Getting the Feeling

“Human Error” in an educational ship-handling simulator

Marcus Arenius
23.10.2009

Department of Computer Science
Linköpings Universitet

Tutor: Oliver Sträter, Universität Kassel
Preface

This thesis could not have been written without the supportive and open-minded personnel at the simulator facility of Fachhochschule Oldenburg/Ostfriesland/Wilhelmshaven (FH OOW) in Elsfleth. Although several persons contributed to making this thesis an exceptional experience, I would first of all like to thank the instructors, Prof. Christoph Wand, Mr. Wilbertz and Mr. Sievers, for providing most of the much needed information and support. Mr. Damm provided excellent help in addressing several of the technical problems associated with the data recording procedures and also shed light on organizational issues. Additionally, Mr. Birnschein resolved many issues concerning the programming of the simulator and also provided a unique insight into the practical issues of seafaring, which I am very grateful for.

Several of the ideas and approaches present in this study took shape during discussions with my tutors Prof. Oliver Sträter and Georgios Athanassiou, at the Fachgebiet für Arbeits- und Organisationspsychologie at the University of Kassel. Their commitment to my thesis proved to be a valuable asset for shaping it into its present form. Thank you.

I would also like to thank all of the students that participated in the study. You provided a very insightful glance into the challenges associated with becoming a mariner. This thesis would not have existed without your participation.

Linköping, October 2009

Marcus Arenius
**Abstract**

In high-risk environments of seafaring, simulators constitute a widely used tool in preparing nautical students for the challenges to be met in real-life working situations. While the technical development of ship bridge simulators continues at a breathtaking pace, little is known on how developments fulfil their intended safety critical purpose during actual simulator training exercises.

In order to investigate this, a mixed-methods quasi-experimental field study (N = 6) was conducted aiming at discerning the systemic causes behind committed human errors and to what extent these causes can be related to the technical layout of the simulator in general and a decision supporting display in particular. The nautical students' performance in terms of committed errors was analysed when the decision supporting display was either inactive or active during two different exercise batches. Drawing upon eye tracking evaluation, interviews and simulator video recordings, systemic causes leading to human errors were identified. Results indicate that all errors occur under the same kind of (stressful) interaction. Based on this design requirements aiming at promoting resilient crew behaviour were proposed.
Table of Content

1. Introduction ........................................................................................................................ 12

2. Theoretic Background ......................................................................................................... 13
   2.1. Two views of human error ............................................................................................ 13
   2.2. Human error in the maritime domain .......................................................................... 13
   2.3. Simulator training .......................................................................................................... 13
   2.4. Conning Display ............................................................................................................ 14
   2.5. Consequences of automation for the maritime domain ................................................ 14
   2.6. Design philosophy: Agile Methods ................................................................................ 14
   2.7. An integrated approach to system safety ....................................................................... 15
   2.8. Cognitive Mill ............................................................................................................... 16
   2.9. Cognitive Dissonance .................................................................................................... 16
   2.10. Cognitive Couplings ................................................................................................. 18
       2.10.1. Type of involvement: Involved versus isolated ...................................................... 18
       2.10.2. Type of task: Active versus monitoring ................................................................. 19
       2.10.3. Type of control: closed loop versus open loop ...................................................... 19
       2.10.4. Number of dimensions: Multidimensional versus one-dimensional ....................... 20
       2.10.5. Necessary operation: Simultaneous versus sequential processing ......................... 20
       2.10.6. Type of presentation: Compensatory versus pursuit ........................................... 21
       2.10.7. Primary Compatibility: compatibility versus incompatibility ................................. 21
   2.11. Resolving mechanisms ............................................................................................... 22
       2.11.1. Fixation ................................................................................................................. 23
       2.11.2. Information ignorance or reduction ...................................................................... 24
       2.11.3. Goal reduction ...................................................................................................... 24
       2.11.4. Goal and information overload ............................................................................ 24
       2.11.5. General remarks ................................................................................................ 24
   2.12. Complexity ................................................................................................................. 24
   2.13. Triangulation ............................................................................................................. 26

3. Purpose .................................................................................................................................. 27
   3.1. Clarifications and restrictions ...................................................................................... 27

4. Method: General .................................................................................................................. 29

5. Methods: Quantitative study ............................................................................................ 30
   5.1. Participants .................................................................................................................... 30
   5.2. Apparatus .................................................................................................................... 30
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.1.</td>
<td>The nautical simulator: An overview</td>
<td>30</td>
</tr>
<tr>
<td>5.2.2.</td>
<td>The instructor station</td>
<td>31</td>
</tr>
<tr>
<td>5.2.3.</td>
<td>The instructor</td>
<td>32</td>
</tr>
<tr>
<td>5.2.4.</td>
<td>Scenario batches</td>
<td>32</td>
</tr>
<tr>
<td>5.2.5.</td>
<td>Bridge 1 “Weser” and its components</td>
<td>37</td>
</tr>
<tr>
<td>5.2.6.</td>
<td>Dikablis eye-tracking tracking system</td>
<td>39</td>
</tr>
<tr>
<td>5.2.7.</td>
<td>Screen capture software</td>
<td>40</td>
</tr>
<tr>
<td>5.2.8.</td>
<td>Screen grabber</td>
<td>40</td>
</tr>
<tr>
<td>5.2.9.</td>
<td>Adobe Premiere Pro</td>
<td>40</td>
</tr>
<tr>
<td>5.2.10.</td>
<td>Dictaphone</td>
<td>40</td>
</tr>
<tr>
<td>5.3.</td>
<td>Design</td>
<td>40</td>
</tr>
<tr>
<td>5.4.</td>
<td>Procedure</td>
<td>42</td>
</tr>
<tr>
<td>5.4.1.</td>
<td>Preparation</td>
<td>42</td>
</tr>
<tr>
<td>5.4.2.</td>
<td>Start of the exercise</td>
<td>43</td>
</tr>
<tr>
<td>5.4.3.</td>
<td>Debriefing</td>
<td>43</td>
</tr>
<tr>
<td>5.4.4.</td>
<td>Start of the next scenario</td>
<td>44</td>
</tr>
<tr>
<td>5.4.5.</td>
<td>Watchkeeping</td>
<td>44</td>
</tr>
<tr>
<td>5.4.6.</td>
<td>Interviews</td>
<td>44</td>
</tr>
<tr>
<td>5.4.7.</td>
<td>Deviations from regular procedure</td>
<td>44</td>
</tr>
<tr>
<td>6.</td>
<td>Qualitative study</td>
<td>46</td>
</tr>
<tr>
<td>6.1.</td>
<td>Interviews</td>
<td>46</td>
</tr>
<tr>
<td>6.2.</td>
<td>Observations</td>
<td>46</td>
</tr>
<tr>
<td>6.2.1.</td>
<td>Addressing the keyhole effect of video observations</td>
<td>47</td>
</tr>
<tr>
<td>6.3.</td>
<td>Data analysis</td>
<td>48</td>
</tr>
<tr>
<td>7.</td>
<td>Quantitative results</td>
<td>49</td>
</tr>
<tr>
<td>7.1.</td>
<td>Descriptive statistics: within-group design</td>
<td>49</td>
</tr>
<tr>
<td>7.1.1.</td>
<td>From “off” to “on”</td>
<td>49</td>
</tr>
<tr>
<td>7.1.2.</td>
<td>From “on” to “off”</td>
<td>49</td>
</tr>
<tr>
<td>7.2.</td>
<td>Descriptive statistics: between-group design</td>
<td>50</td>
</tr>
<tr>
<td>7.3.</td>
<td>Quantitative method reflections</td>
<td>51</td>
</tr>
<tr>
<td>8.</td>
<td>Qualitative Results</td>
<td>54</td>
</tr>
<tr>
<td>8.1.</td>
<td>Interviews</td>
<td>54</td>
</tr>
<tr>
<td>8.1.1.</td>
<td>Maneuvering</td>
<td>54</td>
</tr>
</tbody>
</table>
1. Introduction

Work has become increasingly complex. This development can mainly be attributed to 4 forces; mechanization, automation, centralization of process control and computerization (Hollnagel & Woods 2005). When visiting a modern ship’s bridge e.g. the effects of these forces become clearly visible. Through the manipulation of computer interfaces and handling of automatic mechanisms, the human has gained the ability to control large systems without much physical effort. And while the physical effort has been reduced, the mental demand of controlling all these systems and subsystems has increased equally (Hollnagel & Woods 2005).

This carries certain implications for high-risk environments like seafaring where simulators constitute a widely used tool in preparing nautical students for these complex challenges. While the technical development of ship bridge simulators continues at a breathtaking pace (much like their “real-world” equivalents), little is known on how and if these components fulfil their intended safety critical purpose during actual use. This is especially important since the dependence on technology in complex environments can lead to mistakes resulting in severe consequences (Bainbridge, 1983) for both humans and equipment in the high-risk real-world settings that simulator training is applied to.

It is thus important to examine if the technical equipment in simulators is adjusted to the actual behaviour of the students using them.
2. Theoretic Background

In this study, safety is defined as the absence of human errors. Therefore what I mean with error or more specifically human error will be defined in the following sections and set into context with the other, central concepts for this study.

2.1. Two views of human error

70 to 90% of accident in a variety of domains can in one way or another be attributed to human performance failure (Hollnagel & Woods 2005). The concept of these “human errors” has been under considerable debate. According to Dekker (2006) they are part of what he calls the old view or the bad apple theory. This view states, that it’s very easy to attribute it to the human when an error occurs, as she is the only stakeholder in the system, thus making her the obvious choice to put blame on. What is often disregarded is, that, although the human always acts on the sharp end of a system (in direct contact with the safety-critical process), she is always influenced by the blunt end, the organisation the operator is embedded in, as well (Reason, 1990). This is an important aspect that has to be considered when analysing human performance. Thus, the view on human performance is shifted from something residing within the individual to something that is influenced by a system of several interconnected factors (Hollnagel & Woods 2005). Within this systemic view, the unit of analysis has moved from the individual layer to a system of several individuals/teams and their corresponding technology. Or to put it in other words:

“Human error is not the conclusion of an investigation. It is the starting point.” (Dekker, 2006, P.15)

2.2. Human error in the maritime domain

According to a report published by the IMO in 1994, 75% of all ship accidents worldwide can be attributed to human error. Attempting to combat this issue the International Maritime Organisation (IMO) released a revised version of the Standards of Training, Certification and Watchkeeping (STCW) which requires seafarers to have certain competencies and minimum training standards. The effect of this regulatory body is still debated (Grech, Horberry & Koester, 2008), thus the question of the nature of maritime training is still unresolved and of safety-critical importance.

2.3. Simulator training

The use of virtual environment for training purposes has increased substantially over the last few years. This is foremost due to the combination of the need to practice learned models and theories in a setting similar to real life settings and the increase in computational capabilities (Hettinger & Haas, 2003).

Using a simulator for training purposes is a way to address training restrictions in real-world environments, as practical limitations like risk, cost complexity and the lack of control often reduce applicability of real-world training. Secondly, simulation can enhance certain critical cues that would otherwise not be visible, e.g. by highlighting a part of the ship that is about to fail, thus making it easier to focus training on critical elements. Thirdly, the sheer plethora of different ship bridges and environments could simply overwhelm the students. Therefore a more minimalist approach is feasible, especially since this scaled down learning can lead to better performance on tasks requiring more tools later on. (Hettinger & Haas, 2003)
The ship-handling simulator in this study is of scaled down nature while still leaving room for customization, as the exact layout of some of the components in the maritime simulators are not predefined by national or international standards more than on a rudimentary basis. Therefore, the operators of this maritime simulator facility strive to make the layout of the technical simulator components fit the needs of the nautical students as closely as possible. This is especially true for new technical aids, which leads us to the main object of interest of this study, the Conning (Information) Display that is in the process of being shaped in order to meet student requirements.

