CHALMERS

Simulation and evaluation of screen images for power plant status

Silvia Pfoser

Department of Product and Production Development,
Division Design & Human Factors
Chalmers University of Technology
Gothenburg, Sweden 2010

Simulation and evaluation of screen images for power plant status

Department of Product and Production Development
Division Design & Human Factors
Chalmers University of Technology
SE– 412 96 Gothenburg
Sweden
Telephone + 46 (0) 31 771 1000

Abstract

This master's thesis is part of the research project "Graphical visualization of process status for thermal power plants" performed by the Human-Machine System Research Group at Chalmers University of Technology. The aim of this project is to support proactive work in control rooms of thermal power plants by improving the visualization of power plant status. New screen images for monitoring displays were developed in order to facilitate rapid perception of deviations from a normal state and to support a more proactive behavior. The overall purpose is to help the plant operators and engineers to maintain optimal operational conditions in the power plant over a prolonged period of time.

The main subject of this master's thesis was to further evaluate the developed screen images in a follow-up empirical study. More specifically, it was investigated whether a more graphical representation of process values supports a proactive behavior, which implies that the user acts before alarms appear and thus maintains a prolonged, optimal operation status of the system.

The working procedure of this study is divided into three phases:

1. **Prototyping** – creating an interactive interface prototype for the usability evaluation
2. **Pilot study** – conducting a feasibility study to test the prototype and to refine the test procedure
3. **Usability evaluation** – conducting usability tests with the prototype to collect data about user behavior and experience

The usability evaluation was carried out on an interface prototype built on the basis of the screen images of a feed water system in a thermal power plant. The formulated research questions addressed the user's interaction experience and the cognitive demand to accomplish certain tasks. Three usability methods were applied in this study for answering the research questions: logging the task performance, post-test questionnaires and interviews. The combination of the different data collection methods was aimed at gathering subjective information of the user satisfaction and objective data about the task performance.

In general, it could be determined that the majority of the test participants used the graphical representation as basis for decision-making in the interaction with the prototype. According to the questionnaire responses and comments in the interviews, the process conditions presented graphically contributed to a better understanding of the system's status. This resulted in a better task performance because the participants were able to maintain optimal operating conditions for a longer period of time.

Keywords: human-computer interaction, control room, screen image

Table of contents

1 Introduction

The first chapter in this report introduces the reader to the subject of this master's thesis. A general background to human-machine interfaces in industrial control is presented, including a review of previous research conducted at Chalmers University of Technology. Furthermore, the purpose and the goal of the study are specified followed by the research questions which state what the study specifically investigated. Finally, the limitations of this master's thesis are determined.

1.1 Background

In many industrial control systems, the human-machine interface is the operator's window to highly complex mechanical and physical processes. Independent of the context of use, two fundamental tasks are accomplished by the human-machine interface: communicating information from the machine to the user, and delivering information from the user to the machine. In a complex control system, the specific challenge of the human-machine interface is to enable the users to perceive and manipulate huge quantities of information. Therefore, a user friendly screen design and effective visualization of processes is decisive to minimize the cognitive effort and support the operator even under resource constraints such as time and cognitive workload (Zhang, 2008).

A central issue in operational work is automation that transforms the user's role from actively controlling the system to passive monitoring. Nowadays, control systems autonomously execute long sequences of actions without any interference by the user. Often a more economical and accurate system performance is intended to be achieved by replacing human work. Although automation decreases the user's mental workload, new error potential is created at the same time. If the complexity of automated behavior is hidden, the danger is that the system becomes hard to direct and to observe because the representation of the processes is not sufficient anymore. The additional task of collecting data about the system state may even increase the mental workload. Therefore, it is important that human-machine interfaces of autonomous systems present most salient process changes and events in an appropriate way in order to be quickly perceived and correctly interpreted (Dekker, 2004).

A recent study by the Human-Machine System Research Group at Chalmers University of Technology has focused on improving the visualization of operational status in control rooms of thermal power plants. The study was part of a larger research project financed by the Thermal Engineering Research Institute (Värmeforsk) in Sweden. With the aim to support proactive work in control rooms, new screen images were developed and presented to operators and plant engineers. An interface supports proactive behavior, if it "guides the users in the choice of attentional focus" (Ghaoui, 2006), which means that it suggests or highlights what is expected to be useful for an efficient task completion. In this case, the overall layout of the developed screen images was designed to facilitate pattern recognition so that process information can be interpreted faster and easier. For example, a reappearing element and common object on several screen images is the trend curve that provides a historical view of process values. The so-called mini trends are created to facilitate early and quick detection of deviations from the normal state by showing different shaded alarm levels. Consequently, the operators are able to perceive the current status at a glance on the screen which also allows a quick handling of unanticipated situations. The most positive feedback from the users concerned the more graphical representation of information and the structuring of data which was experienced to contribute to a clearer and faster perception of current plant status (Bligård et al, 2009).

The evaluation of these screen designs was concentrated on gathering subjective, qualitative data by individual feedback from operators on static pictures. The statements of the interviewees confirmed that the main objectives in developing the screen images have been met. However, before the implementation of some of the proposals in a real environment, it was of interest to investigate if the screen designs facilitate a quick detection of deviations from a normal state and thus support a more proactive way of working. This master's thesis addresses this issue and builds upon the previous research by conducting a follow-up empirical study.

1.2 Purpose

The purpose of this master's thesis was to investigate if the screen images, developed by the Human-Machine System Research Group, cause a more self-initiated, proactive behavior. This behavior involves that the user causes changes, rather than just react to changes with the aim to maintain a prolonged, optimal operation of the power plant.

1.3 Goal

The goal of this master's thesis was to further evaluate the existing screen images by studying the user's behavior and reaction to an animated and interactive prototype. The main goal of the study was to address the following research questions:

Question 1: How did the test subjects feel during the tests?
Question 2: How well did the test subjects understood the interface prototype?
Question 3: How well did the test subjects carried out the tasks?
Question 4: How well did the test subjects think they carried out the tasks?
Question 5: How easy was it to accomplish the tasks?
Question 6: Which strategy was pursued to accomplish the tasks?

1.4 Limitations

The usability evaluation is laboratory based which means that the participants acted under controlled experimental conditions. Their attention was focused on the computer screen while scenarios were simulated in fast motion sequences. A higher level of concentration was required than under real operating conditions in thermal power plants; first of all, due to more rapid processes within a short period of time and secondly because no automation was provided which would have supported the participants in carrying out the tasks. Furthermore, it must be noted that this approach does not consider uncontrollable external factors, like noise or time constrains, that may influence the plant operators' capability to act in real working situations.

1.5 Outline of the report

After this introduction to the master's thesis, a general theoretical background of thermal power plant operation and human-machine interfaces in control rooms is presented. Furthermore, the design concept and the screen images developed by the Human-Machine System Research Group are introduced to the reader. The theory chapter concludes with describing the relevant data collection methods of the empirical study.

The next chapter presents the way in which the methods have been applied for the purpose of this master's thesis. At first, type and functionality of the interface prototype are described followed by a detailed test plan for the usability evaluation. Finally, the analysis plan outlines the way how the collected data was compiled and summarized and defines the type of data used to address each research question.

Key findings of objective and subjective data gathered during the tests are summarized in the results chapter. At the beginning, the result of the pilot study is shown followed by the quantitative data on the participants' task performance. Afterwards, significant findings from the post-test questionnaires and interviews are presented to the reader.

The results from the different data collection methods were used for answering the research questions in the analysis chapter of this report. Moreover, different design elements of the interface prototype are analyzed based on the comments from the test participants during the interviews.

This report concludes with a discussion of the data collection methods applied in this thesis project and a review of the pilot study and usability evaluation. Furthermore, it is discussed if the result of the present study supports the aim of the design concept which was basis for the development of the screen images. The conclusion finally summarizes the main inferences drawn from the most significant findings of the study.

2 Theory

This chapter first presents an overview of how heat energy is converted into electrical energy and describes the role of the feed water system in the condensing cycle. Furthermore, it is described how human-machine interfaces in control rooms are connected to mechanical and physical processes in power plants. The screen images of a feed water system which were basis for the creation of an interface prototype are presented thereafter. The chapter concludes with a theoretical background of the data collection methods applied in this project.

2.1 Thermal power plant operation

The conversion of water to steam plays a central role in the cyclic processes of thermal power plants. In a closed loop the steam is recycled, from steam to water and back to steam again (Woodruff, 2005). Different energy sources are used for obtaining heat and generating steam, for example fossil fuels like coal, oil and gas or other sources like bark and wood. The central part of the system, the boiler or steam generator, heats up the water until it changes its form and reaches the desired temperature and pressure. The steam passes through a turbine which drives an electric generator. In the final stage of the turbine the steam is exhausted and pumped through a condenser where it is cooled down and converted back to water. The water leaving the condenser contains dissolved gases which are removed by a deareator before it is reused again as feed water for the boiler (Lindsley, 2000).

The processes in a feed water system take place between the turbine and the boiler where exhausted steam is returned to the boiler in form of feed water. The system collects the condensate and passes it through a filtration process before feed pumps transfer the water back to the boiler. The purpose of feed-regulating valves in the system is to maintain the water level in the boiler.

The schematic presentation of Figure 2.1 shows the cyclic processes in the heating and power plant Riskulla operated by Mölndal Energi near Gothenburg. Heat obtained by burning biomass fuel is used for the steam generation in the boiler. Steam is used for producing electricity on the one hand and on the other for supplying hot water to the district heating network. Finally, the condenser removes the heat and the water is reused again as boiler feed water.

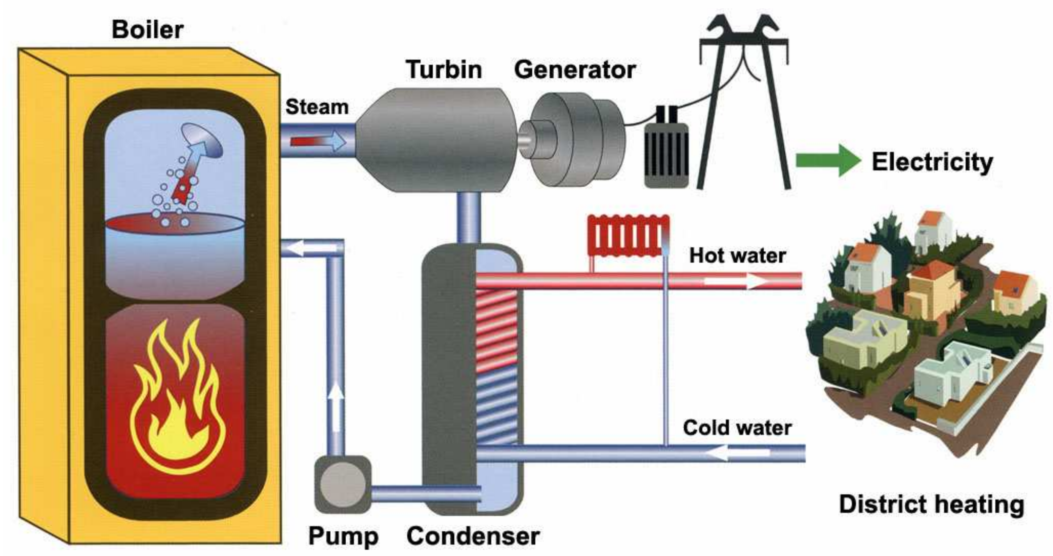


Figure 2.1 Schematic view of the energy production in a heating and power plant (adapted from Mölndal Energi, 2010)

2.2 Interfaces for industrial control

The human-machine interface (HMI) connects the operator to the control system. The process controller as the heart of the control system obtains information about the physical plant status and compares the actual state of the system to the desired state. The graphical user interface (GUI) of a supervisory software provides animated pictures of the processes and enables the operators to manipulate and control them with interactive graphics. Besides the representation of the current status, alarms are activated on the screen if a deviation from the standard value is registered. Sensors detect the process conditions, convert the mechanical and physical properties into signals and transmit them to the central processor unit (CPU). The signals are converted to numeric values and analyzed in terms of deviations from the desired process conditions. The process controller obtains constantly feedback from the sensors which allows a rapid updating of the control room screen contents (Whitt, 2004).

The graphical interfaces nowadays provide images that may look like process flow diagrams, often fully animated and interactive. Figure 2.2 is an example of such a graphical user interface displaying the process conditions of the feed water system in the thermal power plant at Chalmers (Johanneberg campus). As can be seen, this screen image contains different graphical elements like tanks, pumps and valves as well as process values and navigation buttons. According to Whitt (2004), a GUI database configures each element on the graphic screen so that it is linked to the outside world. These so-called action and animation links establish a connection between front-end and the control system. Action links actuated by the operators initiate processes like changing the screen or manipulating local parameters on the HMI. Animation links, on the other hand, give the operator a quick overview of the process conditions and change the appearance according to predefined conditions, for instance a tank that changes to red if the normal tank level is exceeded.

The alarm manager, a standard feature of the HMI, continuously monitors for alarm conditions. It logs and prioritizes alarm events and alerts the operators. Another standard utility program on a HMI system is the historian which records processed data in the background for later retrieval. Historical trends access the log files collected by the historian and display data based on a predefined frequency. The time scale of the trend graph is configurable and can be adjusted by changing the start time. For monitoring purposes, real-time trends are used which are updated much faster than historical trends. Both, the time axes (x-axis) as well as the magnitude (y-axis) scale of the real-time trend graphs can be adjusted (Whitt, 2004).

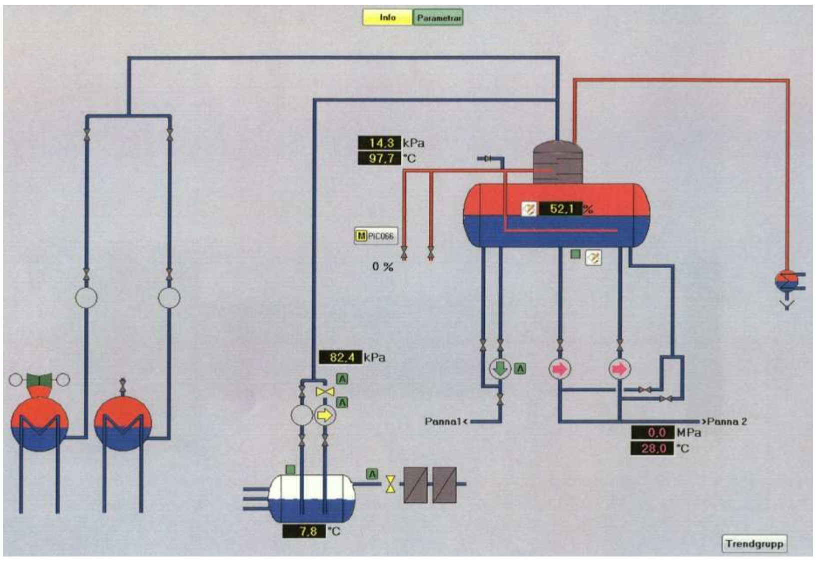


Figure 2.2 Screen image of a feed water system from Chalmers thermal power plant

2.3 Screen images for thermal power plants

Increasing automation minimizes the workload for the plant operators but raises new challenges for human factors engineers in control system design. Automation narrows the plant operator's manual involvement with the processes and shifts the scope of work rather on cognitive tasks. Depending on the level of automation the remaining monitoring task mainly comprises the interpretation of process states and the planning of overall goals for operational activities in the system. An increasing amount of information on the monitoring displays therefore demands for more quality in the visualization of the processes (Hollnagel, 2004).

