores of the SUS and NPS

Measure	Score	Interpretation
System Usability Scale (SUS)	80.7	Above average, Good, ‘B’ (Bangor et al., 2009; Sauro, 2011b).
Net Promoter Score (NPS)	0	Lower quartile performance, below average (Gitlin, n.d.).

2.6.3 Discussion

A low-fidelity CMMS prototype interface was designed for a vessels complex system for a case study at Thales B.V. Usability testing on the prototype interface was performed with stakeholders from a Thales B.V. complex system. The results from the feedback remarks indicated that mainly suggestive remarks (37 in total) were made on the different components of the interface, which might be an indication of the different roles and requirements the stakeholders perceive that the system needs. The results of the questionnaire revealed dissimilar outcomes in terms of the scales used, in which the result of the system usability scale (SUS) was ‘above average’, whereas the result of the net promoter score (NPS) was ‘below average’. Both scales measure different constructs, thus a feasible assertion is that participants were generally satisfied with the usability of the system but would not recommend nor criticise the system to others. In addition, even though the NPS revealed a score of 0, this can be deemed as a good outcome since this means that no participants ‘attacked’ the system, which is relatively seen a satisfactory result for a preliminary prototype.

The SUS is a reliable measure for small sample sizes, since no relationship was found between sample size and questionnaire reliability (Brooke, 2013; Sauro, 2011a). Thus, for this usability test with 11 participants, reliable SUS results were collected. One weakness of both measures (SUS and NPS) can be found in the fact that the outcomes do not give any suggestions on how to improve the system (Sauro, 2011a; Zaki, Muhamad, Ayob, Isa, & Rahim, 2016). However, this research is performed in combination with qualitative feedback from participants, and can thus aid in understanding the quantitative results, and together support a redesign of the prototype.

Even though a small sample of participants was used, and even though difficulties were posed with collecting the quantitative results, useful qualitative feedback has been gathered for a preliminary prototype to be able to make a redesign of some components of the maintenance workset of the prototype. The redesign with explanation can be found in Appendix Q.

3. Overall conclusion

This extensive study focused on the design of a usable low-fidelity computerised maintenance management system (CMMS) prototype interface for vessels (ships) complex systems, as a case study for Thales B.V. The study was conducted in three consecutive phases: (1) a systematic literature review extended by interviews and a questionnaire with stakeholders from Thales B.V., to identify key issues with complex systems maintenance at all levels and CMMSs, (2) a task analysis was carried out on maintenance tasks performed by operators and maintainers working with complex systems, aiding, together with the results from phase 1 in the design of the interface of the low-fidelity CMMS prototype, and (3) usability testing of the prototype.

A list of 14 issues was developed from the literature review. Additional elements to this initial list were added by conducting interviews with stakeholders from Thales B.V. Both lists of issues were merged and resulted in a list of 28 issues with complex systems maintenance and CMMSs. This list was reviewed with stakeholders of a Thales B.V. complex system by means of a questionnaire. The results from the questionnaire display a list of 5 main issues that were considered the most relevant and common issues when dealing with CMMSs and complex systems maintenance. The second phase of the research consisted of carrying out a task analysis for maintenance tasks performed by operators and maintainers working with complex systems, which was reviewed and confirmed by the stakeholders on the representativeness of tasks and identification of problematic tasks. Only one depicted task for operators was deemed as unrepresentative. Furthermore, five tasks were found problematic for operators, and four tasks for maintainers. The results of this and the main issues found from phase 1 of the research were ‘resolved’ in the design of a low-fidelity CMMS prototype as a case study for Thales B.V. The third phase of the research included remote small-scale usability testing on the designed CMMS prototype, with the stakeholders from Thales

B.V. Tasks were given during the usability tests, feedback on the components was asked, and a questionnaire was filled out containing the system usability scale (SUS) and net promoter score (NPS). The task performance measures were not deemed fully reliable; however, they give a good indication of which tasks were difficult to perform for the participants. Most errors were made when participants had to locate the notifications button in the maintenance workset part, which was perceived as being too small by the users. In addition, the troubleshooting form in the maintainer part was most difficult to understand due to its novelty and lack of knowledge on such a form. The results from the usability testing indicated a total of 69 feedback remarks, which mostly consisted of suggestions (37x) on how to improve components on the interface. From the questionnaire, the SUS scored above average, whereas the NPS scored below average. Even though several task performance measures were deemed unreliable, the qualitative feedback has been useful for creating a redesign of parts of the CMMS prototype.