2.4. Conning Display
The word conning refers to “the place in the wheelhouse with commanding vision and which is used by operators when monitoring and directing the ship’s movements” (ISO 8468, 2007, P.4). A display is the means “by which a device presents visual information to the operator, including conventional instrumentation” (ISO 8468, 2007, P.6). The Conning Display may be distributed in workstations for navigation, manoeuvring, monitoring, secondary navigation and docking (ISO 8468, 2007). These are core tasks for the ship crew in general, which means that a lot of information may be integrated into the display. This in turn raises the question of what information that is to be considered relevant in order to improve the training capabilities of the simulator exercises (Hettinger & Haas, 2003).

2.5. Consequences of automation for the maritime domain
The role of the human on the real-world ship’s bridge is shifted from being mainly active to monitoring of automated execution of commands carried out by the technical system (Gauss & Kersandt, 2005). As simulators aim at representing these real-world bridges, they could also be seen as being affected by this change. This in turn would make the simulator environment equally exposed to the irony of automation (Bainbridge, 1983). This concept highlights the irony, that the more advanced an automated control system is, the more crucial the role of the human monitoring the system becomes (which in the case of a ship is its bridge crew). The human is thus not being removed from the system by the automatic system it is rather the case that her role changes. This fact is highlighted in the so-called automation surprises, when the automated system fails or malfunctions and the human has to take over control in order to bring the system back into safe operating limits (Woods & Sarter, 2000). It is therefore of utmost importance for the person monitoring a system to have the skills necessary for performing a task if the without the automated system when it fails.

The criteria for good training could be seen as having changed as well, as the students basically have to have the skill necessary to handle automation surprises and this carries important implications for the design of the Conning Display. The display should not lead the crew into become too dependent on its functioning. This should be taken into account when designing a new technical device.

2.6. Design philosophy: Agile Methods

The aim of this study is not to provide a once and for all design layout of the Conning Display as the exact impact of the proposed design is not known in before and one thus cannot be sure if the layout is correct (or perhaps leads to automation surprises as stated before). This problem is termed the envision world problem (Woods, 2000) and revolves around the notion that we cannot know how the design of an artefact will affect the people using it no matter how much data we collect on the users.
This is due to the fact that introduction of the artefact itself will change the behaviour of the people using it, often in unintended ways (Woods, 2000).

In other words the effect the artefact has on the domain it has been designed for can only be known when the interaction of the users with the new system has been studied extensively. Thus design must be regarded as an iterative process encompassing constant validation of proposed artefact designs against their real-world domain of application. This is especially true for complex domains where errors due to the artefact not being properly adjusted to the needs of the users, can lead to fatal consequences both in terms of equipment and lives (Hollnagel, 2005). Design is thus always a hypothesis about reality that constantly has to be tested in order to prove its validity (Woods, 2000).

This view of design as being an iterative process is a cornerstone of the so called agile methods that are deployed in software development (Dybå & Dingsøyr, 2009). Here, design shifts from being one-time event at some point in product development, to becoming a constant product-shaping force throughout the project, evaluating it by means of real-world challenges. This perspective has gained considerable influence lately (Microsoft programmed the new windows 7 with this technique, Dworschak, 2009).

While design in this study is viewed as an iterative process, this process has to start somewhere. Therefore this study will exclusively aim at producing a design proposition that tackles safety-critical behaviour (Benyon, Turner & Turner, 2005) as this behaviour could result in serious accidents in real-life setting on real ships. Other design issues that do not emerge as safety-critical in this study can be addressed in future development phases, “look and feel” e.g..

In order integrate safety critical behaviour or “human errors” into this iterative design process, a theoretical framework is needed that includes the basic science philosophy presented, the notion that human errors only can be understood from a systemic point of view.

2.7. An integrated approach to system safety

Human performance has to be seen in the relation to underlying systemic factors. Errors only occur if both the human and the working environment (both technical and organisational factors) fail in preventing it, thus errors occur if the man machine system fails in performing (Sträter, 2005). So the unit of analysis is moved from being either the human or the technical component to analysing the interplay between these two parts of a coupled system (Sträter, 2005). This interaction shows properties that are neither fully technical nor completely cognitive but systemic (Hutchins, 1995).

In the following subsections a systemic approach to the assessment of human cognition will be presented. The overarching theory stems from Sträter (2005). First, the cognitive mill will be presented, a concept aiming at explaining how the internal world of and individual and the external world interact. Secondly, the concept of dissonance will be introduced and its driving force in cognition will be illustrated. Thirdly, two of the central concepts for this study will be presented, the cognitive couplings and the resolving mechanisms. The cognitive couplings are used to classify the interaction between internal and external world while the resolving mechanisms describe how the individual mind reacts on this interaction. All of the text is based on Sträter (2005) unless referenced otherwise.
2.8. Cognitive Mill

The division of mind and matter and its implication for thought has fascinated researchers for centuries. One of the principal questions to ask is how these two interact. How does the external world enter our mind and in turn become part of our cognition?

Popper (1997) discerned three worlds, the physical world, the internal world and the conscious world. The conscious world is implied by the physical and the internal world. The external world consists of a set of objectively given information, everything outside the cognitive system in other words, the individual mind. Objectively given information may cue the cognitive system into certain processing actions (Sanders, 1975). The internal system on the other hand refers to the mind part in the Cartesian division. It is the set of subjective or internally given information including all the information that is internal to the cognitive system and is sometimes referenced to as mental model (Gentner & Stevens, 1983).

The constant alignment of the internal and external world as the brain processes and integrates the stimuli from the external world is what Neisser (1976) metaphorically calls the cognitive mill. It reflects the constant and ongoing process of brain activity that shapes the way in which we think. The internal world and objectively given external world are continuously influencing each other via perceptions on the one hand and actions on the other hand. This could be compared to a watermill in that sense, the water flows onto the mill wheel and thus keeps it running. The same holds true with the cognitive processing of the individual. This would not take place if no external stimuli would ever touch it.

2.9. Cognitive Dissonance

But this mill does not in any way turn without variations as the flow of information coming from the external world varies greatly depending on the external circumstances. Additionally, the cognitive system also processes this external information entering the system differently depending on what state it is in (which in turn can depend on goals and wishes). The integration of the external information in the internal world is not thus unproblematic, mismatches or misalignments between the two components can occur. These mismatches are what Festinger (1957) relates to the occurrence of dissonance. Dissonance is created when learned behaviour and the objective situation at hand cannot be matched. This resembles the way in which Hollnagel (2005) uses the term Law of Requisite Variety (Ashby, 1958), in the sense that the variety as presented by the external world cannot be matched by the variety of the cognitive system. When dissonance occurs, the cognitive system naturally strives to establish an equilibrium state again, where the internal and the external world are nicely aligned to each other again. The degree or the amount of cognitive dissonance can lead to respectively different behaviour. A small amount of dissonance can for example result from the unintended dropping of a fork from the dinner table. The hand immediately grabs after the fork to catch it before it reaches the ground. This slip is directly corrected without much cognitive effort therefore the immediate cognitive selection of actions is located on a more subconscious level. Cognitive operation in an equilibrium state also lies on a more subconscious level, for example when executing skill-based behaviour. More conscious processing occurs when some form of abnormality is detected during the coupling of internal and external world.
As stressed before, the alignment of internal and external world is a continuous process. It can be seen as a constantly flowing processing loop where the external world influences the cognitive system via stimuli and the cognitive system in turn reacts to the stimulus with actions resulting from an internal evaluation process.

Three cognitive acts are of central importance for this processing loop: cognitive coupling, cognitive binding and cognitive levelling. A cognitive act describes mental processes involved before executing an action. The cognitive coupling describes how the external world enters the cognitive system, e.g. how many parameters an operator has to monitor. This is the objective mental demand that is placed on an individual from a given situation. The cognitive binding is related to the strategies individuals use in order to cope with the mental demand that is put on him through a mismatch in the cognitive coupling, in other words how the individual binds it to previous experience. The cognitive levelling on the other hand concerns how action selection results from the individual coping strategies used.

For example during a monitoring task an operator has to watch several instruments in parallel and compare them mentally. He detects that one of the instrument shows information that is not congruent with another instrument and experiences dissonance (cognitive coupling). As he knows that this anomaly occurs due to preferences in another display, he does not perceive that the anomaly is particularly demanding for him (cognitive binding, comparing with previous experiences). Consequently the resolution of the dissonance is done quickly by cognitive levelling, an appropriate response pops into the operators mind. In this way an equilibrium state is established in the cognitive system and an action, namely the modification of the settings, is immediately executed without much delay.

This expert performance can be contrasted with the performance of a person that is new to the problem and has not experienced this problem before. The sequence begins, as always, with the cognitive coupling. The person discerns the mismatch between the parameters in the external world and dissonance sets in as this was not something that he expected (cognitive coupling). He does not immediately know how to properly react to the anomaly and therefore perceives the situation as being highly demanding (cognitive binding). To resolve the dissonance, several cycles of levelling are necessary. In this case the operator mentally goes through several procedures to deduce a solution to the problem. This takes a lot of time which means that a large amount of dissonance must be reduced in order to achieve an equilibrium state. While the operator's system is busy with overcoming the dissonance resulting from the cognitive coupling, the technical system which the operator should control, breaches safety limits and has to be shut down.

The individual tries to reduce dissonance that arises during a work by the use of previous experience. This experience is related to memory. The memory can be seen as divided into two layers, the concept layer where objects attributes and options for action are located and the experience layer where the interrelations between the concepts are represented. Usually people that work at the same working place share the same concepts (e.g. from the educational background). But these concepts can have different relations depending on what experiences the individuals have made with them during work. It could be that some persons e.g. usually don't follow certain procedures because they find them cumbersome and inefficient, while others do follow the procedures. So here the
individuals clearly have different experiences with the concept “procedure obedience” in their working context, perhaps due to different working circumstances, like type of shift for example.

What that means in a more general sense is that one cannot expect people to behave in the same manner if they are subjected to the same stimulus (e.g. a procedure). Every person carries certain experiences with him/her that makes him/her more inclined to act in a certain way that is coherent with these experiences, a natural consequence of the constructivist nature of cognition. Therefore a certain task can always lead to some kind of action “j”. But one can never be sure if this action is coherent with the task “i” as it was demanded to be done by e.g. the supervisor. In fact, the person could from experience deem that an action that violates the task “i” would be the best response in that case. Or, in other words, no coherent relationship between sensory information and actions can be postulated, making it hard to predict if the measures introduced in order to improve safety will have any it’s intended effect (Hofstätter, 1973).

To sum it up this model promotes the view of human error as resulting from the natural process of the cognitive mill to align external world and the cognitive system. The external and internal worlds are compared and an alignment between the two worlds is established. This alignment is accomplished through the different cognitive acts; the cognitive coupling, the cognitive binding and the cognitive levelling which finally lead to action execution.

2.10. Cognitive Couplings

When referring to cognitive couplings (Sträter & Bubb, 2003) it should be made clear that the concept neither refers to anything exclusively cognitive nor anything external to the cognitive system. Cognitive couplings classify the principle ways in which the internal and the external world are coupled. Rohmert and Rutenfranz (1975) as cited in Ulrich (2001), propose the concepts stress (Belastung) and strain (Beanspruchung) for explaining this relationship between internal and external worlds. Stress in this context can be seen as something that is “an objective and measurable force affecting the human from the external world ("von außen")” (Ulrich, 2001, P.437). Strain on the other hand is the effect of stress “in the human and on the human” (Ulrich, 2001, P.437). Therefore stress can be seen as affect factor (Einwirkungsgröße) and strain as effect factor on human cognition (Ulrich, 2001, P.437). A reason for this division is that there is no causality between high stress and errors. Some people can cope with high stress e.g. as they are used to it from former, similar jobs while there, on the other hand, may be people that cannot cope with the high stress as they lack the experience for it and consequently produce errors. The two concepts cognitive couplings and resolving mechanisms (Sträter & Bubb, 2003) are associated with the stress and strain and thus always precede performance.

