

Towards consistent SCADA mimic colour practices

UNIVERSITY OF TURKU
Department of Computing
Master of Science Thesis
Computer Science
April 2026
Nelli Rossi

UNIVERSITY OF TURKU
Department of Computing

NELLI ROSSI: Towards consistent SCADA
mimic colour practices

Master of Science Thesis, 79 p., 12 app. p.
Computer Science
April 2026

This thesis investigates colour practices in SCADA mimic displays across the energy and marine industries, motivated by the observation that there is no universal standard governing how colours should be used in these systems. Three relevant standards were examined: ISO 14726 for piping identification colours, ISA-101.01 for human machine interface design principles, and WCAG2ICT for accessibility in non-web software. The alignment between these standards is partial and observed industry practices fill the remaining gaps inconsistently through vendor guidelines and organisational conventions.

Based on the reviewed standards and literature, 12 colour design guidelines were created. Three prototypes per industry were then designed: one with the current colour scheme, one closely following the guidelines, and one as a compromise between the two. The prototypes were tested with 10 industry professionals from Wärtsilä through think-aloud tasks, an interview, and a post-session survey.

The results showed a clear divide between Wärtsilä's industry domains. Full guideline alignment improved perceived usability for marine participants but reduced it for energy participants whilst the compromise prototype was broadly acceptable across both groups. Familiarity with existing colour conventions was a strong influential factor, especially in the energy group. These findings suggest that consistent SCADA colour practices should be pursued through shared principles adapted to industry-specific needs rather than through a single universal colour scheme.

Keywords: SCADA, HMI, colour standards, mimic design, usability

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Research objectives and methodology | 2 |
| 1.2 | Structure | 3 |
| 2 | Supervisory control and data acquisition | 5 |
| 2.1 | Mimics | 6 |
| 2.2 | SCADA in energy and marine industries | 8 |
| 2.3 | Previous research in mimic colours | 11 |
| 3 | Understanding colour | 14 |
| 3.1 | Colour perception | 14 |
| 3.2 | Cognitive processing of colour | 16 |
| 3.3 | Colour associations | 17 |
| 4 | User-centred design | 21 |
| 4.1 | Usability | 23 |
| 4.2 | Accessibility | 24 |
| 5 | Existing standards | 26 |
| 5.1 | ISO 14726 – Physical piping colours | 28 |
| 5.2 | ISA-101.01 – HMI design | 30 |
| 5.2.1 | General colours | 31 |

| | | |
|----------|--|-----------|
| 5.2.2 | Alarm state colours | 32 |
| 5.3 | W3C WCAG2ICT – Accessibility | 33 |
| 5.4 | Observed industry practices | 34 |
| 6 | Prototype development | 37 |
| 6.1 | Guideline creation | 37 |
| 6.2 | Prototypes for energy | 39 |
| 6.3 | Prototypes for marine | 45 |
| 7 | Testing methodology | 50 |
| 7.1 | Participation criteria and recruitment | 50 |
| 7.2 | User testing methodology | 52 |
| 7.3 | Data collection methods | 56 |
| 7.4 | Testing hypotheses | 57 |
| 8 | User testing results | 58 |
| 8.1 | Participant demographics and overview | 58 |
| 8.2 | Colour-specific results | 60 |
| 8.2.1 | Alarm colour perception and severity ordering | 60 |
| 8.2.2 | Yellow colliding in pipes and alarms | 61 |
| 8.2.3 | Pipe colour differentiation | 62 |
| 8.2.4 | Background colour and visual comfort | 64 |
| 8.2.5 | Component state colours | 64 |
| 8.2.6 | The familiarity bias versus observed behaviour | 65 |
| 8.2.7 | Legend and P3 compromise pattern | 65 |
| 8.3 | Other results | 66 |
| 8.4 | Summary of main findings | 68 |

| | |
|---|-------------|
| 9 Discussion | 70 |
| 9.1 Addressing the research questions | 72 |
| 9.2 Limitations | 74 |
| 9.3 Future research | 75 |
| 10 Conclusion | 77 |
| References | 80 |
| Appendices | |
| A Participant background questionnaire | A-86 |
| B Testing introduction | B-89 |
| C Post-test survey (energy example) | C-92 |
| D Consent form | D-97 |

1 Introduction

Supervisory control and data acquisition (SCADA) systems are used to monitor and control industrial processes, and operators interact with them through different pages, including graphical mimics that represent the physical systems. These mimics have many uses for colour, such as pipe media identification, alarm severity indication, component states, and general visual hierarchy. In environments where safety is critical, like in power plants or marine vessels, the correct interpretation of these visual elements directly affects operational safety and efficiency.

Although this thesis is concerned with SCADA systems, its focus is narrower. A SCADA system consists of the full stack of field devices, controllers, servers, and the software that operators run through interfaces. Within that software, a mimic is one specific type of display, which graphically represents physical systems. Other display types are also presented in the software of SCADA systems, yet this thesis is specifically about how colour is used within mimic displays.

Despite the importance of colour use in SCADA mimics, there is no universal standard governing how colours should be used. Multiple standards address aspects of colour use, but these standards operate at different levels of specificity and leave significant gaps. In practice, colour conventions vary between vendors, industries, and projects, with individual designers and organisations filling the gaps left by the standards. This variability means that operators moving between systems or

industries may encounter different colour meanings for the same visual elements, creating potential confusion and errors.

This thesis investigates SCADA mimic colour practices across two industries, *energy* and *marine*, within the context of Wärtsilä, a company that develops SCADA for both industries. The thesis examines existing standards and industry practices, combines these into a set of guidelines, creates prototypes with different levels of alignment to those guidelines, and tests them with industry professionals from Wärtsilä. The goal is to understand both the current state of colour standardisation and the practical effects of implementing that standardisation on perceived usability.

This thesis was commissioned by Wärtsilä, where the author works as a part of the team designing SCADA user interfaces. This provided both the motivation for the research and access to the SCADA systems and professionals involved in the testing. Some knowledge presented in this thesis, particularly regarding current internal colour conventions and design practices, is drawn from this professional experience where published sources are not available.

1.1 Research objectives and methodology

The thesis is guided by two research questions which address what standards exist and how they relate and the practical impact. The two research questions are as follows:

RQ1: What standards and best practices are there for colour use in SCADA mimics for energy and marine industries, and to what extent do they align with each other and with observed industry practices?

RQ2: How does the degree of alignment with identified colour standards and best practices affect the perceived usability of SCADA mimics across energy and marine industries?

RQ1 is addressed through a review of existing standards, an examination of observed industry practices, and colour perception and processing. These are then combined into a set of 12 guidelines. The answer to RQ1 emerges from Chapter 3 through Chapter 6 and is explicitly answered in Chapter 9. RQ2 is addressed through an empirical study in which three prototypes per industry were designed. One represents current practices, one closely follows the created guidelines, and one represents a compromise between these two. The prototypes were tested with 10 industry professionals from Wärtsilä. The testing was done by combining think-aloud observation tasks, interviews, and post-session survey. The answer to RQ2 emerges from Chapter 8 and in Chapter 9 the question is explicitly answered.

1.2 Structure

In Chapter 2, SCADA systems are introduced and it is explained what mimics are and how operators interact with them. Marine and energy are also introduced as the two industries that are studied and previous research on mimic colours is reviewed.

Chapter 3 covers the fundamentals of colour perception, including biological factors, cognitive processing and its implications for interface design, and colour associations from both universal and cultural perspective. This chapter provides the theoretical foundation for understanding why colour choices in SCADA mimics affect usability and provides the theoretical background for the created guidelines.

User-centred design is covered in Chapter 4, and it is introduced as the design philosophy underlying the study. The chapter defines usability as a measurable quality attribute and discusses accessibility as an extension of usability. These concepts are used to frame the evaluation criteria used in the testing.

Chapter 5 reviews existing standards relevant to SCADA mimics: ISO 14726 for piping identification colours, ISA-101.01 for general design principles, and

WCAG2ICT for accessibility criteria. This chapter also examines observed industry practices from SCADA vendors.

Prototype development is described in Chapter 6. First the creation of the 12 guidelines is introduced, after which the development of the three prototypes is described per industry.

Chapter 7 details the testing methodology. This includes participant recruitment, testing procedure, interview, and post-session survey design, and the five hypotheses formulated before testing. The findings from the user testing sessions are presented in Chapter 8. They are organised thematically into colour-specific results, other results, and a summary of main findings including verdicts for each hypothesis.

Chapter 9 discusses the results in relation to the research questions and existing literature. It is followed by a discussion of the study's limitations and potential future research directions are introduced. Conclusions and a summary of the key findings and contributions of the thesis are presented in Chapter 10.

2 Supervisory control and data acquisition

This chapter introduces supervisory control and data acquisition systems at the level of detail needed for the rest of the thesis. The aim is to establish the relation of mimics to these systems, rather than to give a complete account of how the systems work. Mimics are one type of display shown to the operators, and they are the specific focus of this thesis.

Supervisory control and data acquisition is typically shortened to SCADA, and these systems are often used for industrial processes. Instead of being full control systems, they focus more on the supervisory level. A core concept for SCADA is programmable logic controllers (PLC), on which the interface is built. [1] PLCs gather information from sensors and actuators, which they are then able to interchange with supervisory control. The supervisory control is generally a computer with specific software, and it can perform multiple tasks, for example, control all the other connected devices, run algorithms, and send commands. Similarly, there are remote terminal units (RTU) that are connected to the supervisory control. They offer a similar function of collecting data from sensors but are simpler in their logic operations than PLCs. The data provided by PLCs and RTUs can be accessed with the help of the Data Acquisition Server. [2]

Figure 2.1 illustrates a general depiction of how SCADA systems work. On site (e.g., power plant, vessel) there is equipment that collects data, for instance, battery temperature or state of charge. The equipment then sends the data to the PLCs, which again forward it to *human machine interfaces* (HMI). The SCADA system is used by the operator on these HMIs, and so are the mimics. Through the HMIs the operators can manually, or the system itself automatically send control inputs to the PLCs, which send the commands to the equipment.

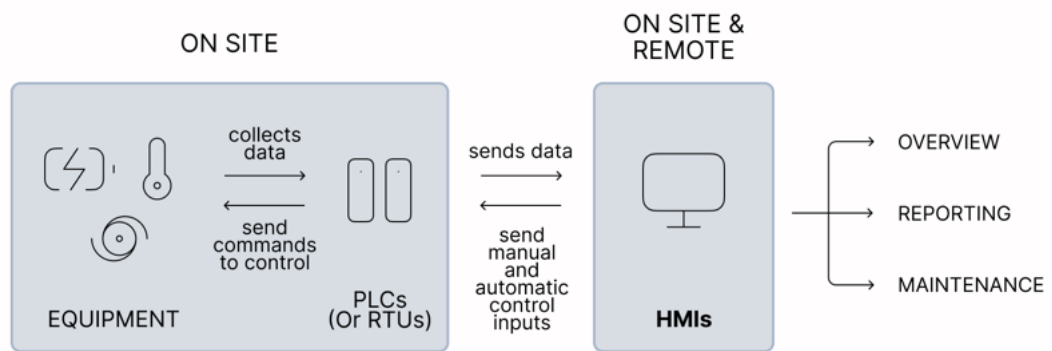


Figure 2.1: Simplified version of how SCADA systems work (adapted from [3])

The aspect that is of most interest for this thesis is the HMI. For SCADA systems, it is the part that will present the data that has been collected to a user. It retrieves data from the servers and presents it to the user as mimics, reports, graphics, trends, alarms, notifications, or other required visualisations. The users, typically operators, will then be able to interact with the process.

2.1 Mimics

Mimics are graphical representations of physical systems or processes of which the status and performance are displayed in real time. They give the user a simplified overview of the system's operations and allows them to monitor and control the

system. The different mimics typically include indicators, alarms, and other physical components of the system. A common use case for a SCADA mimic is in a control room, where an operator can use the system with a computer-based display. They can also be placed as physical panels in cabinets, typically offering a touch-friendly interface. [4]

Other common names for mimics include process pages, system schematics, and schematic diagrams. They are also sometimes referred to as single line diagrams (SLD), but since those cover only electrical systems, it is not a fit term here. Besides SLDs, the mimics can also be piping and instrumentation diagrams (P&ID), which are the most common type for marine. The term *mimic* was chosen for this thesis as it is one of the most common ones and works for both marine and energy industries.

In the marine industry it is not uncommon to find mimics that show most of the vessel's system, whereas in the energy industry the mimics are often more focused on subsystems of a larger entity, like in power plants. As a common convention, SCADA systems in the marine industry open directly to a mimic and in the energy industry functional overviews are preferred. The differences between Wärtsilä's two industry domains are discussed in further detail in Section 2.2.

Figure 2.2 illustrates the difference between a real system and a mimic. The physical marine fuel supply system depicted shows the different components belonging to that system and how they are connected. The system is monitored through the SCADA mimic interface, which is shown next to the actual system. It illustrates the system through symbols, flow paths, and different indicators that an operator can use to assess the system state. How effectively this mimic can communicate the state to an operator, especially through the use of colour, is essential for this thesis.

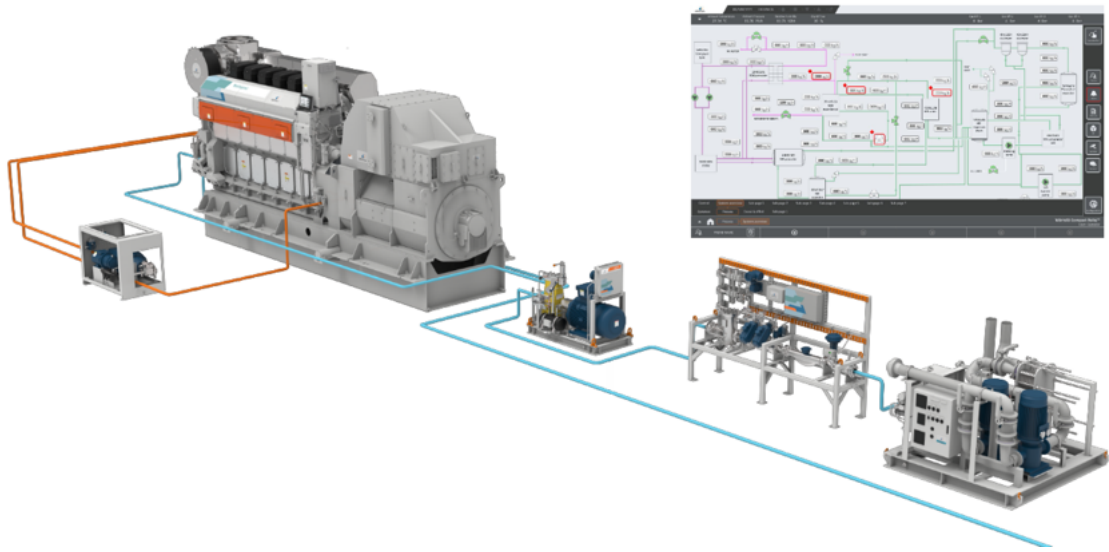


Figure 2.2: A physical marine fuel supply system and its corresponding SCADA mimic diagram

2.2 SCADA in energy and marine industries

A primary function in the energy and marine industries is to monitor engines and generators alongside their output. In the energy industry generators are used to produce electrical energy and thus also the output is monitored. In the marine industry the engines are used to keep the vessels in motion, but they can also work as generators in hybrid vessels, which have batteries for energy storage.

In the energy industry, SCADA systems show SLDs and key performance indicators (KPI), such as oil temperature. In the marine industry, there are P&IDs and KPIs, such as monitoring subsystems that can include for instance engine overview, injection, and exhaust.

In the energy industry, especially in power plants, SCADA systems are divided into packs. This is because the power plants are built in a modular way, meaning that a plant can have different number of similar subsystems in them. These cause the SCADA system to require more overviews of the subsystems and navigating into them. One example of this system is present in Figure 2.3, where the home page of a

plant management system from Wärtsilä is currently open. The different generators can be navigated into by pressing on them to view information specific to it.



Figure 2.3: Plant management system home screen

Another example from the same plant management system is displayed in Figure 2.4. There, the system has been navigated from its home screen to generator 1 and specifically to view its lubrication system. That subsystem is shown as a mimic with P&ID symbols in different states, piping, pipe flow, and different alarms happening in the system. In the energy industry as the mimics are also typically divided into subsystems, it means that only one or very few colours are used per mimic.

In the marine industry the mimics are customarily of the full system, and they typically function as the home screen of the SCADA system. However, that is not the case in every project, and some might prefer to have no mimics at all and use bar graphics instead. Furthermore, there are differences between projects regarding alarm visibility. Certain projects favour reduced alarm visibility in the system to avoid overwhelming operators, but that is not a common approach since alarms are a large aspect of SCADA systems.

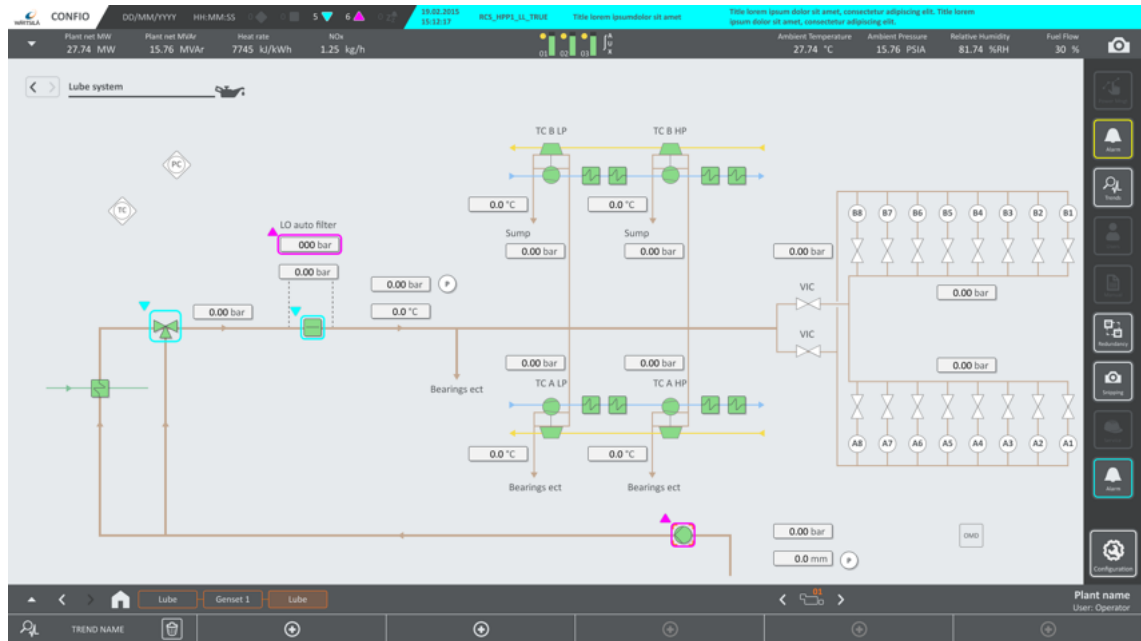


Figure 2.4: Plant management system lubrication system mimic

An example of a marine mimic is illustrated in Figure 2.5. The mimic is operating as the home screen of the system, and the different sections separated by grey lines can be used to navigate into their corresponding submimics. The mimic has P&ID symbols in different states, piping, pipe flow, and alarms happening in the system overall. When comparing Figure 2.4 to Figure 2.5, the visual appearance is slightly different, as in the marine mimic the piping colours are more highly saturated.