The complete research confronts three main limitations in total. These main limitations will be first be discussed, followed by the limitations of the separate phases of methods performed (literature review, interviews and questionnaire, task analysis research, and usability testing). In addition, the consequences of the limitations on the conclusions and generalisability of the separate phases of the research as well as on the complete research will be discussed.

The first main limitation is that due to COVID-19, the researcher was only able to visit Thales B.V. physically for one month. This means that access to useful documents and communication with the stakeholders from Thales B.V. was difficult, which also caused the topic of the thesis to gradually change over the year: Some parts of the research were focused on TACTICOS, and others on complex systems in general, which makes generalisability of the results difficult. During the interviews, questionnaires, and task analysis research, the stakeholders were asked to review complex systems maintenance and CMMSs in general, whereas during the usability testing the stakeholders were displayed a TACTICOS CMMS prototype. These changes between research scope made it difficult for the stakeholders to focus or not focus on TACTICOS, as they were mostly experienced with this system. In addition, due to remote communication (emails, phone calls), it was challenging to grasp the meaning behind what participants were trying to converse, or to understand certain topics related to complex systems maintenance and CMMSs. The second main limitation is that end-users of CMMSs were out of reach throughout the complete study: Participants are merely stakeholders of TACTICOS. They have no direct experience working in real-time with a CMMS, or with performing operator and maintainer tasks. Therefore, they cannot be deemed experts on the topics of complex systems maintenance or CMMSs. However, due to not being able to work with end-users, the stakeholders from Thales B.V. participated as substitutes. From a Human Factors perspective, this should be viewed as highly undesirable, and no firm conclusions can be drawn from this research for generalisability purposes. Nonetheless, Thales B.V. regards these stakeholders as knowledgeable and experienced enough to participate in the different phases of the research and draw valuable information from them. The last main limitation is that only few stakeholders participated during the different phases of the research. This makes it difficult to draw general, overall conclusions. Unfortunately, stakeholders from Thales B.V. were extremely busy due to changes made in their daily work life because of COVID-19. Still, the researcher has tried to contact and email stakeholders as often as possible, to get the best understanding of their input into the research.

Considering limitations of the literature review specifically, one of these is that the keywords used in the search strategy could have been extended to include human factor and CMMS related terms and leave out mathematical related terms. However, while creating the search string, the current terms used were considered important by the researcher, which resulted in a literature review focused on high-level keywords. In addition to that, solely the researcher screened the records, whereas screening performed by multiple researchers might have indicated different, additional, or fewer resulting papers. Additionally, the research was focused on a short time frame of only 5 years, which is uncommon. Even though the literature review was challenged with limitations, a positive aspect of this literature review is that many databases have been examined, which gives the broadest overview of issues that could be found in literature. Future research can focus on adapting the search string, with a detailed focus on CMMSs and human factors related issues, having multiple researchers screen the records, and extending the time frame. When applying these suggestions, the results of the literature review could be generalisable to current issues with CMMSs and be applied to studies focused on resolving issues and/or designing a CMMS prototype.

Following, a limitation of the interviews and questionnaire is, similar to the literature review, the coding and merging of issues which is certainly subjective as the analysis was solely performed by the researcher of the study. Having multiple researchers consider the coding of the interviews and merging of the list of issues might have resulted in different interpretations. Nonetheless, the researcher has been in frequent contact with the stakeholders and tried to get the best understanding of CMMSs and complex systems maintenance as possible. The resulting list of the five key issues from the questionnaire in this study is not generalisable due to the fact that the stakeholders are no end-users of nor experts on CMMSs; the low number of participants; and the discussed limitations of the literature review which also affected the results of the questionnaire. A suggestion for future research is changing the applied interview and questionnaire method into a Delphi study with experts on CMMSs, thus end-users of the system. The Delphi method aims at reaching a consensus of opinion of a group of experts (Dalkey & Helmer, 1963). The application of this method consists of numerous rounds of information gathering using for example interviews, questionnaires, and focus groups. When applying Delphi, a reliable list of issues with current CMMSs from experts' point of view can be drawn.