The Human-Machine System Research Group at Chalmers developed new screen images with the goal to improve the visualization of processes in thermal power plants. In the first place, graphical user interfaces from other process domains were studied. Relevant research work in the domain of nuclear power plants was found in the Halden Reactor Project. In this project, displays for monitoring and controlling a nuclear power plant feed water system were created, based on a design framework called Ecological Interface Design (Welch et al, 2005). In a comparative study it was identified which design guidelines are transferable to the thermal power plants. In the next step, operators and process engineers were interviewed in order to investigate the needs for an improved visualization of plant status. The resulting design principles were basis to create new screen images with the purpose to "clarify the current status, facilitate process understanding for novice users and shorten the learning time and to support handling of unanticipated situations" (Bligård, 2009). The fundamental idea was to shift the focus from numeric process values to a graphic representation of the plant status in order to enable plant operators and engineers to work more proactively. A proactive way of working involves acting before alarms appears, rather than just reacting. The desired result is increased efficiency and productivity in maintaining the optimal state in the system.

2.4 Screen design concept

This master's thesis focuses on the screen images developed for a feed water subsystem in a thermal power plant. As can be seen in Figure 2.3, the lower part of the screen images shows standard graphic symbols for pumps, valves and tanks which are commonly used on displays for process monitoring and control by plant operators and engineers. The upper part of the images visualizes the water flow with elements that support a proactive detection of deviations: mini trends, balance units and level indicators.

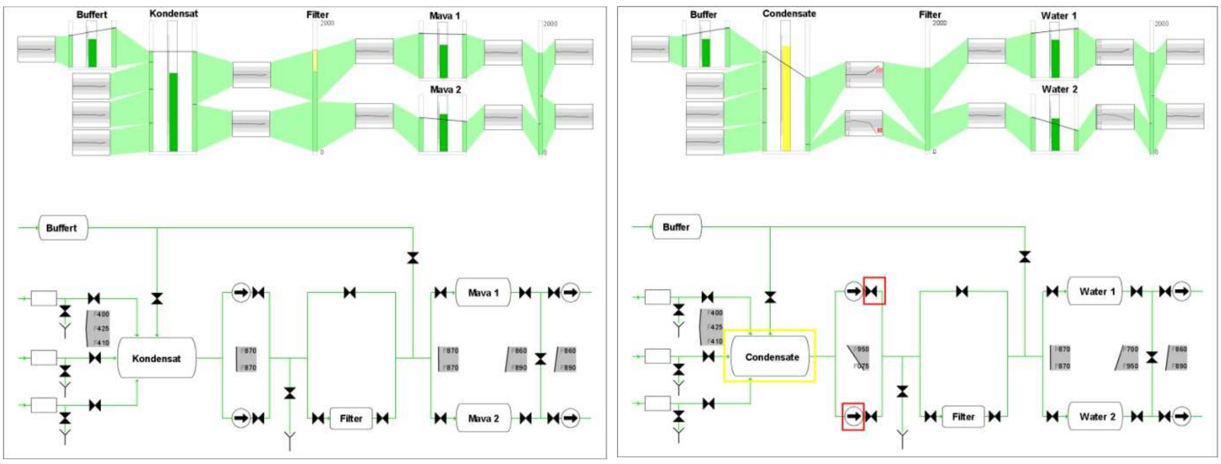


Figure 2.3 Two screen images in normal and alarm state from Bligård (2009)

Mini trends are graphic elements that provide a historical view of process values and disturbances. The different shaded areas in the trend boxes represent an optimum level in the middle and two alarm levels in the outermost areas; alarm level one is indicated in yellow and alarm level two in red. The white areas provide the maximum contrast for the red alert phase as shown in Figure 2.4.

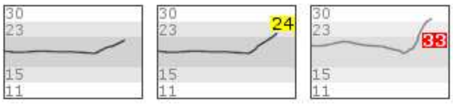


Figure 2.4 Mini trends from Bligård (2009)

Another graphic element in the upper part of the screen image is the balance unit, indicating the inflow and outflow at the tanks. As demonstrated in Figure 2.5, irregularities can easily be detected by the slope of the balance line which reveals whether the inflow is higher or the outflow.

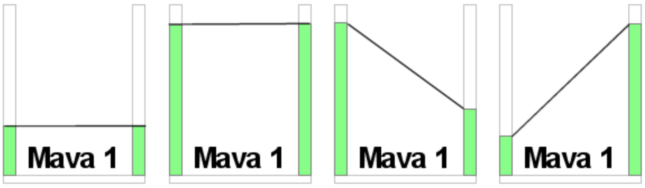


Figure 2.5 Balance units in different stages from Bligård (2009)

The increasing and decreasing level indicators show the changing volume in the tanks and trigger alarm by turning yellow in alarm level one and red in level two. As shown in Figure 2.3, alarm messages are displayed at several places on the screen images: at the tanks, pumps, valves and mini trends.

2.5 Usability evaluation

Usability evaluation techniques are important research tools in the development of interactive systems. They range from complex test designs with large sample sizes to informal qualitative studies with only a few participants. Different testing approaches require different time effort and expenses depending on the objectives of the usability evaluation (Rubin, 2008).

Nielsen (1993) distinguishes between formative evaluation and summative evaluation. Formative evaluation identifies usability problems and provides designers guidance how to improve the design during the design process. In contrast, summative evaluation often takes place at the end of the design process in order to assess the overall quality of a system and to investigate whether the usability objectives have been met.

In summative and formative evaluations, researchers measure factors which are dependent from the user’s behavior. Lazar et al (2010) categorize frequently measured dependent variables into five groups:

- efficiency and accuracy
- subjective satisfaction
- ease of learning and retention rate
- physical or cognitive demand

Experimenters specify how dependent variables are measured in the study and may choose several of them in order to examine a broad range of user behavior.

2.5.1 Prototyping

User interface prototypes are partial implementations of a design and used to explore design ideas and to discover or refine user requirements. In usability engineering it is most common to use prototypes for gathering test data in order to explore usability issues during the development of a system. User interface prototypes vary in fidelity, cost and effort to create them. Selecting the right prototyping technique depends on the goals and resources of the project and the potential users or other stakeholders the prototype is presented to. A rough and sketchy prototype is fast to create and can provide important and useful input for redesign activities. Moreover, low-fidelity prototypes demand smaller investments and have a flexible format which can easily be adapted to changing situations. On the other hand, more realistic prototypes provide sufficient functionality so that potential users interact in a relatively unhindered way. Hi-fidelity prototypes are considered to be useful to get feedback on the system performance or aesthetic characteristics of a visual design (Rosson and Carroll, 2001).

2.5.2 Usability testing

In usability tests, representative users are asked to interact with prototypes in order to study their subjective reactions. A usability evaluation examines the user's performance or satisfaction with the current version of the system or prototype and can take place at different stages in the development process of a system. The goal is to identify aspects of the design that can be improved or to explore if usability objectives have been met (Rosson and Carroll, 2001).

The way in which participants are exposed to the testing condition is described in between- and within-subject testing. In the latter, also called repeated measures, each participant completes the same tasks under multiple test conditions. The advantage of this method is that observing the same participants in all kinds of situations allows direct comparison of the performance. Additionally, a smaller size of participants is needed for the tests compared to between-subject tests. However, the disadvantage is the possible impact of learning effects on the test results. The participants learn from experience how to use the interface prototype and may get better at completing the tasks. On the other hand, in between-subject tests each participant is exposed to one test condition which allows for more original experience and reduces the time for the experiment (Lazar et al, 2010).

2.5.3 Logging the use

Logging as supplementary method during usability testing is used to collect information about the detailed use of the system. An interface log shows statistics about the frequency in which events, for example alarm or error messages, occur. Based on the automatically collected data, it is possible to explore which system features are frequently or rarely used. This may be relevant for further improvements, for instance to make features more accessible or remove them completely from the system. Combining logging data with questionnaires and follow-up interviews has the advantage to investigate not only the actual use of the system, but also the users' motivation to interact with the prototype in a certain way (Nielsen, 1993).

2.5.4 Post-test questionnaires and interviews

Subjective satisfaction, opinions and feelings about the system's ease of use and ease of learning can be assessed with questionnaires and interviews. According to Nielsen (1993), these data collection methods are considered to be indirect from a usability point of view because they do not study the user interface itself, but rather the user's understanding of the product's strength and weaknesses.

The most efficient way to capture preference data in a questionnaire is using a rating scale to state the opinion. Semantic differential scales present a pair of opposite adjectives, like weak and strong or beautiful and ugly, on a 5-point or 7-point scale. In the Likert scale technique, the level of agreement from "strongly disagree" to "strongly agree" to a positive or negative formulated statement is rated. Collecting this data is recommended at the end of each task or at the end of the entire session.

Quick post-task ratings can help to pinpoint aspects of the design or tasks which are found to be problematic. Post-study questions on the other hand are answered with a better understanding of the system which may provide a more effective overall evaluation (Tullis, 2008).

Interviews are a frequently used technique to gather user requirements. Conducting interviews allows for collecting data which cannot be captured by surveys because follow-up questions to clarify the participant's opinion are not possible. However, one-to-one interviews can be very time-consuming and they are not suitable to cover a larger geographic region or to obtain information from a large sample of population.

Courage et al (2005) distinguish three types of interviews: unstructured, structured and semi-structured interviews. In an unstructured interview, the questions to be discussed are open-ended and the interviewee is free to talk and do not need to cover certain topics. A structured interview consists mainly of closed-ended questions and the interviewee has to choose from the provided options. The topics for discussion are predefined which means that the interviewer does not ask more questions as listed in the script. Semi-structured interviews are a combination of both, structured and unstructured interviews. The interviewer covers a set of open-ended and closed-ended questions but also encourages the respondent to talk freely.

2.5.5 Pilot study

A pilot or feasibility study should be performed before the actual usability tests are conducted in order to refine the experimental procedure. Usually, pilot tests can be carried out quite informal with people who are easily available. Colleagues for example can act as test users in a pilot study to find out severe problems in the test plan. The feedback of only a few pilot subjects can be valuable to modify materials and procedures (Nielsen, 1993).

During a pilot study, the experimenter explores for example if the instructions or questionnaires are comprehensible for the pilot subjects in order to avoid misunderstandings beforehand. It may also happen that the planned time for the test session will be exceeded. In this case, either the time-frame for the test or the tasks need to be changed. The experimenter may also find that the tasks are too difficult or easy than expected so that a revision of them is necessary before the usability test. All these measures based on the feedback of the pilot tests help to improve the quality and efficiency of the usability study.

2.5.6 Techniques for data analysis

The collected data must be compiled and summarized into a form that enables to see patterns. Two techniques are common for analyzing data from rating scales in questionnaires. In the first one, the average of the numeric values are calculated which are assigned to each scale position. For example, on a 5-point Likert scale the value 1 is assigned to "strongly disagree" and the value 5 to "strongly agree". The other technique is to look at the top-2 and bottom-2 scores and to present the amount or percentage of participants who somewhat or strongly agree or disagree.

Summarizing the responses from open-ended questions in interviews or questionnaires is more challenging. One technique is to collect all the responses in a document and to group similar comments together. It must be considered, however, that participants express similar opinions and suggestions in various ways. Therefore, a logical grouping performed by more than one rater is recommended in order to increase the reliability (Tullis, 2008).

For analyzing the user's performance, the experimenters define activities for measuring quantifiable data, for example the time taken to complete specific tasks or to recover from errors. The experimenters decide when to start and stop the timer and they choose the timing precision. Other typical measurements are for instance the number of completed tasks within a certain period of time and the number of user errors (Nielsen, 1993).

3 Method

This chapter presents the methods applied in this master's thesis for evaluating the screen images of the feed water system. In the first step, an animated interface prototype was built based on the screen images. The aim was to simulate processes in the system and to run pre-defined scenarios that require interaction between the prototype and the test participants. The first section in this chapter informs about the type and fidelity of the interface prototype and the used technology. Moreover, the scenarios are presented that motivated the test participants to interact with the prototype. The second section contains a test plan for the usability evaluation of the interface prototype covering the methodology used, the test procedure and details about the test participants.

3.1 Interface prototype

In this thesis project, a prototype of a simplified feed water system was needed, which was able to run several pre-defined scenarios. It must be noted, that the actual physical reality and the realistic manipulation of the processes in the system were secondary for this study. The purpose of this interface prototype was to enable the test participants to act freely within predetermined parameters. The interaction with the prototype was kept simple and efficient because the focus should remain on the visualization of the processes. The result was an interface prototype reduced in the number of features and with no underlying physical processes.

Adobe Flash CS 4 was chosen to be an appropriate tool for creating the interface prototype and to run more or less challenging scenarios. The original screen images presented in chapter two had to be slightly modified to adjust them to the purpose of the test setup (3.3 Test plan). It was planned to conduct two usability tests with each participant. In one of the tests, the whole screen image was shown and in the other one the flow visualization in the upper part of the image was removed. In the following, the two parts of the screen image are named differently: flow visualization (upper part), schematic diagram (lower part). In order to provide sufficient information in each of the tests, it was necessary to make modifications in both parts as listed below.

Schematic diagram:

- Values were added to show the amount of liters pumped through the system.
- A small volume bar in the tank symbols was included as water level indication.
- Sliders as control elements were attached to the pumps and valves. The sliders at the pumps show the amount of water (in liter) pumped out of the Kondensat tank whereas the sliders at the valves indicate as a percentage how much they are open.

Flow visualization:

- Unfiltered and drained water was colored in blue tones to highlight the difference to filtered water.
- The alarm zones in the outermost areas of the trend boxes were downscaled in order to provide more room for adjusting the water flow within the alarm limits.

The following figure presents the interface prototype after the modifications mentioned on the previous page. The elements of the static images have been transferred to vector based objects animated in 2D. These objects were drawn and controlled dynamically at runtime using ActionScript 3.0 as programming language.