As SCADA systems are for supervisory control and data acquisition, there are also differences in the balance of control and monitoring between industries and systems. In the energy industry, monitoring has priority and control comes second. In the marine industry monitoring and controlling are more balanced.

SCADA systems can be located in different places, and they are normally in more than one location, between which the control can be transferred. For example, in the marine industry, it is typical that the system is visible on the bridge and in the engine control room (ECR). In the energy industry the corresponding common locations are the engine room (ER) and ECR. Because marine mimics are also visible

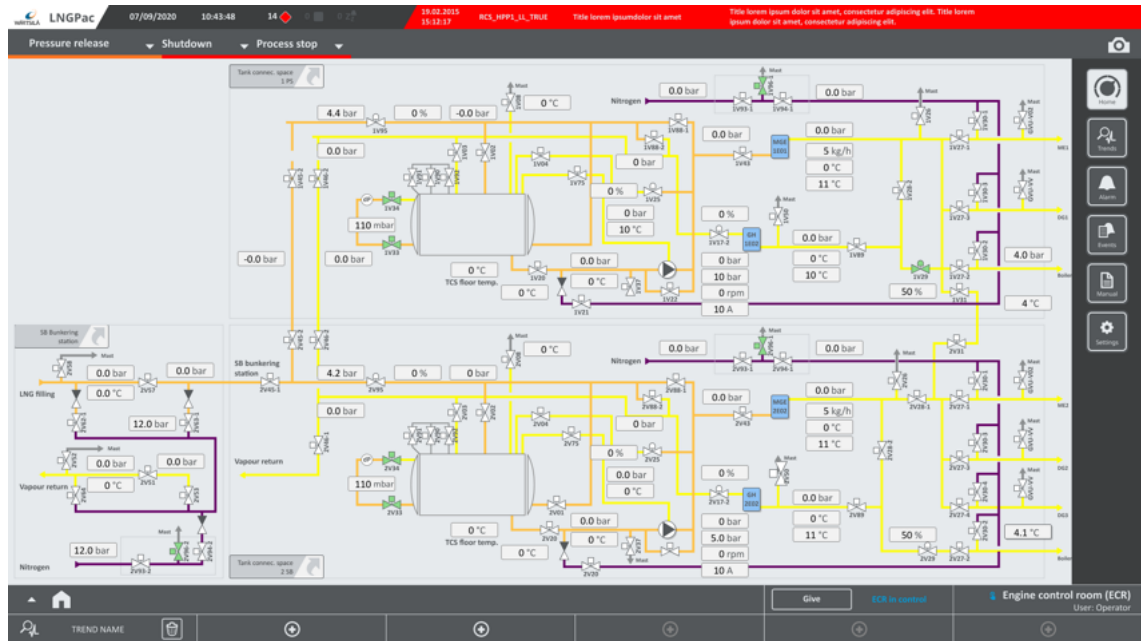


Figure 2.5: Fuel supply system home page mimic

on bridges, the systems need to accommodate to the lighting during different times of day. This essentially means that mimic colours must be adjusted for both a light and dark mode in the marine industry.

In the two industries the levels of control are different based on the user level (e.g., commissioning, operator). Some functionalities are only available and controllable for the higher-level users. Moreover, most values are controlled automatically by the PLC, yet they can be switched to manual mode, where the user can control them. However, there are some values, which can only be in auto mode and no manual mode is available.

2.3 Previous research in mimic colours

Research done specifically on SCADA mimic colours is extremely scarce and the closest existing study yielded from a systematic search is by Boulhic et al. [5]. The study directly tested whether ISO 14726 could be applied to marine supervision HMI.

For the study, the authors created two colour schemes for marine vessel ER supervision HMIs. The first colour scheme was called Normative pipe identification colour code (NPICC), which applied the colours from the ISO 14726:2008 standard directly. The second scheme, called Ergonomic recommendations colour code (ERCC), used single colours per fluid, which were chosen based on common associations and general ergonomic guidelines. They performed two experiments, one with 20 novices and one with 19 experts and both experiments used search tasks on mock HMIs. The experts were maritime college students with experience in sailing on vessels, while the novices were university students with no maritime experience. The groups were further divided into NPICC and ERCC groups. [5]

In the novice group, the NPICC increased mental workload when compared to the ERCC. However, specifically the bicolour-coded fluids in the NPICC had lower error rates than both the single-colour fluids within the same group and the ERCC group, meaning that the dual-colour approach from ISO 14726 was harder to use but more reliable for correct identification. They also tested the colours with 100% and 50% saturation and found desaturated pipe colours for the ERCC group to improve alarm detection without negatively affecting fluid identification. [5]

In the expert group, the NPICC had lower mental workload and faster search times when compared to the ERCC, which is the opposite of the novices. Although it was not a direct result of finding the ISO colours easier, but more about falling back on professional knowledge, system familiarity, and text labels when faced with unfamiliar bicolour coding. With the ERCC, the colours felt reliable and intuitive, therefore the participants actively used the colour codes, which increased their mental workload. [5]

The authors concluded that they would not recommend implementing the ISO standard's bicolour system on marine supervision HMI. Even though it improved

identification accuracy for novices, it was too cognitively demanding for experts and they actively avoided it. [5]

Their study has a strong connection to this thesis as it will also test ISO 14726:2008-based pipe colours. However, this thesis includes participants from two industries, energy and marine, which Boulhic et al. did not include. This thesis will also be introducing a prototype with single colours for pipes inspired by the ISO standard. Boulhic et al. treated alarm colours separately in their study; however, in this thesis, they are included alongside pipe colours. Furthermore, the experts and novices in their study were of different experience level, since in this thesis the most experienced participants are professionals with many years of experience and the least experienced still had some familiarity with their respective industry's SCADA systems.

3 Understanding colour

Judd & Wyszecki (1975), as cited in [6], estimate that the number of colours is as high as ten million. All these different colours can be perceived differently based on many factors both human and environmental that are present. Technically colour itself can be described just as the visual property of light that is not related to things such as lightness, saturation, texture, glossiness or translucency. In reality, colour holds great impact in our daily lives by adding richness and immediacy to visual information, while also affecting mood and performance. [6] With such strong effect, colours are often subject to different guidelines or stricter standards.

The visual design guidelines, especially for colour that are used while designing are subject to the context in which they are used. The designer's job is to make sense of the guidelines and understand the product and its users. A successful interface communicates clearly to the users which content is available and how it is possible to interact with. In addition to following general guidelines, designers will also have to keep in mind the organisation's branding goals. [7] Even though guidelines are helpful tools, it takes more to understand how colour truly works and the differences in how people perceive it.

3.1 Colour perception

Colour perception varies between individuals based on biological factors, either innate or through age. For example, as we age, the lens in our eye yellows and as a

result blues become duller [6]. Colour blindness is also something that affects many people. Approximately 1 in 12 men are affected by it, but only 1 in 200 women, with estimation of about 300 million people worldwide [8]. The effect of gender to colour perception goes even deeper than just colour blindness.

Bimler et al. [9] have discovered that males and females have some differences in their average responsiveness to specific dimensions of colour space. These include such as males placing less weight on the red-green axis, while also placing more weight on the separation in the lightness axis. Of course there are also differences between each gender group, and these generalisations should be taken critically. [9] However, Abramov et al. [10] have performed a study that shows males require longer wavelengths to experience the same hues as females do, which affirms a biological difference between the genders. Hence, despite these variations being small and colour perception being remarkably similar between individuals, differences do exist. [10] This is an important notion when considering SCADA systems, since the operators might be of either gender and for a safety-critical system it suggests that colours should be paid close attention to.

Additionally, Iriguchi et al. [11] have discovered that hormones have some effect on colour perception between males and females. The effects of menopause were studied, and while the results cannot be fully excluded from aging as of yet, they show some potential for pre- and post-menopausal differences in colour perception [11]. Also, testosterone and estrogen have been hypothesised by Da Silva et al. [12] to be linked to the perception of colours, especially on the red-green axis. Their finding that decreased perception in greyscale and on the red-green axis was observed in females who used hormonal contraceptives also suggests to the fact that hormones play a role in colour perception. [12]

3.2 Cognitive processing of colour

As colour is something most people see similarly, it is also interesting to approach it from a cognitive processing perspective. For designing interfaces, it is important to know if some colours cause cognitive strain or if other choices make using the system easier.

The design of the interface can lead to different levels of engagement in the user based on the level of information that is presented there and how. If the visual experience offers too little information to the user, it will feel understimulating and the brain will try to reject that information, rather focusing on something else. On the other hand, information that is too abundant causes a similar reaction, where the brain aims to reject what it cannot organise or understand. Therefore, it is important to keep the information harmonious and avoid under- and overstimulating designs. [6]

In addition to the amount of stimulating information, there are other factors that affect the attention of the users. Humans are constantly forced to make decisions between what attracts our attention and what does not. The human eye also has a fixational reflex, meaning that typically it notices and focuses on objects that are bright and feature movement, and usually those objects are deemed significant. [13]

As established, using bright colours can draw the attention of users to specific elements on a screen and help them to direct focus on them. Using too many colours, however, will weaken any of the colours used. [14] Other than bright colours, contrast is also something that draws attention quickly and easily, making it a good choice when elements need to stand out [7].

Colour and contrast are important aspects for distinguishing detail and depth while also helping to catch and divert the attention of users. As such, they are important to consider as tools for indicating key areas in design. [13] Besides key elements, colour and contrast are important for the background as well. Wang et al.

[15] observe that background colour has some impact on the accuracy of users, but it has not shown to affect response speed. From the tested colours, green, red, and blue, green showed the highest accuracy with higher brain activation. [15] However, it is important to keep in mind how such things as saturation and brightness affect tests like these, where different colours are compared to each other.

Wang et al. performed the tests so that the luminance contrast was kept the same across all different combinations, but the colour difference was altered. They found that higher saturated colours with large colour difference impaired cognitive performance. Hence, even if higher saturated elements might be easier to recognise for the user, the high saturation makes it harder to process what is being read. [15] To summarise, readability does not always equal to comprehension. It is important to check luminance contrast to ensure legibility of text and key elements and to keep the background colour desaturated to support comprehension. However, with some key elements, like alarms, it makes sense to rely on high saturation, since those rely heavily on fast acknowledgment and colour associations.

3.3 Colour associations

As colours affect us both physically and psychologically, they also have some set meanings for most people. They are associated with different things and concepts, while also evoking varying emotions in people. It is also important to note that countries and cultures play a significant role in the associations that the majority of the population have about colours. Common associations include such as green meaning freshness, white meaning purity, and purple meaning royalty. More common colour associations can be found in Table 3.1. [6]

Such as in traffic lights, the red, yellow, and green are the most interesting ones from these for the SCADA alarm colours. In Table 3.1 the negative association for red includes danger, for yellow there is caution and warning, and for green the

Table 3.1: Colour associations [6]

| Colour | Positive association | Negative association |
|--------|--|--|
| Red | Sense of power, strength, action, passion, sexuality, and love | Anger, forcefulness, danger, impatience, violence, and revenge |
| Yellow | Brightness, warmth, and joy | Caution and warning |
| Blue | Cool and tranquil, patience, comfort, loyalty, and security | Fear, coldness, passivity, and depression |
| Orange | Warmth, friendliness and cheerfulness, excitement, and energy | Ignorance |
| Purple | Royalty, sophistication, and religion | Bruising or foreboding |
| Green | Nature, growth, soothing, and freshness | Envy, greed, guilt, and jealousy |
| Black | Dramatic, sophisticated, stylish, and soothing | Evil and death |
| White | Pure, fresh, easy, cleanliness, and goodness | Blind, winter, cold, and distant |

positive association includes a soothing emotion. These colours also have the opposite associations from these, but as they are so commonly used to indicate different states of systems or even to direct traffic, the correct associations come more naturally.

Even with this said, there are still cultural differences with the actual associations for colours that are commonly used for alarms or statuses. An extensive look at different colours and their associations in different cultures can be seen in Figure 3.1 [16]. Explicitly red is associated with danger in only two cultures, Western/American and Japanese. To look at it more broadly, association with anger is in four cultures: Western/American, Japanese, Eastern European, and African. In South American culture it is also associated with death.

Besides cultural and learned associations, colours are also fundamentally contextual. They are dependent on the information that comes before and what is seen simultaneously from the surroundings. Given this contextual nature of colour, it is important to note the fact that using only colours to serve as a reliable signal for visual tasks might not be the best course of action. Continuing from this, it has been

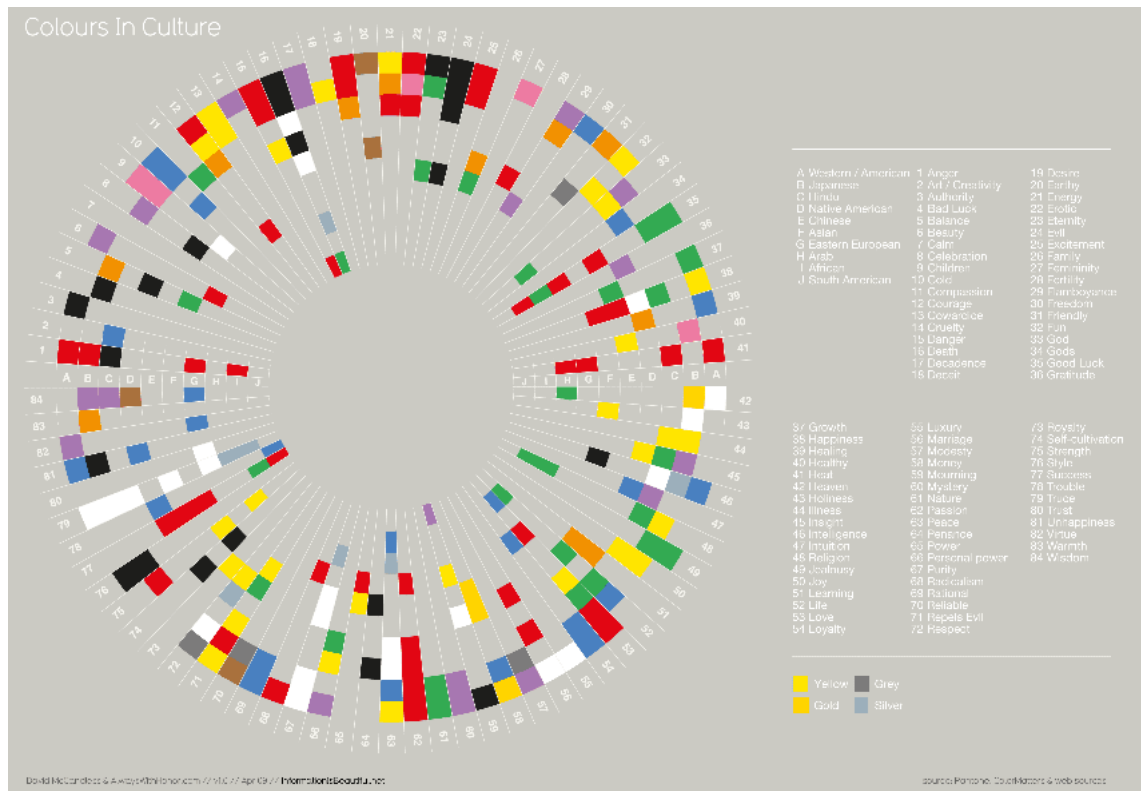


Figure 3.1: Colour associations in different cultures [16]

also suggested that colour could not convey the same level of information that form can. Therefore, even for a colour like red, there could be some difference in how the colour is perceived based on if it is in the foreground or background, or if the colour is used for an object or for illumination. [6]

Another contextual factor for colours comes from professional backgrounds. In some professions, colours have specific meanings and can be even taken advantage of. [7] Thinking about SCADA operators, it is also good to know the industry standards and customs as well as more common associations for colours. In addition to having meanings in professions, colours can also establish a company's brand identity. Using specific colour schemes in the interfaces of their systems can create a recognisable visual identity [14]. In some cases, the brand colours might even overrule some other

association or contextual factors in decisions. It is important to find balance between all these elements without losing usability.

4 User-centred design

All systems and their user interfaces (UI) should be easy to learn as well as useful, while remaining easy and pleasant to use [17]. Since computers have become more common in homes, UI design has been following many similar requirements. Still, many methods misestimate the value of the interaction with users during the design process [17].

Some designers believe that rational thinking is the solution to creating functioning UIs. With rational analysis of tasks and the best way to perform them, they believe that there is no need to involve users and their existing ways of thinking into the process. Another belief is that design guidelines are sufficient to create functioning UIs. As they provide general recommendations for the designs, they can be helpful in the first stages of designing yet in the end, they do not make up for actual user involvement. [17]

As Gould and Lewis [17] have suggested initially in the 1970s, there are three design principles that should be followed. First, there should be early focus on users and the tasks that they will be performing. By different methods, the designers should first aim to understand the users of the system they are designing. Second comes empirical measurement, meaning that the intended users should be able to use a prototype of the system early in the process to determine the performance of the system at that stage. Lastly, iterative design is highlighted, which means that

designers should utilise user testing to find problems and fix them and repeat this cycle of design as many times as necessary. [17]

These principles were not often followed by designers directly as they were intended, either deliberately or by mistake, and they were often undervalued. User diversity can be underestimated if the designers have limited contact with the users, resulting in incomplete understanding of their differences. Alternatively, user diversity can be overestimated, resulting in the feeling that user testing would not be sensible with only a few users, but in reality, testing with many users might lead to useless results. One could also believe that users might not know what they need. This view surfaces the issue that the actual acquisition of useful information from users might be a difficult task, since direct questioning will most likely not give the results needed. Some designers might also think that their job does not require or even permit contact with users, placing the responsibility of the usability completely elsewhere. [17]

This focus on users is the essence of user-centred design (UCD), and as UI and user experience (UX) design have become more researched and studied, it has become increasingly popular as a design method. Nonetheless, there is the question if UCD practices truly address the parts that are important for users in a system.