The limitation of the task analysis research is the main limitation mentioned before regarding the stakeholders not being experts since they are not directly involved with performing operator or maintainer tasks. This means that the results drawn from this part of the research are not conclusive, nor generalisable to overall identified and problematic tasks for operators and maintainers. Therefore, future research should focus on including experts in carrying out the task analyses to identify all significant main tasks for the end-users of CMMSs. In addition, to understand the operators' and maintainers' thought processes, a cognitive task analysis (CTA), instead of hierarchical task analysis could be performed, as explained in Schaafstal et al. (2000), who used pre-liminary observational studies. "*Cognitive Task Analysis is the extension of traditional task analysis techniques to yield information about the knowledge, thought processes and goal structures that underlie observable task performance. It captures information about both ... overt observable behavior and the covert cognitive functions behind it [to] form an integrated whole*" (Chipman, Schraagen, & & Shalin, 2000, p. 3). CTA could be applied to future research, to generate a deeper understanding of the actions performed and the decisions made by operators and maintainers working with complex systems maintenance.

The first limitation to be considered from the usability testing is that many participants gave feedback while performing the tasks, went off-track and stopped thinking aloud. The researcher tried to keep the participants on the right track, but this was found difficult. From the perspective of the participants, it might have been the case that the researcher was unclear with regards to the tasks and what she needed from the participants. In addition, the interface of the prototype had many components, and the tasks to be performed included multiple sub-tasks, but there was relatively little time to perform the usability testing. Therefore, many components of the interface were not discussed by every participant and the tasks were difficult to understand and perform for participants. Accordingly, the task performance measures were unreliable and should not be taken into account in drawing the main conclusions. Consequently, the results of this research are more qualitative than initially aimed for. However, collecting qualitative data for a preliminary prototype is deemed useful in order to create multiple redesigns of the system. Even though the execution of usability testing was flawed in this study, contributing to the limitations as mentioned before, Thales B.V. has expressed satisfaction with the designed interface, and appreciated the effort of conducting usability testing. Something that the researcher noted, after performing usability testing, is that Thales B.V. does not conduct usability testing on (newly) designed products or systems. Therefore, this research is deemed remarkably relevant in advising Thales B.V. on the importance of usability testing when developing products. End-users of systems feel less motivated to learn and understand a new system when they feel estranged by the system (Wandt et al., 2012). This feeling can be abated when the system is better adapted to the end-user, which can be realised by performing usability testing. In addition, usability testing makes the system more intuitive for the end-users, which means less training on new systems will be needed to get users to work with it. Also, the usability of systems increases product quality and can be financially profitable for corporations such as Thales B.V. (Simões-Marques & Nunes, 2012). Thus, usable systems require less training, and positive experiences with corporations' products increase their sales. Therefore, future research at Thales B.V. should focus on conducting usability testing of systems with end-users. For future referral, tasks should include only one main task, to make it understandable for the participants, which will also produce reliable task performance measures. This, in combination with performing multiple design reiterations after gathering qualitative feedback can generate reliable quantitative data. This quantitative data can help to evaluate the results and is useful for companies such as Thales B.V. in order to indicate, for example, system satisfaction data to customers. If a usable CMMS is designed, Thales B.V. can promote and sell the system as all-encompassing and complete, due to the system being used by maintainers, operators, and maintenance managers.

All in all, this study faced three main limitations, considering remote communication, end-users, and few participants, as well as several limitations in the different phases of the research. Due to the amplitude of the limitations and the consequences these have on the overall results, this study should not be generalised to CMMSs, complex systems maintenance, or other domains. Nonetheless, a CMMS has been designed following design principles, but can be improved upon by taking the limitations into account and replicating the study, in order to design a usable CMMS specified at the end-users of the system. Thus, future research at Thales B.V. should focus on employing these end-users during data collection, performing usability testing, and looking into the concept of CMMSs. The baseline for maintenance management in complex systems is already set with the development of CMMSs. A software that fits Thales B.V. might already exist, but merely needs to be adapted to the companies' and end-users' needs. In conclusion, this study was a huge contribution to Thales B.V., and they were very happy with the work performed by the researcher. The overall research showed its usefulness in aiding Thales B.V. with a method to design a usable CMMS. Thales B.V. has been working on the design of a 'maintenance workset' for a long period of time now, in which this research gave new insights from a Human Factors viewpoint. They are currently working on a similar maintenance management system, and still use this research and its results, and will continue to do so in the future.