Cognitive Couplings revolve around the concept of stress, that is, what the nature of external stimuli that enter the cognitive system is. There are several different cognitive coupling types that classify how the task context puts stress on the individual (Sträter, 2005). The couplings are not mutually exclusive and every cognitive coupling is divided further into a stressing mode and a mode that is comparatively less stressing for the individual(s) engaged in interaction. These modes, however, are mutually exclusive in the sense that a person can not be in two modes of the same coupling type at the same point in time.

2.10.1. Type of involvement: Involved versus isolated
The types of involvement can be loosely associated with Piaget's basic constructivist constructs of accommodation and assimilation of cognitive schematas (Piaget, 1947). It has to be ruled out if the cognitive system is interacting with the technical system at all. "Involved" means that the cognitive system, an individual, takes in external stimuli from the working context and reacts on them, in other words that it constantly assimilates the information into a mental model of the situation. Being isolated from the system is per se not bad as benefits can stem from retreating from the work place and letting the experiences made settle (Sträter, 2005). There are e.g. records of power plant maintenance personnel that remembered errors during work on their way home from work. The information experienced during work had time to be integrated into the cognitive system and its safety critical significance unveiled (Sträter, 2005). The individual must be in an involved interaction mode in order to be in any of the other couplings. It is more stressful for the individual to be in an involved cognitive coupling mode than in an isolated mode.

2.10.2. Type of task: Active versus monitoring

A person standing in an active cognitive coupling mode with its working environment gives input to the external world via actions. This requires the individual to focus on the task at hand and interferes with any monitive activity that should be performed simultaneously. This is to be contrasted with the monitive cognitive coupling mode in which the cognitive system has to divide its attention in order to monitor the working environment (without engaging in active manipulation of the external system).

The monitive cognitive coupling occurs in complex environments, where tasks often consist of determining the system status and consequently, if it is within safe operating limits. The monitive type of task is more demanding than the active task as the human has to keep track of many different parameters at the same time and additionally has to discern their significance in the current working context and goals.

2.10.3. Type of control: closed loop versus open loop

The type of control that an individual exerts over the external environment while interacting with it can be divided into closed loop control and open loop control. Closed loop control is characterized by a continuous tight coupling between action and feedback that is necessary to perform a task correctly. A typical tracking task could for example be the continuous adjustment of an aircraft’s positioning and the bank while approaching the landing strip of an airport. The pilot has to readjust the aircraft continuously to ensure a safe landing. Another tracking task could be the adjustment of a ship's position according to a map when passing through dangerously shallow waters. The ships position has to be matched against the map constantly and adjusted appropriately to avoid grounding.

The tight coupling shared by these tasks is typically present when feedback from actions appears in a range between 200 and maximally 2000 ms after execution. If feedback is delayed more than 2000 ms, actions appear to be decoupled from the environment (Sträter & Bubb, 2003).

The amount of information that can be used for closed loop control also varies depending on which context a task is performed in. While research indicates that a person normally remembers 7+2 items (Miller, 1956) these items are reduced to 3 chunks in a highly dynamic task like e.g. car driving (Sträter & Bubb, 2003). The time span for tracking behaviour is also reduced from to 200-600 ms if the environment is dynamic. In situations where the time frame for closed control execution is at 200
ms, the amount of information that is remembered is even reduced to 1 unit, an effect called tunnel vision, which is often observed in emergency situations (Sträter, 2005). It can therefore be concluded that memory is highly context dependent and artefacts mediating tracking behaviour should take these findings into consideration.

Open loop control is exerted when actions are carried out without or with delayed feedback. If the feedback is delayed more than 2000 ms, then the feedback is typically seen as decoupled. Open loop control therefore denotes what is perceived as a one-time execution of an action and is in principle possible without checking the outcome of it. (Sträter, 2005)

Closed loop control puts more stress (demand) on the individual than open loop control as it requires the operator to constantly adjust his/her actions according to timely situational demands. Open loop control only requires a one-time execution and is independent from timely demands in that sense. (Sträter, 2005)

2.10.4. Number of dimensions: Multidimensional versus one-dimensional
Monitoring tasks consist of the surveillance of several different instruments or displays of a technical system that are relevant for the task at hand. Often, the independent parameters displayed only make sense when their interdependence is made clear (Sträter, 2005). When navigating a ship e.g. one has to make sure to take the water current into consideration when setting speed and heading. Otherwise the ship will drift away from the originally intended destination.

A display can represent this information in two ways, integrated or separately. Separately means that the two independent parameters have to be mentally combined to make sense for the task at hand. This should be contrasted with displays in which the parameters are represented integrated in a way that makes their task-relevant interdependence clear. The current’s effect on a ship’s heading could for example be displayed by one graphical arrow that shows the ship’s heading and one arrow that shows how the heading is influenced by the current. In that way the discrepancy between the two headings would be visually visible from the distance between the two arrows. As it is this discrepancy that really is in the crew's interest when traversing waters with strong currents, this is what should be displayed. So the arrow representation means that the dimensionality has been reduced from two (mentally comparing heading with affecting current) to one (seeing the graphical difference between arrows).

Greater stress is put on a person that has to imagine the interrelation between parameters mentally than on a person working with an integrated, one-dimensional version of the same parameters.

2.10.5. Necessary operation: Simultaneous versus sequential processing
In order to successfully manage a task an operator has to perform certain manipulations of the system parameters. These manipulations can vary depending on what kind of manipulations the task promotes. Some tasks, like for example checking an instrument for correct functionality requires the operator to perform certain actions in a step by step pattern that is predefined by the system. Another example would be the way a tire has to be installed on a car. The bolts can only be screwed after the tire is in the right position. That is, the necessary operations are sequential. Other tasks can promote parallel, simultaneous manipulations. The tire installation task for example may be sequential, but the order in which the four tires are installed is up to the individual. The necessary
operations therefore must be carried out in parallel. In parallel does not necessarily mean, that the operations have to be executed at the same point in time but rather that the sequence of the actions can be determined by the individual him-/herself. Simultaneous operations puts more stress on the cognitive system as the progress of the different operations has to be kept in mind and coordinated respectively. (Sträter, 2005)

2.10.6. Type of presentation: Compensatory versus pursuit
The external world can provide the cognitive system with information concerning potential critical situations in different ways. Some displays for example present alarms. This information presentation puts high levels of stress on the operator as he/she must figure out why the alarm sounded. This can be hard work especially if the alarm only provides the information that something has gone wrong without specifying how large this difference between the current state and the desired (presumably safe) system state is. Information thus is presented in a compensatory way in the sense that the operator himself/herself has to discern the exact meaning in relation to goals and the task at hand. Another example would be the blinking of a red light indicating that the pressure in a container is beyond safe limits. Nothing in this red light per se gives the operator a hint about what is wrong with the system, the operator reinterpret this generic symbol for its specific meaning. (Sträter, 2005)

This can be contrasted with pursuit presentation of information. Here the current and desired system states are visible at one glance making the difference between them obvious. A typical display leading to pursuit tasks is e.g. an analogue gauge display, indicating current system state with the help of an arrow and the critical area that is highlighted with red colour e.g.. The operator does not have to reinterpret the display to discern its specific meaning in the working context as intensely as would have been the case if the same information being had been presented in a compensatory way.

Pursuit presentation puts less stress on the human as it directly provides the operator with information on how big the difference between current and critical system state is. This is not the case with information that is presented is a compensatory way and forces the operator to mentally rule out what the alarm means, in other words he/she has to find out how the current system state deviates from the desired system state.

2.10.7. Primary Compatibility: compatibility versus incompatibility
The information displayed during man-machine interaction can either be incompatible or compatible and the compatibility itself can be either in internal or external mode. External compatibility refers to the extent in which information in the external world is congruent with other information in the external world in a certain situation (Sträter, 2005). The easiest way to explain external compatibility is to put it in contrast to external incompatibility. E.g. two analogue instruments relevant for a task could differ in the direction the arrow rotates e.g., one instrument displays counter-clockwise and the other clockwise rotation. The instruments would thus lack external compatibility as the expected behaviour from one instrument can’t be transferred to another. So the term external compatibility refers to the extent to which the meaning of instruments and display can be discerned within a system.

Internal compatibility refers to the extent to which external world is compatible with the expectations a user has from other systems when using a device. For example, if a car driver is used to manual transmission then he/she could continue to do move his/her left foot when shifting gears.
even if the car has automatic transmission until the internal image about the cars functioning has been modified to the new demands, e.g. by experience. Norman (1993) coined the term *affordance* for describing the way an artefact is created according to the mental model the user has of its functioning (Sträter, 2005).

It is important to stress the difference between internal and external compatibility. While external compatibility refers to the extent in which instruments within a system can be mapped to similar functioning, internal compatibility refers to the extent to which the behaviour of a system or a subcomponent of the system resembles other systems. The domain of behavioural transfer is different, for external compatibility it is the behaviour as promoted by the system at hand, for internal compatibility it is all the behaviour that has been internalized from other systems.

What this means for designers, is that the concept of meaning becomes an essential component of designing an artefact as the artefact's meaning gets established by comparisons to other artefacts in the same or other systems. This strongly resembles the way DeSaussure (1983, P.114) thought meaning to be established in language:

> Each of a set of synonyms like redouter ('to dread'), craindre ('to fear'), avoir peur ('to be afraid') has its particular value only because they stand in contrast with one another. No word has a value that can be identified independently of what else is in its vicinity.

Meaning in this sense is structural or relational and not referential. Meaning can only be discerned by seeing a sign in relation to other signs.

> Concepts [... ] are defined not positively, in terms of their content, but negatively by contrast with other items in the same system. What characterizes each is being whatever the others are not. (Saussure, 1983 P.115, as cited by Chandler, 2007, my italics)

For internal compatibility the difference of the meaning of an object emerges from objects in other systems, while in external compatibility the meaning emerges the difference of an object to others within a system.

Incompatibility denotes the situation that the functioning of an object in a system does not match the operators understanding of it. External incompatibility refers to the fact that the system at hand promoted this misunderstanding while internal incompatibility refers to the fact when other systems promoted the misunderstanding.

### 2.11. Resolving mechanisms

Resolving mechanisms are strategies developed by an individual to cope with stress resulting from the cognitive couplings. They are coping mechanisms within the individual and therefore naturally bound to the experiences and concepts of a person and are located within the "strain" part of the stress/strain division. The resolving mechanisms do not necessarily lead to errors as they are just natural mechanisms for aligning the internal and external world by reducing dissonance.

In practice this means that while the stress in terms of Cognitive Coupling modes may be equal for all individuals provided that the task and corresponding circumstances stay constant, the actual effect
of stress on an individual may vary depending on various factors. Strain represents this variable effect on the individual mind.

For example, if the captain of a ship has the task of setting the heading for the entrance of a harbour, then this task puts certain stress on him/her via Cognitive Coupling modes. He must e.g. make sense of multiple dimensions (current, wind, speed, etc) and coordinate all the different actions necessary for task achievement (give rudder commands, monitoring, etc). While this poses less of a challenge for a person that has been trained for this situation regularly, it probably would be harder to manage for a complete novice. The novice would be placed under great strain, not being able to infer the correct combined effect of wind and current on future ship heading, leading to a situation in which an accident could occur due to degraded performance (especially if no technical aids are simplifying the task by reducing stress). So, although the stress is equal in both cases, the strain is bigger for the novice due to infrequent experience with this (or a similar) situation. It should be noted, however, that high stress correlates with a rise in strain levels therefore impaired performance, provided that everything else stays constant (Sträter, 2005). The right kind of training can thus be regarded as something that may aid the individual in coping with stress (resulting from the cognitive couplings) by supporting resolving mechanisms (strain) and thereby improving performance (e.g. in terms of committed errors per person, Figure 1). Several different resolving mechanisms can be discerned.