Keinonen [18] addressed this topic in a paper in 2008, where he examined UCD and the different terms that were in use at the time and how well they answer to user needs. As established, UCD can be seen as an umbrella term, which covers approaches including but not limited to human factors and ergonomics, usability measurements and inspections, and design for user experience. As design systems and methods have developed, UCD has moved towards UX, which highlights more the emotional aspect of the design, meaning how it affects the users and how pleasurable it is to use in contrast to just plain usability. Another shift is that the term UCD itself has been moving towards human-centred design, since some feel that the concept of

‘user’ does not fully represent the nature of humanity. [18] Even the ISO standard 9241-210 is titled ‘Human-centred design for interactive systems’ [19]. However, in this thesis it is still defensible to refer to UCD, since the term ‘user’ differentiates the role from designers. Furthermore, the user of a SCADA system is a very specific user, even more separating it from the more general human-centred design.

4.1 Usability

As UCD covers more the philosophy behind the design method, usability offers the measurable quality. The ISO standard 9241-11:2018 defines usability 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.’ Jakob Nielsen alternatively defined usability in 1993 as having multiple components, for instance learnability, efficiency, memorability, errors, and satisfaction [20].

Learnability means that the system should be easy to learn and users could easily start working with the system. Efficiency more specifically means that the system should be efficient to use in a way that after learning the system, users can be productive. Memorability refers to the system being easy to remember, therefore casual users can return to the system without having to relearn everything. Errors means that users should make few errors while using the system and if errors happen, users can easily recover from them. Satisfaction refers to the pleurability of the system, and that users feel subjectively satisfied while using the system. [20]

Usability can be evaluated through testing, but also by going over the design while checking heuristics. In 1994, Nielsen refined the 10 usability heuristics that he had been working on [21]. Nielsen’s 1994 article presented nine heuristics as the most relevant, to which a tenth was added in the later popularised version [22]. The nine from his article have remained the same for over 20 years already, proving their

relevance. The popularised version of Nielsen's heuristics as presented in the cited source is displayed in Figure 4.1.



Figure 4.1: The popularised version of Nielsen's heuristics [22]

Some of these heuristics provide good guidelines for creating the prototypes for the mimic colours that will be introduced in the following chapter. Not all are applicable to this thesis; however, 1, 2, 4, 6, 8, and 9 are all adaptable.

4.2 Accessibility

Usability and accessibility are closely related concepts. The accessibility of a product addresses discriminatory aspects of UX for people with disabilities. [23] This means

that accessibility can be seen as an extension of usability, whilst being broadened to include users with impairments [24].

The four main types of impairments can be categorised as visual, hearing, motor, and cognitive impairments. These manifest as problems such as difficulties reading text with low contrast between text and background or problems navigating with a mouse. [24] Contrast issues are directly related to the topic of this thesis, since it is focused on colour use in SCADA mimics.

Related concepts such as ‘design for all’ or ‘inclusive design’ are related to accessibility, but those focus on designing for the widest range of users without a direct focus on users with impairments. However, this might be a difficult task to complete since some user groups might have needs that contradict each other, leading to a compromise in the design. [24]

Implementing accessibility guidelines has been said to provide benefits for non-disabled users as well [23], [24]. Considering accessibility, such as sufficient contrast and not relying on colour alone, all operators would additionally benefit from it, not only those with colour vision deficiencies.

5 Existing standards

Standards and guidelines in UI design offer many benefits for both end-users and designers. These benefits include increased productivity, reduced mental and physical stress, and reduced training expense. When designs follow the same standards, they will also have improved user-system interoperability across different applications and improved overall product quality. [25]

There are multiple different resources that designers can use when creating interfaces, including principles, guidelines, and standards. Books that contain so-called principles are considered as the most informal, since they do not involve undergoing a systematic consensus-building process. These principles provide mostly general guidance that is applicable to most UIs. Style guidelines are more formal, since they usually include a systematic review process done with representatives from consortia members and other interested parties. However, guidelines lack a formal process to achieve consensus among all affected parties. Actual standards, such as the ones by ISO or ANSI, have documented development procedures and they ensure that all companies and individuals that are affected by the standard have an opportunity to represent their interest and take part in the development process. This means that before standards are approved, the procedures need a consensus-building process, which is lacking in style guidelines. [25]

The way that standards work is that they have requirements and recommendations. A requirement can be recognised by the word 'shall', meaning that the if the system

is stated to follow a standard, it needs to conform that point. Recommendations use the word ‘should’ and systems that follow the standard are not required to conform these statements. [25]

When creating HMIs, the foundation is in the statutory and regulatory requirements. After these requirements in the hierarchy are standards, which should offer a sound basis for designing corporate guidelines and specifications. HMIs can look different based on where they were designed, but they should all fulfil at least the two foundational layers of standard and guideline hierarchy. At the top is HMI philosophy, which dictates consistency between systems and technologies inside companies. This HMI design hierarchy is illustrated in Figure 5.1. [26] Although standards like ISO are intended to be used as voluntary, sometimes they can also be cited in legislation [25].

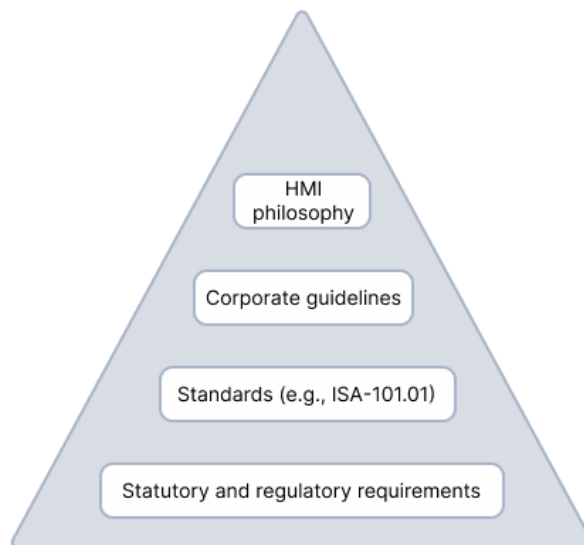


Figure 5.1: Hierarchy of rules when designing HMIs (adapted from [26])

An example of how standards are created from ISO is their working group 5 (WG5) that works on software ergonomics and human computer dialogs. WG5 includes professionals from government, academia, and industry. In addition, it includes

user groups that represent approximately twelve countries. Within this committee, standards are created by using consensus to agree on proposed material. [25]

Standards are not a perfect solve-all solution and have their limitations. There are no guarantees that following a standard will make the product usable. UI development also evolves all the time, and sometimes standards are not caught up, leaving designers to improvise. Standards are also not platform-specific, meaning that they are designed to be independent from any platform, thus making the guidance general and sometimes difficult to interpret. Even if they are made to be general, they still do not always apply universally in all cases and require the designer to use their own judgement. [25]

5.1 ISO 14726 – Physical piping colours

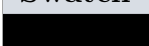











The International Organisation for Standardisation (ISO) is a global federation that is composed of national standards bodies, referred to as ISO member bodies. ISO technical committees typically draft standards and each interested member body has a right to be represented in that committee. These drafts need approval of at least 75% of the voting members before it can be published as an international standard. This represents the consensus that is required in the process of developing standards. [27]

ISO 14726:2008 is an international standard for ships and marine technology and specifically identification colours for the content of piping systems. It was prepared by subcommittee SC 3, which is responsible for piping and machinery. SC 3 is a subcommittee of technical committee ISO/TC 8, which is for ships and marine technology. [27] Withdrawn versions of the standard are ISO 14726-1:1999 and ISO 14726-2:2002, and the current published version ISO 14726:2008 has been last reviewed in 2024 and will be reviewed every five years [28].

The standard specifies a main colour and an additional colour for identifying piping systems in ships and marine technology. Main colour is meant to indicate a group of similar media and additional colour is used in combination with the main colour to indicate a specific medium. These colours are not only for physical piping but can also be used for drawings and diagrams representing piping systems. Additionally, the standard can be used for land installations, but that is limited to marine structures. [27]

The main colours and the media they represent are in Table 5.1. These are the colours that are the same for similar media, yet there are differences inside that group. For example, the main colour for flammable gases is yellow, therefore the colours for hydrogen are yellow-blue-yellow, for acetylene they are yellow-grey-yellow and for liquid gas yellow-white-yellow. The additional colour that separates media is seen in the middle and the main colour that separates the medium into groups of similar media is seen on the outsides. [27] Not every possible medium is described in the standard; however, following its logic it is possible to add the missing media with appropriate adaptation.

Table 5.1: Main colours of pipes from ISO 14726:2008 [27]. Colour swatches are approximate visual representations.

| Swatch | Main colour | Media |
|---|-------------|----------------------------|
|  | Black | Waste media |
|  | Blue | Fresh water |
|  | Brown | Fuel |
|  | Green | Sea water |
|  | Grey | Non-flammable gases |
|  | Maroon | Air and sounding pipes |
|  | Orange | Oils other than fuels |
|  | Silver | Steam |
|  | Red | Fire fighting |
|  | Violet | Acids, alkalis |
|  | White | Air in ventilation systems |
|  | Yellow | Flammable gases |

Since mimics in the marine industry require many individual colours to distinguish between different pipes, following ISO 14726:2008 might make the recognition of pipes and their media easier for the users. Not only for users who are familiar with the physical piping colours, but also for others as a faster way to notice differences between pipes. If a designer must choose a single solid colour to represent pipes, there might be issues with colours appearing too similar to each other. However, using main and additional colours might cause the mimic to become too crowded and overstimulate the user with information.

5.2 ISA-101.01 – HMI design

The International Society of Automation (ISA) is a non-profit professional association, and it develops widely used global standards. ISA's industry standards are developed in consensus by thousands of professionals worldwide. [29] The world's primary international standard organisations are the International Electrotechnical Organisation (IEC) and ISO, which are both sanctioned by the United Nations [30]. ISA differs from these organisations by accepting individual people to join and participate, whereas IEC and ISO accept countries as members and each country speaks with a unified voice. Since IEC and ISO only accept countries, ISA cannot directly contribute and must go through a country. ISA goes through the United States represented by American National Standard Institute (ANSI). That is why sometimes the ISA-101.01 standard is also referred to as ANSI/ISA-101.01. [30]

ISA-101.01-2015 is a standard for human machine interfaces for process automation systems, which covers designing, implementing, operating, and maintaining HMIs. The standard highlights human factors engineering to improve usability by taking into account the users' sensory and cognitive limits, situation awareness, and ergonomic principles. It covers interaction methods and display styles as well as hardware interface considerations. Some of the main benefits of following a standard

like ISA-101.01 include improved safety, enhanced operator effectiveness, consistent HMI design, and optimised situation awareness. Having different vendors' HMIs following a standard is important for operating industrial processes safely, efficiently, and effectively. [31]

5.2.1 General colours

General principles mentioned in ISA-101.01-2015 include that failures of a display or items on the display should be immediately apparent to the operator and the HMI should aid in detecting and responding to abnormal situations. Regarding the consistency of the design, the HMI should have a consistent look and feeling for the user. For situation awareness specifically it is stated that if the process is functioning as expected, then the display should show minimal sensory stimuli. Highlighting deviating from the expected should be signalled by visuals and/or audible signals. Situation awareness is especially important to consider, as it has been identified as one of the main factors in accidents caused by human error. [32]

ISA-101.01-2015 includes a specific subsection for colour within the broader subsection for visual considerations. It is stressed that colour vision deficiencies (e.g., colour-blindness or age-related) should be considered during design, and lists red-green, green-yellow, and white-cyan as common impairments. To accommodate for colour vision deficiencies contrast and brightness can be utilised yet should still be considered within the context of the operating environment. Age-related deficiencies, such as dulling blues and purples to grey, should be taken into consideration if they are applicable for the environment. [32]

Furthermore, in the colour section it is mentioned that colour coding should be used as an addition to other means of presenting information (e.g., shape, text, brightness, size), which should be the main means of conveying information to the user. Important information should be perceptually more notable than less

important information, thus the use of colour should represent the importance of the information being presented. Other colour coding could be used for clarity and emphasising elements; however, colour should still be used conservatively and consistently throughout the HMI to indicate information. [32]

In addition, background and foreground interactions are defined in the standard. For the background an unsaturated or neutral colour should be selected in order to limit any distortion to colours and ensure that main information stays visible. The colour combinations for the background and foreground should provide sufficient contrast so that the elements would be clearly visible. The background colour should also account for possible ambient conditions that the HMI will be used in. Rendering images in greyscale is mentioned as a helpful tool to ensure that all elements have sufficient contrast to remain detectable to users with colour vision deficiencies. [32]

The standard considers visual dynamics, and it states that only a limited number of colours can be used effectively on displays because of the limits to perception and cognition. This results in designers having to rely on motion, like changes in position or visibility. Excessive motion can be distracting, which means that the use of visual dynamics should be reserved for highlighting abnormal situations. [32]

5.2.2 Alarm state colours

ISA has a standard ISA-18.2-2009 for alarm management, and in accordance with it the ISA-101.01-2015 states that the colours used for alarms should only be used for those and not used for any other purpose to strengthen their cognitive meaning. Colour should not be the only used method for differentiating between alarms, which could be resolved by, for instance, using symbols to help distinguish between alarms. As noted above about the most important information being presented with the most noticeable colours, it naturally applies to the alarms and should be considered when designing colour schemes for alarm states. [32]

An abnormal situation can be visualised through blinking or flashing; however, it should only appear on elements that are intended to appear and disappear. An example of that would be symbols or borders on elements when there is an unacknowledged alarm. Once the alarm has been identified, the blinking or flashing should stop and return to a solid state without any changes in visibility. This should also be limited to just form and not text and numbers, since it would make reading them difficult for the user. [32]

5.3 W3C WCAG2ICT – Accessibility

The World Wide Web Consortium (W3C) is an international non-profit organisation where web standards are developed by member organisations, a full-time staff, and the public [33]. Through the Web Accessibility Initiative (WAI), W3C has published the Web content accessibility guidelines (WCAG) [34]. The current version is the WCAG 2.2, which is an approved ISO standard ISO/IEC 40500:2025. Previous versions of WCAG 2 include 2.1, 2.0, and currently WCAG 3 is being developed. The newest version of WCAG 2 does not deprecate or supersede the previous one, but W3C encourages the use of the latest version of WCAG. [34]

WCAG explains how to make web content more accessible, and it is mostly intended for people who need a standard for web accessibility, like web content developers [34]. Since WCAG was designed for web content, it assumes the presence of a browser which would not be accurate for non-web software, like SCADA HMIs. To address this, W3C published a document called Guidance on applying WCAG 2 to non-web information and communications technologies [35]. It supports accessibility solutions across different technologies, and it was first published in 2013. The latest version of WCAG2ICT covers WCAG 2 levels A and AA. [35]

As described in WCAG 2.0, level A is the minimum level of conformance where the web page satisfies all level A success criteria. Both level A and level AA success

criteria need to be satisfied for level AA conformance. For level AAA conformance, success criteria for levels A, AA, and AAA need to be satisfied. However, it is not recommended that level AAA conformance be set as a requirement, since it is not possible for some content to satisfy the criteria required. [36]

WCAG2ICT addresses the use of colour specifically, and one level A criterion is that colour should not be used as the only visual means of conveying information. In addition, there is a level AA criterion which is for non-text contrast. It states that the visual presentation of user interface components (excluding inactive components) and graphical objects should have a contrast ratio of at least 3:1 against adjacent colours. [37] These two are the criteria from WCAG2ICT that apply to pipe, alarm, and state colours in SCADA mimics.

5.4 Observed industry practices

While the standards introduced in the previous sections provide general frameworks, there is still considerable variability in SCADA systems between vendors, projects, and industries due to the lack of further formal standardisation. An industry trend appears to be to include mainly greyscale or muted colours in HMIs. Rockwell Automation has a white paper which is a style guide for process HMIs, where they directly give guidance about colour conventions [38]. They instruct to use colour mindfully in HMIs and be careful about overusing it only to make the display appear less dull. They state that alarms should have bright and intense colours, which are not used elsewhere in the system. Live data should be visualised with less intense and cool colours and examples of dark blue or dark green are provided. Display background should be non-saturated, such as light grey and foreground colour should be minimised. [38]

For alarm colours, Rockwell suggest magenta as low priority, yellow as medium, orange as high, and red as urgent priority. Furthermore, they use distinct forms for

each alarm, which are illustrated with the colours in Figure 5.2. [38] A different vendor, AVEVA, has forms and colours for alarm indicators organised differently than Rockwell. AVEVA has red for highest severity, yellow for second highest, cyan for second lowest, and magenta for lowest severity [39]. The figures and specific colours are illustrated in Figure 5.3.



Figure 5.2: Rockwell guideline on hierarchy of alarms (adapted from [38])



Figure 5.3: AVEVA guideline on hierarchy of alarms (adapted from [39])

This shift towards greyscale is also referred to as ‘high-performance HMI’. Instead of using green and red for active and inactive components, a high-performance HMI should use white and greys. White should be for running, dark grey for stopped and medium grey for a component that is not sending signals. [40] There has also been discussions whether this shift is good for HMIs and if minimalism or realism should be sought after, but overall, the trend seems to have wide support [41]. In Wärtsilä currently, there is a compromise between the two trends at use, where green is for active, white for inactive, and only outlines if the component is disabled. For the energy industry, however, there is also the difference that in the US red colour is used for system state to indicate active and green inactive, argued by cultural difference.

These guidelines from vendors demonstrate that while companies have developed their own colour conventions, there is no universal agreement between them. The

differences in alarm hierarchies alone illustrate the challenge that operators have to navigate when moving between systems.

6 Prototype development

The preceding chapters reviewed colour theory, standards, and industry practices, which are brought together into design guidelines in the following section. Thereafter, the design process of each prototype is described.

6.1 Guideline creation

From the standards, research, and industry practices introduced in the preceding chapters, a set of 12 guidelines was created. Common principles were identified across the reviewed standards and literature, and where recommendations overlapped or reinforced each other, they were consolidated into a single set of guidelines. The guidelines are intentionally more general rather than detailed and prescriptive to define the principles which could then be applied to prototypes in both marine and energy industries that require different colours to be used.

The most influential sources were the ISA-101.01 standard for general HMI colour principles, the ISO 14726 standard for piping colours, and the WCAG2ICT guidance for colour use related to accessibility. The guidelines are the following:

1. Colours from the ISO 14726 should be used as reference colours.
2. Background colours should be kept desaturated and neutral (e.g., light grey), while highly saturated colours should be limited to small, attention-critical elements for maximum comprehension.