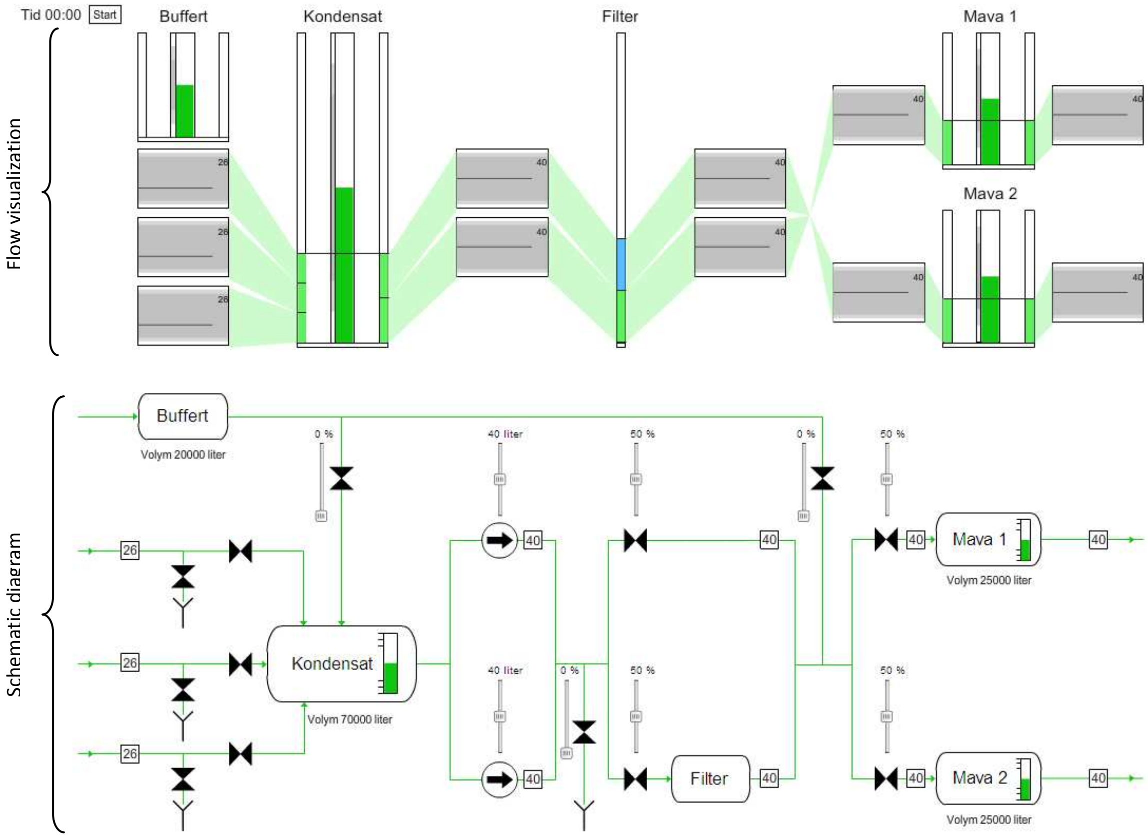


Figure 3.1 Interface prototype based on the screen images of the feed water system (Appendix A and B)

3.1.1 Functionality of the prototype

The overall purpose in the tests was to ensure that the feed water tanks Mava 1 and Mava 2 are supplied with sufficient quantities of water from the Kondensat tank. The test participants were able to control the amount of water by adjusting two pumps (pump one and two in Figure 3.2). Furthermore, seven valves could be opened and closed for regulating the water flow within the system. As shown in Figure 3.2, valve number one could be opened to let water run out of the system. Valve two and three were used to regulate how much water flows through the filter and how much is bypassed which means that the water flows uncleaned to Mava 1 and 2. The water inflow to the feed water tanks was controlled by valve number four and five. If needed, buffered water could be added to the tanks by opening valve six and seven. The water supply for the Kondensat tank (“Inflow” in Figure 3.2) and the amount of water flowing out of Mava 1 and 2 (“Outflow” in Figure 3.2) were pre-determined parameters which could not be changed by the test participants.

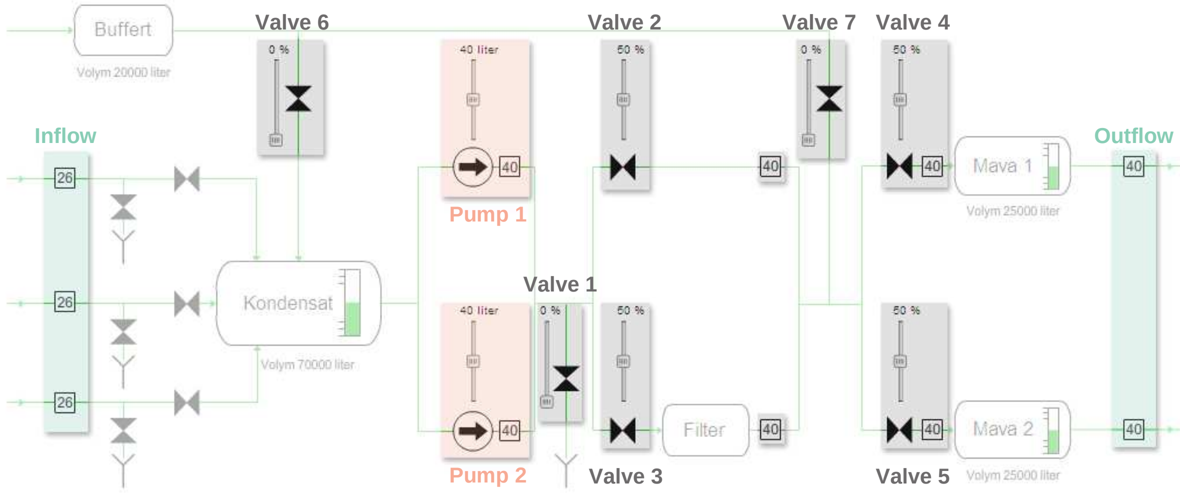


Figure 3.2 Valves and pumps in the schematic diagram

The interface prototype generated alarm messages when the water level in the tanks was getting too high or too low. The volume bars in both parts of the screen image turned yellow when reaching the first alarm limit and red if the second limit was exceeded. The limits were indicated as lines in the schematic diagram and as grey-shaded areas in the flow visualization (Figure 3.3).

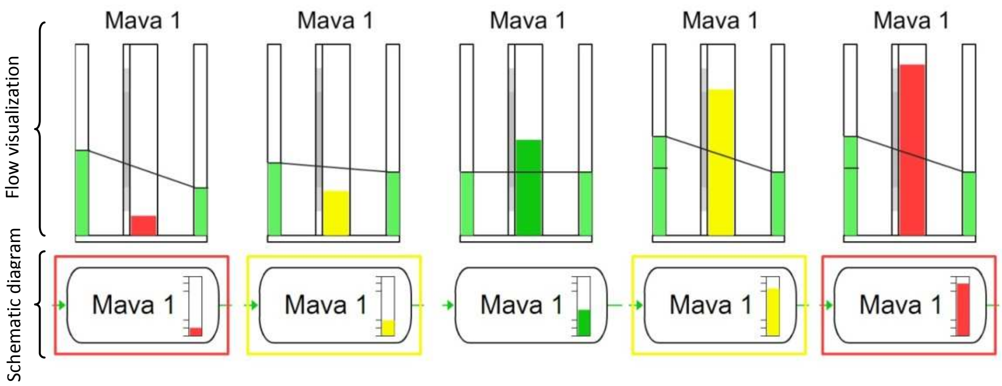


Figure 3.3 Yellow and red alarms at the tanks

Since the valves regulate the water flow within the system, alarms at the pumps occurred if too much water was pumped out of the Kondensat tank. Either the valves had to be opened more so that the pumped water could unhindered flow through the system, or the pumps had to be adjusted accordingly. The alarm messages appeared as yellow or red frames around the pumps, depending on how critical the situation was. Examples for such alarm situations at the pumps are illustrated in Figure 3.4.

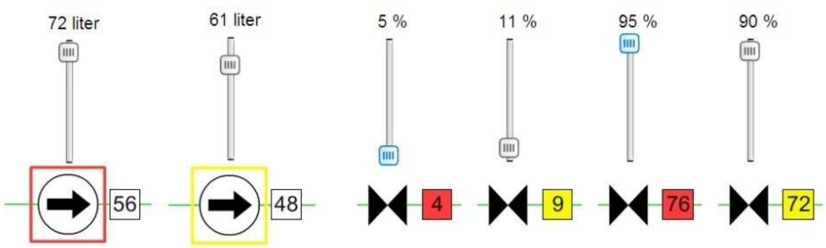


Figure 3.4 Yellow and red alarms at pumps and valves

Occasionally, alarms at the valves appeared in the form of yellow or red marked values (Figure 3.4). If the water flow passed 70 liters, the value turned yellow and changed to red if the amount of water exceeded the limit of 75 liters. Yellow and red alarms occurred also if the water flows decreased to 10 liters and below. The only valve which could be closed completely without initiating alarm was the bypass valve, highlighted as valve number two in Figure 3.2. Closing this valve allows to filter the maximum amount of water before it flows to the feed water tanks.

3.1.2 User interaction scenarios

As mentioned previously, the water inflow to the Kondensat tank and the outflow of Mava 1 and 2 were controlled by the interface prototype (Figure 3.2). A sequence of events was created where inflow and outflow parameters were updated every 30 seconds. One test run with these changing events lasted for 14 minutes and included an introduction (two minutes) and three scenarios (four minutes each). During the first two minutes, the water supply was quite balanced before scenarios started where inflow and outflow parameters changed differently so that more water was accumulated in the system than at other times. These challenging scenarios were developed in order encourage the participants to interact with the prototype and to vary their mental workloads.

Figure 3.5 illustrates the changes of the total inflow and outflow of the system in one test run. As shown in the line diagram, the inflow and outflow is quite balanced during the intro. In the first scenarios, there is a sharp increase of water inflow at the Kondensat tank whereas the outflow shows only a slight upward trend. The second scenario causes a situation with the reversed effect. Within the next four minutes the inflow is continuously lower than the outflow. In the last scenario, the total outflow is nearly as high as the inflow, but there are very unbalanced and unsteady fluctuations in the outflow of the feed water tanks Mava 1 and Mava 2 which are not visible in the diagram.

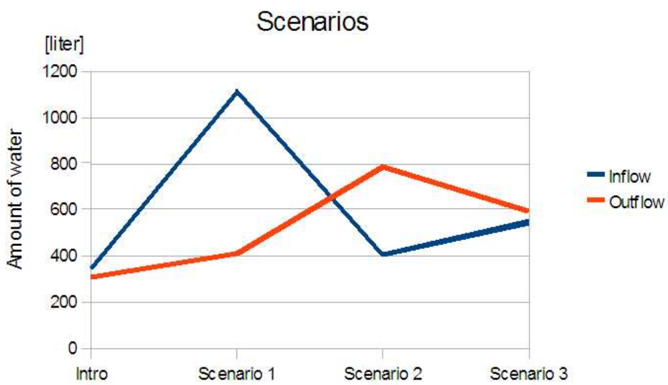


Figure 3.5 Test scenarios with an overall duration of 14 minutes

The interface prototype generated log files as a record of the participants' interactions during the tests. Every second, the dynamic parameters in the system were logged in a database, for instance the amount of filtered and unfiltered water, volume levels in the tanks, the amount of water pumped out of the Kondensat tank, etc. Within 14 minutes, the overall duration of one test run, 840 data sets were stored. This data allowed to assess the individual performance of each test participant and to compare the result with others.

3.2 Test plan

In this master's thesis, empirical methods were applied in form of usability tests, post-test questionnaires and interviews. The combination of these data collection methods was aimed at gathering subjective and objective data: user statements towards the interface prototype in the interviews, questionnaire responses and task performance during the usability tests.

The test plan included following three stages:

1. Pilot-tests with six students
2. Usability tests with six students
3. Usability tests with six plant operators

In the first stage, informal pilot tests were conducted with six students; firstly, to work through the tasks, instructions and questionnaires, secondly, to determine how much time is required from each user and thirdly to test the correct operation of the flash prototype and the database. After the pilot tests, the actual usability evaluation of the interface prototype was carried out with students and plant operators. The number of plant operators available was limited to six in total. For that reason, also six students were chosen in order to meet the number of plant operators who acted as participants in the usability tests.

3.2.1 Methodology

For this study, a within-subject test setup was chosen to be appropriate since only a small number of power plant engineers were available for the entire evaluation. This means that each participant was exposed to two tests, one with the flow visualization and one with the schematic diagram only. With this test setup it was also easier to detect significant differences in the performance of the same participant in both tests.

The three scenarios which generated different quantities of water in the system (3.4.1 User interaction scenarios) were shuffled in each test run. To achieve the required number of 12 tests, the sequence of the scenarios was repeated a second time as illustrated in Figure 3.6.

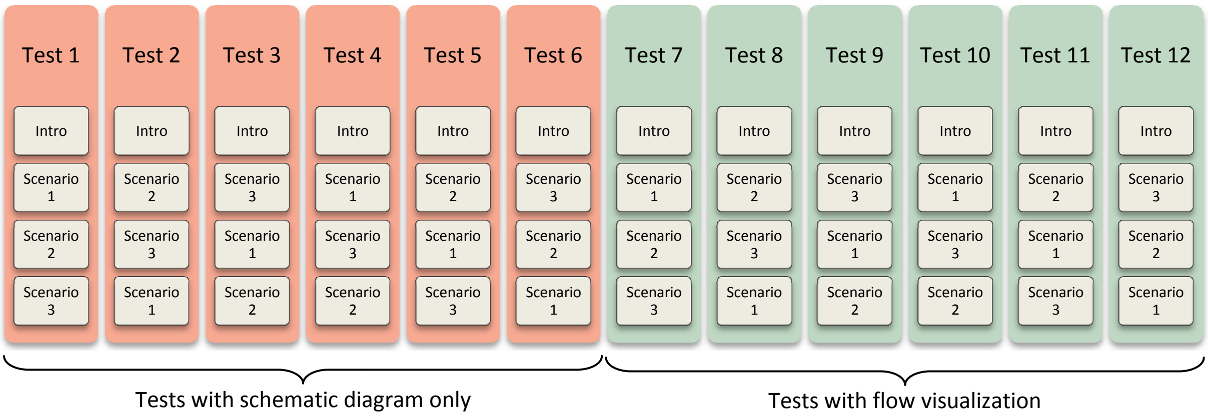


Figure 3.6 General test setup

In the next step it was ensured that no participant performs two tests with the same sequence of scenarios. As can be seen in the Figure 3.7 on the next page, the test participants were divided into pairs, one student and one plant operator each, who carried out the same tests but locally and temporally independent from each other. Additionally, two groups were formed, both exposed to two different test procedures. Group A started with the test that contained the schematic diagram only, whereas in Group B the water flow visualization was shown in the first test. The reason for changing the order of the tests was to distribute a possible learning effect across both tests. The participants were placed randomly in Group A and B with the constraint that each group consisted of three students and three plant operators.