4. Reference list

Abele, S., & Weyrich, M. (2017). Decision Support for Joint Test and Diagnosis of Production Systems based on a Concept of Shared Knowledge. *Ifac Papersonline, 50*(1), 15227-15232. doi:10.1016/j.ifacol.2017.08.2374

Adobe. (2020). Retrieved from https://www.adobe.com/nl/products/xd.html

Annett, J. (2003). Hierarchical task analysis. In E. Hollnagel (Ed.), *Handbook of Cognitive Task Design* (pp. 17 - 35). Boca Raton, Florida: CRC Press.

ATLAS.ti Scientific Software Development GmbH. (2020). *Update History*. Retrieved from https://atlasti.com/update-history/#1600702454811-ceb033dd-9fcf

Bagida, K. (2010). *Computerized Maintenance Management Systems Made Easy: How to Evaluate, Select, and Manage CMMS*. New York, NY: McGraw-Hill.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: adding an adjective rating scale. *The Journal of Usability Studies, 4*(3), 114–123. doi:10.5555/2835587.2835589

Barnum, C. (2020). Preparing for usability testing. In *Usability Testing Essentials: Ready, Set ...Test!* (pp. 197-248). doi:10.1016/B978-0-12-816942-1.00006-X

Braaksma, A. J. J. (2012). *Asset information for FMEA-based maintenance.* University of Groningen, Groningen. Retrieved from https://www.rug.nl/research/portal/publications/asset-information-for-fmeabased-maintenance(ef38be1a-5fb1-4ecd-b39b-1b2dc0d39978).html

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

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry, 189*(194), 4-7. Retrieved from https://liacs.leidenuniv.nl/~verbeekfj/courses/hci/SUS-questionaire.pdf

Brooke, J. (2013). SUS: a retrospective. *The Journal of Usability Studies, 8*(2), 29–40. Retrieved from https://uxpajournal.org/sus-a-retrospective/

Chaparro, A., & Groff, S. L. (2002). Human Factors Survey of Aviation Maintenance Technical Manuals. Paper presented at the *16th Human Factors in Aviation Maintenance Symposium*. Retrieved from https://commons.erau.edu/publication/969

Chipman, S. F., Schraagen, J. M., & & Shalin, V. L. (2000). Introduction to cognitive task analysis. In J. M Schraagen, S. F. Chipman, & V. J. Shute (Eds.), *Cognitive Task Analysis* (pp. 3-23). Mahwah, NJ: Lawrence Erlbaum Associates.

Dalkey, N., & Helmer, O. (1963). An Experimental Application of the DELPHI Method to the Use of Experts. *Management Science, 9*(3), 458-467. doi:10.1287/mnsc.9.3.458

de Noord, R. (2019). *Creating a maintenance workset for TACTICOS.* (Master thesis). University of Twente, Enschede.

del Amo, I. i. F. n., Galeotti, E., Palmarini, R., Dini, G., Erkoyuncu, J., & Roy, R. (2018). An innovative user-centred support tool for Augmented Reality maintenance systems design: a preliminary study. *Procedia CIRP, 70*, 362-367. doi:10.1016/j.procir.2018.02.020

Department of Defense. (1980). *PROCEDURES FOR PERFORMING A FAILURE MODE, EFFECTS AND CRlTlCALlTY ANALYSIS (MIL-STD-1629A)*. Washington, DC: Government Printing Office

Department of Defense. (2012). *Design Criteria Standard: Human Engineering* (MIL-STD-1472G). Retrieved from http://everyspec.com/MIL-STD/MIL-STD-1400-1499/MIL-STD-1472G_39997/

Fisher, N. I., & Kordupleski, R. E. (2019). Good and bad market research: A critical review of Net Promoter Score. *Applied Stochastic Models in Business and Industry, 35*(1), 138-151. doi:10.1002/asmb.2417

FMEA-FMECA. (2006). *FMEA RPN*. Retrieved from https://www.fmeafmeca.com/fmea-rpn.html

Furda, A., Cholette, M. E., Ma, L., Fidge, C., Hill, W., & Robinson, W. (2015). Making Optimal and Justifiable Asset Maintenance Decisions. In P. W. Tse, J. Mathew, K. Wong, R. Lam, & C. N. Ko (Eds.), *Engineering Asset Management - Systems, Professional Practices and Certification* (pp. 253-263) doi:10.1007/978-3-319-09507-3_23.