![Figure 1. Schematic: The effect of experience on performance. The line with smaller quadratic data points denotes the performance (in terms of committed errors) of people with training, the line with the bigger quadratic points denotes performance of individuals without training.](image)

**2.11.1. Fixation**

As described earlier, the cognitive system strives for equilibrium. The cognitive system can only perform when a stable state is reached. Therefore the stable state represents an “action-enabling” state that should not be abandoned too quickly if actions are to be performed at all. Else, the case then the cognitive system would experience constant dissonance and would not be able to act at all. It therefore requires quite strong cues to offsets a dissonance state in the cognitive system once
equilibrium has been established. While this enables the cognitive system to perform, it can lead to dangerous situations as well. Festinger (1957) called the strife for keeping the equilibrium intact by the cognitive system for one of the mayor error mechanisms in cognition, as important cues hinting to the occurrence of an error simply are not taken into account. (Sträter 2005)

2.11.2. Information ignorance or reduction
A mismatch may be perceived but the dissonance is levelled out as the operator deliberately chooses to leave out certain information in the external world, as he/she has a certain hypothesis about why an error occurred and how to tackle it (Sträter, 2005). If an operator e.g. knows from experience that an alarm is oversensitive, he/she may chose to ignore it in a situation when it is actually critical. This is usually the case in stressing situations, when a lot of external information is present. The operator reduces this load by leaving out certain information (Sträter, 2005).

2.11.3. Goal reduction
Sometimes the internal world implies more decision possibilities than the external information would give rise to. When an operator encounters conflicting goals and he/she for example chooses the goal efficiency over safety, although the information in the external world would imply otherwise. So the internal world enriches the decision process with alternatives that are not viable in the current situation. It is highly difficult to change goals as they are a manifestation of experiences (that in turn represent the relation between concepts as described earlier). Therefore the abandonment of a goal in favour of another is accompanied with high levels of dissonance. Once the new goal has been established, it often leads to more negative statements about the previous goal. This usually occurs during situations associated with low levels of stress. (Sträter, 2005)

2.11.4. Goal and information overload
A mismatch between the internal and external world can lead to the uncomfortable situation that an operator cannot decide what to do, due to a level of dissonance that is too high. The dissonance can result from inappropriate information collection or due to goal reduction. The information from the external world leads to two or more alternatives of action and the operator mentally jumps between both of them without processing them in depth (Reason, 1990). This situation usually arises when the operator is in a stressful coupling mode. It leads to the fact that a necessary action is performed too late and an error may occur.

2.11.5. General remarks
The resolving mechanisms do not necessarily lead errors. They are just the natural way of the cognitive system to react upon dissonance by reducing it until a stable state is reached. Hollnagel (2005, P.22) refers to the cognitive system as consisting of at least a human and an artefact as a system that can “modify its behaviour on the basis of experience so as to achieve specific anti-entropic ends”. The term entropy is referring to the amount of disorder within a system prohibiting the achievement of a task. The concept of dissonance gives a biologically founded view on what this entropy is. As long as the individual experiences dissonance, no actions can be executed and hence no manipulation of the external environment occurs, that could lead to errors.

2.12. Complexity
In order to illustrate the concepts of cognitive couplings and resolving mechanisms it is useful to put them in perspective to other research. In that way similarities can be pointed out and differences be spelled out.

Overload comes not from the number of planes that a controller is working as from the complexity of the interactions (Weick, 1987)

This quote highlights the fact that a task does not have to be stressing for a person only because he/she has to manage many interactions. Experienced controllers could for example manage lots of flight simultaneously without having the feeling of handling a complex task. For a novice, on the contrary, this would probably appear to be an overwhelming task, to the point of goal and information overload. So how come then, that Carl Weick had to emphasize this point? The answer is that there appears to be a conceptual ambiguity regarding the nature of complexity. Weik placed complexity in the phenomenological realm in the sense that it is something the individual has to experience in order for it to be present.

But, while complexity certainly is something that an individual experiences, it also stands in relation to the task that is to be mastered. There are tasks that evoke feelings of complexity for the majority of the individuals doing it while it does not for other individuals that were specifically trained for the task. In other words, the task that is to be mastered poses certain demands (e.g. in terms of complexity) and the issue seems to be the relationship between this demand (as something the task puts on the individual) and demand as something the individual experiences a task to be.

The aforementioned stress/strain model is a conceptual tool for dealing with this ambiguity. It encompasses the division between the factors affecting the human and the effect they finally have on the human. The link connecting the two concepts is the cognitive mill (alignment internal/external world). Within the mill, the cognitive couplings function as a sort of hypothesis on how demanding the external world (task) will be for individuals, by classifying the stressfulness of the required interactions. However, to experience complexity is the dissonance actually evoked by the task that a person has to level (resolving mechanisms) in order to complete it and is always dependent on the internal world (concepts and experiences) of a person. If we integrate this conclusion with the prediction from the theory, that performance degrades with rising stress levels, we get the following illustration (Figure 2).
Figure 2. Schematic: Stress/strain. Performance in terms of committed errors is dependent on the stress levels and the experiences and concepts (e.g. in terms of training) of a person.

The theory predicts a decline in performance if stress levels rise (difference between grey and black lines, Figure X). And while the performance in tasks associated with high stress levels (black lines) will improve when experiences and concepts are changed positively (e.g. training, black line on the left), the baseline of errors/person will still be higher than in less stressful tasks that have been addressed with training (grey line on the left). The cognitive couplings aim at denoting this task-associated effect of stress on performance that is independent of the individuals conducting the task. Strain, on the other hand, denotes the actual individual effect on the individual.

Thus these concepts are both of importance for investigations of why errors occurred and for design tasks aimed at preventing these errors from happening.

2.13. Triangulation

The concept of triangulation as described by Thurmond (2001) is central to this study. Triangulation involves the “combination of two or more data sources, investigators, methodologic approaches, theoretical perspectives [...] or analytical methods [...] within the same study” (Thurmond, 2001, p.253) and leads to the respective triangulation (investigator triangulation, theoretical triangulation, etc.). The intent of using triangulation is “to use two or more aspects of research to strengthen the design to increase the ability to interpret the findings (Campbell & Fiske, 1959 as cited by Thurmond 2001, p.253). In this study quantitative and qualitative designs have been employed in order to enhance interpretation (across-method triangulation, Thurmond, 2001).
3. Purpose

The general purpose of this study is to see how well the Conning Information Display fulfils its intended safety critical purpose in actual simulator training exercises. The quantitative part of this study will provide insight into the question whether the display influences performance at all (Heiman, 2001) both in comparison to other students (between group comparison) and in comparison with the same student's performance on other exercises (within group comparison).

- The dependent variable performance describes how the combination of stress (through cognitive couplings) and the corresponding strain (resolving mechanisms) influences behaviour. It is operationalized by the number of errors a crew commits.

- The independent variable determines if a crew can use the Conning Information Display for an exercise or not and is operationalized by switching it “on” or “off”.

The qualitative part of this study concerns the nature of performance as operationalized by committed errors. Central question are

- What is the characteristic (Johansson, 2003) of performance in this context?
- Is this characteristic of performance influenced by the conditions of the Conning Display (on/off)?

For this several methods will be employed (within-methods triangulation, Thurmond, 2001).

- Indirect, naturalistic observations will be used to infer the cognition of the captain on the bridge

- Semi-structured interviews will be conducted in order to uncover the captain's intentions when performance deviated of events thus providing further insight into the nature of the causes behind the error

After analyzing if the Conning Information Display influences student performance and what the characteristic of this performance is, a design proposition for the decision supporting display integrating these findings will be presented, in terms of design requirements.

3.1. Clarifications and restrictions

The issue of whether the training scenarios are representative for real world settings is not of interest in this study. In this study, the focus does not lie on the question of the external validity of the simulator training to real-world settings in general. Rather this study aims at uncovering if the introduction of the Conning Display makes performance as applicable to real-world settings as performance without it. Thus it is the relative difference between performance with and without Conning Display that is of interest.

Furthermore the aim of this study is to show what causes that underlie the basic and most severe errors as pointed out by the instructor monitoring the simulator exercises. This means that other
actions that could be regarded on errors based on different criteria are excluded from analysis. These could be addressed in future studies.
4. Method: General

In order to rule out the influence of the Conning Display on student performance it is necessary to compare the performance with it and without it. While quantitative experimental designs can yield insight into the question of whether differences exist at all, qualitative studies can rule out what the characteristic (Johansson, 2003) of these differences are. It is thus feasible to combine these two approaches in order to get a richer picture of the impact the Conning Display has on student performance. This should open up for a deeper insight into the quantitative and qualitative characteristics of performance especially since the number of participants is limited in real-life settings and thus a measured outcome is more susceptible of being caused by variations in behavior within the participating sample.

The quantitative part of this study consists of comparing performance in terms of committed errors within groups conducting exercises with Conning Display and without. The qualitative part consists of semi-structured interviews that were conducted with the participants. Furthermore the participants’ behavior was observed live as I was present in the instructor room during the exercise and ad-hoc, after the exercises were recorded. The ad-hoc observation was done using eye-tracking, the simulator-inherent video recording of the bridge and the instructor station overview camera. The recordings were used as supplementary material during the interviews when deemed necessary (more on that in the section “interviews”).

Finally, combining the results from the quantitative and the qualitative study design, requirements (Benyon et al., 2005) on a new presumably safer layout of the system will be elaborated.

The quantitative and the qualitative part of the study will be described separately although the “apparatus” section in the quantitative study section also incorporates descriptions of equipment that was used in the qualitative analysis. This was done in order to enhance readability as the eye-tracking device and the other video capture software ran in parallel to the experiment and thus could seen as being part of it.
5. **Methods: Quantitative study**

The method section of the quantitative part of this study will be discussed. The apparatus section also includes equipment employed in the qualitative part of this study.

5.1. **Participants**

7 individuals participated in this study (N=6). The average age of the participants was 22. The participants were exclusively male with the exception of one female that had to quit the experiment prematurely and will therefore be excluded from analysis. The participants were students of Marine Science studying at the FH Oldenburg/Ostfriesland/Wilhelmshaven. All participants had a major course assessment (Leistungsnachweis) in system monitoring (Systemüberwachung). They had completed the first maneuvering exercises (3 scenarios) previous to the study. The participants were selected by means of *administrator selection* (Shadish, Cook & Campbell, 2002) as the instructor for the exercises delegated the teams. All the participants were verbally given the opportunity to give informed consent prior to the conduction of an experiment session. No rewards (in terms of money, etc) were handed out for participation in the experiment.

5.2. **Apparatus**

The material used and its function for the experiment will be described.

5.2.1. **The nautical simulator: An overview**

The simulator facility at FH Oldenburg/Ostfriesland/Wilhelmshaven consists of several simulators and an instructor station. The simulators are all behavioral-based simulators (verhaltensbasiert), their layout is generic and not an exact replica of an existing real-world ship bridge. Despite that, they incorporate a great part of the minimum of equipment present on state-of-the-art commercial ships.

The simulated bridges differ in terms of equipment and space. Bridge 4, 3 and 2 are smaller and have a more restricted view on the simulated outside than bridge 1. Hence Bridge 1 is considered to be more realistic in terms of similarity to real-word bridges by the personnel operating the simulated bridges. The instructor supervises the students from the instructor station which is located adjacent to bridge 1. It provides a clear and direct view into bridge 1 through mirrored-glass. Briefing, the initial instruction of the students before the exercise starts, and Debriefing, the review of simulator exercises, takes place in the debriefing room or the instructor station. Both debriefing room and instructor station are equipped with projectors used for displaying the recorded students’ performance in the debriefing.
5.2.2. The instructor station

The instructor station is equipped with several instruments for monitoring student performance. The surveillance is done visually and by means of communication, as several displays mediate the events on the bridges in terms of video and audio. The CCTV-display for instance shows video signals from the different bridges in one integrated split-screen. The video signal shows overview over the entire bridge and the present crew (The crews conducting the simulator exercises are aware that they are being recoded).