3. Background should be selected to provide acceptable and sufficient contrast in possible ambient lighting.
4. Using too many colours could weaken any of the colours used, hence sometimes it is enough to rely on contrast. However, excessive contrast should be avoided to prevent eyestrain.
5. Converting the interface to greyscale is recommended to check the contrast for legibility.
6. Colour choices should not make the interface overstimulating but also not understimulating for maximum understanding.
7. Colour should be used conservatively and consistently.
8. Colours used for alarms should not be used for other purposes to avoid weakening their cognitive meaning.
9. The use of visual dynamics (e.g., flashing or blinking) should be reserved for highlighting abnormal situations.
10. For age-inclusive design, lens discoloration should be considered, which tends to dull blues and purples to greys.
11. Colour should not be the only used method for conveying information or distinguishing a visual element, as this excludes users with colour vision deficiencies. Possible alternatives or additions can be such as form or descriptive text.
12. Interface components (e.g., visual information) should have a contrast ratio of at least 3:1.

In a more detailed mapping, the ISA-101.01 is the heaviest contributor as guidelines 2, 3, 4, 7, 8, 9, and elements of 11 are based on it. ISO 14726 directly informs

guideline 1, and WCAG2ICT contributed to guidelines 11 and 12. From Chapter 3 the sections about colour perception and cognitive processing contributed to guidelines 5, 6, and 10. The observed industry practices from Section 5.4 also align with several of the guidelines.

6.2 Prototypes for energy

Three prototypes were created for the energy group based on an exhaust gas system mimic from a Wärtsilä power plant management system. All three prototypes shared the same layout, component placement, and data values, with only the colours of the elements varying between them.

Prototype 1 (P1) portrays the mimic with the current colours at use in Wärtsilä. P1 is represented in Figure 6.1. It contains three types of pipes: yellow for gas and exhaust, green for water, and blue for air. There are different types of P&ID symbols and values showing readings from sensors, which are set to 0 as placeholder values. There are alarms which can be seen on the mimic and on the top of the UI, where the different amounts are visible. The alarms are modelled after the ones by AVEVA introduced in Section 5.4. Each alarm is present in the mimic once and the places were chosen at random. The states of the different components follow the logic introduced in Section 5.4, meaning green for active, white for inactive, and only outlines for disabled. Grey dashed lines are used to represent grouping logic and some groups are titled. The prototype was designed with the other UI elements since that is how operators would perceive it. As the prototypes were presented as static, the different blinking logics of the alarms could not be implemented, therefore they are overlooked in this case.

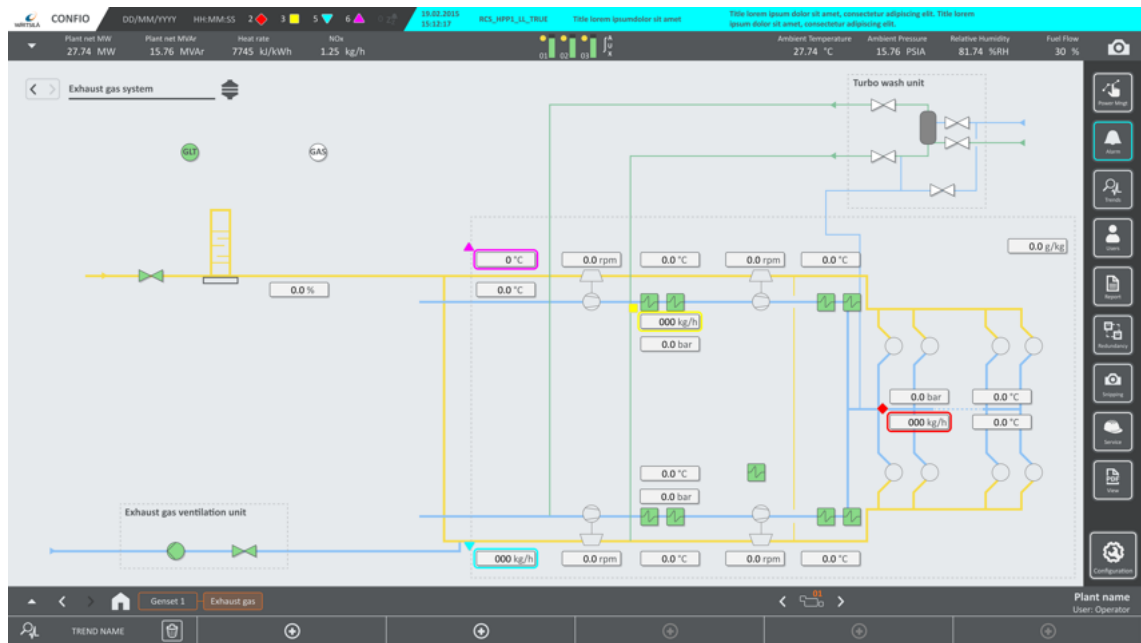


Figure 6.1: Prototype 1 for energy

Prototype 2 (P2) is presented in Figure 6.2 and it portrays the mimic closely following the guidelines introduced in Section 6.1. The pipes were changed from P1 to be thicker to offer more visual clarity with additional colours. In P1, the pipes are 2 pixels wide for regular and 4 pixels wide for double walled, whereas in P2 they are 4 pixels for regular and 6 pixels for double walled. In addition, the pipes in P2 have an outline which is 2 pixels wide on both sides and represents the main colour of the medium. The colours are referenced from ISO 14726; however, their hue and saturation have been altered to appear more calming to the eye. Exhaust gas is represented by black-grey-black, natural exhaust air by white-brown-white, and gas-turbine wash water by blue-orange-blue. Each of these combinations has at least a contrast ratio of 3:1 with each other, excluding the blue-orange-blue combination in favour of preventing eyestrain by choosing a more desaturated orange colour. The outer main colours were very close to having a 3:1 contrast ratio with the background, except for the white-brown-white, where the brown was the more contrasting colour.

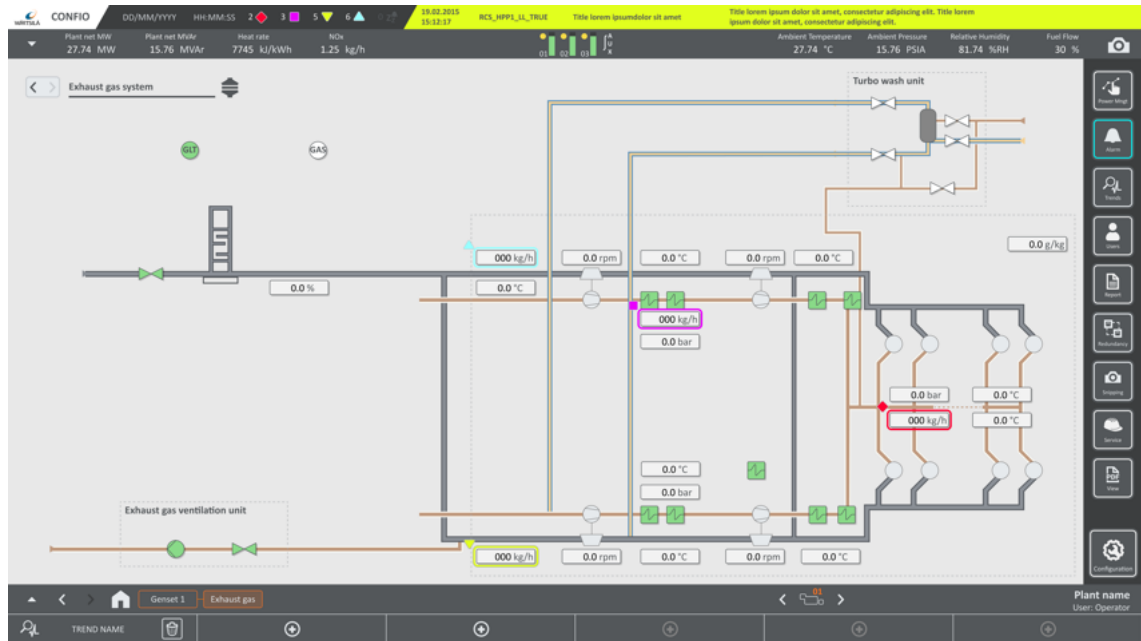


Figure 6.2: Prototype 2 for energy

The alarm colours and the hierarchy of them was also changed when compared to P1. In P1 the alarms are red, yellow, cyan, and magenta from most severe to least severe. However, when converting those to greyscale, the luminance contrast is not following a logical hierarchy as red and magenta are the most visible colours. Ordering the colours based on the luminance contrast results in having them as red, magenta, cyan, and yellow. That is concerning regarding associations with yellow for caution causing confusion if it could appear as least severe. Each colour was then slightly altered, as the red was changed to be slightly pinker, yellow more towards green with an added darker border, cyan to be more a light blue, and magenta to be slightly darker. These changes made the luminance contrast follow the ordering of red, magenta, yellow, and blue from most to least severe. The alarms and their greyscale versions are illustrated in Figure 6.3. Red was left as the most severe since it has a common association to danger, but yellow was moved from second highest severity based on the contrast, even if it has an association to caution. The contrast

was checked and red reached a contrast ratio of over 3:1, but all others were lower as contrast ratio gradually decreased when the alarm became less severe. Colours with higher contrast ratio were tested, but they became too dark to invoke any feeling of alarm any longer.

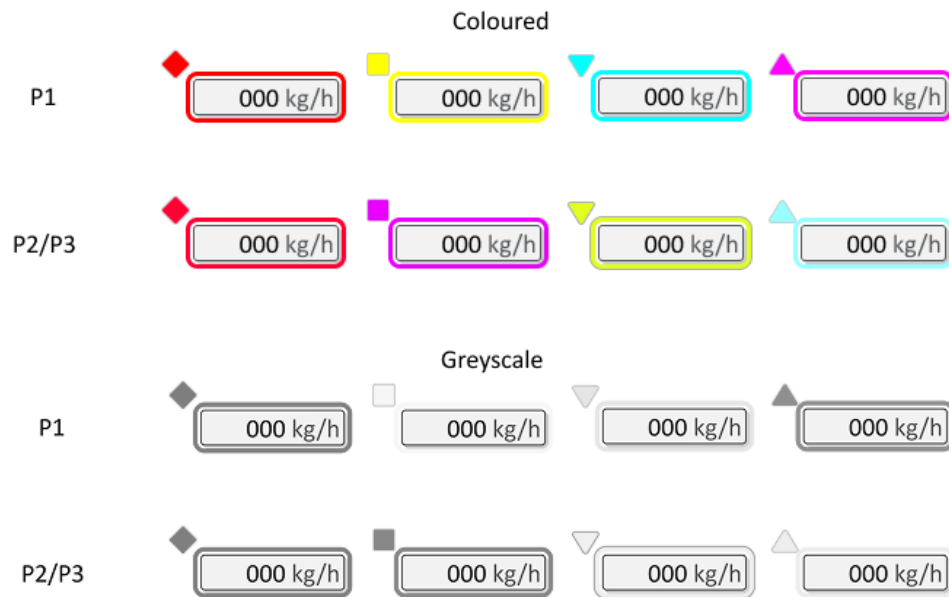


Figure 6.3: Alarms for P1 and P2/P3

The state colours were not changed to match the common convention in the industry, since the pipes are not without colour in these mimics, which already breaks the basics of the greyscale ‘high-performance HMI’ approach. The background was changed from a slightly blue-tinted grey to a greyscale version of that exact colour. This follows the created guidelines and offers a more neutral background for the mimic. The two different background colours are presented in Figure 6.4.

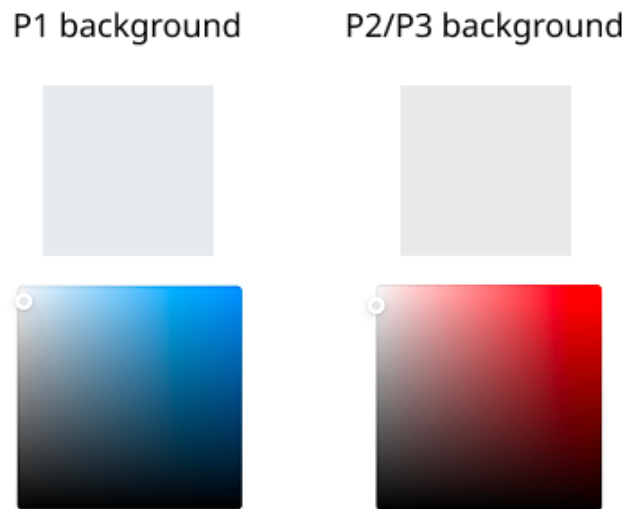


Figure 6.4: Background colour for P1 and P2/P3

For prototype 3 (P3) that is presented in Figure 6.5 the changes made to the alarms and background were kept the same as in P2. The only change was done to the pipe thickness and colouring. The pipes are the same size as in P1, meaning 2 pixels wide for regular and 4 pixels wide for double walled. The colours for the pipes were taken from P2, but the main and additional colours were combined to create a single colour for each piping medium. That was done by overlaying the colours at 50% opacity over a white background and colourpicking the combination. Visual inspection was used to determine whether the main or additional colour was on top during this phase, since in some combinations reversing the order downplayed the effect of the other colour significantly. The combined colour was then altered for increased contrast, and all except the blue reached the 3:1 ratio limit. This method and the resulting colours is displayed in Figure 6.6.

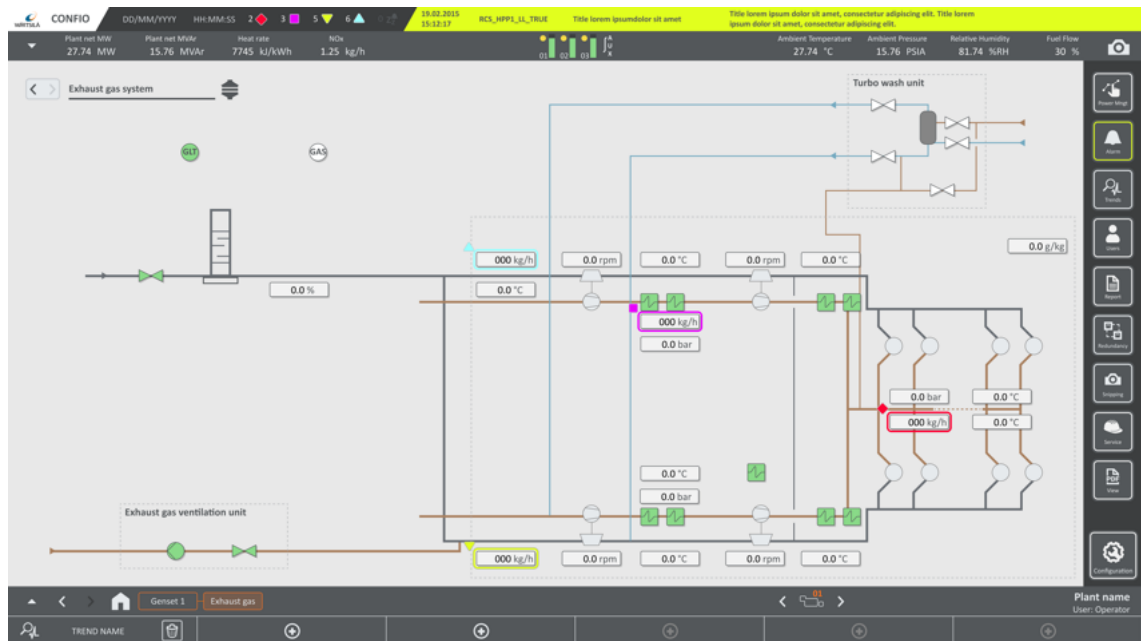


Figure 6.5: Prototype 3 for energy

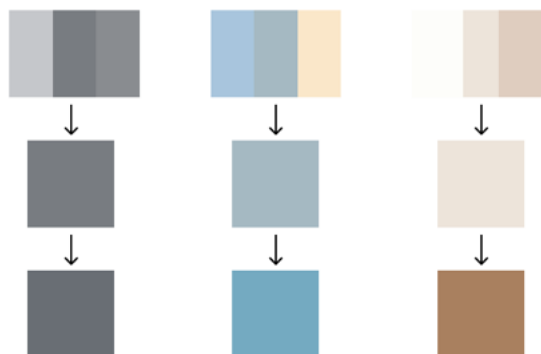


Figure 6.6: How colours were combined for P3 and resulting piping colours

Both P2 and P3 were converted into greyscale during different stages of the design process to check for legibility of the mimic. It was clear that in P2, differentiating between the pipes was easier than in P3 when in greyscale, since the combination of two colours offers more possibilities of luminance contrast. Both P2 and P3 are represented in greyscale in Figure 6.7.

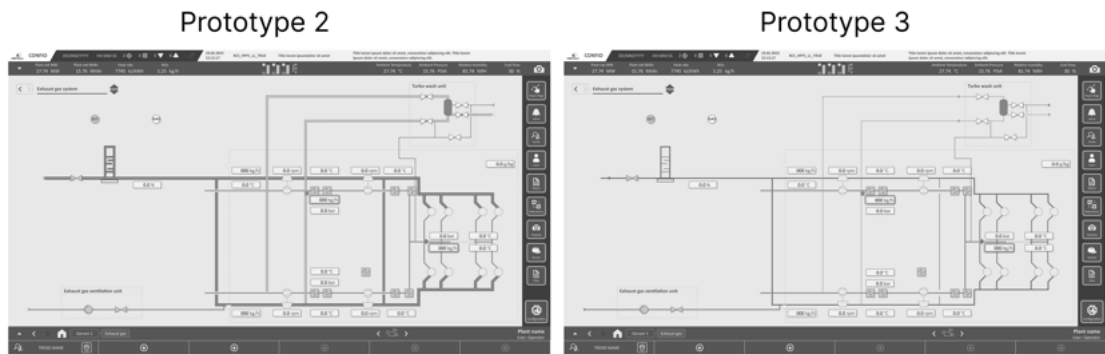


Figure 6.7: Completing guideline 5 by converting both created energy prototypes to greyscale and checking for contrast

6.3 Prototypes for marine

Three prototypes were created for the marine group based on a mimic from a Wärtsilä liquefied natural gas fuel supply system. Similar to energy, all three prototypes for marine shared the same layout, component placement, and data values, with only the colours of the elements varying between them. The creation of the prototypes follows the logic of the ones in Section 6.2 closely.