Galar, D., & Kumar, U. (2017). Chapter 7 - Maintenance Decision Support Systems. In D. Galar & U. Kumar (Eds.), *eMaintenance* (pp. 371-474) doi:10.1016/B978-0-12-811153-6.00007-5

Gallab, M., Bouloiz, H., & Tkiouat, M. (2015). Decision support for occupational risk overcome in maintenance activities. Paper presented at the *Science and Information Conference (SAI)*, 426-434. doi:10.1109/SAI.2015.7237177

Gitlin, J. (n.d.). *What is a good Net Promoter Score? And how does it vary across industries?* Retrieved from https://www.surveymonkey.com/curiosity/what-is-a-good-net-promoter-score/

Hassan, T. B., & Barnett, D. B. (2002). Delphi type methodology to develop consensus on the future design of EMS systems in the United Kingdom. *Emergency Medicine Journal, 19*, 155-159. doi:10.1136/emj.19.2.155

Hugo, J. V., & Gertman, D. I. (2016). A Method to Select Human–System Interfaces for Nuclear Power Plants. *Nuclear Engineering and Technology, 48*(1), 87-97. doi:10.1016/j.net.2015.10.004

Huijben, J. A., Wiegers, E. J. A., de Keizer, N. F., Maas, A. I. R., Menon, D., Ercole, A., . . . Lingsma, H. F. (2019). Development of a quality indicator set to measure and improve quality of ICU care for patients with traumatic brain injury. *Critical care, 23*(95). doi:10.1186/s13054-019-2377-x

International Electrotechnical Commission (IEC). (1999). Dependability management – Application guide. In *Reliability centered maintenance* (Vol. IEC60300-3-11). Geneva.

International Organization for Standardization (ISO). (2018). Ergonomics of human-system interaction. In *Part 11: Usability: Definitions and concepts* (Vol. ISO 9241-11).

Iung, B., Levrat, E., Marquez, A. C., & Erbe, H. (2007). E-MAINTENANCE: PRINCIPLES, REVIEW AND CONCEPTUAL FRAMEWORK. *IFAC Proceedings Volumes, 40*(19), 18-29. doi:10.3182/20071002-MX-4-3906.00005

Johnson, J. (2021). Chapter 15 - We Make Errors. In J. Johnson (Ed.), *Designing with the Mind in Mind (Third Edition)* (pp. 259-274) doi:10.1016/B978-0-12-818202-4.00015-5.

Jorgensen, E. L. (1994). *The Interactive Electronic Technical Manual Overview "Setting the Stage"*. Paper presented at the AFEI CALS Expo International, Florida. https://www.navsea.navy.mil/Portals/103/Documents/NSWC_Carderock/Cals94j.pdf

Kans, M. (2009). The advancement of maintenance information technology. *Journal of Quality in Maintenance Engineering, 15*(1), 5-16. doi:10.1108/13552510910943859

Karim, R. (2008). *A service-oriented approach to e-maintenance of complex technical systems.* (Doctoral thesis). Luleå tekniska universitet, Luleå. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:ltu:diva-17241

Kobbacy, K. A. H., & Murthy, D. P. (2008). *Complex System Maintenance Handbook*. London, England: Springer.