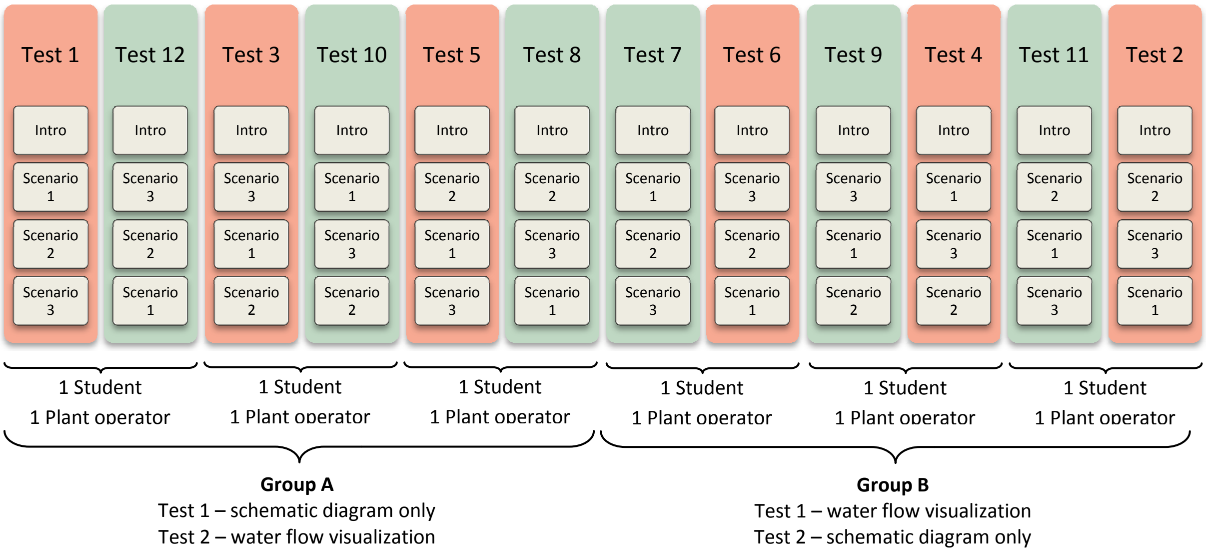


Figure 3.7 Test setup Group A and Group B

3.2.2 Test procedure

In total, 12 individual usability sessions were conducted. Figure 3.8 shows the test procedure of each session with a duration of approximately one hour. The first 10 minutes were used for an introduction to the interface prototype and for explaining some background information. Afterwards the test participant performed two training scenarios of 7 minutes each. The first test with the schematic diagram was presented to participants of Group A whereas Group B started with the additional water flow visualization. The test scenarios ended after 14 minutes and the participants were asked to fill out the first questionnaire. Group A continued with the test that contained the supplementing flow visualization and Group B performed the test with the schematic diagram only. After 14 minutes the test scenarios were finished and the second questionnaire was handed over to the test participants. The whole session concluded with an interview of about 10 minutes which was audio recorded for a later review.

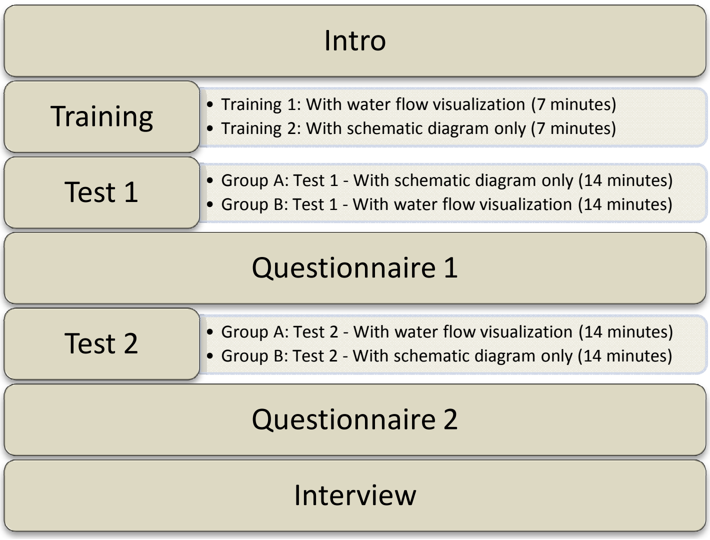


Figure 3.8 Test procedure for the usability evaluation

3.2.3 Participants and location

As representative users, six plant operators were chosen to act as test participants in this usability study. The test sessions were performed at their workplace in the thermal energy plants Sävenäs in Gothenburg and Fortum in Stockholm. Furthermore, six students without relevant qualifications participated in the study. The intention was to gather feedback from unbiased participants and to compare the test results of the student group with those of the plant operators. The usability tests with the students took place at Chalmers University of Technology.

Participant type	Characteristic	Number of Participants
Plant operators at Sävenäs	2 service technicians 1 plant engineer	3
Plant operators at Fortum	2 service technicians 1 shift leader	3
Students at Chalmers	5 Engineering science 1 Natural science	6
	Total number of participants	12

Table 3.1 Participants in the study

3.2.4 Tasks

The test participants were told to balance the water level in the tanks at an optimum level, thus within the alarm limits. This was presented as the first and most important task. The second one was to avoid alarms at the pumps and valves during the test scenarios. As third task, most of the water should be filtered before it flows to the feed water tanks. The last one was to drain as less water as possible. If too much water was already collected in the Kondensat tank, the drain valve could be used to let water flow out of the system. However, it was not intended leave the valve constantly open. For this reason, a fourth task was formulated so that the drain valve was used only if necessary.

The following list that summarizes the tasks according to their priorities was created for the test participants as a reminder:

1. Keep the volume in the tanks within the alarm limits
2. Keep the valves and pumps within the alarm limits
   High high 75
   High 70
   Low 10
   Low low 05
3. Filter as much water as possible
4. Dump as little water as possible

The reminder with the tasks and their priorities was placed near to the test participants on the table or was displayed on another computer screen. Training and test scenarios were played as Adobe Flash animation in a standalone player on a Laptop or PC.

3.2.5 Data logging

During the usability tests, logs of selected parameters were generated as a record of the participants' interactions with the interface prototype. The automatically collected statistics included following data:

- Time of activation and duration of alarms
- Red and yellow alarms at the tanks (Kondensat, Mava 1 and 2)
- Red and yellow alarms at the pumps and valves
- Amount of filtered and unfiltered water
- Amount of drained water

3.2.6 Post-test questionnaires and interviews

To complement the data logging of the user performance, it was decided to gather preference information from the participants after the prototype evaluation. For that purpose questionnaires were created to collect opinions about different aspects of the design. Furthermore, interviews were carried out after each session to better understand the issues raised during the usability tests.

Two questionnaires with 5-point rating scales were developed for measuring the participants' feelings and experience after each test. The first questionnaire included 10 questions which are repeated in the second one as well. Six additional questions were added to the second questionnaire for an overall evaluation of both tests (Appendix C).

In general, the questionnaires covered four topics which are listed below:

> ***Emotional state***
> The first three questions captured the feelings of the participants. Specifically, it was asked to which degree they felt confident, stressed and satisfied during the test. The rating scale was labeled with "not at all" to "very much".
>
> ***Understanding***
> In the following three questions the participants rated the understanding of the interface prototype from "difficult" to "easy".
>
> ***Task performance***
> Four questions related to the given tasks and the test participants were asked to evaluate how difficult or easy it was to accomplish them.
>
> ***Comparison of test one and two***
> Six multiple choice questions after the second test asked for an overall evaluation of the first and second test.

The interviews at the end of each session were semi-structured and based on the individual ratings in the questionnaires. To identify the reason behind the ratings, the interviewees were encouraged to clarify their opinions about the interface prototype.

Mainly following issues were covered in the interviews:

- Feelings during the tests
- Understanding of the interface prototype
- Key events or critical incidents that happened during the tests
- Parts of the interface prototype which were particularly helpful in both tests
- Aspects of the design the participants liked or disliked in both tests

3.3 Data analysis

In the first step, the collected data during the usability tests was compiled and summarized. Handwritten notes and interview responses were transferred to a master sheet on the computer which was kept as a running summary during the data collection process. The transcript of the interviews was used to group similar responses into categories which helped to scan the results more quickly. Furthermore, the questionnaire responses were summarized to identify the number of participants who selected each possible choice. Due to the small amount of test participants it was not necessary to calculate average scores in order to detect trends. Therefore, the focus was more on significant differences in the top-2 and bottom-2 scores on the 5-point rating scales.

The log files of the test participants already provided summaries of the interaction with the interface prototype. The data sets were separately collected in tables according to following criteria:

1. Performance of group A and B
2. Performance of students and plant operators
3. Performance of the tests that contained the schematic diagram only and performance of the tests with the supplementing flow visualization

These statistical summaries enabled to see differences in the performance of the groups, but also differences in the overall performance of the two tests.

In the second step, the data was analyzed and the research questions introduced in the first chapter of this thesis report were answered. The type of data used to address each of the questions is presented below.

Question 1: How did the test subjects feel during the tests?
The participants described their experience and feelings in the interviews and commented the answers from the questionnaires.

Question 2: How well did the test subjects understood the interface prototype?
How quickly and easily the individual participant understood the interface prototype was discussed in interviews based on the questionnaire responses. The log files additionally provided information about a possible learning effect in the second tests.

Question 3: How well did the test subjects carried out the tasks?
The recorded data during the tests was used to address this research question. The analysis of the data was based on certain criteria: number and duration of all alarm events, number and duration of yellow and red alarms, amount of filtered and drained water.

Question 4: How well did the test subjects think they carried out the tasks?
The participants assessed their own individual task performance in the questionnaires and their statements and opinions were further discussed in the interviews. Additionally, the recorded data allowed to objectively evaluate the task performance and to compare the results with the answers from the test participants.

Question 5: How easy was it to accomplish the tasks?
The questionnaire responses and the statements from the interviews were used to answer this research question. The test participants evaluated the tasks in the questionnaires and discussed their opinions in the interviews.

Question 6: Which strategy was pursued to accomplish the tasks?
This question was answered based on the comments during the interviews. The participants talked about the strategies they developed to accomplish the tasks and how they implemented them during the tests.

4 Results

This chapter highlights key findings from objective and subjective data gathered during the test sessions with the students and plant operators. At first, a small-scale pilot study was conducted in order to identify potential problems with the interface prototype and the planned test procedure. Based on the feedback from this study, parts of the test procedure and scenarios were modified. These changes are described at the beginning of this chapter followed by the results of the empirical study. The results include quantitative data from the log files for an objective assessment of the participants' task performance and significant findings from the evaluation of the questionnaires. Moreover, comments from the interviewees about their feelings during the tests and their understanding of the system are summarized. The presented results from the different data collection methods were basis for the analysis in the next chapter.

4.1 Pilot study

The feedback in the pilot study conducted with six students was basis for changing the initial length of the training and test scenarios. It was found that the training scenarios with an overall duration of four minutes were too short. In the next step, the practice time was increased to 10 minutes and later on to 14 minutes in total. At the same time, the test scenarios were reduced to a total length of 28 minutes in order to keep the whole test procedure, including questionnaires and interviews, within an acceptable time frame for the participants.

The test scenarios with different fluctuations in the water flow were found to be occasionally challenging, but the participants were absolutely able to cope with these kinds of situations which means that no changes had to be made in the scenarios. The test procedure, however, was modified during the pilot study. In some pre-tests, the training and test sessions were not separated instead the procedure began with a training unit, continued with the first test and started again with training as introduction for the second test. This procedure seemed to confuse some participants which resulted in more questions during the actual tests. When they finally talked about their experiences during the interviews it turned out, that they sometimes mixed together training units and test scenarios. For these reasons, it was decided to let the participants complete the training scenarios first, before they start with the two tests.

A recurring issue in the pre-study was the usage and the positioning of the sliders. The sliders caused problems when participants wanted to set certain values by dragging the slider up and down because of the relatively small slider area. Additionally, the positioning of the sliders was criticized in the test with the flow visualization. Some participants found the placement not to be optimal because adjusting the sliders and monitoring the effects of the input happened for them in two parts of the interface.

Difficulties in monitoring the increasing and decreasing volume values could be reduced to a certain degree by slowing down the changing rate of the values. Changes every second were easier to perceive and helped the participants understand faster if the volume in the tank was increasing or decreasing.

The amount of information given by moderator has increased steadily during the pilot study. It has become clear that the interface cannot be operated completely intuitively especially if background information is missing. Moreover, the understanding and correct interpretation of the different interface elements was required in order to accomplish the tasks.

Every misunderstanding would have had considerable impact on the individual and overall result. Therefore, the moderator tried to cover most of the issues beforehand in the introductory talk to avoid misunderstandings. Additional questions by the participants and experiencing the system prototype in the training scenarios helped to develop a basic understanding for the test session.

As a result of a more extensive introduction at the beginning, the participants began to forget about details regarding the tasks and priorities. In most cases the moderator had to remind them to filter the water and to drain as less water as possible. Consequently, a to-do list was created presenting the tasks in the order of their priorities. This reminder was placed on the table or displayed on an additional screen next to the test subjects.

4.2 Data logging

This section presents the results of the data recorded in each test based on the in chapter two defined criteria: number and duration of all alarms, number and duration of yellow and red alarms, amount of filtered and drained water. According to the test plan, every test participant was exposed to two tests for evaluating two interface prototypes. Both displayed a schematic view of the feed water system but one of them included the additional visualization of the water flow. In the following, the tests are named differently to make a clear distinction between the interface prototypes to be evaluated. Therefore, “test with the flow visualization” describes the evaluation of interface prototype with the schematic view and the supplementing visualization whereas “test with the schematic diagram” simply refers to the prototype that displays the schematic view only.

4.2.1 Alarms at tanks, pumps and valves

The participants kept in all the tests the three tanks within the red alarm zone. Only the yellow upper limits triggered alarms, which means that most of the participants had more water in the system than required. In both tests, two students and two plant operators were able to balance the water inflow and outflow at the tanks so that no alarm appeared at all. Eight participants in total kept the tank volumes within the alarm limits in the test with the flow visualization. Only four test subjects were able to avoid alarms at the tanks in the test with the schematic diagram. Regarding pumps and valves, nearly all participants caused red and yellow alarms in both tests. Only one plant operator managed it to avoid red alarm at the pumps and valves in the test with the flow visualization.

As can be seen in Table 4.1, the plant operators generated almost twice as many alarms at the tanks. Consequently they spent a longer period of time in the yellow alarm zone compared to the student group. The opposite is the case when looking at the alarms at the pumps and valves. Here, a considerably higher number of red and yellow alarms was recorded in the student group. With 377 registered alarms the students remained more than twice as long (63 min 20 sec) in alarm state compared to the plant operators (29 min 40 sec).