Prototype one (P1) portrays the mimic in the current style in use at Wärtsilä and it is presented in Figure 6.8. The mimic in P1 was adapted from an older system only for the purpose of this thesis, therefore it is not perfectly accurate to the system it represents (i.e. some valves being different or values missing due to lack of space). P1 is mainly similar to energy P1, yet the pipes are different and the mimic is more complex due to it representing a whole system. It contains five types of pipes: yellow for natural gas vapour, orange for liquid natural gas, white for ventilation, blue for cooling water, and purple for nitrogen. In addition to these, electrical lines are

also included with solid grey. Alarms and component statuses are the same as in energy P1.

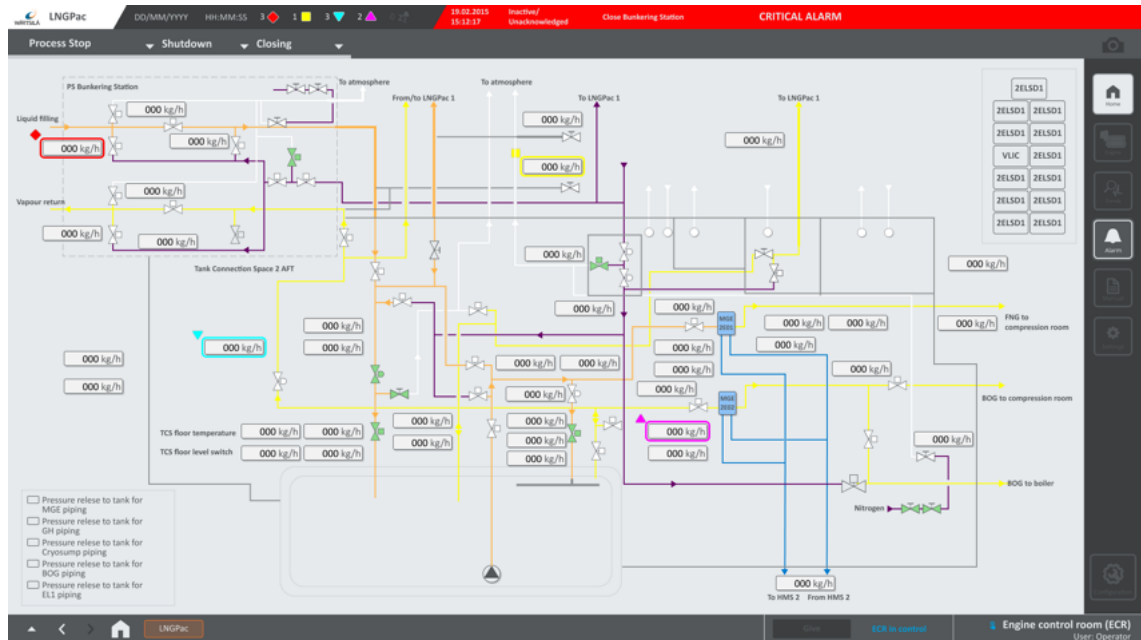


Figure 6.8: Prototype 1 for marine

Prototype 2 (P2) is presented in Figure 6.9 and it portrays the mimic closely following the guidelines introduced in Section 6.1. Same changes to alarm colours and hierarchy, as well as the background colour, were made as in energy P2. Furthermore, in the same manner that was done for energy P2, here the pipes were changed from P1 to be thicker to offer more visual clarity with additional colours. Colours are from ISO 14726 but altered for a more soothing appearance. Liquefied natural gas is represented by yellow-purple-yellow and natural gas vapour by the same colours, but with a lighter hue of the purple, since the convention in current Wärtsilä systems is to represent liquid and vapour with similar colours but different hues. The exact type of air in ventilation systems was unclear, so ventilation is portrayed by a fully white colour in both middle and outlines. Cooling fresh water is represented as blue-purple-blue and nitrogen as grey-green-grey. The same purple was used for additional colour

in natural gas vapour and cooling fresh water. It was done intentionally to avoid adding multiple new colours if unnecessary, since the main colour as the outline helps separate the pipes. The contrast ratios between the main and additional colours of the pipes were between 2:1 and 3:1, since following the guideline directly caused the mimic to appear too visually demanding. The contrast ratio between main colours and the background colour was approached similarly.

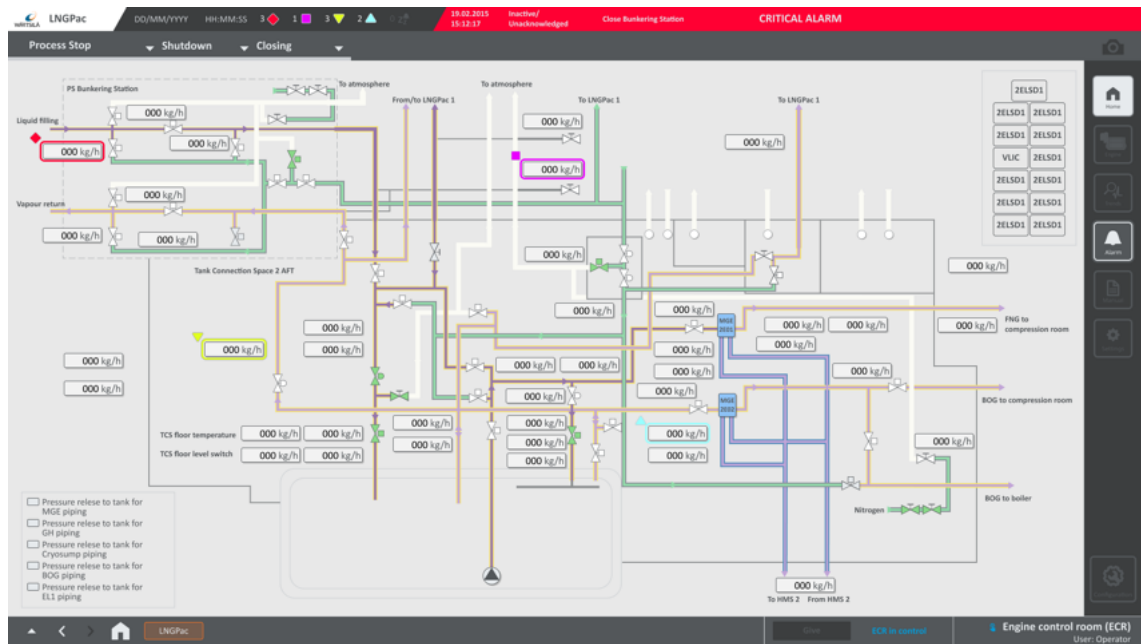


Figure 6.9: Prototype 2 for marine

For prototype 3 (P3) that is presented in Figure 6.10, the changes made to the alarms and background were kept as the same as in P2. The pipe thickness was changed to the thinner one that was used in P1, and the colours were changed to a combination of the main and additional colour in P2. The colours were picked following the same pattern as was done in Section 6.2 for energy P3. The resulting colours are illustrated in Figure 6.11. The white for ventilation and grey for electrical lines required no combining since they were singular colours in P2 as well. Contrast ratios of the pipe colours and background failed to meet the 3:1 guideline in most

and on average it was closer to 2.5:1. Colours with higher contrast ratio were tested, but they became too dark and took attention away from the alarms.

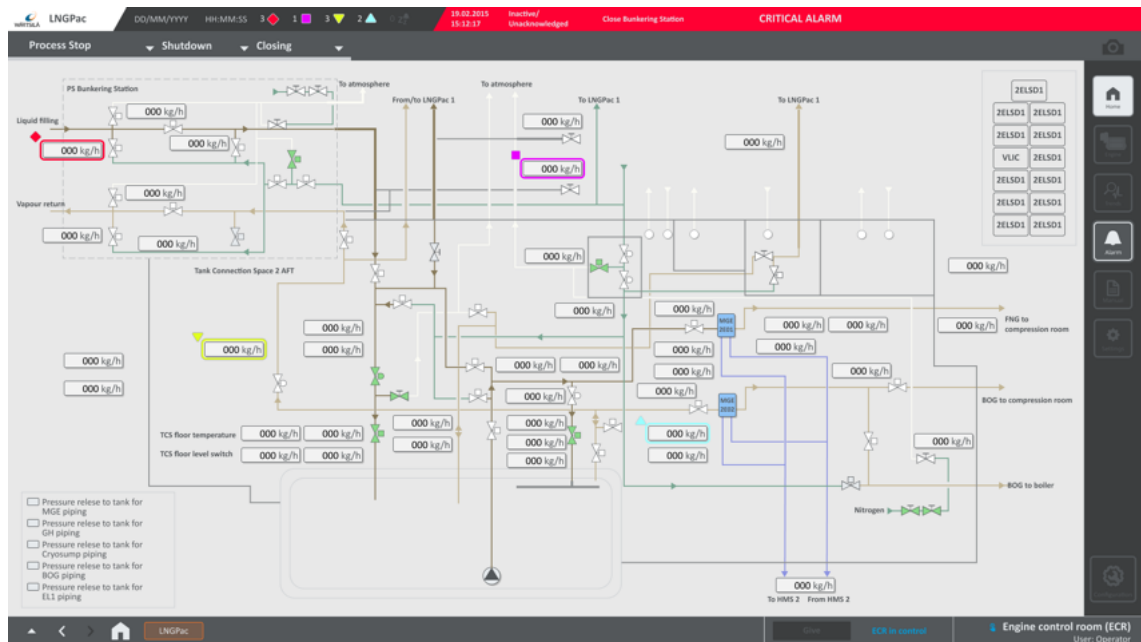


Figure 6.10: Prototype 3 for marine

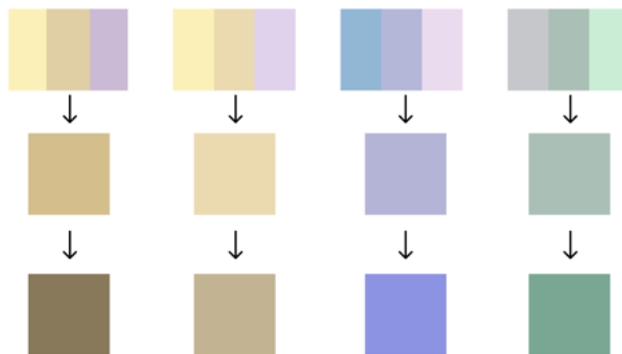


Figure 6.11: How colours were combined for P3 and resulting piping colours

As done before, both P2 and P3 were converted into greyscale during different stages of the process. With a more complex system and more piping colours used, it reinforced the observation from before that in P2 differentiating between the pipes

was easier than in P3 when in greyscale. Both P2 and P3 are presented in greyscale in Figure 6.12.

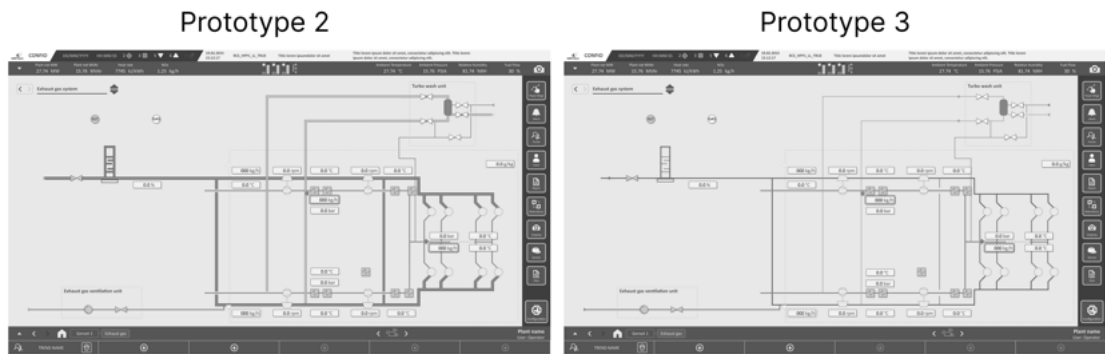


Figure 6.12: Completing guideline 5 by converting both created marine prototypes to greyscale and checking for contrast

7 Testing methodology

To evaluate the prototypes introduced in Chapter 6, user testing was conducted with energy and marine industry professionals. This chapter describes the participation overview and recruitment, structure of the testing sessions, data collection, and hypotheses for the results.

7.1 Participation criteria and recruitment

The participant background questionnaire that was sent out to potential participants is presented in its entirety in Appendix A. The questionnaire was used to match experience levels to the participation criteria and already provide some background information that would be valuable for drawing conclusions from the possible test results.

Participation in the study required professional experience with SCADA or HMI systems in varying degrees in either the energy or marine industry. Participants for each group were selected to have either low, medium, or high experience, though the distribution was not strictly moderated. Besides potential participants reporting their experience with SCADA in years, they also had to evaluate their familiarity with SCADA systems by selecting from poor/fair/average/good/excellent, where responses from different ends of the spectrum were sought. In addition, familiarity with SLDs or P&IDs was asked with the same scale in case some participants felt

more familiar with those and thus that professional knowledge could also be taken into consideration.

Furthermore, as a non-excluding criterion it was asked if the participants had worked with SCADA systems, which vendor they had worked with the most. This was important for understanding what kinds of familiarity bias the participants might have towards Wärtsilä products or if they were carrying mental models from different vendors. Moreover, primary role type was asked for to try and find people with different backgrounds. This study did not focus on testing with operators and focused more on other industry professionals. The list that participants could choose their primary role type from included operator/control engineer/process engineer/designer/HMI developer/other.

In addition to professional background, age and biological sex were included in the criteria, but not as strongly as professional knowledge. Mainly participants were aimed to represent different age groups, which were selected from categories 18-24/25-34/35-44/45-54/55-64/65+. Biological sex was selected from male/female and there was no option to leave that information out, since background research pointed to some differences in colour perception between the sexes. Similarly for findings about how age can affect colour perception, which is another reason it was asked for. It was hoped to have an even split between male and female participants, yet it was also acknowledged beforehand that participants pool might skew towards male.

Participants were furthermore asked about any possible colour deficiencies, which were not an excluding factor. Any colour vision deficiencies would be accepted within the criteria to get a more diverse group of participants. As it appeared in background research, colour blindness and especially protanopia is common and is reasonable to include in testing which focuses on colours.

Initially the participation count was aimed at 20 participants total, but due to pragmatic access constraints it was lowered to 10. The 10 participants would be

further divided into two groups consisting of five participants, one group for the marine industry and one for the energy industry. The split was for comparison between the groups and in addition it allowed to provide participants with prototypes of mimics that would fit better in the industry they have more experience with.

The participants were found via internal contacts at Wärtsilä as purposive sampling. It should be taken into consideration that all participants are Wärtsilä employees and how this may have affected the results of the testing. Potential participants were approached by introducing the topic, which included briefly explaining that it was about SCADA mimic colours and improving usability, after which they were provided with the participant background questionnaire if they expressed interest in participating.

7.2 User testing methodology

Each testing session followed a protocol consisting of an introduction, task-based evaluation, post-session interview, quantitative survey, and questions from participants. The testing was conducted remotely, and it was moderated. The sessions were done via Microsoft Teams, and the prototypes were presented via Figma. From Microsoft Teams a function called ‘Give control’ was used to give the participant a possibility to move their mouse cursor on the screen and point at relevant areas that the facilitator could see them.

The whole session was planned to last 45-60 minutes which was informed to the participants. In addition, it was mentioned, that if they wished to, it was acceptable to take longer. The average time of the executed testing sessions was 70 minutes, and the shortest session lasted approximately 26 minutes whilst the longest took approximately 105 minutes.

Each participant was shown three prototypes from their respective industries. Between each participant the order in which the prototypes were presented was

randomised. Randomisation of the ordering was used to minimise the effect of the presentation order on the test results.

The introduction that was provided to the users was shown as presentation slides to keep it as similar as possible between sessions. The introduction slides are presented in Appendix B. First, the participants were introduced to the concept of usability testing. Then, their role as the participant was explained to them, including instructions as there being no right or wrong answers, thinking aloud, and confirming completion of tasks. It was also mentioned that the facilitator cannot answer task related questions or guide the participant. Then, they were presented the session timeline and asked for consent confirmation, which will be introduced in more detail in Section 7.3. Lastly, they were prompted to ask any questions they might have and then reminded of everything with a short checklist and told to make sure their screen was bright and did not have any tonal settings on that would alter the colours. After the session with second energy participant (second participant overall), a mention about the values being placeholders was added into the introduction section.

After the introduction, participants moved to the task section. As noted above, the order of the prototypes was randomised for each participant so the prototype they saw first might not have been P1. The tasks that the participants were asked to perform were the following:

- T1:** Take a moment to look at this screen. Tell me what you understand about how this system is organised.
- T2:** Based on what you can see, what does the overall state of this system tell you and why?
- T3:** What does the visual difference between the elements mean to you?
- T4:** Something in the system has changed. Looking at the display as a whole, what do you think is different and why?

T1 was used to assess whether the colour changes affected system readability and the initial understanding of it. To guide their attention gently towards the alarms, states, and pipe media and their understanding of those, T2 was introduced. In T3, the goal was to try and prompt the users to identify different pipe media or further explain the differences in alarms and states if not already discussed. Before reading T4 to the participants, the current prototype was changed to an empty one (only mimic was removed, other UI elements stayed the same) for just a few seconds, after which an altered version was shown. The only changes in each time that the T4 was performed were the states of the P&ID symbols, like valves. After the testing session with the 1st marine participant, the altered versions of the marine prototypes were changed, since they were unintentionally left without changes for that session.

Facilitator approach was to be involved as little as possible but still stay active and encourage the participant to keep going. This included saying small remarks without indicating in tone or meaning if their observations were correct or not. One thing to note is that on T3, some participants struggled to understand the meaning of ‘visual difference’ therefore to any participant asking about it, the alternative explanation ‘differences in how they look’ was provided. Additionally, if the participants remained quiet for a significantly long period of time, they were asked if they were confident to move on or if there was something else they would like to add to the current task. After completing the tasks for all three prototypes, the participants were interviewed about their experience. During the interview all three prototypes were visible to them side by side on the same screen. The questions were designed to have the participant reflect on the colour differences and how those affected them. Some participants were prompted with ad hoc questions during the interview and those were mainly related to the alarm colours. The interview questions were the following:

1. Across the three interfaces, which felt most natural to read and why?

2. Was there anything that confused you consistently across all three prototypes, or was there something that was confusing only in a certain prototype?
3. How do you feel that the use of colour affected your ability to understand the system?
4. Were there any colours or colour combinations that felt intuitive or familiar from your professional experience?
5. Were there any colours or colour combinations that felt unexpected or conflicting with what you would normally associate them with?
6. In a real operational situation, which of the three prototypes would you feel most confident using and why?
7. Is there anything you felt that was missing visually that would help you understand the system state faster?
8. Is there anything else about any of the prototypes that you would like to comment on that we have not covered?