Labib, A. W. (2004). A decision analysis model for maintenance policy selection using a CMMS. *Journal of Quality in Maintenance Engineering, 10*(3), 191-202. doi:10.1108/13552510410553244

Lang, D., Wunderlich, P., Heinz, M., Wisniewski, L., Jasperneite, J., Niggemann, O., & Röcker, C. (2018). Assistance system to support troubleshooting of complex industrial systems. Paper presented at the *14th IEEE International Workshop on Factory Communication Systems (WFCS)*, 1-4. doi:10.1109/WFCS.2018.8402380

Lee, J., & Wang, H. (2008). New Technologies for Maintenance. In K. A. H. Kobbacy & D. N. P. Murthy (Eds.), *Complex System Maintenance Handbook* (pp. 49-78) doi:10.1007/978-1-84800-011-7.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., . . . Moher, D. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLOS Medicine, 6*(7). doi:10.1371/journal.pmed.1000100

Lopes, I., Senra, P., Vilarinho, S., Sá, V., Teixeira, C., Lopes, J., . . . Figueiredo, M. (2016). Requirements Specification of a Computerized Maintenance Management System – A Case Study. *Procedia CIRP, 52*, 268-273. doi:10.1016/j.procir.2016.07.047

Lundgren, C., Skoogh, A., & Bokrantz, J. (2018). Quantifying the Effects of Maintenance – a Literature Review of Maintenance Models. *Procedia CIRP, 72*, 1305-1310. doi:10.1016/j.procir.2018.03.175

Martinetti, A., Rajabalinejad, M., & van Dongen, L. (2017). Shaping the future maintenance operations: reflections on the adoptions of Augmented Reality through problems and opportunities. In R. Roy, T. Tomiyama, A. Tiwari, K. Tracht, E. Shehab, J. Mehnen, J. A.

Erkoyuncu, & N. Tapoglou (Eds.), *Proceedings of the 5th International Conference in through-Life Engineering Services* (pp. 14-17) doi:10.1016/j.procir.2016.10.130.

Martinetti, A., Schakel, E. J., & van Dongen, L. A. M. (2018). Flying asset: Framework for developing scalable maintenance program for unmanned aircraft systems (UAS). *Journal of Quality in Maintenance Engineering, 24*(2), 152-169. doi:10.1108/jqme-12-2016-0073

Mindomo. (n.d.). *Mindmaps om mee samen te werken, concept maps, outlines en Gantt grafieken*. Retrieved from https://www.mindomo.com/nl/

Nielsen, J. (1994a). *Enhancing the explanatory power of usability heuristics*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, Massachusetts, USA. https://doi.org/10.1145/191666.191729

Nielsen, J. (1994b). Heuristic evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods*. New York, NY: John Wiley & Sons.

Norman, D. A. (1983a). Design principles for human-computer interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1–10. doi:10.1145/800045.801571

Norman, D. A. (1983b). Design rules based on analyses of human error. *Communications of the ACM, 26*(4), 254–258. doi:10.1145/2163.358092

Notebook Service Center. (n.d.). *Software Manager*. Retrieved from https://www.nsc.utwente.nl/software/

Peeters, J. F. W., Basten, R. J. I., & Tinga, T. (2018). Improving failure analysis efficiency by combining FTA and FMEA in a recursive manner. *Reliability Engineering & System Safety, 172*, 36-44. doi:10.1016/j.ress.2017.11.024

Pintelon, L., & Parodi-Herz, A. (2008). Maintenance: An Evolutionary Perspective. In K. A. H. Kobbacy & D. N. P. Murthy (Eds.), *Complex System Maintenance Handbook* (pp. 21-48) doi:10.1007/978-1-84800-011-7.

Polisena, J., Castaldo, R., Ciani, O., Federici, C., Borsci, S., Ritrovato, M., . . . Pecchia, L. (2018). HEALTH TECHNOLOGY ASSESSMENT METHODS GUIDELINES FOR MEDICAL DEVICES: HOW CAN WE ADDRESS THE GAPS? THE INTERNATIONAL FEDERATION OF MEDICAL AND BIOLOGICAL ENGINEERING PERSPECTIVE. *International Journal of Technology Assessment in Health Care, 34*(3), 276-289. doi:10.1017/S0266462318000314

Qualtrics. (2021). Retrieved from https://www.qualtrics.com/

R-3.6.2 for Windows (32/64 bit). (2019). Retrieved from https://cran.r-project.org/bin/windows/base/old/3.6.2/

Rastegari, A., & Mobin, M. (2016). Maintenance decision making, supported by computerized maintenance management system. Paper presented at the *Annual Reliability and Maintainability Symposium (RAMS),* 1-8. doi:10.1109/RAMS.2016.7448086