	Tanks[1]		**Pumps and valves**[2]	
6 Students	9	08:23	377	63:20
6 Plant operators	16	13:46	248	29:40
	25	22:09	625	93:00

[1]Yellow alarms only
[2]Yellow and red alarms

Table 4.1 Alarms caused by students and plan operators

Table 4.2 on the next page shows the recorded alarms in the test group A and B. Both groups consisted of three operators and three students, but the first test in group A included the schematic diagram only and group B started with test that contained the additional flow visualization. As can be seen in this table, a significant improvement took place in the second tests of Group A.

In these tests they were able to reduce the number and the duration of alarms at the tanks to a great extent. Group A generated six alarms less in the second try with the flow visualization and decreased the alarm duration by more than five minutes. A slight improvement can also be determined in the number and duration of alarms at the pumps and valves. Group B on the contrary caused more alarms at the tanks, pumps and valves in the second tests that contained the schematic diagram only.

Comparing the data from group A and B, it can be determined that the highest number of alarms at pumps and valves was recorded in the first tests of group A (182), however, the longest alarm duration appeared in the second tests of group B (28:36). In both groups, these peak values emerged in the tests with the schematic diagram only.

Group A[1]	**Tanks**[2]		**Pumps and valves**[3]	
	Number of alarms	Alarm duration	Number of alarms	Alarm duration
Test 1 – schematic diagram	8	07:50	182	20:55
Test 2 – water flow visualization	2	02:43	171	20:23
	10	10:33	353	41:18

Group B[1]	**Tanks**[2]		**Pumps and valves**[3]	
	Number of alarms	Alarm duration	Number of alarms	Alarm duration
Test 1 – water flow visualization	7	05:43	113	23:06
Test 2 – schematic diagram	8	05:53	159	28:36
	15	11:36	272	51:42

[1]Each group consisted of three students and three operators
[2]Yellow alarms only
[3]Yellow and red alarms

Table 4.2 Number of alarms and alarm duration (mm:ss) within group A and B

Table 4.3 illustrates that in the tests with the water flow visualization, considerably fewer alarms at the tanks, pumps and valves were generated. Moreover, the participants reduced the duration of all alarms by 11 minutes and 30 seconds regardless whether the flow visualization was presented in the first or second test.

Group A and B[1]	**Tanks**[2]		**Pumps and valves**[3]	
	Number of alarms	Alarm duration	Number of alarms	Alarm duration
Water flow visualization	9	08:26	284	43:29
Schematic diagram	16	13:43	341	49:31
	25	22:09	625	93:00

[1]Each group consisted of three students and three operators
[2]Yellow alarms only
[3]Yellow and red alarms

Table 4.3 Number of alarms and alarm duration (mm:ss)

The alarms generated by all participants in both tests are presented as bar diagram in Figure 4.1. The diagram gives an overview of red and yellow alarm messages in the duration of their occurrence. As can be seen, yellow alarms caused longer alarm states in both tests than red alarms. In the test with the water flow visualization, each test subject generated alarms of 4 minutes and 19 seconds at an average. In comparison to that, 5 minutes and 16 seconds is the average duration of yellow and red alarms in the test with the schematic diagram only.

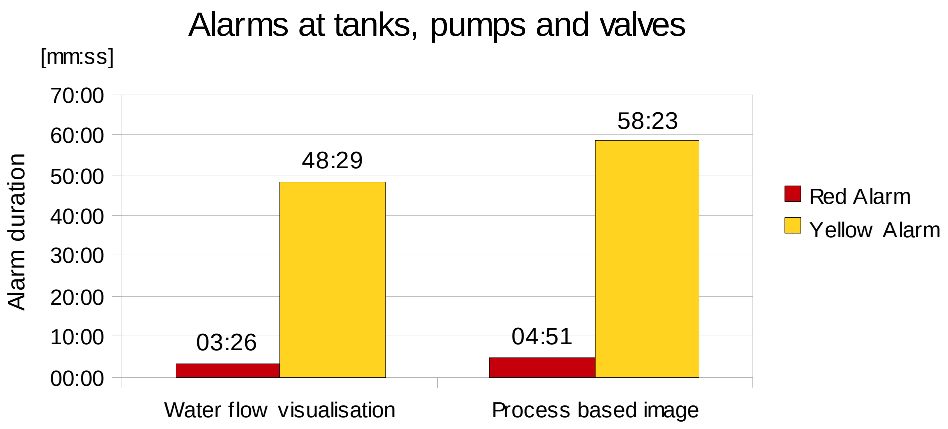


Figure 4.1 Duration of yellow and red alarms at the tanks, pumps and valves

4.2.2 Filtered water

Table 4.4 presents the total amount of filtered water in both tests. As shown in this table, less water was cleaned in test with the water flow visualization. No significant difference, however, can be determined when comparing the amount of filtered water by the students and plant operators.

Group A and B[1]	Water filtered	Water not filtered
Water flow visualization	589,385	39,470
Schematic diagram	600,284	32,198
	1,189,669	71,668

[1]Each group consisted of three students and three operators

Table 4.4 Total amount of filtered and unfiltered water in both tests

More prominent is the increase of cleaned water in the second tests in group A and B (Table 4.5 on the next page). Regardless whether the second tests contained the schematic diagram only or the supplementing flow visualization, an increase of 1.25 % in group A and 4.7 % in group B was recorded. Moreover, most of the water in the system was filtered by group B in the second tests with the schematic diagram only.

Group A[1]	Water filtered	Water not filtered
Test 1 – schematic diagram	291,007	16,390
Test 2 – water flow visualization	294,704	13,197
	585,711	29,587

Group B[1]	Water filtered	Water not filtered
Test 1 – water flow visualization	294,681	26,273
Test 2 – schematic diagram	309,277	15,808
	603,958	42,081

[1]Each group consisted of three students and three operators

Table 4.5 Amount of filtered and unfiltered water (liter) within group A and B

4.2.3 Drained water

As can be seen in Table 4.6, less water was drained during the tests with the schematic diagram only. Comparing the amount of water dumped by students and plant operators a contrary result was obtained. The students drained less water in the tests with the flow visualization. The least amount of water, however, was drained by plant operators in the tests that contained the schematic diagram only.

	Water flow visualization	Schematic diagram
6 Students	67,588	70,033
6 Plant operators	71,997	50,957
	139,585	120,990

Table 4.6 Amount of water (liter) drained by students and plant operators

Taking a closer look at the two groups, it can be determined that most water in the system was dumped in the first tests (Table 4.7). The maximum amount of water that ran through the drain valve was achieved in the first tests with the schematic diagram only by the participants in group A.

Group A[1]	Drained water	Group B[1]	Drained water
Test 1 – schematic diagram	74,577	Test 1 – water flow visualization	73,263
Test 2 – water flow visualization	66,322	Test 2 – schematic diagram	46,413
	140,899		119,676

[1]Each group consisted of three students and three operators

Table 4.7 Amount of drained water (liter) within group A and B

4.3 Post-test questionnaires

The majority of the test subjects (10 out of 12) answered that they felt quite or very confident during both tests. Though, more students than plant operators felt confident during the tests with the flow visualization. Furthermore, the stress level was rated by three respondents to be quite high in the tests that contained the schematic diagram only. However, just one participant felt quite a lot of stress in the test with the additional flow visualization. In the ratings about feeling stress, no significant differences can be found between students and operators.

The overall understanding of the processes was rated to be almost equally good in both tests. However, it can be determined that the current status and the effects in the system were more intelligible in the tests with the flow visualization. In this case, more plant operators than students found it easier to understand the current status in the tests with the flow visualization.

Some more respondents found the tasks of keeping the tanks, pumps and valves within the alarm limits to be easier in the tests with the flow visualization. Especially avoiding alarms at the pumps and valves was rated to be more difficult in the tests with the schematic diagram only. Recovering from critical alarm states and developing a strategy was evaluated to be almost equally easy in both tests. Though, twice as many students as plant operators found it easier to develop a strategy in the tests with the flow visualization.

The majority of the respondents (10 out of 12) evaluated the overall experience to be best in the tests with the flow visualization. The overview of the running processes and controlling them was also found to be better in these tests by the same number of respondents. Furthermore, most of the participants (8 out of 12) thought that they have worked more proactive in the tests with the flow visualization.

The question in which of the two tests the participants think they had a better task performance was answered as follows: seven evaluated their overall performance to be better in the tests with the additional flow visualization, four participants though they performed better in the tests with the schematic diagram only and one expected no difference in the results. As can be seen in Table 4.8, there is no considerable difference between students and plant operators in the answers.

In which of the two simulations do you think was your task performance better?			
	Water flow visualization	Schematic diagram	Equally good in both
6 Students	3	2	1
6 Plant operators	4	2	
	7	4	1

Table 4.8 Performance evaluation of students and plant operators

Comparing the answers within the two test groups the result becomes clearer. As presented in Table 4.9, almost all participants in group A and B evaluated their performance to be better in the second test. In group A, the second test contained the additional water flow visualization and in group B this test displayed the schematic diagram only.

In which of the two simulations do you think was your task performance better?			
	Water flow visualization	Schematic diagram	Equally good in both
[1]**Group A**	5	1	
[2]**Group B**	0	5	1
	5	6	1

[1]Group A (3 students, 3 plant operators): Test 1 – schematic diagram; Test 2 – water flow visualization
[2]Group B (3 students, 3 plant operators): Test 2 – water flow visualization; Test 2 – schematic diagram

Table 4.9 Performance evaluation within group A and B

4.4 Post-test interviews

Based on the transcript of the semi-structured interviews, categories were identified for grouping similar responses together. This categorization enabled to scan the result quicker in the further analysis. The following tables contain illustrative quotes that represent each category of response which should give the reader an overview of the range of comments. However, it must be noted that for a holistic view not only this summary was used for answering the research questions but the whole transcript of the interviews.

The first Table (4.10) summarizes the comments about the participants' feelings during the tests. The majority of the test participants felt less stressed during the second tests because their understanding of how to control the processes in the system was perceived to be better. Other ones stated that they felt either less stressed or more confident in the test with the supplementing flow visualization because the overview was better and changes in the system were discovered faster. When asked what has caused stress in each of the tests, the opinions of the interviewees were divided. However, the permanent changes of the water inflow and outflow, as well as the appearing alarm messages, were stressful to a greater or lesser extent for many participants. In this context, four participants reported in the interviews about challenging situations during the tests and a couple mentioned that they missed a kind of signal to indicate when the parameters have changed. Especially monitoring the process values and reacting accordingly was perceived to be stressful during the tests.

Feelings during the tests	Reasons	Stress factors
bit less stressed in the 2nd test	I knew how I can get out of critical situations	stressed when alarm appeared
bit less stressed in the 2nd test	overall understanding was better	
less stressed in the 2nd test	get more used to the system	
less stressful in the 2nd test	better control over the processes	
less stressed with the flow visualization	I could immediately see what is going on everywhere	stressed when inflow and outflow changed rapidly
more stressed in the 2nd test	the numbers were not so clear about the water level in the tanks	difficult to see if the volume is increasing or decreasing
more confident in the 2nd test	fixed number of situations; each time you encounter a new one you learn what you have to do	looking at the volume numbers stressed me most
felt more confident in the test with the flow visualization	noticed changes very quickly and easily in the flow visualization	if much more water flows in than out
quite stressed in the 1st test	I felt I had no overview	
equally stressed in both tests	you know that alarms pop out in both tests	stressed when alarm appeared; you always try to avoid this
did not feel stressed at all in both tests	used to deal with these kinds of situations	very unusual to do so much manually; in reality a lot is controlled automatically
frustrated in the trainings scenario without the visualization	had to sit and wait and to watch the volume numbers	had to focus really hard to get to know if the volume was increasing or decreasing

Table 4.10 Interview comments about feelings during the test

The following Table 4.11 summarizes comments about parts of the interface prototype which caused confusions during the tests. The incidents described by the interviewees were mainly individual cases. Nevertheless they are presented in the following in order to provide a basis for further analysis.

Key events/critical incidents	Reasons
confused about the alarms at the valves	did not realize at the beginning that I can open them to 100% without having alarms; confusing liter and percentage
graphic translation of the flow	valve closed but a green area visible, value was zero
pumping water out of the Kondensat tank	never really got the effect I wanted
buffer tank did not tell me how much water flows out	I guessed and watched the volume values.
how fast the tank volume is approaching the alarm levels	it was not so clear where the alarm limits are and how long I can continue until alarms appear
missed alarms at the pumps	looking at the visualization takes away the attention from the alarms at the pumps and valves
at the beginning I thought it is good to have blue water	blue water is very pure and very clean
did not understand the height of the filter	feel that you can just continue to raise the level

Table 4.11 Interview comments about key events or critical incidents during the test

The interviewees were asked which strategy they pursued to accomplish the tasks and which aspects of the interface helped particularly to balance the water level. As can be seen in Table 4.12, the balance units and the water level indicators at the tanks were found to be most useful during the tests with the supplementing flow visualization. In the tests with the schematic diagram only, the focus was mainly on the volume numbers at the tanks and the changing inflow and outflow values. Three test participants used the process values as main source of information in both tests. The following quote illustrates the reason for their behavior: “If you know how to run things, why should you suddenly begin to change your behavior? There is no reason for that.” After they had established a strategy, they kept it until the end of the test sessions without considering alternatives.

Test with supplementing flow visualization	Test with schematic diagram only
could easily adjust the pumps to the balance line	changes of the volume numbers and counted inflow and outflow numbers
balance line was helpful and I liked the volume bars	looked at the volume numbers and checked if the inflow and outflow values were increasing or decreasing
focused a lot on inflow and outflow bars and the line between	volume numbers were most helpful and I looked at inflow and outflow numbers
inflow and outflow numbers at MAVA 1 and 2 were more direct information	increasing or decreasing volume numbers were main source of information
the balance line gave quick idea what is going on	looked very much on the volume bars
balance lines at the tanks helped me a lot	looked at the volume numbers and how fast or slow they are increasing or decreasing
the volume bars showed how fast the volumes are increasing and decreasing	sum up the numbers and subtract
orientation of the balance line; you can directly see if the volume is increasing or decreasing	mainly looked at the volume numbers, compared the inflow and outflow numbers
balance lines and inflow and outflow bars helped a lot	looked at the volume bar; calculated inflow and outflow values and looked at the volume numbers
balance line helped me to see the trend of the flow; was direct information and the best source	

Continued from page 26 - Both tests
focus was on the inflow and outflow values
sum up the values because they were very easy to calculate
looked preliminary at the volume values; also compared inflow and outflow numbers at the tanks

Table 4.12 Parts of the interface prototype which were particularly helpful

Table 4.13 shows different opinions of the participants about proactive working during the tests. The participants who thought that they worked equally proactive in both tests found on the one hand that they have reacted faster in the tests with the additional flow visualization but on the other hand they assumed that quicker responses to changes have not led to better results in their task performance. Several other participants stated that they worked more proactive in the second tests; either they thought this was connected to more practice and a better understanding in the second try or they stated that the flow visualization was decisive for more proactive working.