The outer view of the bridges on the simulated environment is also displayed in the instructor station. Several flat screens show the visual information from the simulated outside that is available to the crews on the bridges during exercises.

It is also possible to monitor the radar present on the different bridges from the instructor station by tuning in to them. So in order for an instructor to check if radar settings are done correctly on bridge 1, he/she simply presses the corresponding button in the instructor station and immediately gets an exact replica of the radar display on bridge 1.

The arguably most important device in the instructor station is the main display. Here, the instructor may manipulate different parameters of an exercise and monitor the progress of the different bridges; e.g. weather conditions can be changed, equipment malfunctions on the bridges can be simulated and the location of the ships can be altered. It is also possible to monitor the important ship parameters, like e.g. rudder commands, engine rating, speed, rate of turn and drift. The main display also shows an overview window, which displays the crews’ ship position on a map, thereby enabling the instructor to follow the ships’ movement from a “bird-eye” view. It is also possible to see the history of a ship’s movement. This is displayed by means of trails that each ship leaves behind it (similar to footprints in the snow). A ship’s heading (its future course) is indicated by two arrow-shaped vectors in front of the ships. These two vectors display the course over ground and the course through water and hence also implicitly the discrepancy between the two in terms of how far
the vectors are apart from each other. Course over ground is where the ship would head if the course-influencing factors are taken into account (like current, wind or sea condition).

The main window additionally displays all the information that would be present on a state-of-the-art nautical chart. This information includes the intensity and direction of the current(s), the water depth, buoy information and terrestrial information, where landmasses are e.g.. The electronic chart also displays radar corridors that should not be crossed. The instructor room contains two main displays, instructor station 1 and instructor station 2. During the experiment instructor station 2 always displayed the same information as instructor station 1 on the main window. Apart from that station 2 was as operable as station 1.

The instructor is also able to communicate verbally with the bridges. Several communication channels exist for this purpose for instance the hand-held communication device and a normal, standard telephone. The instructor may also enter the bridges personally, if necessary.

Another important feature is the loading of the prefabricated scenarios that the students will have to challenge. These scenarios can, depending on purpose, vary greatly in complexity and difficulty. Depending on scenario type the instructor can choose if the bridges are able to see each other visually in the simulated environment or not.

5.2.3. The instructor

The instructor is the person responsible for simulator training execution and surveillance. The instructors participating in the experiment all have nautical experience from either being a pilot or captain. Two instructors participated in the experiment, one of them a captain of a vessel and the other a pilot.

5.2.4. Scenario batches

As argued before, it is always important to study cognition in its natural context. Thus, this study specifically includes the existing scenarios that the nautical students would have to master during their education. Therefore, the scenarios were not modified.

The simulator training in Elsfleth consists of 2 batches of scenarios (exercises). They both incorporate 3 scenarios each.

Scenarios

Every batch consists of different scenarios with objectives that have to be met by the students. The maps that form the basis for the scenarios always simulate real world locations. This means that the virtual environment as seen by the students is intended to look and behave like the real world equivalent that is simulates (the simulated harbor in Dover has the same currents and moles as its real world equivalent).
Batch 1
The first two scenarios aim at giving the students an impression on maneuvering with current. However the third scenario, "passage", does not confront the student with current and the crews also see each other in the virtual environment which is not the case in the other two scenarios.

Scenario 1 Maneuvering with current (Fahren im Strom)
The goal in this exercise is to stay in the corridor marked out by the buoys. All ships start at the same position without seeing each other. The current drifts the ships in south-eastern direction (Figure 4). The exercise is over when the ships have entered the southern section as marked out in Figure 5.

Figure 4. The red line marks one possible track through the exercise. The ships start out in the western section of the map
Figure 5. The exercise usually stops when the ships (yellow dot) enter the southern part of the map, highlighted by the black ellipse in the map

**Scenario 2 Entrance of the port of Wilhelmshaven (Einlaufen Wilhelmshaven)**

The objective for this scenario is to enter the harbor and traverse through the lock that is located further into the harbor (Figure 6). The ships start out in the northern part of the map and are initially oriented southwards. Initially, the current drifts the ships in southern direction, however, there is no current in the water in the harbor itself. The ships operated by the students are not able to see each other.

![Figure 6. Entrance of the port of Wilhelmshaven. Red line denotes one way to master the scenario](image)

**Scenario 3 Passage (Passieranöver)**

The objective of this scenario is to conduct a safe passing maneuver. The students have to evade each other while keeping a safe distance to the channel wall (Figure 7). The ships start out on collision course with one positioned in the north of the scenario oriented southwards and the other ship positioned in the south oriented northwards (Figure 7). There is no current in this exercise and the crews see each other.
Batch 2
The following scenarios all include current except for the berthing/cast off scenario. The last scenario also includes wind.

Scenario 1 Entrance of the port of Dover (Einlaufen in Dover)
The objective in this scenario is to enter the harbor and to berth at the eastern mole (Figure 8). All the ships start out at the same position in direction towards the harbor. The bridges do not see each other. The current drifts the ships in south-western direction, and slightly changes direction more southwards in the proximity of the eastern mole (Figure 8). The wind (not displayed on the map) comes from the east.
Figure 8. Entrance of Dover – The ships start out in the eastern part of the map (red dot), enter the harbor and berth at the end of the red arrow.

Scenario 2 cast off and berthing (Anlegen und Ablagen)
The objective in this scenario is to cast off and to berth somewhere in the marked out area in front of the big ship (Figure 9). The crew has to call the instructor when they think that they want to cast off or berth. The instructor then conducts the necessary actions on the hawsers. The ships all see each other.

Figure 9. Cast off and berthing – The observed crew (yellow dot) have to cast off and then berth somewhere in the vicinity of the black ellipse.

Scenario 3 turning and berthing with current
The objective in this scenario is to turn the ship around somewhere in the channel and to berth.
between the 2 ships on the eastern quay as highlighted in Figure 8. The ships all start out at the same position in south-eastern direction. The current also originates from that direction. The eastern part of the channel is marked out with buoys that indicate where to expect shallow water.

![Figure 10](image)

*Figure 10. The crews start out in the northern part of the map (yellow dot) and berth in the vicinity of the black ellipse after turning the ship.*

### 5.2.5. **Bridge 1 “Weser” and its components**

Bridge 1 consists of the standard equipment found on most vessels. The equipment is realistic in the sense that it is physically present and a replica of existing equipment in terms of general functionality and, to a certain degree, also appearance. The technical director referred to this simulator as being behavior-based (generic), that is, not an exact replica of an existing bridge type. In subsequent sections, some of the equipment that was used most frequently during the exercises is explained. Explaining all equipment with associated functions would be beyond of the scope of this thesis.

**Conning Display**

The conning display consists of several views or modes. The mode of the display that was mainly used during the exercises was the *combo mode*. Therefore only this mode will be described in detail. The other modes are harbor mode, fairways mode and berthing mode.

The combo-mode was created in order to incorporate all the relevant information from the other modes in one integrated view. The view was introduced as the persons responsible for the display design noticed that a lot of the information had to be represented in all views anyway, making one integrated view more sensible (The combo-mode was the view that was always active when the students entered the simulator for participation in the field experiment).
The view displays the ship’s type, length, breadth and draught (forward and after). Date and time are displayed as well as heading and speed (over ground and through water). An echograph showing water depth is present as well as thruster information, showing engine Rotation per Minute (RPM) and Propeller Pitch. Furthermore the ship’s movement is displayed (moving gain). The compass and the current rudder angle are shown in the center of the screen (Figure 11).

![Conning Display in combo-mode](image)

**Figure 11.** The Conning Display in combo-mode.

**The RADAR**

The Radar is an object detection system that identifies the range, altitude, direction, or speed of both moving and fixed objects such as aircraft, ships, motor vehicles, weather formations and terrain. It is operated with a trackball. There are two radars present on the bridge, one on the left and one on the right side of the bridge.

**Doppler Log**

The purpose of the Doppler log is to display longitudinal speed ahead and distance sailed through water (water track). It also shows the longitudinal speed ahead and distance sailed over ground (ground track) and the bow and stern lateral speed to starboard and port (water or ground track).

**Echo sounder**

The purpose of the Echo Sounder is to provide depth readings from a bow or stern transducer, displayed on the bridge in fathoms, feet or meters. It includes a depth alarm function. It also gives realistic readout of keel clearance.

**Throttle**

The throttle is used for adjusting the ship’s engine output.
**Steering system**
The steering system digitally displays the ship's heading and its turn rate. It allows the user to set a new course, to give a new rudder command and to adjust the ship's turn radius and the rate of turn. It enables steering in different modes (remote, track, auto, manual, joystick control and NFU).

**VHF DSC**
The DSC is a world-wide system for ship-to-shore, shore-to-ship, and ship-to-ship calling. Digital selective calling is an integral part of the GMDSS and is used primarily for transmitting distress messages from ships, associated acknowledgements from coast radio stations and relaying distress alerts from either ships or coast radio stations. (Technical Manual Section 5c - Polaris Radar/ARPA, POLARIS Ship’s Bridge Simulator)

5.2.6. **Dikablis eye-tracking tracking system**
The Dikablis eye-tracking system was provided by the department of Arbeits- und Organisationspsychologie of the University of Kassel and is used in order to capture, record and evaluate a person's eye-movement. It is divided into several hardware and software components that will be described in the following sections.

**Hardware**
The hardware consists of the head unit holding the cameras, the transmitter, the wireless recording station and a laptop capturing the video signal. The head unit holds a camera pointing at the eye and recording the pupil movement. The eye is spotlighted with infrared light to ensure that the pupil is detected in situations with impaired lighting. The second camera, the field camera, is directed at the surroundings and thus records everything that is in a person’s field of vision. The glasses are affixed to the participants head with ribbons. A cable transfers the video signals from the head unit to a wireless transmitter that is attached to the person wearing the head unit. There is a display on the transmitter indicating the remaining battery power in percent. The battery-driven transmitter forwards the video signals to the wireless reception station that in turn forwards the video signals into a laptop. A standard laptop with installed Dikablis software is then used to process the incoming video signals from the eye- and field-camera.

**Software**
The software bundle installed on the laptop for editing and displaying the video signal consists of the dikablis player, the dikablis recorder, the marker detector, dikablis analysis and D-lab. The dikablis player is simply used to play the files recorded with the dikablis recorder. The dikablis recorder incorporates several different functions, most prominently the adjustment of the eyecamera to the individual carrying the glasses. This must be done to ensure that the combined video of the person’s field of view and the person’s gaze crosshair are synchronized properly.

The dikablis analysis software was only used for exporting videos from recorded projects in this study and will thus not be explained further. D-lab hosts functions enabling the processing of recorded projects. However, this program was not used.
5.2.7. Screen capture software

The screen capture software AutoScreenRecorder 3.0 was used to record the display signal of Instructor station 2. The resulting file was saved on the hard drive of instructor station 2 in avi-format.

5.2.8. Screen grabber

This video grabber enables audio and video signal capturing. This is done by plugging in the appropriate plugs to the composite video and stereo audio jacks. This signal enters the computer through an USB-port. In this way a replica of the video signal going to the CCTV from the overview camera on bridge 1 can be transferred to a computer for further processing. This processing is done with the software CyberLink PowerDirector 6.5 that enables the transformation of the incoming video/audio signal into a format that can be played by most standard video players (.mpeg).

5.2.9. Adobe Premiere Pro

This software was used to combine and synchronize the exported video from a dikablis gaze recording session, the instructor station and the CCTV-station. The videos were arranged in a split-screen view, allowing the simultaneous observation of the whole bridge (CCTV), the location of the ship in the scenario (instructor station) and the gaze pattern produced by the captain (exported dikablis video).