Rausand, M., & Vatn, J. (2008). Reliability Centred Maintenance. In K. A. H. Kobbacy & D. N. P. Murthy (Eds.), *Complex System Maintenance Handbook* (pp. 79-108) doi:10.1007/978-1-84800-011-7.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review, 81*(12), 46–124. Retrieved from https://pubmed.ncbi.nlm.nih.gov/14712543/

Rooney, J. J., & van den Heuvel, L. N. (2004). Root Cause Analysis For Beginners. *Quality Progress, 37*(7), 45-53. Retrieved from https://asq.org/quality-progress/articles/root-cause-analysis-for-beginners?id=0228b91456514ba490c89979b577abb4

Sauro, J. (2011a). *Measuring usability with the system usability scale (SUS)*. Retrieved from https://measuringu.com/sus/

Sauro, J. (2011b). *A practical guide to the System Usability Scale: Background, benchmarks, & best practices.* Scotts Valley, CA: CreateSpace Independent Publishing Platform.

Schaafstal, A., Schraagen, J. M., & van Berlo, M. (2000). Cognitive task analysis and innovation of training: the case of structured troubleshooting. *Human Factors, 42*(1), 75-86. doi:10.1518/001872000779656570

Seel, N. M. (2012). Problems: Definition, Types, and Evidence. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2690-2693) doi:10.1007/978-1-4419-1428-6_914.

Simões-Marques, M., & Nunes, L. I. (2012). Usability of Interfaces. In I. L. Nunes (Ed.), *Ergonomics: A Systems Approach* doi:10.5772/37299. London, United Kingdom: IntechOpen.

Stone, D., Jarrett, C., Woodroffe, M., & Minocha, S. (2005). *User Interface Design and Evaluation.* (1 ed.). Massachusetts, MA: Morgan Kaufmann.

Su, Y., Liang, X. R., Wang, H., Wang, J. J., & Pecht, M. G. (2019). A Maintenance and Troubleshooting Method Based on Integrated Information and System Principles. *Ieee Access, 7*, 70513-70524. doi:10.1109/ACCESS.2019.2915327

Tagliente, D. A., Lospinuso, M. D., & DiRosa, F. (2017). Dynamic fault monitoring and fault-based decision making in vehicle health management systems. Paper presented at the *2017 IEEE AUTOTESTCON*, 1-5. doi:10.1109/AUTEST.2017.8080503

Thales. (n.d.). *TACTICOS - Combat Management System*. Retrieved from https://www.thalesgroup.com/en/tacticos-combat-management-system

Tretten, P., & Karim, R. (2014). Enhancing the usability of maintenance data management systems. *Journal of Quality in Maintenance Engineering, 20*(3), 290-303. doi:10.1108/JQME-05-2014-0032

usability.gov. (n.d.). *Planning a Usability Test*. Retrieved from https://www.usability.gov/how-to-and-tools/methods/planning-usability-testing.html

van den Haak, M., De Jong, M., & Jan Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour & Information Technology, 22*(5), 339-351. doi:10.1080/0044929031000

Wahid, A.-S. B. A., Ahmad, M. Z. B., Ahmad, K. A. B., Taylor, J. P., Abdullah, A. B., Ali, A.-A. B. A.-S. A. B., & Kitagawa, K. (2018). DEMYSTIFYING SHIP OPERATIONAL AVAILABILITY - AN ALTERNATIVE APPROACH FOR THE MAINTENANCE OF NAVAL VESSELS. *Journal of Engineering Science and Technology, 13*(12), 4326-4346. Retrieved from http://jestec.taylors.edu.my/Vol%2013%20issue%2012%20December%202018/13_12_30.pdf

Wandt, K., Tretten, P., & Karim, R. (2012). *Usability aspects of eMaintenance solutions*. Paper presented at the International Workshop & Congress on eMaintenance, Luleå. fudinfo.trafikverket.se

Zaki, M. F. M., Muhamad, W., Ayob, A., Isa, N. F., & Rahim, M. (2016). A Descriptive analysis of civil engineering maintenance in government hospital buildings. *Matematika, 32*(2), 103-111. doi:10.11113/matematika.v32.n2.925