Proactivity	Reasons
equally proactive	flow visualization made me react faster; easier to see when things were changing; no difference in the results
equally proactive in both tests	probably acted faster in the test with the flow visualization
equally proactive in both tests	probably acted faster with the flow visualization; found it less demanding than looking at the values
more proactive in the 2nd test	proactive working was not connected to flow visualization; felt more confident in the second try
reacted more proactive in the 2nd test	had a better understanding
more proactive in the 2nd test (test with the flow visualization)	saw quicker if the flow was balanced without looking at the numbers
more proactive in the 2nd test (test with the flow visualization)	reacted better and faster; felt easier although alarms appeared as well
reacted more proactive in the test with the flow visualization	acted much more in advance; easier to see where are the problems
reacted faster and more proactive in the 1st test	in the second one the volume level went quite high
more proactive in the test without the flow visualization	paid more attention; at the line you can't see small changes easier with numbers

Table 4.13 Opinions about proactive working during the tests

As can be seen in Table 4.14, the majority of the participants assessed their task performance to be better in the second tests, mainly because of a better overall understanding due to more training. Others mentioned that the visual representation of the processes was decisive since they were able to detect problems faster. Two participants expected their task performance to be better in the tests with the schematic diagram only. One of them found that the interface with the supplementing flow visualization contained too many parameters to look at. The other participant assumed that he missed alarms at pumps and valves because he was too focused on the graphical elements in the upper part of the screen.

Overall performance	**Reasons**
performed better in the 2nd test (with flow visualization)	took a few seconds to look and understand what is going on
2nd one was better	had more training
performed better in the 2nd test	easier to implement my strategy
better in the 2nd test	had more training
better in the 2nd test (with flow visualization)	visual input is very direct
end result better in the 2nd test	got more used to the system
easier to control the flow in the 2nd test	better understood the system and had training
was better in the 2nd test	had more training
was better in the 2nd test	had more control over the water flow
performed a bit better in the 2nd test (with schematic diagram only)	in the 1st test I lost the focus and certainly missed alarms
performed a bit better in the 1st test (with schematic diagram only)	too many parameters in the 2nd one
result is the same in both tests	more satisfying when you had the flow to look at

Table 4.14 Opinions about overall performance

5 Analysis

In this chapter, the research questions for this study are answered individually based on the results of the data logging and post-test questionnaires presented in the previous chapter. Moreover, the comments from the interviews were used to provide information why the participants chose their answers in the questionnaires on the one hand, and on the other and to understand their individual behavior during the tests. This chapter concludes with analyzing the different design elements of the interface prototype discussed in the interviews.

5.1 Research questions

Each research question is stated again and individually addressed based on the results of the data collection methods and according to the analysis plan in chapter three.

Question 1: How did the test subjects feel during the tests?

The ratings in the questionnaire showed that most of the test subjects felt quite or very confident in both tests. A bigger difference could be determined in the ratings about feeling stress during the tests. More respondents felt quite a lot of stress in the tests with the schematic diagram only. It was mentioned in the interviews that observing the constantly changing inflow and outflow values and acting accordingly was a decisive stress factor, mainly because the test participants had no control over these parameters. Some participants stated that they had to be less concentrated in managing challenging situations in the tests with the supplementing flow visualization. The better overview of the running processes made them feel either more confident or less stressed. Most participants, however, believed that the better overall understanding of the prototype in the second tests was the decisive factor that reduced stress and enhanced confidence in completing the tasks.

Plant operators tended to transfer the interface prototype to their working environment and discovered that the processes in the system were more unsteady than in reality. The inconsistent water flow and the absence of automation were considered as (minor) stress factors. Some participants felt stressed when certain actions have not led to the desired effects or if aspects of the system were misinterpreted. In these cases the lack of control and understanding were decisive factors that caused uncertainty, confusion and even frustration. The reported incidents are presented in detail later in this chapter, when the different parts of the interface prototype are analyzed. In the present study it could not be determined to which degree the identified problems diminished the task performance but it is most likely that they had an impact on the level of mental workload.

Question 2: How well did the test subjects understood the interface prototype?

The test participants used the training sessions to discover the scope of functions provided by the interface prototype. They activated pumps, opened and closed valves and watched the effects of their actions. In the interviews, the majority of the participants stated that they acquired a general understanding of the prototype during the training session and started to develop and implement strategies in the beginning or middle of the first test.

The biggest difference between plant operators and students was that the latter required more detailed information and asked questions more frequently during the training sessions. The plant operators needed fewer explanations, most likely because their professional background helped them to understand the fundamental functionality of the interface prototype. Nevertheless, this does not mean that they automatically required less training. Interacting with the prototype helped both, students and plant operators, to learn how to response to the certain events in the simulation. The basic understanding obtained in the training session was further developed each time they encountered new situations and learnt how to deal with them.

The majority of the participants stated in the questionnaires that they understood the processes in both tests almost equally good. However, more participants evaluated the effects of their actions and the current status to be more comprehensible in the test with the supplementing flow visualization. The reasons given in the interviews were that the graphical representation of the water flow provided a better overview over the running processes. Changes could immediately be recognized because the visual feedback to the input was perceived to be direct and unambiguous. According to the interviews, these statements represent the opinion of the majority of the participants.

Question 3: How well did the test subjects carried out the tasks?

Summarizing the data from the log files, it can be determined that less alarms at the tanks, pumps and valves were generated during the tests with the supplementing flow visualization. Moreover, twice as many participants were able to keep the tank volumes within the alarm limits compared to the tests with the schematic diagram only. Additionally, it was found that the students caused in total more alarms and remained longer in alarm state than the plant operators. However, students and plant operators filtered and drained approximately the same amount of water in both tests which implies that they basically worked under the same conditions.

In group A and B, two peak values occurred in the tests that contained the schematic diagram only. In the first tests of group A, the highest number of alarms at the pumps and valves was recorded. Group B, on the other hand, registered the longest alarm duration in the second test caused by alarms at the pumps and valves.

The results from the data logging suggest that the graphical representation of the water flow as additional source of information supported the participants to maintain an optimal operation status for a longer period of time. Taking a closer look at the groups, it is remarkable that peak values of the recorded alarm messages appear only in the tests with the schematic diagram. Comparing students and plant operators, it can be stated that the latter were more successful in keeping the water level within the alarm limits. However, due to the rather small amount of students and plant operators as test participants, the overall performance could significantly be influenced by each individual.

Question 4: How well did the test subjects think they carried out the tasks?

In the interviews and questionnaires, the majority of the test participants (10 out of 12) evaluated their overall performance to be better in the second tests. The main reasons given were more training and experience in interacting with the prototype. However, this self-assessment could only be confirmed by the data logging of the participants in group A. In total, they caused less alarms and remained a shorter period of time in alarm state during the second tests. According to the questionnaire responses, more participants in this group found it easier to understand the current status of the system and the effects of their actions during these tests. By contrast, in group B no improvement regarding the number and duration of alarms could be determined in the second try. The learning effect therefore was particularly overestimated by the participants in the group B. This allows the conclusion that more practice in the interaction with the interface prototype did not automatically lead to an improved task performance during the second tests. On the other hand, it was found, that most tests with an improvement in number and duration of alarms contained the supplementing flow visualization. Thus, it appears that the graphical representation of the processes was the decisive factor for a better task performance in the second tests.

According to the questionnaire responses, eight participants felt they reacted more proactive in the test with the flow visualization. The opinions of the interviewees about the reasons for their proactive behavior in the tests were divided.

In the interviews, three test subjects related a proactive way of working to the learning effect and thus a better understanding of the interface prototype in the second tests. Other three participants stated that the graphical representation of the water flow made them acting more in advance because they had a better overview of the current status. Two participants believed that faster reactions in the tests with the flow visualization was connected to a more proactive behavior, while others felt that they probably reacted quicker, but worked equally proactive in both tests. The latter thought that quicker responses in a matter of seconds did not significantly influence the overall result.

One participant with a different opinion stated that the test with the schematic diagram only caused automatically a more proactive working. He was aware that he had to be more concentrated to follow the continuous changes in the system and to react accordingly. This was the reason for him to evaluate his behavior to be more proactive and the performance to be better in the test with the schematic diagram only. Contrary to his opinion, fewer alarm messages with a shorter duration were registered in the test with the additional flow visualization. Although he made greater efforts in maintaining an effective task performance, he failed in achieving an optimal system status for a longer period of time. It is likely that the mental workload was basically higher during the test with the schematic diagram only and the tasks demanded more attentional resources than he expected.

Question 5: How easy was it to accomplish the tasks?

According to some participants, they actually have not been in really critical situations during both tests. Although yellow alarms occurred at the tanks, no test subject reached the red alarm zone. Moreover, the water flow could be adjusted easily within a few seconds to let alarms at pumps and valves disappear. In the questionnaires, most of the respondents therefore stated that recovering from critical alarm states was equally easy in both tests. However, more participants evaluated the task of keeping the tanks, pumps and valves within the alarm limits to be easier in the second tests. During the interviews, some participants mentioned that especially the balance units helped them to monitor the changes in system and to adjust the pumps and valves accordingly. Others stated that the better understanding of the system in the second tests facilitated completing the tasks.

Question 6: Which strategy was pursued to accomplish the tasks?

The basic strategy to accomplish the tasks was the same for nearly all participants. Based on the prioritization it was intended to keep tanks, pumps and valves within the alarm limits and to filter 100 % of the water. In most situations the bypass valve stayed closed or was occasionally opened while excess water was directed through the drain valve. Only one student did not drain any water at all which resulted in exceptionally long-lasting alarms at the pumps and valves compared to the other participants.

The test participants performed different individual strategies to control the water level during the first and second tests. Based on the comments from the interviews, the following paragraphs give an overview of which kind of information was used for decision-making.

Strategy 1 - test with the additional flow visualization

In the test with the water flow visualization, the majority of the test participants (9 out of 12) found the balance lines most helpful for controlling the water level in the system. The main reason given in the interviews was that the lines were perceived as clear and direct source of information which helped to observe the changing water levels in the tanks. Most of these participants found the way of representing the inflow and outflow to be very intuitive and used them right from the beginning. Besides the orientation of the balance lines, the increasing and decreasing volume bars were considered to be important water level indicators. Two participants additionally obtained more accurate information from the values that displayed the water quantities in the system.

Strategy 2 - test with the schematic diagram only

In the tests without the flow visualization, 8 out of 12 test participants used the volume numbers as main source of information whereby the majority of them additionally looked at the inflow and outflow values in order to adjust pumps and valves accordingly. Only two participants, one student and one plant operator, used the increasing and decreasing volume values as the only information source for decision-making.

Strategy 3 - both tests

Two students and one plan operator (3 out of 12) did not look much at the flow visualization, but rather compared the inflow and outflow values at the tanks. One of these participants watched additionally the volume numbers and the changing rate to get to know how fast the water level was increasing and decreasing. This strategy was basis for decision-making in both tests. Asking for the reason of this practice, all three answered that after once they had developed a strategy they felt captured in their behavior pattern during the entire test session. Regardless of whether the graphical flow visualization was shown in the first or second test, there was no reason for them to consider other information sources than the values.

Two plan operators reported that they used the process values as basis for decision-making in the training sessions but began to pay more attention to the balance units during the tests. After awhile they understood how to use this information and noticed that it was easier for them to react on the basis of the graphical representation than on the process values. The comments from these participants suggest that the mental effort to recognize the trend of the water flow was lower when using the graphical representation of the water flow as main source of information.

In general, it can be said that the professional experience of the plant operators did not lead to a uniform approach for accomplishing the tasks. On the contrary, they developed and implemented different strategies during the tests just like the students. It appeared, however, that the participants tended to stick to their initial strategies which seemed to work for both tests. Those who considered other alternatives and rethought their actions sometimes changed their behavior and positively reflected on it in the interviews.

5.2 Interface prototype

In the interviews, the participants were encouraged to talk about the overall experience with the prototype and the user interface elements in particular. In the following paragraphs, the comments about the different parts of the interface prototype are summarized and analyzed.

5.2.1 Buffer tank

Additional water from the buffer tank was used to a different degree by each participant. The difficulty in adding water from the buffer was that no value indicated the quantity of water supplied to the tanks. Test participants who used the water in the buffer tank mainly guessed and watched the volume values, if they were increasing or decreasing and how fast. In the test with the flow visualization, the graphical representation of the water flow made it possible to see more accurately how much water was added to the tanks. It is likely that some participants deliberately left out the buffer tank in their strategies because the usage was not intelligible for them. Other interviewees reported that they completely forgot about this feature during the test session because they were too focused on other information presented. It is remarkable that none of the plant operators used the buffer tank although additional water would have helped to cope with some of the challenging situations they encountered. However, the plant operators generated fewer alarms in the system and remained a shorter period of time in alarm state, which means that using this feature did not automatically provide an advantage in accomplishing the tasks successfully.

5.2.2 Process values

The rapid changes of the increasing and decreasing volume values at the tanks required a high degree of attention. A test participant, who used these numbers as main source of information in the schematic diagram, reported that watching the continuous changes at the tanks was the main stress factor for him during the tests. In contrast, the flow visualization showed him a better overall picture of the running processes which enabled a quick detection of fluctuations in the water quantities.

One test participant decided to calculate and compare the process values in both tests in order to get an accurate basis for decision making. The system was found to be very simple so that comparing the values during the tests was an easy and simple way for him to balance the water level in the system. Although the flow visualization was considered to be useful to get an overview of the current status, a precise calculation was found to be decisive for adjusting pumps and valves. Another test subject just did not consider other information sources as basis for further actions than the process values. Although the permanent calculations was perceived as an additional mental effort and a considerable stress factor, no other information was found to be equally or even more helpful than the process values. After a short reflection in the interview, however, he realized that it might have helped to look at the balance units as well.

The usage of different units for activating pumps (liters) and opening and closing valves (percentage) raised several questions. A test participant reported that once he opened a valve to 100 %, the displayed value turned yellow. Later on he had never fully opened this valve again until he realized in the second test that he misunderstood the alarm mechanism caused by the different units of percentage and liters.

5.2.3 Water level indicators

The majority of the participants considered the volume bars in the schematic diagram as not being accurate enough to control the constantly changing water levels in the tanks. For most participants, they primarily showed an overview of the tank levels whereas the process values were used as main source of information for adjusting pumps and valves. Additionally, it was reported that these rather small level indicators caused uncertainty because they provided for a couple of participants insufficient information about the remaining time until alarms occur. It is likely that this was one of the reasons that a higher number of alarms and a longer alarm duration were generated during the tests that contained the schematic diagram only.