5.2.10. Dictaphone

The first few interviews were recorded with Digital Voice Recorder, ICD-P620, a Dictaphone. As the battery-power of the dictaphone posed problems, the remaining interviews were recorded with the laptop-inherent microphone of a Dell Inspiron 1525 Laptop Computer.

5.3. Design

A quasi-experimental within group and between group field study was conducted (n=6). The dependent variable is termed performance (as defined in the purpose section) and operationalized by the number of errors committed by the participants in the scenarios as pointed out by the instructor (expert evaluation). The independent variable is the presence of the Conning Display and consists of the conditions "on" or "off" during the exercise.

An experimental design was chosen in order to determine if the new Conning Display influences performance at all. If this study would have been carried out without a control condition (Conning Display off) one would not have been able to draw conclusions if errors are related to the design of the display or the difficulty of the task the crew has to complete.

The design makes within and between group comparisons between the experimental condition and control condition possible. The reason for the deployment of two designs is that field experiments are bound to the normal flow of work as present at the location of experimental conduct. Thus
unpredictability, that is present in normal working conditions, will be present in the field experiment as well and potentially pose a threat to internal validity (Sternberg, 2006).

One risk factor was e.g. the fact that the instructors never know in advance how many scenarios that would be mastered during one trial. It could be, that one crew manages to go through all 3 scenarios with only one try each. But it could also happen, that the crew has to repeat a scenario several times, not leaving enough time for the last scenario that should take place in that trial. When considering a between groups design for example this could mean that a control (or experimental) condition is no longer available when evaluating the data as the crew that should have performed this exercise did not have time to do so. Or, in the worst case, no control/experimental pairs would be complete at all, rendering the design ineffective. That is why two experimental and two control conditions were created.

The within groups design is a good choice when the number of participants is low, as each participant serves as his/her own control, which was the case in this study (Shadish, Cook & Campbell, 2002). Unfortunately within group designs carry with them the problem of practice effects and comparability (Heiman, 2001). The comparability issue poses a threat to validity when control and treatment condition differ in terms of the nature of the underlying task. Possible errors in one of the conditions could thus be attributed to differing difficulty levels between the different tasks tested in the conditions and not to the experimental manipulation per se. In this study, it was unknown to what degree the scenarios would be comparable and an elaborate schema for evaluating the exercises in terms of comparability could not be constructed despite several interviews with the instructors.

However, discussions with the instructors resulted in the insight that at least two scenarios would take place per batch. Thus, these 2 scenarios could be divided into a control and an experimental condition, provided that they are comparable to each other, securing 1 comparison. Additionally, if more scenarios would be completed, then these could also be used by switching the Conning Display on and off respectively. E.g. if the crew A manages to go through 2 scenarios and crew B manages to finish 3 scenarios, then the third scenario is not lost due to the missing control condition as would be the case when using solely a between scenario, as the performance of crew B in the third scenario can be compared with the performance of crew B in the second scenario.

However, it is also important to compare the performance of crews conducting the same scenarios as this makes the comparison problem between conditions, that is inherent to the within design, obsolete. In order to enable a between group design there has to be at least one control group in every batch, which would probably be the case as the instructor stated, that 2 scenarios usually were completed in each batch.

Thus, combining the within and between group design resulted in following design that makes both comparisons possible (Table 1). "Group" denotes which group of individuals that were manning the bridge. The within comparison between experimental and control condition can be seen in Table 1

Table 1. within group design.
5.4. Procedure

5.4.1. Preparation

The author was the only researcher present with the exception of one recording session when a colleague assisted. Regardless, the procedure stayed the same. In the morning the Dikablis equipment was arranged before the arrival of the students. A laptop was plugged in to the video stream of the CCTV-station. The recording software on the laptop was also started and tested to make sure that everything was functioning normally. After that the screen capture software on the instructor station 2 was started and placed in standby mode to ensure rapid handling when the stations display should be recorded later on. The battery charger for the transmitter was powered up in the instructor room to ensure that the spare battery would be charged and ready when the main battery in the transmitter was low on power. Finally, a name list was placed on the bridge for the participants to fill out.

The briefing of the students started, which the researcher initially recorded. The students were briefed about the purpose, the content and the assessment criterion of the first scenario. After the briefing the students were verbally informed about the purpose of the present study by the instructor, divided into to groups of 2 to 3 students each and assigned to one of the bridges. They were also informed that the crew on bridge 1 would be divided into roles (captain, helmsman, lieutenant). A volunteer from the group on bridge 1 was assigned the role as captain and briefed
verbally on the way to the bridge about their right to abort the experiment at any time and about the fact that their data would be treated anonymously. After the arrival at bridge 1, the preparation phase started in which the students conducted their passage planning, adjusted their equipment and, if necessary, entered waypoints or other data relevant to the scenario objectives.

The captain put on the glasses and was briefed about the calibration that had to be done in order to adjust the eye-tracking glasses to his physiognomy. The participant put on the glasses, the researcher switched on the transmitter and the receiver station and conducted the calibration. During the calibration the captain was able to hear everything on the bridge and to communicate with the rest of the crew. The calibration continued until the researcher deemed it to be satisfactory with the upper time limit being the point in time when the instructor wanted to start the exercise. Before that, the instructor checked if the crews were ready and the researcher could request some extra calibration time for the researcher if he was not finished yet. The Conning Display was switched “on” or “off” depending on the design scheme and the participants were informed about the fact that the display had to stay that way for the remainder of the scenario.

5.4.2. Start of the exercise

After that, the eye-tracking recording was started in the Dikablis recording software before the simulation started, in other words in the preparation phase. Subsequently, the recording of the CCTV station in the instructor room was initiated. Therefore, the eye-tracking video and the CCTV-video start out close to each other in time. This is not necessarily the case for the record of the instructor station as this was initiated after the preparation phase. More precisely that recording was started exactly when the instructor said “now” in the phrase “the exercise starts now”. This was important for the synchronization of the different recordings later on.

The instructor informed the researcher about errors that were made by the students online, during the course of the scenario while the researcher made field notes on these errors. The first scenario was over when the instructor asked the students to come for debriefing. Then, the transmitter was switched off and the recording of the recording software was stopped. After this the recording of the CCTV camera was stopped and the file saved. The recording of the screen capture software on instructor station 2 was also terminated. All of this was done as quickly as possible in order to be able to participate in the debriefing.

5.4.3. Debriefing

The Debriefing could take place in the instructor room or in the debriefing room, depending on available time. If the debriefing took place in the debriefing room the Dictaphone was used for recording purpose (if the students gave their consent). The students were informed verbally about the purpose of the recording and that it would be treated anonymously.

If the recording took place in the instructor room then the same consent question as stated before was posed. The recording per se was done with the Dictaphone for the first few participants and afterwards with the laptops’ inbuilt microphone as the Dictaphone posed technical problems.
5.4.4. **Start of the next scenario**

After the first debriefing the crews were directly briefed about their objectives in the next scenario. The crews returned to their bridges for preparation. The captain on bridge 1 had usually kept the glasses on during debriefing and briefing and the calibration was now tested to see if the glasses had moved out of place (thus making the previous calibration imprecise). If they had moved out of place another calibration was conducted. Again, the Conning Display was switched “on” or “off” depending on intended design and the participants were once again informed that the display had to stay that way for that scenario. The battery was also checked, if the power on the display was below 15%, then the battery was replaced by the fully charged one and consequently put in the charger. Again, all the recordings were started at the right time (CCTV, instructor station and eye-tracking) and the researcher monitored the scenario with the instructor, again taking notes when errors were committed until the crews where called for debriefing.

This procedure was repeated until all exercises were over. A check was performed after every exercise to ensure that the Conning Display was in the same mode as when the exercise started. Additionally, the recorded videos were checked to see if the display mode had been changed.

5.4.5. **Watchkeeping**

After the last scenario was finished, the participants were asked if they wanted to sign themselves up on a list with their name, exercise type (maneuvering or watchkeeping), e-mail and mobile phone number for further contact concerning the interview on the experiment.

The procedure for the watchkeeping scenarios was the same as the procedure for the maneuvering scenarios. The files produced by the screen capture software were transferred to an external hard drive and later on copied to a PC with the other video files for further processing.

5.4.6. **Interviews**

Some days after the participants had attended to the experiment an e-mail was sent to them asking them for an interview appointment. The date for the appointment was usually set for Thursday the week after the experiment. The interview took place on different locations. The first interview was conducted in the (empty) debriefing room. All the interviews that followed were conducted in another room in the simulator building. They lasted approximately 1 hour, during which the researcher asked questions and followed up on the answers if they were perceived as unclear.

5.4.7. **Deviations from regular procedure**

Originally 8 participants were to be included for the maneuvering scenarios but, due to organizational factors, this was not the case. Thus, 2 participants had to be excluded from the design leaving 6 participants. This happened when the experiment was already underway, therefore the impact on the design, leading to only 3 participants in for the between comparisons per scenario, had to be accepted.
Additionally, the first batch for Person 1 deviated from procedure in the sense that only the first exercise was recorded without Conning Display, all others were recorded with it.
6. Qualitative study

In the following section the qualitative methods employed in this study will be described consisting of interviews and observations.

6.1. Interviews

A semi-structured interview was carried out with the experiment participant. The aim of the interview was to generally discern the systemic causes of an error and to specifically connect them to the concept of the cognitive couplings. In order to create questions that already have had proven insightful in accident investigations the questions proposed by Klein (1998) as cited by Dekker (2006) were adapted. They were combined with questions the researcher generated from the descriptions of the cognitive couplings by Sträter (2005). If the researcher deemed the questions to cover the same topic, then the version of Klein was kept, as those questions have, according to the author, already proven themselves useful in several accident investigations. If no equivalent was found, then the question was added in order to see if it would extract useful information that the other questions missed (See Appendix A: Interview questions for a complete account of the questions). The understandability of the questions was examined in a pilot study (See Appendix B: Pilot study).

The researcher also incorporated Dekker’s (2006) approach of only confronting the interview participant with video recordings after they had recalled the exercise and not before they had recalled, in order to keep their first account as undistorted as possible. Additionally, the researcher also summed up what the participant said at regular intervals to ensure that a correct understanding of the participant’s account had been established.

6.2. Observations

The participant behavior was observed indirectly and online during the exercise. This was done by looking through the window in the instructor station onto the bridge and by monitoring the CCTV-camera and instructor station display. Discussions were held with the instructor about what he deemed to be an error.

Naturalistic observations count as a widely used technique for observing cognitive performance in real-life situations (Sternberg 2006). The advantage of using observations is that one can observe what the participant really does and thus will not be susceptible to the problem of persons not recalling actions accurately as may be the case when only interviews are employed (Breakwell, Hammond & Fife-Schaw, 2000). Thus, observations and interview data can complement each other particularly well (Benyon et al.2005), leading to within-methods triangulation (Thurmond, 2001).

For this study formal observations were conducted in the sense that only the behavior sequence leading to an error was of interest. Basically 2 rounds of observations were conducted. The first round consisted of attending the exercise from the instructor station. During these observations field notes were taken on student performance, especially when errors occurred.

However, the problem with errors is that their occurrence cannot be predicted accurately. Thus, the researcher is often somewhat surprised when an error occurs. This can hamper the resolution at
which field notes are taken, potentially leading to the leaving out of important details previous to the error. Thus, a second round of observations was conducted, using only the video recordings and with special focus on the sequence where the errors occurred. Thus, the second round of observations may be regarded as a bit more formal in the sense that the researcher knew exactly when an error would happen and thus could analyze these sequences with greater detail, by e.g. freezing videos or artificially speeding the exercise up to understand the movements of the ships better.