After completing the interview, participants were given the survey, which was separate for both groups. The main reason for that was for the data to be easier to differentiate yet moreover that the survey could include images of the prototypes which were different for both groups. The survey consisted of Likert-scale questions, ranking questions, questions with yes/no/partially answers, and an open word question where participants could suggest improvements. The full survey is presented in Appendix C, which is for the energy group, but the one for marine participants only differed for their respective prototype images.

After finishing with the survey, the participants were given the option to ask questions from the facilitator. If they wished to ask nothing, the participants were thanked and the session was concluded.

7.3 Data collection methods

Both qualitative and quantitative data were collected during each testing session for a more systematic analysis of participant responses. Data from each session was collected through facilitator notes, audio and screen recordings, and post-session survey responses.

During the introduction the participants were asked to confirm consent, and the basics were covered verbally with each participant. They were then given a form to sign where information was more detailed, and they had to fill in their names and check a box confirming their consent to continue. Mainly it contained the actions that were expected from them, like completing tasks, an interview, and a post-session survey, and data confidentiality. It was mentioned that the session will be recorded, that the recording and notes will be visible only to the researcher, and all data will be anonymised, and participants will not be identifiable in any published findings. The full form is presented in Appendix D. The recording of the session was only started after the form was submitted as confirmed. The sessions were done via Microsoft Teams which allows recording meetings directly in the platform.

Notes of each participant consisted mostly of verbal points that they made during completing the tasks. However, it was also written down if the participant did not notice some changes, like the alarm colours changing between prototypes or if changes in T4 were noticed correctly or not. Behaviour was not easy to monitor since the testing was done remotely, but some mouse cursor movements were noted, like following the piping lines.

The qualitative data was later organised in a table where observations were labelled into different categories covering alarm perception, colour associations, system comprehension, and interface preferences. Each participant's notes were gone through, anonymised, and filled into the table accordingly to see which categories had the most weight by being mentioned more often. The quantitative data was also

filled into a table and anonymised in the same manner. The similar format allows for easy comparisons between participants but also between qualitative and quantitative data from the same participant.

7.4 Testing hypotheses

The following hypotheses were formulated before testing based on standards analysis and prototype design rationale introduced in the previous chapters. They reflect both predicted user behaviour based on research done on colour conventions and anticipated differences between the energy and marine participant groups.

- H1:** Participants will express a need for a legend or labelling support for piping colours.
- H2:** Marine participants will rate P2 more favourably than energy participants.
- H3:** Participants will associate red with highest and yellow with second-highest alarm severity, regardless of prototype.
- H4:** P2 will be rated most favourably by possible participant(s) with colour vision deficiencies.
- H5:** P3 will offer the best usability balance across both industries.

8 User testing results

This chapter presents the findings from the user testing sessions described in Chapter 7. The results are organised thematically and the chapter concludes with a summary of the main findings and the evaluation of the five hypotheses.

8.1 Participant demographics and overview

Ten participants took part in user testing and five were from the energy industry and five from the marine industry. The participants' demographics are presented in Table 8.1. Participants have been anonymised and given IDs, where MP stands for marine participant and EP for energy participant. The numerical value in the ID indicates the order that participants were tested in inside the group. The overall sequence of testing was mixed between groups due to participant availability rather than completing one group before the other.

A notable point about the participant demographic is how only one in ten participants was female. This seems to indicate a heavy male dominance in these industries; however, it cannot be stated with certainty given this sample size. In addition, there was one participant who has slight protanopia and this was not surprising given the male-dominance of the field and how common colour vision deficiencies are in men, as stated in Section 3.1.

The participants were mainly designers, with a few HMI developers and other unspecified roles. Years of experience in SCADA systems ranged from 0 to 19,

Table 8.1: Participant demographics collected through pre-survey

| ID | Group | Age group | Sex | Role categ. | SCADA* exp. (y.) | SCADA familiarity | CVD |
|-----|--------|-----------|-----|-------------|------------------|-------------------|--------------|
| MP1 | Marine | 35-44 | M | HMI dev. | 19 | Excellent | No |
| MP2 | Marine | 35-44 | M | Designer | 8 | Good | No |
| MP3 | Marine | 18-24 | M | Other | 0.5 | Fair | No |
| MP4 | Marine | 35-44 | M | Other | 0 | Fair | No |
| MP5 | Marine | 25-34 | M | Designer | 1 | Fair | No |
| EP1 | Energy | 25-34 | F | Designer | 5 | Excellent | No |
| EP2 | Energy | 45-54 | M | Designer | 15 | Excellent | No |
| EP3 | Energy | 25-34 | M | Designer | 0 | Fair | No |
| EP4 | Energy | 25-34 | M | HMI dev. | 6 | Excellent | Protanopia** |
| EP5 | Energy | 35-44 | M | Other | 14 | Good | No |

Note. CVD = Colour vision deficiency. *Years of experience with SCADA/process control. **Slight protanopia.

therefore different levels of experience were well covered. However, there is a noticeable difference in the experience levels between the groups. In the marine group, there were more participants with less experience, but one with 19 years which skews the average heavily. The energy group's experience distribution was more balanced. Without MP1, the marine group has a significantly lower average of experience than the energy group does, and the marine group's lower experience may have influenced prototype preference ratings.

The self-evaluated familiarity with SCADA systems ranged from fair to excellent, although the option of poor was also provided. Even amongst participants who registered 0 years of experience, they rated their familiarity as fair rather than poor. This could reflect adjacent experience with SLDs and P&IDs rather than direct SCADA exposure.

Before initiating testing, the order in which the prototypes were presented to each participant was randomised. The ordering for each participant is presented in Table 8.2.

Table 8.2: Prototype presentation order per participant

| ID | Prototype order |
|-----|-----------------|
| EP1 | P2 → P3 → P1 |
| EP2 | P2 → P1 → P3 |
| EP3 | P1 → P2 → P3 |
| EP4 | P3 → P2 → P1 |
| EP5 | P3 → P1 → P2 |
| MP1 | P3 → P1 → P2 |
| MP2 | P2 → P1 → P3 |
| MP3 | P3 → P2 → P1 |
| MP4 | P1 → P2 → P3 |
| MP5 | P2 → P3 → P1 |

8.2 Colour-specific results

This section presents the colour-specific findings from the user testing sessions organised thematically based on the categories that emerged during qualitative analysis.

8.2.1 Alarm colour perception and severity ordering

The most consistently emerging topic across all 10 participants was alarm colour perception and severity ordering. Regardless of which prototype the participant was shown, the alarm severity ordering was not intuitive for most. Several participants (EP2, EP3, MP1, MP2) referenced traffic light associations: red meaning danger, yellow caution, and green safe. However, the actual alarm hierarchy in the prototypes caused confusion. EP3 intuitively placed magenta as the most critical alarm yet felt surprised when presented with the current alarm ordering in P1 (red, yellow, cyan, magenta) which the participant was already familiar with from experience. This is also interesting because that instinct aligns with some industry practices which were mentioned by MP1 and MP2. Industry conventions are split and there is no clear one way to organise alarm severities.

The alarm ordering from P2/P3 (red, magenta, yellow, blue) was found confusing by EP2, yet EP3 found the same ordering to be perhaps more logical than the current ordering. Therefore, even amongst experienced energy professionals there is no shared mental model for where magenta should be placed in the alarm hierarchy. The cyan in P1 and light blue in P2/P3 was problematic across many participants from both groups. EP4 found both the cyan and light blue to be difficult to differentiate from the background. MP4 also noted independently that the light blue alarm colour was easy to miss, and MP3 characterised it as a ‘weird alarm colour choice’. This connection across different experience levels and colour vision deficiency suggests that the low-severity alarm colour is a genuine usability concern.

8.2.2 Yellow colliding in pipes and alarms

Yellow emerged as the single most problematic colour in P1, in both energy and marine groups, and 8 participants mentioned it explicitly. Besides visibility, there was a semantic collision between the yellow pipes and the yellow alarms. MP2 felt that the yellow pipe made the yellow alarm disappear and MP3 noted that they were so similar in colour that they could be associated with each other. EP1 reported the yellow colour, especially in the pipes, being not clearly visible from far away, while EP2 observed that the yellow pipe and yellow alarm might be confusing for some people.

The yellow problem is not just about the pipes as it also intersects with the background brightness difference between the prototypes. In P2/P3 where pipes had colours based on the ISO standard or inspired by it, the yellow alarm was more visible. However, in the interview EP1 specifically noted that the yellow alarm was more visible in P1 from far away when compared to P2/P3. The yellow alarm colour does sit against a brighter background colour in P1, where in P2/P3 the colour has

been changed to a more lime shade and it looks slightly more toned down, making it slightly less conspicuous, even when the luminance contrast is higher.

8.2.3 Pipe colour differentiation

The alarms and the yellow colour were more unified between both groups, whereas for the pipe colour differentiation the split between the groups was more pronounced. In the Figure 8.1 the Likert averages are presented from the post-session survey with also the highest and lowest responses. The averages show that the energy group rated P2 readability at 2.8, while the marine group rated it at 4.2 which is 1.4 points higher than the energy group. The overall average for P2 in the energy group is 2.5 and in the marine group it is 4.0, making the divergence between the groups at 1.5 the highest with P2. The average for P1 in the energy group is 4.6 whilst for the marine group it is 3.4. Lastly, the averages for P3 are 3.4 for the energy group and 3.9 for the marine group. These averages can be heavily influenced by single outliers, since $n=5$ per group. For example, EP2 gave P2 and P3 mostly ‘strongly disagree’ answers, which lowers the energy averages down significantly. Opposing responses came from EP4, who mostly rated P3 with ‘strongly agree’.

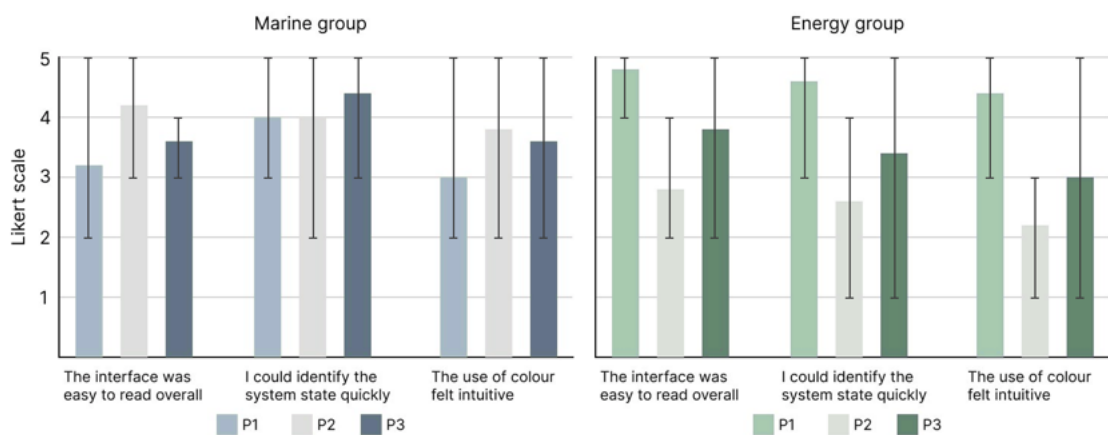


Figure 8.1: Likert averages with highest and lowest responses

In Figure the distribution of answers to whether colour distinctions between system components were clear are presented. These answers clearly show the pattern of the energy group rating P2 the lowest, whereas the marine group rated P1 the lowest. The positive response for P1 from the energy group was equal to the positive response for P2 from the marine group.

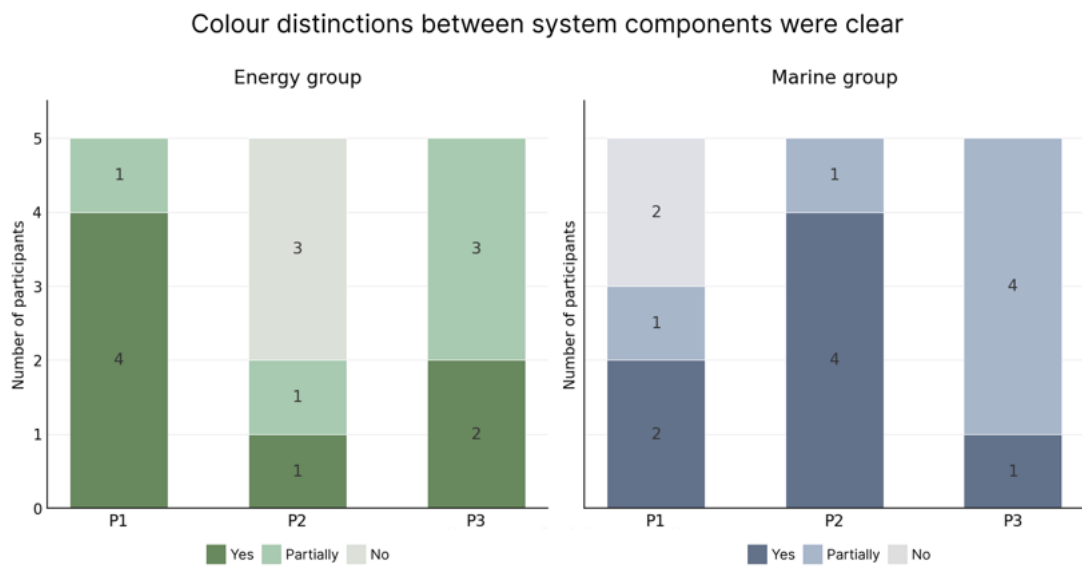


Figure 8.2: How users rated the clarity of colour distinctions between systems

The qualitative results offer some insight to these quantitative results. As marine participants deal with P&ID mimics of the full system with many different pipe media running simultaneously, the pipe colouring based on the ISO 14726 standard in P2 gave them useful information. MP4 even described the difference between P1 and P2 as huge and MP3 felt that the pipe differentiation between the pipes in P2 was clear and a huge improvement. Energy participants, however, found the extra colour information overwhelming and unnecessary, since they are used to seeing less pipe media per mimic, causing P2 to be rated as even extremely confusing by EP2.

There were also some colour associations about the pipe colours that appeared between groups and participants. Brown was associated as a fuel colour by EP1,

EP2, and MP1, and blue was instinctually associated with water by EP1 and EP4. MP3 alternatively associated blue with electrical rather than water which shows that the associations are not universal.

8.2.4 Background colour and visual comfort

Multiple participants across both groups commented on the background brightness of P1 when compared to P2/P3. EP1 specifically preferred the darker background in P2/P3 for operational use, since ‘operators are looking at a screen a lot’. MP1 made a distinction between P1 being ‘easier on the mind’ and P2/P3 being ‘easier on the eyes’. MP2 found the background colour in P2/P3 to be more comfortable and P3 ‘immediately more calming’.

This has a direct link to the ISA standard which recommends using neutral, desaturated backgrounds. This appears validated here and the participants that preferred the P1 background still acknowledged that the brighter colour was harder for sustained use.

8.2.5 Component state colours

All participants who commented on the component states correctly identified green indicating active or running and this was consistent across both groups and all prototypes. However, between the inactive and disabled states there was some confusion. EP4 specifically noted that the difference between disabled and inactive was unclear. In addition, MP4 was not sure what the grey (disabled) and white (inactive) meant. In the energy prototypes there were two components that were more separate from the mimic and EP3 mistook them as a ‘legend’ showing the meaning of green versus white.

MP3 raised a point about green valve states and the green nitrogen pipe in P2/P3 colliding. This highlights the point made in the ISA standard which warns against

semantic overlaps, especially for alarms. The same principle can furthermore be argued for state colours when pipes share the same hue.

8.2.6 The familiarity bias versus observed behaviour

There is a disconnection between stated preference and observed preference, especially within the energy group. 4/5 energy participants chose P1 for operational use based on familiarity, yet qualitatively several of them acknowledged strengths in P3. EP1 ranked P3 first in the post-session survey while other participants favoured P1 and EP4 mentioned that P3 could be considered if the greys were ‘toned down’. EP2, the strongest advocate for P1, actually failed to notice that the familiar alarm colours had been changed when first presented with P2 and only realised they were ‘wrong’ after seeing P1. This does indicate towards learned expectation rather than truly being about usability, suggesting that familiarity bias was a significant factor.

In the marine group, MP5 was an outlier for ranking P1 first despite others in the group favouring P2 and P3. Also, the fact that in the marine group the overall level of experience was lower on most participants could mean they did not have such strong mental models about the current system.

8.2.7 Legend and P3 compromise pattern

7/10 participants expressed a desire for some form of legend, labelling, or additional indicator for pipe colours. This appeared in both groups and all experience levels which directly supports H1. The participants’ wishes ranged for legends explicitly to indirect wishes for flow indicators and the possibility to see if the pipes were active. MP3 actively looked for a legend on the screen while EP4 wanted more information about alarm states in the top bar of the UI.

In the post-survey rankings P3 rarely came first but was almost never last either. It appeared as first or second choice for 8/10 participants when evaluating operability.

The Likert data supports this pattern that emerged from rankings as P3's averages sit consistently between P1 and P2. Since P3 was acceptable for all and even optimal for a few, the cross-industry implications of the design are worth noting.

8.3 Other results

Besides colour-specific observations, the testing sessions also surfaced several findings related to the broader usability of the prototype mimics. The fourth task asked participants to identify what had changed after a brief interruption (the mimic content was briefly hidden). Detection rates differed notably between groups, and the effect of the different prototypes was not noticeable. Excluding MP1 (who had no visual changes due to a testing error), energy participants detected changes in 12/15 prototypes and marine participants in 5/12. Inside the energy group the third shown prototype had a detection rate of 7/9, while the first one had 5/9. Since the prototypes were shown to the participants in a randomised order, this cannot be understood as learning bias, but it also does not support that any prototype specifically would support state change detection. In the marine group MP2 and MP4 detected nothing across all three prototypes and MP4 explicitly attributed it to there being too many elements on the screen to notice any changes. Meaning that regardless of colour scheme, more elements means that state changes are harder to notice.

Many participants also reported changes that did not happen when prompted to look for changes in T4. EP1 thought that placements had changed, EP3 hallucinated a blue line changing and mentioned noticing alarm indicator forms that they believed were not there before, although they were present in each prototype. MP1 thought that a blue alarm had changed, and MP3 thought that the top-left layout had changed. This pattern suggests that when participants were asked to look for changes it primed them to perceive differences where none existed.

Although flow arrows were present in all prototypes, their recognition was inconsistent across participants. Most energy participants, MP1, and MP2 noticed and understood their meaning. Contrastingly MP3 stated twice that they perceived there to be no flow and they did not connect the arrows to flow direction, explaining that they are more familiar with animation used as indication for flow. EP1 found the arrows clearly visible in P1 and P3 yet not in P2 due to the pipes being represented in the bicolour style.