On the other hand, one test participant found it also hard to identify the critical stages of the much larger level indicators integrated in the water flow visualization. The alarm levels shown as grey-shaded areas were not prominent enough for him. He suggested using red and yellow colors in the alarm scale for a more salient alarm indication. However, there is a danger that a lot of vibrant colors hide the key information so that one may struggle to find the needed information. This is taken into account in the design concept of the screen images since only alarm messages pop out of the background whereas the basic color scheme is rather bland. Difficulties in recognizing the alarm zones could be eliminated by adjusting size and contrast of the grey-shaded areas. It must be noted that the interface prototype was presented to the plant operators on a Laptop with poor contrast and a smaller screen size compared to the monitors used during the tests with the students. It is likely that none of the students found it therefore difficult to recognize the alarm limits on the screen.

5.2.4 Balance units

The balance units, especially the lines between inflow and outflow at the tanks, were highlighted as most valuable source of information by the majority of the test participants. During the interviews, it was frequently mentioned that watching the lines helped to control the water flow in the system because changes at the tanks could easily be detected. Many participants used the balance units and the water level indicator as only information source for adjusting pumps and valves. Some others combined the information from the process values with the graphical representation of the system status. For example, a plant operator considered the balance line to be "the most direct and interesting information". This test participant further noticed that the flow visualization was useful to get an overall picture. However, he obtained more detailed information from the values in the schematic diagram which was basis for further actions.

5.2.5 Filter

The graphic translation of the water flow was perceived to be confusing in the situation where the bypass valve was completely closed. In this case, a green area was still visible, starting from the trend graph and pointing to the figure of the filter (Figure 4.1). One participant remarked that at a first glance he assumed that there is still water running through an alleged closed valve. It became clear that there was no water flow when he checked the values in the trend box and the valve.

Another test participant interpreted the blue color to be an indicator of clean water. However, the opposite is the case, since the amount of water colored in blue flows unfiltered to the feed water tanks as shown in Figure 4.2. This misunderstanding could be cleared up at an early stage during the training scenario and did not appear again in other sessions. Furthermore, the size of the bar that gives an overview of filtered, unfiltered and drained water caused a misinterpretation as well. The height of the bar was designed in such a way that it could display the maximum amount of water in the system. Since it was intended to keep the water flow at an optimal level, the illustration gave the impression that just half of the capacity was used the whole time. This invited one test subject initially to let more water run through the filter until he realized the negative effects in the appearing alarm messages.

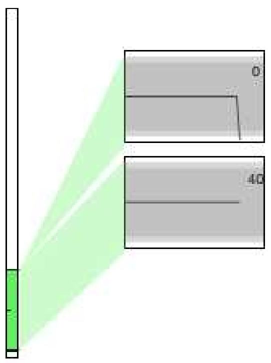


Figure 4.1 Closed bypass valve

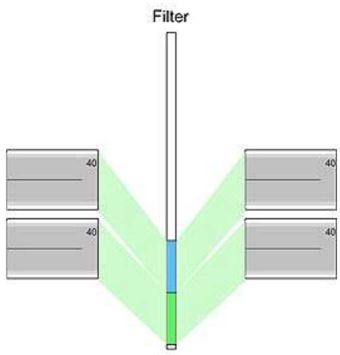


Figure 4.2 Filtered and unfiltered water

5.2.6 Mini trends

All test subjects considered the mini trends as being not helpful during tests. They basically understood that the trend curves displayed the amount of water in the system over a period of time. However, this additional data was ignored because the focus was on other, for the participants more important information. The values in the trend box that turned yellow or red in critical alarm situations were not noticed at all. Since the interaction with the pumps and valves took place in the schematic diagram, the alarms were primarily recognized in this part of the interface prototype.

In the interviews, the students and plant operators represented different opinions about the mini trends. A plant operator did not pay any attention to the trend curves because “it was too much information to look at” and “they were too small to see something”. Another one remarked that “the curves are not direct information which you need the whole time.” He explained that he occasionally uses trend diagrams in his work to get an overall picture of the processes during the last 12 hours or a shorter period of time. It is not surprising that the students on the other hand could not see any usefulness in showing a historical view at all. As a result, two of them suggested to remove the trend boxes completely and to put more emphasis on the visualization of the water flow within the system.

6 Discussion

Within this chapter, the advantages and disadvantages of the data collection methods applied in this project are discussed. Furthermore, issues with the interface prototype are briefly addressed followed by a review of the pilot study and usability evaluation. Finally, the design concept of the screen images is discussed under consideration of the results presented in the previous chapter.

6.1 Methods

The combination of three usability methods, logging the actual use, post-test questionnaires and interviews, enabled to obtain a good overall picture of the user behavior and experience. Based on the data logging, it was possible to review and objectively assess the participants' task performance. It must be noted that there was no right or wrong way to control the water level in the system, neither was a time specified when to actuate the pumps or valves. The task performance refers only to the number and duration of alarms without identifying individual responses to particular events. With this method, it could be determined how successful the participants were to keep the water level within the alarm limits.

The focus in the post-test interviews and questionnaires was rather on the subjective user experience and level of understanding. Subjective data therefore provided explanations for the behavior and a basis for a reasonable interpretation of the objective performance data. In the interviews, the test participants were encouraged to talk freely about their reactions towards the interface prototype. The post-test questionnaires were helpful tool to start a conversation with the participants. However, due to the small amount of participants the questionnaire responses were often spread on the 5-point scales. Therefore, it was difficult to make clear statements about the answers on the rating scales. On the other hand, the results of the multiple choice questions, with only three possible answers, were clear and provided a good basis for further reasoning.

Rosson et al (2001) state that the combination of subjective and objective data collection methods also raises a challenging aspect of usability testing when subjective opinions do not correspond to the performance data. In the present study, nearly half of the participants wrongly assessed their own task performance. The learning effect was particularly overestimated in the second tests which caused this controversial result of objective and subjective data. On the other hand this finding suggests that not only experience and training contributed to a better task performance, but also a better information representation in form of the graphical representation of the water flow.

The attention of most participants was focused on the screen and the tasks were carried out without communicating actively with the moderator. One student and one plant operator were thinking aloud while interacting with the interface prototype. To verbalize the behavior seemed to help them in managing the constantly changing situations. Encouraging other participants to think aloud would probably have helped to receive clues about preference and performance information at the same time. On the other hand, forcing a thinking aloud technique could have been distracted in completing the tasks as well.

During the usability tests, it was optional for the participants to use the graphical representation of the water flow. The flow visualization suggests and highlights what is expected to be useful to the test participants. However, some parts of the interface, designed for improving efficiency, were not recognized or considered to be helpful by some of the participants. Choosing which sources of information, process values or graphics, are used for decision-making was influenced by the individual preference of the test subjects. It can be assumed that the educational or professional background of the participants was reflected in the decision-making process. However, the aim of the present study was not to investigate the reasons for a certain preference. But to control the effect of individual variability, a within-subject test setup was chosen so that each participant was exposed to two different experimental conditions.

Furthermore, a distinction must be drawn between proactive and reactive actions by the test participants and it needs to be clarified what was investigated in the present study. Insulated from normal work distractions, the participants focused only on the interface prototype and the tasks presented. The simulated fluctuations of the water level motivated the participants to intervene in the processes. Under these experimental conditions, the interactions with the interface prototype involved to a great extend reactive actions: waiting for changes and reacting with the aim to adjust to new situations. However, it was not intended to test the responsiveness of the participants to different visual stimuli. The speed of reaction in a matter of seconds did not lead to a significant difference in the results of the two tests, but rather the overall strategy of self-initiated and guided actions supported by the interface design.

6.1.1 Interface prototype

Building the interface prototype at the beginning of this thesis work became a major programming effort itself. It was required to provide enough functionality so that the participants were able to intervene freely within predefined parameters. Moreover, a certain dynamic of the processes and complex interactions within the system had to be simulated in order to create conditions that motivate the test participants to interact with the prototype as much as possible within a short period of time. Therefore, the processes in the system were simulated faster and more unsteady than under realistic conditions in a thermal power plant. Additionally, no automation that facilitates the work was attached to the interface prototype in order to challenge the test participants. The plant operators were confronted with unusual scenarios and the fact that any change in the system had to be done manually and very quickly. The students on the contrary seemed to experience the interface prototype more as a computer game rather than an interface of a real system.

Another issue was that different computer monitors created different opinions about the interface prototype. The Adobe Flash animations were displayed on a laptop and PC with screens of different size and contrast. On the laptop, the content was smaller and the grey-shaded areas were harder to recognize. Both issues were addressed by two plant operators independently of each other. The students did not mentioned any of them, however, it could not be determined that these factors significantly influenced the task performance.

6.1.2 Pilot study

Overall, it can be determined that the small pilot study was very useful to identify potential problems with the interface prototype, the questionnaires and the test procedure. The difficulty in this phase of the project was to find out how much training is required by the test participants in order to address the individual learning ability. During the pilot tests, the training sessions steadily increased from initially 4 minutes to 10 minutes and finally to 14 minutes. On the one hand, the participants should have enough time to get used to the interface prototypes and on the other, the whole usability evaluation should not be extremely lengthy so that the participants may become fatigue. A compromise was found which takes in account both issues and resulted in 14 minutes training at the beginning and a duration of 28 minutes for the test scenarios. Including the time for the introduction, the post-test questionnaires and interviews, the usability evaluation took approximately one hour for each participant. The overall duration was found to be appropriate, but the individual time necessary to achieve the same level of knowledge could not be provided. It must be noted that dealing with variability is not an issue of especially this study, but of all empirical work in Human-Computer Interaction. Lazar et al (2010) state that the major issue of empirical studies is to create exactly the same conditions for each experiment in order to control all potential factors that may influence the validity of the results. In this present study, it was intended to reduce the influential factors as far as possible by placing the participants randomly into two groups. This random assignment is an effective method to distribute variables such as background or general motivation across the experimental conditions (Rosson and Carroll, 2001).

The pilot-tests were also useful to experiment with different test procedures. Completing the training session with the two interface prototypes first, provided in most cases enough time to clear up misunderstandings and to answer unresolved questions right at the beginning. Since a lot of information was given in the introduction, the to-do list as reminder for the tasks was appreciated and occasionally used from all participants during the test session.

Not all suggestions from the pilot tests, especially those concerning the positioning of the sliders, were implemented in the actual study. The decision to place the sliders in the schematic diagram was made because this image was used in both interface prototypes. Moreover, the visualization of the water flow was intended to be an additional source of information in one of the tests without any interactivity attached to it. The sliders became an issue for discussions with some participants who felt that they missed alarm messages at the pumps and valves while they were monitoring the processes in the upper part of the interface. In fact, however, these participants caused less alarms and remained a shorter period of time in alarm state during the tests with the supplementing flow visualization. This implies that the positioning of the sliders in the schematic diagram has not significantly influenced the results as one might assume. During the tests they have proven to be an appropriate tool for the purpose of adjusting the valves and pumps quickly within a finite range.

6.2 Usability evaluation

The test participants identified aspects of the prototype that caused uncertainty in accomplishing the tasks. Although they addressed specific interface elements, the problems found do not fundamentally conflict with the design concepts the mini trends, level indicators and balance units, are based on. For example, the water level indicators at the tanks were considered to be too small and the alarm limits were not clearly recognizable. This problem could simply be resolved by displaying the images on a larger monitoring display with better contrast or by increasing the size of the different elements. The presence of the mini trends was of secondary importance during the tests. Therefore, the usage of these interface elements should be reviewed under conditions where the information presented is necessary and important for the users. Significant problems with the balance units, on the other hand, could not be identified. On the contrary, the processes displayed were considered to be unambiguous and easy to read by the majority of the users. However, the visual feedback when closing a valve could be improved by reconsidering the graphic translation of the water flow. For instance, the green area as representation of the water flow could completely be eliminated in order to prevent the misleading impression that water is still flowing through a closed valve. Furthermore, the graphical representation of the filter needs to be reviewed in detail especially its size and the color scheme of unfiltered water. It is likely, however, that a representation of a more realistic and complex water feed system would demand for other solutions of displaying the processes at the filter units.

The problems identified in the schematic diagram are of importance if this image is used for further research activities. In this case, it is recommended to slow down the change rate of the volume values and to display the amount of water flowing out of the buffer tank. Additionally, for improving the visual clarity of the schematic diagram, the size of the water level indicators and alarm limits at the tanks should be reconsidered. It might be useful to display the schematic diagram and the flow visualization on two different screens in order to reduce the information density and to provide more space for designing and arranging the individual interface elements. However, under realistic conditions most of the issues discussed in the interviews would probably not occur, for example the usage of percentage and liters at the tanks and valves which caused confusions among some of the students. Plant operators are used to deal with much more values that show not only proportionality but also different physical units defining temperature, mass, time, pressure, etc.

6.3 Design concept

In the present study the participants' task performance was assessed in a different ways based on the data logging, post-test questionnaires and interviews. It could be determined how successful the participants were in keeping the water flow within the alarm limits and it was investigated what initiated their actions in the tests.

The results of the data collection methods suggest that the flow visualization guided most of the participants' attention and supported their decision-making to a great extent. This implies that the main purpose of the design concept, to facilitate a quick detection of deviations from the normal state and to support handling of unanticipated situations, was fulfilled. Additionally, it can be stated that the graphical representation of the processes in the system caused a better task performance and thus supported a more proactive way of working.

It can be concluded that an interface design based on the developed design concept can help to balance the mental workload by making the processes within the system more visible. According to Stanton et al (2010), mental workload plays an important role in designing human-machine interfaces for control rooms of power plants. Inappropriate workload levels, caused either by excessive or low task demands, diminish the user's performance. On the one hand, operators become mentally underloaded if automation reduces the situational awareness and consequently slows down the operators' responsiveness. On the other hand, operators experience mental overload if the given tasks exceed their attentional capacity. Good interface design is therefore a "key element in the safety, reliability and efficiency of complex systems" (Stanton et al, 2010).

6.4 Future work

This thesis report concludes with two suggestions for further research. At first, the parts of the interface prototype which caused confusions among the test participants need to be revised. Secondly, it would be interesting to analyze the user's behavior in a more realistic setting with a realistic imitation of the controls. A simulator that represents processes of a shift period would allow to study user behavior and experience in longer experimental sessions with extended training periods. As a first step, some of the design proposals could also be implemented in the control systems to evaluate them under real operational conditions in control rooms of thermal power plants. The usage on a daily basis would be subject for follow-up studies with plant operators and engineers.