6.2.1. Addressing the keyhole effect of video observations
Strangely, not only operators of complex control rooms experience what is termed as the keyhole effect (Voshell & Woods, 2005). In this study, or probably in every study that involves video observations, the researcher also has to struggle with this effect.

The keyhole effect occurs when a display cannot mediate the richness of the physical environment it is recorded in as it only displays a small portion of all the information that is of interest for the task at hand (Voshell & Woods, 2005). Thus, a lot of the variety that is present in the original environment is lost. In this study, a method to combat this issue consists of within-method triangulation (Thurmond, 2001).

Indirect observations make it hard to discern what the individuals are focusing on by just looking at them. This can be traced back to several physical reasons, the participant can be facing the opposite direction of the camera and even when the person is looking in direction of the camera it can be hard to discern what he looks at due to the fact that the resolution of the video is too low or the object of interest is out of camera view. Thus, eye-tracking videos can be of value solving these problems by making the actual gaze visible.

With eye-tracking on the other hand, it is not always possible to see what for instance a participant does with his hands when looking out of the window, as this may happen outside of the participant’s field of vision. So the field of vision of the eye-tracking camera may render it impossible to capture all the relevant information for at task by reducing the richness of relevant information (like looking through a key-hole) and therefore possibly hampering the analysis.

However, the simulator contains an inbuilt CCTV-camera that always records the student’s performance during exercises and shows what would be hidden from view for the eye-tracking camera. According to the responsible technical personnel at the simulator, this camera was in a tedious trial-and-error process placed in the spot on the simulator that provided the best overview of the bridge and its crew. Additionally, the audio signals from the crew on the bridge are also recorded with several microphones enabling the capturing of sound and therefore verbal communication.

Therefore, when observing these videos in parallel nearly everything on the bridge may be observed at any time regardless of what the participant looks at and how he/she is positioned. These two videos thus encompass everything on the bridge. However, may still be quite hard to put these observations in the general context of the exercise as the researcher, being a novice in the domain, may lack the professional vision for what the students actually are doing, how they affect the ship and what the effect in the broader context of the scenario is.
Thus, the recording from the instructor station, that is especially designed for monitoring and analyzing the behavior of the crews, was analyzed. Finally, all these videos where integrated into a synchronized single video making it possible to

- Observe what the individual is looking at
- Observe what is the participant does without looking at it and what is generally visible on the bridge (other crew members e.g.)
- Observe how the actions of the crew stand in relation to the exercise, e.g. in terms of objectives

The third bullet is not an observation in the classical sense as it is a graphical display of the result of the crews' actions. However, it is an observation of the behavior of a system, consisting of the individuals (crew) and the technical component (the bridge) and how this system behaves in respect to scenario objectives. This observation of system behavior fits very well to the underlying systemic model of human cognitive performance and to the fact that it is hard to analyze behavior in isolation without the context in which it occurs.

Furthermore the synchronized video recording enables the analysis of the difference between these levels of observation:

- What is the individual looking at and what should he have been looking at according to the general course of the ship in the scenario (instructor recording/eye-tracking)?
- Does the participant notice what happens around him, outside of his field of vision (eye-tracking/CCTV)?
- How does the whole crew react when critical events occur (CCTV/instructor recording)?

While it can be argued, that this method encompasses high explanatory power in terms of what may be observed, it can at best give hints on the intentions behind actions. Therefore, interviews are necessary to gain a more complete picture of events (within-methods triangulation).

6.3. Data analysis
The recorded video data and the interview results were analysed by comparing what participants stated in the interviews with what happened during actual exercises and what the instructor had pointed out as being the error. This was a hermeneutic process requiring the material to be reanalyzed several times as e.g. the students themselves sometimes did not exactly know why an error occurred.
7. Quantitative results

While the sample of participating people was quite large (17 participants) only 7 of them were included in the result section and furthermore 1 of the 7 persons had to take off the eye-tracking glasses due to discomfort leaving another person, that was helmsman, not captain, with the glasses. 6 persons constitute a very small sample of the population that these finding should be applied to. Thus, due to the low number of participants, no inferential statistics will be employed. Instead descriptive statistics will be presented.

7.1. Descriptive statistics: within-group design

The students’ performance as operationalized by the number of committed errors under the two conditions Conning Display on/off will be presented.

7.1.1. From “off” to “on”.

2 errors were committed when the condition of the Conning Display was “on” (N=6). 7 errors were committed when the condition of the Conning Display was “off”. Thus 5 errors less were committed when changing the Conning Display from “off” to “on”.

Table 3. Descriptive statistics concerning the within group conditions control/treatment.

<table>
<thead>
<tr>
<th>Participant</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Errors in total</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>1.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Standard dev.</td>
<td>0.90</td>
<td>0.37</td>
</tr>
</tbody>
</table>

When the Conning Display was turned on the crews committed 6 errors less.

7.1.2. From “on” to “off”.

3 errors were committed when the condition of the Conning Display was “on” (N=5). 1 error was committed when the condition of the Conning Display was “off”. Thus, 2 errors less were committed when changing the Conning Display from “on” to “off”.

Table 4. Descriptive statistics concerning the within group conditions treatment/control.

<table>
<thead>
<tr>
<th>N = 5</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant</td>
<td>On</td>
<td>Off</td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Errors in total</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>----------------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Mean</td>
<td>0.60</td>
<td>0.20</td>
</tr>
<tr>
<td>Standard dev.</td>
<td>0.49</td>
<td>0.40</td>
</tr>
</tbody>
</table>

When the Conning Display was turned off the crews committed 2 errors less.

### 7.2. Descriptive statistics: between-group design

Of these 6 participants 3 participated in the scenarios of batch 1 and 3 participated in the scenarios of batch 2. For every scenario conducted the performance of the group with Conning Display and the group without was compared. As there were only 3 people in every scenario this led to the fact that there were always a 1 to 2 distribution in terms of participant assignment to conditions.

It can be seen that the errors occurring in the scenario seem to be roughly equally distributed over the conditions of the Conning Display in batch 1 (Table 5).

**Table 5. The number of errors that were committed in total for batch 1.**

<table>
<thead>
<tr>
<th>Batch 1</th>
<th>Conning Display</th>
<th>Errors/person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>On</td>
<td>Off</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

The number of errors is equally high (3).

However, if we take into account that the crews were not distributed equally over the conditions and weight the errors thereafter, the “on” condition is associated with 25 percentage points less probability for an error to occur (Table 6).

**Table 6. Error probabilities batch 1.**

<table>
<thead>
<tr>
<th>Batch 1</th>
<th>Conning Display</th>
<th>(Errors / Try) / Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>On</td>
<td>Off</td>
</tr>
<tr>
<td>1</td>
<td>1/3</td>
<td>1/1</td>
</tr>
<tr>
<td>2</td>
<td>1/2</td>
<td>1/1</td>
</tr>
<tr>
<td>3</td>
<td>1/1</td>
<td>1/2</td>
</tr>
<tr>
<td>Errors / tries</td>
<td>3/6</td>
<td>3/4</td>
</tr>
<tr>
<td>Error Probability</td>
<td>0.50</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The probability for committed errors is 25 percentage points larger without the Conning Display than with it. 1/2 thus signifies that 1 of 2 crews conducting the scenario made one or several errors. “Try” refers to how often crews conducted a scenario.

When comparing the errors per person it turns out that the crews without the Conning Display committed 4 errors while the crews in the other condition produced 0 errors.

**Table 7. The errors per persons for batch 2.**
### Table 8. Error probabilities batch 2.

<table>
<thead>
<tr>
<th>Errors/Person</th>
<th>Conning Display</th>
<th>Errors/person</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On</td>
<td>Off</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td><strong>total errors /condition</strong></td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

4 errors were committed with Conning Display "off" and 0 with the Conning Display "on".

The probability for committing an error per exercise was 50% for the crews without the Conning Display.

Table 8. Error probabilities batch 2.

<table>
<thead>
<tr>
<th>Batch 2</th>
<th>Conning Display</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors/Person</td>
<td>(Errors / Try) / person</td>
</tr>
<tr>
<td>4</td>
<td>0/3 1/1 1/4</td>
</tr>
<tr>
<td>5</td>
<td>0/1 0/3 0/4</td>
</tr>
<tr>
<td>6</td>
<td>0/2 3/2 3/4</td>
</tr>
<tr>
<td><strong>Errors / Tries</strong></td>
<td>0/6 4/6</td>
</tr>
<tr>
<td><strong>Probability</strong></td>
<td>0.00 0.50</td>
</tr>
</tbody>
</table>

The probability for committed errors was 0 when the Conning Display was “on”.

Putting it together, we get the probabilities 60 % for errors without the Conning Display and 25% for scenarios with the Conning Display.

Table 9. Error in a scenario in relation to the condition of the Conning Display.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Conning Display: Off</th>
<th>On</th>
<th># of crews that made errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch 1</td>
<td>1</td>
<td>2/2</td>
<td>1/1 3/3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1/1</td>
<td>1/2 2/3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0/1</td>
<td>0/2 0/3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>XXX</td>
<td>1/1 1/1</td>
</tr>
</tbody>
</table>

| Batch 2  | 1                    | 0/1| 0/2 0/3                     |
|          | 2a                   | 2/2| 0/1 2/3                     |
|          | 2b                   | 1/2| 0/1 1/3                     |
|          | 3                    | 0/1| 0/2 0/3                     |
| **Tries associated with error / total tries all batches** | 6/10 | 3/12 |
| **Probability** | 0.60 | 0.25 |

### 7.3. Quantitative method reflections

Inference methods like e.g. the independent and dependent t-test were intentionally excluded as the number of participants that are available for between group comparisons is very low (3 per scenario). The within group t-test was also intentionally left out as the scenarios seem to differ too much in terms of difficulty level and the number of participants was very low here too (N = 6). Additionally,
carry-over effects could be present in both between and within group comparisons from being exposed to both conditions of the Conning Display (practice effects). Thus, differences in terms of number of errors have to be interpreted with caution as they could be heavily influenced by intervening variables.

When looking at scenarios associated with errors in this study differences emerge. It can be noted that every crew commits errors in the first scenario in the first batch, regardless of whether the Conning Display is turned “on” or “off”, suggesting that the error occurs due to factors not taken into consideration when designing the display. However, it appears that the error occurs more frequently when taking the number of “tries” into account for every condition.

3 of 3 crews make errors in the first scenario of batch 1, one crew committed the error with Conning Display and one committed the error without the aid of the Conning Display. This also suggests that the Conning Display is not effective in preventing these errors.

None of the crews committed errors in the third scenario indicating that scenario 3 could differ from the other two scenarios in some way, which could be true: The crews can see each other and there is no current in this scenario compared to the other 2 scenarios. In a general sense this scenario tests the student’s capacities of avoiding other ships in a small passage (channel) while the other 2 scenarios in batch 1 center around how the students cope with environmental factors like currents and obstacles (e.g. berthed ships). So the comparability of this scenario with the others is likely to be questionable, something that has to be considered when evaluating the within group comparison as these were done under the assumption of complete comparability between the scenarios.

Table 10. Error in a scenario set into relation to the condition of the Conning Display

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Conning Display:</th>
<th># of crews that made errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Off</td>
<td>On</td>
</tr>
<tr>
<td>Batch 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2/2</td>
<td>1/1</td>
</tr>
<tr>
<td>2</td>
<td>1/1</td>
<td>1/2</td>
</tr>
<tr>
<td>3</td>
<td>0/1</td>
<td>0/2</td>
</tr>
<tr>
<td>4</td>
<td>XXX</td>
<td>1/1</td>
</tr>
<tr>
<td>Batch 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0/1</td>
<td>0/2</td>
</tr>
<tr>
<td>2a</td>
<td>2/2</td>
<td>0/1</td>
</tr>
<tr>
<td>2b</td>
<td>1/2</td>
<td>0/1</td>
</tr>
<tr>
<td>3</td>
<td>0/1</td>
<td>0/2</td>
</tr>
<tr>
<td>Tries associated with error / total tries all batches</td>
<td>6/10</td>
<td>3/12</td>
</tr>
</tbody>
</table>

"Try" refers to a crew conducting an exercise.