Several participants noticed varying pipe thicknesses but interpreted them differently. Six participants (EP1, EP2, EP3, EP4, MP2, MP3) noticed the varying thicknesses and EP1 and MP2 correctly guessed or knew that it related to physical pipe size but did not specify that thicker pipes represent double walled. MP3 thought that thicker pipes meant that they were crossing area boundaries and initially thought they appeared only in a specific prototype yet later realised they had been present in all of them. There was a singular comment from MP1 about P2 whether it was ‘just the colour coding’ or if all the piping was double walled. In addition, later in the interview portion he brought up the fact that if the colours do not indicate double-walled piping, then indicating those might become a challenge.

In the marine prototypes solid grey lines were used to represent electrical lines. However, these were not recognised correctly by multiple participants. MP2 interpreted the solid grey line as a separator for different areas and MP3 assumed it indicated the main area. In marine group’s P2 and P3 where pipe media colours were based on ISO 14726, the electrical lines remained grey as in P1. Mainly because the standard does not define a colour for electrical systems and grey is the current colour used in Wärtsilä systems. As this left them visually unchanged, the participants attempted to decode their meaning incorrectly.

8.4 Summary of main findings

The user testing indicated several patterns about how colour affects the usability of SCADA mimics. The most prominent pattern was an apparent division between energy and marine participants in their response to increased pipe colour differentiation. P2 was more favoured in the marine group whilst the energy group felt that extra colour information was overwhelming and unnecessary. In the energy group's survey P1 was always the highest ranking in Likert answer averages, whereas in the marine group P1 was never highest.

P3 appeared as a compromise between both groups, as it was rarely rated last in the post-survey rankings by either group. P3 appeared to be acceptable for most and even optimal for a few.

For both groups yellow appeared as the most problematic colour in P1. The issue was with the visibility of the colour, but also with the semantic collision between the yellow pipe and alarm. The problem with the yellow alarm visibility was also slightly related to the background colours, since some felt that P1's brighter background made it more visible.

Alarm colour perception was something that emerged as a topic with all participants. Regardless of prototype, most participants did not find the alarm colour ordering intuitive. Several participants referenced associating red with danger and yellow with caution through traffic lights, meaning that they thought red should be the most severe alarm. However, some participants had an existing mental model where the most severe alarm should be magenta, followed by red and yellow.

Detecting changes in the mimic component states could not be found related to the colour scheme. Although there was a divide between the groups, since the energy group performed consistently better with the task. Noticing and understanding the arrows in the pipes representing the flow was inconsistent, and especially P2 seemed to interfere with the arrow visibility. No participant correctly stated that thicker

pipes specifically meant that they were double walled, and the difference in thickness was noticed by 6/10 participants. Grey dashed lines used for grouping also caused some confusion especially in the marine group, since there was also a solid grey line that represented electrical lines.

Of the five hypotheses formulated prior to testing, they were supported in varying degrees by the collected data. H1 appears to be supported, as 7/10 participants across both groups, and all experience levels expressed a need for a legend or colour labelling. H2 also appears supported, as marine group's average Likert score for P2 is 4.0 versus energy group's 2.5 average for P2. H3 appears partially supported, since several participants did rank red as highest and yellow as second, but some mentioned magenta as highest being more intuitive. H4 does not appear supported, as EP4 ranked P2 last for both intuitiveness and operational use. H5 also appears partially supported as 8/10 participants placed it in the top two for operational use and it never received the lowest group average on any Likert question, and it could be more accurately described as the least divisive rather than the most effective.

These findings raise several points for discussion, particularly regarding the role of familiarity in shaping colour preferences and whether a single colour scheme could be feasible for both industries. The following chapter covers the emerging discussion.

9 Discussion

One of the most significant findings of this study is that the same degree of colour alignment with standards produced opposite effects depending on the industry context. Given the definition of usability that was introduced in Section 4.1, which mentions it being dependent on specified users in a specified context of use, the findings from the testing demonstrate that even within the same company colour usability is not transferable. As energy mimics show less complicated subsystems when compared to increasingly complex marine mimics, adding more colour created unnecessary visual noise which is one underlying explanation for the split between the industries. These findings suggest that an approach that would be suitable for more than one industry may be counterproductive.

The strong preference for P1 amongst energy participants appears to reflect learned expectations rather than measured usability. This is an important distinction, as it implies that an attempt to introduce improved colour practices will face resistance from experienced users regardless of the objective quality of the change. This is consistent with the findings from Boulhic et al., where experienced participants avoided unfamiliar colour coding and favoured strategies built through professional practice instead. The clearest evidence lies in EP2 failing to notice that alarm colours had been changed until seeing P1, therefore what was framed as ‘correct’ was actually ‘familiar’. On average, the marine group’s level of experience was lower, therefore they had weaker mental models and were more open to alternative colour schemes.

This connects back to the ISA-101.01 standard, where consistency is emphasised to benefit experienced users but sometimes it could lead to suboptimal practices.

The main focus when testing started was in the piping colours, but the alarm hierarchy problem highlighted itself in the results. Even within the same company professionals had different expectations about the hierarchy, especially where magenta should be placed. This suggests that alarm colour conventions are learned through exposure to specific systems rather than being directly derived from universal cognitive associations. While some participants mentioned the association with traffic lights (red for danger, yellow for caution), there was no common consensus for colours outside of this familiar set. This demonstrates that the absence of a direct standard for alarm colours has resulted in fragmented practices. Alarm colours are essentially based on mental models of individual professionals rather than shared industry conventions.

The collision between the yellow pipe colour, which is derived from ISO 14726's main colour for flammable gases, and the yellow alarm colour poses a broader challenge for interface design. Individual standards are internally consistent but can conflict when applied simultaneously to the same UI. This suggests that colour guidelines for mimics cannot treat pipe colours and alarm colours independently as they occupy the same visual space.

In UCD the goal is often not to find the optimal solution for one user group, but to find the solution that is acceptable across the widest range of users. P3 achieved this as it was the least divisive prototype, since it consistently appeared in the top two rankings across both industries while rarely being the first choice. It was optimal for few, but it was acceptable to most and it was not rated strongly negatively by either group. This pattern of broad acceptability aligns with the inclusive design principle of accommodating the widest range of users. These findings suggest that a compromise approach of using single muted colours inspired by ISO 14726 instead of the bicolour system could offer a viable path for cross-industry colour practices.

The contrasting responses from the two different background colours supported ISA-101.01's recommendation for neutral, desaturated backgrounds. Participants across both groups preferred the darker, more neutral background in P2/P3 for sustained use, even when some participants felt that the brighter background in P1 was initially easier to read. This also directly connects to Wang et al.'s finding about saturation impairing cognitive performance introduced in Chapter 3.

The experience from EP4 provides a limited yet relevant data point for accessibility. The cyan alarm was problematic not only for EP4 with slight protanopia, but for multiple participants without any CVD. This suggests that some accessibility concerns overlap with general usability which connects to Sauer et al.'s point that implementing accessibility guidelines can benefit non-disabled users as well. However, with just $n=1$ for CVD, no strong accessibility conclusions can be drawn, yet it can offer direction for future research.

Connecting these findings back to the previous research done by Boulhic et al., there are similarities yet some differences. While Boulhic et al. recommended against implementing ISO 14726 on marine supervision HMIs, the findings from this study suggest that a simplified application of colours inspired by the ISO standard (P3) can be positively received by marine professionals. This distinction is significant, as it implies that the issue is not with the ISO-based colours, but more with the cognitive demands of the bicolour coding mechanism. However, participants in this study also preferred the ISO bicolour system in P2, which contradicts with the findings from Boulhic et al.

9.1 Addressing the research questions

The research questions that these findings ought to answer are repeated here:

RQ1: What standards and best practices are there for colour use in SCADA mimics for energy and marine industries, and to what extent do they align with each other and with observed industry practices?

RQ2: How does the degree of alignment with identified colour standards and best practices affect the perceived usability of SCADA mimics across energy and marine industries?

Regarding RQ1, the analysis of existing standards revealed that while there are standards that address aspects of colour use in SCADA systems, they have different levels of specificity and leave significant gaps that are filled inconsistently by vendors and individual designers. Industry practices are closely related to the standards, like favouring greyscale for most of the mimic and ISA-101.01's recommendation for neutral backgrounds and intentional use for colour. ISO 14726 covers pipes but only for marine, ISA-101.01 gives general principles but no specific colour recommendations, and WCAG2ICT provides contrast thresholds. The alignment between these standards is partial and they leave large gaps for vendors to fill, like the alarm colours. This study has shown that even professionals within the same company disagree on alarm hierarchy which is evidence of an existing consensus gap.

The answer to RQ2 is not straightforward, as the degree of alignment with the identified standards affected perceived usability differently depending on the industry. Full ISO alignment (P2) was beneficial for marine but detrimental for energy and partial alignment (P3) was the least divisive. Energy participants were on average more experienced and they showed resistance driven by familiarity, whilst the less experienced marine participants were more receptive to new colours.

9.2 Limitations

With the sample size being only $n=10$ and $n=5$ per industry group, the study cannot support statistical generalisation. Due to the size, individual responses heavily influence group Likert averages and rankings, even if they appear to indicate trends. This is demonstrated most clearly by EP2 being responsible for the lower P2/P3 scores for energy. Any quantitative patterns should be interpreted as indicative rather than conclusive.

The study focused on first impressions, since each participant saw the prototypes only during the singular testing session, which lasted approximately 70 minutes on average. This is insufficient for assessing how preferences might shift after days or weeks of operational use. Especially in the energy group where familiarity was an issue, participants might adapt over time, and the initial resistance might change.

All the participants were Wärtsilä employees which introduces organisational bias and their familiarity with the specific colour conventions in Wärtsilä shaped their mental models. In addition, the participant roles were skewed heavily towards designers and HMI developers rather than actual operators. Designers most likely respond to design decisions more consciously whilst operators probably respond more instinctively based on daily use patterns.

Only one participant was female and based on the research on sex-based differences in Chapter 3, the findings cannot account for potential differences in colour perception being related to sex. EP1 was furthermore the only participant in the energy group who ranked P3 first, yet it cannot be stated whether that relates to sex differences in colour perception or simply individual preference based on this data. In addition, there was only one participant with CVD, which means that accessibility conclusions cannot be drawn from a single case.

Remote testing posed some limitations as the facilitator could not observe body language and could only track mouse cursor movements. Screen brightness, colour

calibration, and ambient lighting all varied between participants and those could only be controlled to the extent that participants were instructed to turn screen brightness high and verify they did not have tonal settings on that would alter colours.

The prototypes themselves were presented as static images, meaning that visual dynamics like blinking for alarms could not be tested. The absence of animation could have influenced how the participants perceived alarm visibility, especially for the alarms with lower severity. The first marine participant (MP1) was shown prototypes without any state changes for T4 due to an oversight which was afterwards corrected. That limits marine group's T4 answers from 5 to 4. The mimics between industries were also different, and while it was necessary to provide content that is appropriate for each industry, it means that some observations are influenced by the different complexity and layout of the two mimics.

The post-session survey was completed after the interview, which means that responses might have been influenced by the reflection and articulation from the interview. If something had not been noticed by a participant during the testing, they might have formed an opinion about it during the interview, directly affecting the survey results. This could make the agreement between qualitative and quantitative data seem higher.

Several of the prototypes deliberately did not follow the created guidelines as they were written. For instance, the 3:1 contrast ratio was not achievable for all pipe colours or all alarm levels. If guideline compliance had been achieved the results might differ, yet full compliance was sometimes incompatible with other design goals like visual comfort.

9.3 Future research

To evaluate whether the learning curve for new colour conventions is acceptable in practice, a study with operators using the different colour schemes over an extended

period in a simulated operational environment would be needed. This would address the familiarity issue by seeing if operators could adapt to new colours.

A study with a larger sample size would enable statistical testing and allow for more confident generalisations. Participants should be recruited from multiple different SCADA vendors and include operational staff in addition to designers and developers. The divide should be more skewed towards operators, since they might have different priorities. A dedicated study with a meaningful sample of participants with CVD would properly address the possible accessibility issues.

Marine SCADA systems typically require both light and dark themes because lighting can change throughout the day. It would be important to test colour schemes dedicated for dark surroundings as well. Testing with animated prototypes could evaluate whether the animation behaviour adequately compensates for lower contrast in the less severe alarm colours.

The P3 compromise showed promise as a cross-industry solution, yet it was only tested once in initial form. An iterative design process refining P3 and retesting would follow the UCD principles outlined in Chapter 4 more closely and could produce a more polished design.

The finding that there is no shared mental model for alarm severity ordering, even within one organisation, suggests a need for focused research on alarm colour conventions. A survey across multiple vendors and regions could map the existing alarm colour practices and identify whether any hierarchy has sufficient consensus to serve as a basis for standardisation.

10 Conclusion

This thesis examined colour practices in SCADA mimics across the energy and marine industries. It identified relevant standards (ISO 14726, ISA-101.01, WCAG2ICT) and created a set of 12 guidelines based on them and research on colour perception. Three prototypes per industry were designed with different degrees of alignment with the guidelines and they were tested on 10 industry professionals through a combination of tasks, interviews, and surveys. The study was motivated by the observation that SCADA colour practices vary widely across vendors and industries despite the existence of standards that could inform them.

For RQ1, the standard landscape is partial and fragmented. The relevant standards address different aspects of colour use at different levels and they do not contradict each other but leave significant gaps. Those gaps are then filled by industry practices inconsistently through vendor guidelines and organisation conventions. For RQ2, the effect of standard alignment on perceived usability was dependent on the industry. Full alignment (P2) improved perceived usability for marine participants and impaired it for energy participants, while a compromise approach (P3) was broadly accepted across both groups.

The most significant finding is that identical colour changes produced opposite usability responses depending on the industry. Boulhic et al. tested ISO 14726 colours only in a marine context with students and maritime college participants and in contrast, this thesis extends that work by comparing energy and marine

professionals' responses to the same design changes. Operational context, especially the complexity of the mimic, determines whether additional colour information helps or hinders. A single standard applied uniformly across industries risks reducing usability in at least one context.

The P3 compromise approach of using single muted colours inspired by ISO 14726 and its broad acceptability suggests that a simplified application of ISO-aligned colours may be more practical than a full ISO implementation or entirely custom colour schemes. This had not been tested before, as Boulhic et al. only tested the full bicolour ISO implementation against a completely different colour scheme and did not test a middle ground.

This study also demonstrated how experienced professionals felt that familiar colours are equal to usable colours. This reinforces Boulhic et al.'s finding regarding expert resistance to unfamiliar colour coding. This resistance should be anticipated and managed through transition support rather than interpreted as evidence that the change is not beneficial.

The methodology presented, comprising the 12 guidelines and the three prototypes, could be used for other industries wanting to test colour scheme changes. It offers a starting framework and can be built upon for other industrial HMI contexts.

Based on these results, organisations maintaining SCADA systems across multiple industries should not attempt to implement a single universal colour scheme. Instead, the findings suggest establishing shared principles that can be adapted to the needs of different industries. The P3 compromise approach is a starting point for cross-industry compatibility, but industry-specific adjustments are still needed. The desire for legends or labelling being expressed by 7/10 participants suggests that adding text alongside colour coding could reduce reliance on colour alone, thus benefiting general usability and accessibility. The lack of consensus on severity ordering for

alarm colours points to a need for explicit documentation rather than assuming that colour associations will be intuitive.

The path forward is in establishing shared principles that acknowledge universal colour perception aspects and factors specific to the context over creating a single colour scheme. The guidelines, prototypes, and testing methodology in this thesis offer a starting point for that process.

References

- [1] A. Daneels and W. Salter, ‘What is SCADA?’, in *Proceedings of the International Conference on Accelerator and Large Experimental Physics Control Systems*, Trieste, Italy: CERN, 1999.
- [2] A. Ujvarosi, ‘Evolution of SCADA systems’, *Bulletin of the Transilvania University of Brasov. Engineering Sciences. Series I*, vol. 9, no. 1, 2016.
- [3] L. Kulpaviciute. ‘What is SCADA?’, Opoura, Accessed: Jan. 14, 2026. [Online]. Available: <https://opoura.com/educational-content/what-is-scada/>.
- [4] L. Nygren, ‘Adaptive mimic: Mimic to HMI panel’, Degree Programme in Electrical Engineering and Automation, Bachelor’s Thesis, Novia University of Applied Sciences, Vaasa, Finland, 2023.
- [5] L. Boulhic, A. Bignon, F. Silone, T. Morineau, J. Rechar, and J.-F. Bouillon, ‘Effects of color codes used on marine supervision HMI on mental workload and information retrieval: Experimentations with novices and experts’, *International Journal of Industrial Ergonomics*, vol. 67, pp. 180–191, Sep. 2018, ISSN: 01698141. DOI: 10.1016/j.ergon.2018.05.010. Accessed: Mar. 28, 2026.
- [6] J. Best, Ed., *Colour Design: Theories and Applications* (Woodhead Publishing Series in Textiles 128), 1st ed. Cambridge: Woodhead Publishing, 2012, ISBN: 978-0-85709-553-4. DOI: 10.1533/9780857095534.