7 Conclusion

In this study, research questions were formulated to evaluate the screen images of the feed water system developed by the Human-Machine System Research Group at Chalmers University of Technology. The aim was to investigate if the graphical representation of the system's status facilitates a fast detection of deviations from the normal state and supports handling of unanticipated situations. The intended proactive behavior involves to think ahead and to initiate changes in the system before alarms appear.

Overall, it can be confirmed that the flow visualization changed the behavior of the majority of the participant so that less alarms were generated. More specifically, it was found that not only the learning effect contributed to a better task performance, but also the graphical elements, designed to support a more effective decision-making. The test participants identified the balance units and level indicators as valuable sources of information for controlling the water level. Furthermore, the statistics confirmed that the participants caused less alarm messages and remained a shorter time in alarm state during the tests with the supplementing flow visualization. These findings suggest that the graphical representation of the system's processes guided most of the participants' attention and choices and thus supported a more proactive behavior.

The complex working environment of plant operators and engineers in control rooms requires an intuitive screen design that provides an effective visualization of information. With increasing automation, the operators need to perceive and manipulate large amounts of data under constraints such as time and cognitive workload (Zhang, 2008). This raises new challenges for human factors engineers which makes further research on improving the visualization of information in the process industry necessary and important.

Acknowledgements

A number of people have generously given time, advice and valuable information during this master's thesis. I would like to thank the people who supported me in my work:

I am thankful to my supervisors, Lars-Ola Bligård and Jonas Andersson, for the guidance and provided assistance during this project.

Many thanks to Johan Bengtsson, service technician at Sävenäs Thermal Energy Plant, and Per Nordin, plant engineer at Fortum in Stockholm, for organizing the test sessions with their colleagues.

Finally, this thesis would not have been possible without all the students, plant operators and engineers who participated in pre-tests and the usability study. Thank you all for your contribution and valuable feedback.

References

Ahlén, F. (2009), *Control Room User Interfaces – proposals for screen images used for visualization of plant processes*, Chalmers University of Technology, Gothenburg

Bligård, L.-O., Andersson, J., Thunberg, A., Osvalder, A.-L. (2009), *Graphical visualization of process status for thermal power plants*, Proceeding HFES-EC Annual Meeting

Courage, C., Baxter, K. (2005), *Understanding Your Users: A Practical Guide to User Requirements Methods, Tools, and Techniques*, 1st edition, San Francisco: Morgan Kaufmann

Dekker, S. (2004), *To Engineer is to err* – Chapter 7 from Sandom, C. & Harvey, R. (2004), Human Factors for Engineers, London: The Institution of Electrical Engineers

Ghaoui, C., (2006), *Encyclopedia of human computer interaction*, Hershey Pa.: Idea Group Reference

Hollnagel, S. (2004), *Automation and human work* – Chapter 6 from Sandom, C. & Harvey, R. (2004), Human Factors for Engineers, London: The Institution of Electrical Engineers

Lazar, J., Feng, J. H., Hochheiser, H. (2010), *Research Methods in Human-Computer Interaction*, John Wiley & Sons Ltd

Lindsley, D. (2000), *Power-plant control and instrumentation: the control of boilers and HRSG systems*, London: Institution of Electrical Engineers

Nielsen, J. (1993), *Usability engineering*, Boston: Academic Press

Noyes, J., Bransby, M. (2001), *People in control: human factors in control room design*, London: Institution of Electrical Engineers

Rosson, M.B., Carroll, J.M. (2001), *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*, 1st edition, San Francisco: Morgan Kaufmann

Rubin, J., Chisnell, D. (2008), *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*, 2nd edition, Indianapolis: Wiley Publishing, Inc.

Salovaara, A., Oulasvirta, A. (2004), *Six Modes of Proactive Resource Management: A User-Centric Typology for Proactive Behaviors*, [Online], Available: http://portal.acm.org, [23 July 2010]

Stanton, N. A., Salmon, P., Jenkins, D., Walker, G. (2010), *Human factors in the design and evaluation of central control room operations*, Boca Raton: CRC Press

Tullis, T, Albert, B. (2008), *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*, San Francisco: Morgan Kaufmann

Whitt, Michael (2004), *Successful Instrumentation and Control Systems Design*, Durhan, NC: ISA

Woodruff, E. B., Lammers, T. F. (2005), *Steam plant operation*, New York: McGraw-Hill

Zhang, P. (2008), *Industrial control technology: a handbook for engineers and researchers*, Norwich, NY: William Andrew

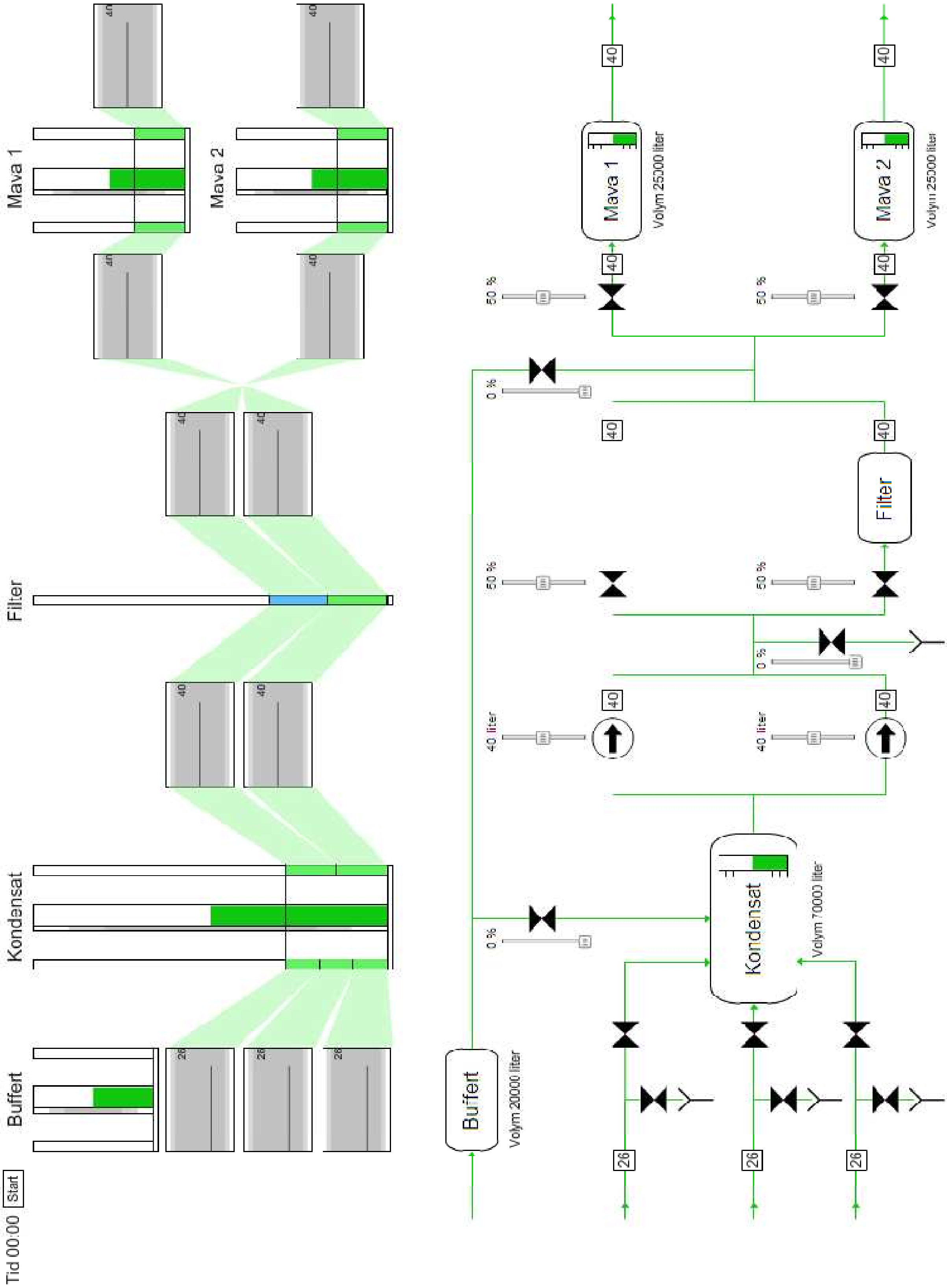
Welch, R., Braseth, A. O., Nihlwing, C., Skraaning Jr., G., Teigen, A., Veland, Ø. (2007), *The 2005 Ecological Interface Design Process and the Resulting Displays*, OECD Halden Reactor Project, Halden

Welch, R., Friberg, M., Nystad, E., Teigen, A., Veland, Ø. (2005), *EID - Prototype Design & User test 2004*, OECD Halden Reactor Project, Halden

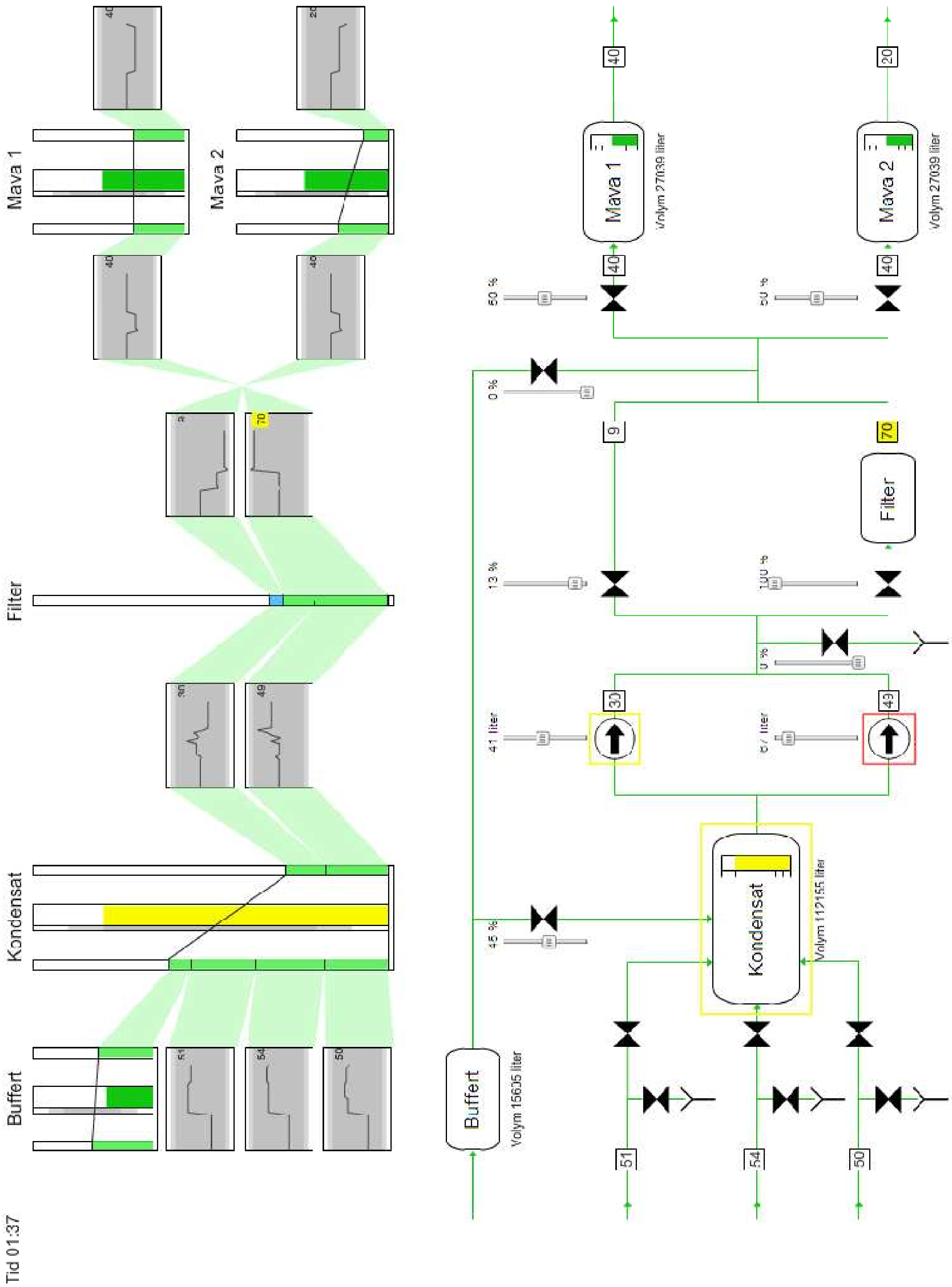
Web resource

Mölndal Energi, [Online], Available: http://www.molndalenergi.se/, [23 July 2010]

Appendix A – Interface prototype, normal state



Appendix B – Interface prototype, alarm state



Appendix C – Post-test questionnaire

Feelings during the test

1. I felt confident during the test.

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very much

2. I felt stressed during the test.

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very much

3. I am satisfied how I completed the tasks.

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very much

Understanding of the processes

4. How easy was it to understand the processes in the system?

	1	2	3	4	5	
Difficult	○	○	○	○	○	Easy

5. How easy was it to understand the system's current status?

	1	2	3	4	5	
Difficult	○	○	○	○	○	Easy

6. How easy was it to understand what the effects will be in the system?

	1	2	3	4	5	
Difficult	○	○	○	○	○	Easy

Self-assessment of the task performance

7. How easy was it to keep the volumes in the tanks within the alarm limits?

	1	2	3	4	5	
Difficult	○	○	○	○	○	Easy

8. How easy was it to keep the pumps and valves within the alarm limits?

	1	2	3	4	5	
Difficult	○	○	○	○	○	Easy

9. How easy was it to recover from critical alarm states?

	1	2	3	4	5	
Difficult	◯	◯	◯	◯	◯	Easy

10. How easy was it to develop a strategy to accomplish the tasks?

	1	2	3	4	5	
Difficult	◯	◯	◯	◯	◯	Easy

Preference questions (additional questions after the second test)

11. In which of the two simulations was the overall experience best?

☐ Simulation **without** flow visualization
☐ Simulation **with** flow visualization
☐ Equally good in both simulations

12. In which of the two simulations to you think was your task performance better?

☐ Simulation **without** flow visualization
☐ Simulation **with** flow visualization
☐ Equally good in both simulations

13. Which of the two simulations provided the best overview about the running processes?

☐ Simulation **without** flow visualization
☐ Simulation **with** flow visualization
☐ Equally good in both simulations

14. In which of the two simulations did you better understand the effects of your input?

☐ Simulation **without** flow visualization
☐ Simulation **with** flow visualization
☐ Equally good in both simulations

15. In which of the two simulations did you have better control over the running processes?

☐ Simulation **without** flow visualization
☐ Simulation **with** flow visualization
☐ Equally good in both simulations

16. In which of the two simulations do you think you worked more proactive (acted in advance before alarms appeared, rather than just reacted)?

☐ Simulation **without** flow visualization
☐ Simulation **with** flow visualization
☐ Equally good in both simulations