-
- [7] A. Cooper, R. Reimann, D. Cronin, C. Noessel, J. Csizmadi, and D. LeMoine, *About face : the essentials of interaction design*, Fourth edition. Indianapolis: John Wiley & Sons, 2014, ISBN: 1-118-76657-1.
- [8] Colour Blind Awareness. ‘About colour blindness’, Accessed: Nov. 12, 2025. [Online]. Available: <https://www.colourblindawareness.org/colour-blindness/>.
- [9] D. L. Bimler, J. Kirkland, and K. A. Jameson, ‘Quantifying variations in personal color spaces: Are there sex differences in color vision?’, *Color Research & Application*, vol. 29, no. 2, pp. 128–134, Apr. 2004, ISSN: 0361-2317, 1520-6378. DOI: 10.1002/col.10232. Accessed: Dec. 7, 2025.
- [10] I. Abramov, J. Gordon, O. Feldman, and A. Chavarga, ‘Sex and vision II: Color appearance of monochromatic lights’, *Biology of Sex Differences*, vol. 3, no. 1, p. 21, 2012, ISSN: 2042-6410. DOI: 10.1186/2042-6410-3-21. Accessed: Dec. 7, 2025.
- [11] M. Iriguchi, H. Koda, and N. Masataka, ‘Colour perception characteristics of women in menopause’, in *Proceedings of the 2018 International Joint Workshop on Multimedia Artworks Analysis and Attractiveness Computing in Multimedia*, Yokohama Japan: ACM, Jun. 6, 2018, pp. 20–25, ISBN: 978-1-4503-5798-2. DOI: 10.1145/3209693.3209694. Accessed: Dec. 7, 2025.
- [12] M. A. Da Silva, T. E. D. A. Anfe, A. B. Matos, and G. F. Vieira, ‘Influence of gender, anxiety and depression symptoms, and use of oral contraceptive in color perception’, *Journal of Esthetic and Restorative Dentistry*, vol. 27, S1 Mar. 2015, ISSN: 1496-4155, 1708-8240. DOI: 10.1111/jerd.12127. Accessed: Dec. 7, 2025.
- [13] Z. O’Connor, ‘Colour, contrast and gestalt theories of perception: The impact in contemporary visual communications design’, *Color Research & Application*,

- vol. 40, no. 1, pp. 85–92, Feb. 2015, ISSN: 0361-2317, 1520-6378. DOI: 10.1002/co1.21858. Accessed: Nov. 4, 2025.
- [14] M. Chauhan and R. Agarwal, ‘Impact of screens on how users think’, in *2023 3rd International Conference on Intelligent Technologies (CONIT)*, Hubli, India: IEEE, Jun. 23, 2023, pp. 1–5, ISBN: 979-8-3503-3860-7. DOI: 10.1109/CONIT59222.2023.10205565. Accessed: Dec. 7, 2025.
- [15] X. Wang, L. Xu, J. Li, C. Cheng, and L. Yu, ‘How background colour shapes digital text-information processing: A fNIRS-eye tracking study’, *International Journal of Human–Computer Interaction*, pp. 1–15, Jul. 3, 2025, ISSN: 1044-7318, 1532-7590. DOI: 10.1080/10447318.2025.2524491. Accessed: Dec. 7, 2025.
- [16] Information is Beautiful. ‘Colours in cultures’, Accessed: Nov. 26, 2025. [Online]. Available: <https://informationisbeautiful.net/visualizations/colours-in-cultures/>.
- [17] J. D. Gould and C. Lewis, ‘Designing for usability: Key principles and what designers think’, *Commun. ACM*, vol. 28, no. 3, pp. 300–311, Mar. 1985, ISSN: 0001-0782. DOI: 10.1145/3166.3170.
- [18] T. Keinonen, ‘User-centered design and fundamental need’, in *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, Lund Sweden: ACM, Oct. 20, 2008, pp. 211–219, ISBN: 978-1-59593-704-9. DOI: 10.1145/1463160.1463183. Accessed: Dec. 10, 2025.
- [19] International Organization for Standardization. ‘ISO 9241-210:2019 - ergonomics of human-system interaction — part 210: Human-centred design for interactive systems’, Accessed: Feb. 2, 2026. [Online]. Available: <https://www.iso.org/standard/77520.html>.

-
- [20] J. Nielsen, *Usability engineering*. San Francisco (CA): Academic Press, 1993, ISBN: 0-12-518405-0.
- [21] J. Nielsen, ‘Enhancing the explanatory power of usability heuristics’, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Boston Massachusetts USA: ACM, Apr. 24, 1994, pp. 152–158, ISBN: 978-0-89791-650-9. DOI: 10.1145/191666.191729. Accessed: Feb. 2, 2026.
- [22] Nielsen Norman Group. ‘10 usability heuristics for user interface design’, Accessed: Feb. 2, 2026. [Online]. Available: <https://www.nngroup.com/articles/ten-usability-heuristics/>.
- [23] W3C Web Accessibility Initiative (WAI). ‘Accessibility, usability, and inclusion’, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.w3.org/WAI/fundamentals/accessibility-usability-inclusion/>.
- [24] J. Sauer, A. Sonderegger, and S. Schmutz, ‘Usability, user experience and accessibility: Towards an integrative model’, *Ergonomics*, vol. 63, no. 10, pp. 1207–1220, Oct. 2, 2020, ISSN: 0014-0139, 1366-5847. DOI: 10.1080/00140139.2020.1774080. Accessed: Mar. 28, 2026.
- [25] P. Reed et al., ‘User interface guidelines and standards: Progress, issues, and prospects’, *Interacting with Computers*, vol. 12, no. 2, pp. 119–142, Nov. 1999, ISSN: 09535438. DOI: 10.1016/S0953-5438(99)00008-9. Accessed: Mar. 28, 2026.
- [26] I. Nimmo, ‘Introduction, how HMI has been evolving’, in *Sensemaking in Safety Critical and Complex Situations*, 1st ed. Boca Raton: CRC Press, Jul. 13, 2021, pp. 1–12, ISBN: 978-1-003-00381-6. DOI: 10.1201/9781003003816-1. Accessed: Mar. 28, 2026.
- [27] International Organization for Standardization, ‘Ships and marine technology — identification colours for the content of piping systems’, International Orga-

- nization for Standardization, Geneva, Switzerland, Standard ISO 14726:2008, 2008.
- [28] International Organization for Standardization. ‘ISO 14726:2008’, ISO, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.iso.org/standard/44744.html>.
- [29] International Society of Automation. ‘About the international society of automation (ISA)’, isa.org, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.isa.org/about-isa>.
- [30] International Society of Automation. ‘ISA’s relationship to IEC standards’, isa.org, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.isa.org/intech-home/2021/june-2021/departments/isa-s-relationship-to-iec-standards>.
- [31] International Society of Automation. ‘ISA-101.01-2015, human machine interfaces for process automation systems’, isa.org, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.isa.org/products/ansi-isa-101-01-2015-human-machine-interfaces-for>.
- [32] International Society of Automation and American National Standards Institute, Eds., *Human machine interfaces for process automation systems: American national standard: ANSI/ISA-101.01-2015*, Research Triangle Park, North Carolina: ISA, 2015, 62 pp., ISBN: 978-1-941546-46-8.
- [33] W3C Web Accessibility Initiative (WAI). ‘About W3C WAI’, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.w3.org/WAI/about/>.
- [34] W3C Web Accessibility Initiative (WAI). ‘WCAG 2 overview’, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.w3.org/WAI/standards-guidelines/wcag/>.

-
- [35] W3C Web Accessibility Initiative (WAI). ‘WCAG2ICT overview’, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.w3.org/WAI/standards-guidelines/wcag/non-web-ict/>.
- [36] W3C. ‘Understanding conformance | understanding WCAG 2.0’, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.w3.org/TR/UNDERSTANDING-WCAG20/conformance.html>.
- [37] W3C. ‘Guidance on applying WCAG 2 to non-web information and communications technologies (WCAG2ict)’, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.w3.org/TR/wcag2ict-22/>.
- [38] Rockwell Automation, ‘Process HMI style guide’, Rockwell Automation, Milwaukee, WI, White Paper PROCES-WP023A-EN-P, May 2019.
- [39] AVEVA. ‘Alarm indicator symbol component’, Accessed: Mar. 28, 2026. [Online]. Available: <https://docs.aveva.com/bundle/sp-salibrary/page/62305.html>.
- [40] RealPars. ‘What is high-performance HMI?’, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.realpars.com/blog/high-performance-hmi>.
- [41] J. R. Koelsch. ‘The great graphics debate: Is going gray good for HMIs?’, Automation World, Accessed: Mar. 28, 2026. [Online]. Available: <https://www.automationworld.com/products/control/article/13316377/the-great-graphics-debate-is-going-gray-good-for-hmis>.

Appendix A Participant background questionnaire

SCADA Interface Design Study - Participant Background Questionnaire

This questionnaire helps me match participants' experience levels for my thesis research on SCADA mimic colour standardisation. It takes approximately 5 minutes. Your responses will help me ensure a representative sample of users with different familiarity levels with process control systems. All data will be anonymised in the final thesis.

* Required

* This form will record your name, please fill your name.

1. Age *

⋮

- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- 65+

2. Sex assigned at birth *

- Male
- Female

3. What is your familiarity with SCADA systems? *

- Excellent
- Good
- Average
- Fair
- Poor

4. How many years of experience with SCADA/process control do you have? *

The value must be a number

5. How frequently do you work with SCADA/process control work? *

- Daily
- Weekly
- Monthly
- Less frequently

6. If you have worked with SCADA systems, which vendor's systems have you worked with most (e.g., Wärtsilä, ABB, Siemens, Schneider)?

Enter your answer

7. What is your familiarity with Single line diagrams (SLDs) or Piping and instrumentation diagrams (P&IDs)? *

- Excellent
- Good
- Average
- Fair
- Poor

8. Do you have any colour vision deficiencies (e.g., colour blindness)? If yes, what?

Enter your answer

9. Are you familiar with work in the energy sector and/or marine sector? *

- Energy sector
- Marine sector
- Both
- Neither

10. Primary role type *

- Operator
- Control engineer
- Process engineer
- Designer
- HMI Developer
- Other

Appendix B Testing introduction

Introduction

SCADA mimic design usability testing

Thank you in advance!

I appreciate you taking the time to participate in this testing and supporting my thesis study.

What is usability testing?

In a usability testing (user testing) session a facilitator asks a participant to perform tasks, usually on one or more user interfaces. While the participant completes the tasks, the facilitator observes the participant's behaviour and listens for feedback.

Your role as participant

Testing the design, not you

There are no right or wrong answers

Think aloud

Narrate what you're looking at, what you expect, what confuses you etc.

Don't second guess

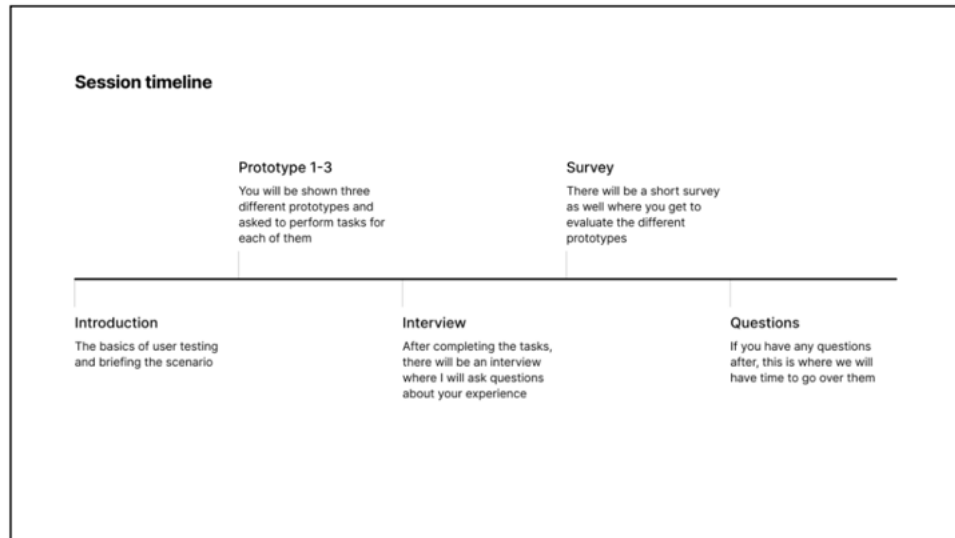
First reactions are valuable and it's okay to trust your gut

Confirm completion

After you feel you have completed a task or have given up on it, let me know

What if you don't know how to complete a task?

As the facilitator I cannot answer task specific questions or guide you. If you're unsure about your question, feel free to ask and I will judge if I'm allowed to answer it. If a task feels like you cannot complete it, that is completely okay and you can just say that.



Consent confirmation

Participation involves:

- Interacting with three SCADA interface prototypes and completing assigned tasks
- A post-task interview discussing your impressions and experience
- Completing a short usability survey

Data and confidentiality:

- The session will be recorded and notes will be taken for research purposes
- Recordings and notes will be stored securely and accessible only to the researcher
- All data collected will be anonymised and you will not be identifiable in any published findings

Any questions before we start?

Remember to think aloud as you go and tell me what you're looking at, what you expect to happen, and anything that catches your attention.

I won't be able to answer questions during the tasks, but you're welcome to ask anything afterwards. There are no right or wrong ways to approach this.

Appendix C Post-test survey (energy example)

Post-test survey - energy

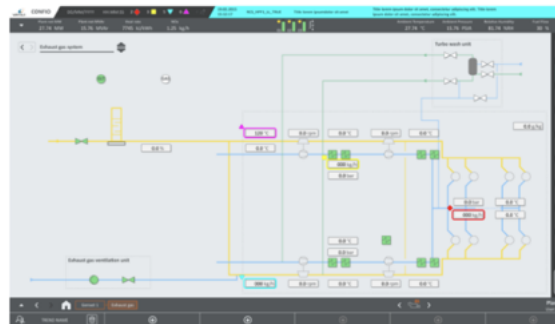
* Required

* This form will record your name, please fill your name.

Prototype ratings

1

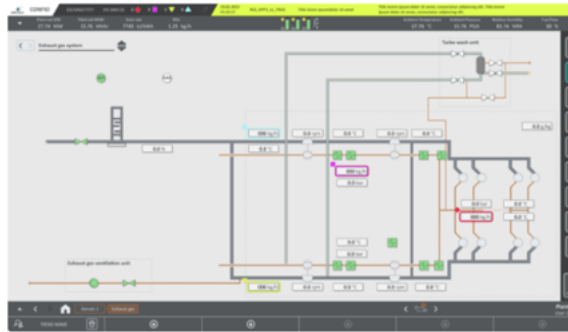
Prototype 1 *



| | Strongly disagree | Disagree | Neutral | Agree | Strongly agree |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| The interface was easy to read overall | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I could identify the system state quickly | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| The use of colour felt intuitive | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

2

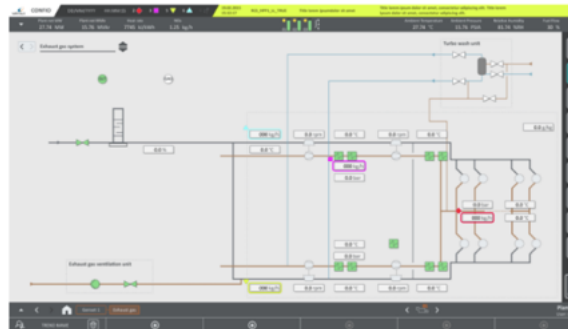
Prototype 2 *



| | Strongly disagree | Disagree | Neutral | Agree | Strongly agree |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| The interface was easy to read overall | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I could identify the system state quickly | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| The use of colour felt intuitive | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

3

Prototype 3 *



| | Strongly disagree | Disagree | Neutral | Agree | Strongly agree |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| The interface was easy to read overall | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I could identify the system state quickly | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| The use of colour felt intuitive | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Comparative ranking



4

Rank the three prototypes from most to least intuitive to read *

| |
|-------------|
| Prototype 1 |
| Prototype 2 |
| Prototype 3 |

5

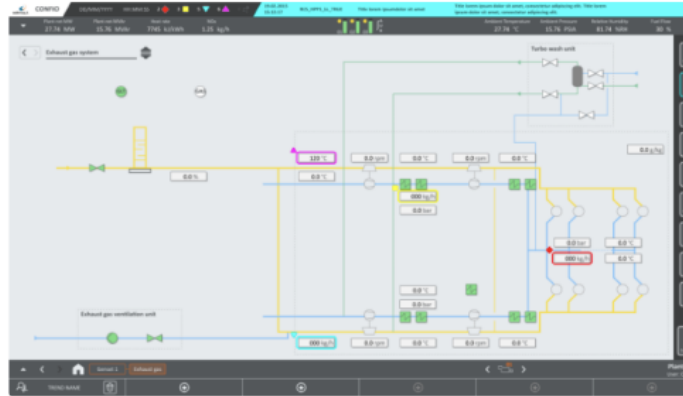
Rank the three prototypes from most to least suitable for real operational use *

| |
|-------------|
| Prototype 1 |
| Prototype 2 |
| Prototype 3 |

Colour clarity

6

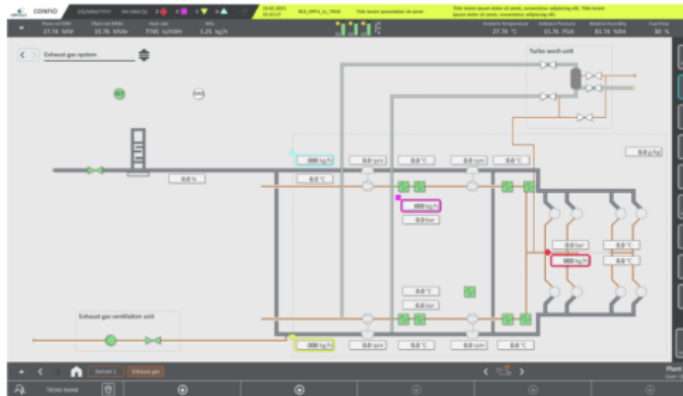
The colour distinctions between system components were clear in Prototype 1 *



- Yes
- Partially
- No

7

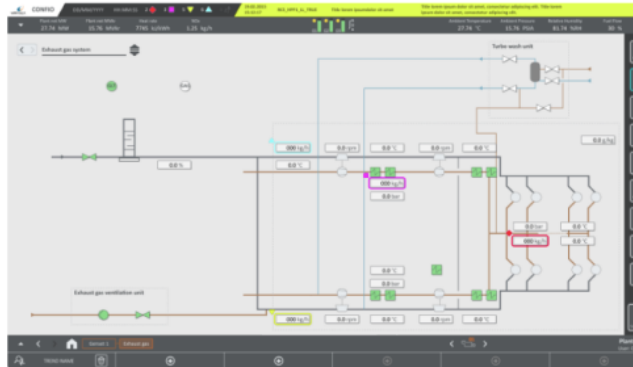
The colour distinctions between system components were clear in Prototype 2 *



- Yes
- Partially
- No

8

The colour distinctions between system components were clear in Prototype 3 *



- Yes
- Partially
- No

Open feedback



9

Is there any colour, element, or visual aspect you would change in any of the prototypes? If so, what and why?

Appendix D Consent form

Consent Form

Participation involves:

- Interacting with three SCADA interface prototypes and completing assigned tasks
- A post-task interview discussing your impressions and experience
- Completing a short usability survey

Data and confidentiality:

- The session will be recorded and notes will be taken for research purposes
- Recordings and notes will be stored securely and accessible only to the researcher for the duration of the thesis project and deleted thereafter
- All data collected will be anonymised and you will not be identifiable in any published findings

Your participation is entirely voluntary and you may withdraw at any time during the session. If you withdraw, any data collected up to that point will be discarded upon your request.

If you have any questions feel free to ask them during the session or later by contacting nelli.rossi@wartsila.com

* Required

* This form will record your name, please fill your name.

1. First name *

2. Last name *

3. By selecting below, I confirm that:

- I have read and understood the information above
- I understand that my participation is voluntary
- I consent to the session being recorded
- I agree to participate in this study

*

Confirm