When looking at the errors in Batch 2, it becomes evident that the second scenario is associated with the most errors. 2 of 3 crews made errors in this scenario both in the first and the second part of it. The third scenario was not associated with any errors although this scenario seems to resemble the scenarios in the first batch, scenario 1 and 2. But unlike scenario 1 in batch 1 this exercise is conducted with a smaller and hence more maneuverable ship. And in contrast to scenario 2 in batch 1, this scenario does not confront the students with current breakpoint water (see qualitative
results). Scenario 2a in batch 2 was conducted with a very small ship and is the only scenario in which the crews have to cast off which can cause pose challenges of its own and thus limit its comparability. For the within design this means that differences in terms of errors could thus be attributed to the scenarios being different, limiting comparisons between them.

However, although no statistical inference method has been used it, seems as if the Conning Display influences performance in a positive way (although performance seems to degrade when the display is switched "off" from being "on"). Therefore it could be assumed that real performance improvement is due to good information display in the Conning Display. Thus, if these improvements stem from the Conning Display, the associated functions that improved behavior should be preserved in the new design proposition.

This was originally envisioned to be done by analyzing the eye-tracking data with analysis software in order to quantify what functions were looked at. However, due to technical difficulties this proved to be impossible. Therefore, the interviews will be used as indicator for what functions that could have a positive effect on performance.
8. Qualitative Results

8.1. Interviews
In order to extract the functional requirements that the students regard as useful, interviews were conducted with the participants of the maneuvering scenarios (N=6) and with some of the participants of the watchkeeping scenarios (N = 5). See Appendix A for further information concerning the interview questions.

8.1.1. Maneuvering
In the maneuvering exercises 4 participants stated that they used the display. These 4 were questioned further about which functions they used in the display. The number in front of a function shows how many of the 4 participants mentioned it.

*Which functions did you use in the Conning Display? (N=4)*
- 4x rate of turn (Drehgeschwindigkeit)
- 3x speed (Geschwindigkeit)
- 3x Course over ground (Kurs über Grund, Fahrt des Schiffes)
- 3x Course through water (Kurs über Wasser, Fahrt des Schiffes)
- 2x Rudder angle (Ruderlage)
- 1x machine output (Maschinenleistung)
- 1x echo sounder (Echolot)

8.1.2. Watchkeeping
Only some of the interview participants in the watchkeeping exercises used the Conning Display (N=3). Again the number before the function indicates how many participants found it useful.

*Functions the participants found useful (N = 3)*
- 2x Wind
- 2x Current (Strom)
- 2x Type of ship (Schiffstyp)
- 1x Compass (Kompass)
- 1x Course over ground (Kurs über Grund)

8.1.3. Combined requirements
By adding the requirements gathered in the maneuvering exercises to the ones collected from watchkeeping, following order emerges.

*Useful functions (n = 7)*
These are the functions that should probably stay in the new layout of the Conning Display as they have proven themselves useful in watchkeeping and maneuvering exercises.

- 4x rate of turn (Drehgeschwindigkeit)
- 4x Course over ground (Kurs über Grund)
- 3x speed (Geschwindigkeit)
- 3x Course through water (Kurs durchs Wasser)
- 3x drift from current and wind (Versetzung: Strom, Wind)
- 2x rudder angle (Ruderlage)
- 2x ship type (Schiffstyp)
- 1x engine output (Maschinenleistung)
- 1x echo sounder (Echolot)
- 1x compass (Kompass)

\section*{8.2. Observations}

3 different persons conducted the first batch consisting of 3 scenarios while 3 persons took on the second batch, consisting of 3 scenarios. For each of these persons the Conning Display was switched on and off for each scenario and the results will be structured thereafter in the following section. In this analysis, the characteristics of the errors committed under both conditions were of interest and if these characteristics differ or are similar on any dimensions. Thus, the errors and their systematic explanation will be presented according to Conning Display condition.

\subsection*{8.2.1. General approach}

The error sequences will be presented in 3 steps. First, the error that occurred will be made explicit. Then, performance-shaping factors will be presented (factors that influenced the crew’s behavior) and the cognitive couplings and resolving mechanisms will be set into relation to these factors. Finally a possible design solution will be presented, addressing the issues identified, will be proposed.

The design solutions always aim at preventing an error as encountered by the crew in the simulator exercises from happening. However, it is always possible to recommend that the crew could have chosen an alternate route altogether, thus avoiding the specific error completely. This is not regarded as a solution in this study, as error situations often arise from what were perceived as reasonable decisions beforehand. That is why a (design) solution must be as close to the encountered situation as possible. Alternative solutions tackling the problem further down in the chronological chain are only considered if no solution can be found in the immediate situation. This approach should ensure that the amount of possible solutions is reduced greatly, thus focusing the analysis of error causes on the immediate situation as (reasonably) encountered by the participants.

Additionally, all of the behavior observed in the following section can be classified as being involved rather than isolated (in terms of cognitive couplings) as the captain always is present at the working environment, therefore this classification will not be mentioned. Furthermore, no indications for external compatibility or incompatibility could be deduced from the data. Thus this cognitive coupling will not be mentioned in the cognitive coupling tables in the following text either.

\subsection*{8.2.2. Conning Display “off”}

\textit{Batch1, scenario 1, person 1}

\textbf{Error}

After the first turn the crew initiates another turn towards the middle of the waterway. When approaching the last turn it can be seen that the crew has given a starboard rudder command but that the ship does not react to it immediately as it still turns towards portside as a result of the previous rudder command (to get more central in the waterway). The ship leaves the designated track (Figure 12).
Figure 12: The ship leaves the waterway. This constitutes a violation of exercise instructions (white circle).

Performance shaping factors
The passage through the designated waterway should be characterized by closed-loop control by the captain as he continuously has to adjust the ship’s heading in order to stay within the limits posed by the radar corridor (Figure 13).

Figure 13. The captain initiated rudder commands towards the central radar line (yellow line in the middle). The ship (black circle) reacted too late to this command making it impossible to take the necessary turn.
However, the ship’s slow reaction to rudder commands complicates matters. As stated in Sträter (2005), tracking tasks requires the system to provide feedback within 2 seconds of action execution. Otherwise the operators actions seem to be decoupled from his environment which, in other words, means that the captain does not feel that the ship reacts as directly to the rudder commands given (as the ship reacts with too much delay). To further complicate matters, the crew is not yet completely familiar with the ship’s reaction to rudder commands as they handle the ship type for the fourth time and for the first time under the conditions of this scenario.

In the order to master the scenario, the crew has to turn at least twice, each time with the current affecting from a different direction. This makes the prediction of the turn radius very difficult. Thus, it is not surprising, that the captain stated that the biggest challenge in this exercise was to figure out the interdependency between rudder angle and the ship’s manoeuvring behaviour. The current is a dimension that has to be put into relation with other parameters mentally in order to discern how the ship will turn (e.g. speed, rudder angle and type of ship). Therefore, the captain interacts with a system that is multidimensional in relation to the task of turning the ship and thus correspondingly puts a high level of stress on him/her.

When the rudder command has been given, the captain has to monitor the ship system in order to determine if the command will have the expected effect, in other words perform an evaluation of the actual state with the expected state (Hollnagel, 2005) (cognitive coupling: monitor, stressful). This evaluation is complicated by the fact that the system does not display actual and desired state in itself. The captain has to compensate for this by deducing these states (compensatory information display, stressful). This is more demanding than if the system would provide both states in relation to each other for the task in the technical system.

The necessary operations are parallel as the captain may choose freely when to initiate the turn and when to monitor the ship’s reaction (parallel, stressful). The lack of system displays showing actual and desired state leads to the fact that the captain has to deduce how the ship will react to rudder commands. Additionally, the captain has not had any experience with this ship type under these conditions (e.g. current) and therefore probably has not had time to develop an attuned feeling for the cause-effect relation of rudder commands and ship reaction (situation with low internal compatibility, stressful).

This scenario is also a good example for how delayed feedback on a slightly incorrect prediction may contribute to errors. If the prediction is not accurate, then the user should get feedback directly in order to correct it in time, otherwise corrective actions may occur too late. According to Hollnagel (2005) the ability to react proactive in a system “[...] means that the feed forward control every now and then is interrupted by an evaluation of the current state, more specifically of the actual state with the expected state. Any differences lead to a correction – possibly also a correction of the underlying model – after which the system can proceed.” (Hollnagel, 2005, P.138).

As stated before, this is not the case in this scenario as the system does not provide the right form of information for this comparison of current and desired state and the student lacks the experience to predict the ship’s behavior without external aids.
Another factor that should be considered is the scope of the prediction, how far in time it should stretch out. It is not viable to initiate an action that may lead to the accomplishment of a sub-goal but endanger the main goal.

In this exercise both accuracy and scope of the prediction proved to be problematic. The captain does not get immediate feedback on how well the rudder commands will translate into the desired sub-goal (turning the ship) and is thus not immediately able to discern in what way actual system state differs from the expected state. Additionally, one of the sub-goals that could be observed during the exercise was to position the ship further in the middle of the waterway by means of rudder commands. According to the instructor, this leads to a better starting point for subsequent turning. However, when this was done by the crew, the next sub-goal, to initiate the subsequent turn, was harder to fulfill as the ship already had a rate of turn in the opposite direction stemming from the previous turn towards the middle of the waterway.

This has to do with the fact that it is hard to correctly predict if the rudder command, when given, will lead to the desired state. The rudder commands, that the crew execute, will take some time to have effect (it can take up to 5 minutes for the effect of a rudder command to be observed), therefore the crew “lags behind” the actual events.

Actions that are executed in this way, without fast feedback on correctness, are by definition open loop. Unfortunately for the novices, they often lack the experience with the ship necessary to perform the correct open loop action. Thus, they face the difficult situation of controlling a task that should be of closed loop but instead is of open loop nature. Open-loop control is the opposite of what the situation demands, namely continuous reaction on external threats and conditions. However, if they would get immediate feedback on their actions (rudder commands), they could find the correct action in a step-by-step manner through continuous adjustment of the ship’s rudder (closed loop).

As mentioned before, the feedback of the system is only visible after a while as it takes time for the rudder command to show effect. When the effect is visible, however, it takes as much time to reverse it by another rudder command. Therefore wrong predictions cannot be reversed easily as the ship reacts too sluggish for effective counter measures. It is thus not surprising, that the captain during this exercise stated that the biggest challenge was to imagine how the ship would react to rudder commands.

When setting these events into relation with the resolving mechanisms, the fact that the cognitive system requires stability to operate has to be considered. Actions can only be executed if the cognitive system is in an equilibrium state. This equilibrium state means that the captain has levelled all dissonance and therefore is able execute an action. This entire error sequence results from the fact that the rudder command was given too late, an error that is typically attributed to the resolving mechanism information and goal overload that is characteristic for high demanding/stressful cognitive coupling modes as summarized in Table 11.

Table 11. Cognitive Coupling types.

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Involvement</td>
<td>(+) Monitory</td>
<td>Monitors effect of Rudder</td>
</tr>
<tr>
<td>Type of Control</td>
<td>(-) open-loop</td>
<td>Giving rudder commands with delayed feedback in order to adjust the ship's heading to the desired heading</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, Current, Wind, ship inertia, etc</td>
</tr>
<tr>
<td>Necessary Operation</td>
<td>(+) Simultaneously</td>
<td>Can choose freely when to do what</td>
</tr>
<tr>
<td>Information Presentation</td>
<td>(+) Compensatory</td>
<td>Difference between actual state and desired state is not displayed</td>
</tr>
<tr>
<td>Internal Compatibility</td>
<td>(+) Internal Incompatibility</td>
<td>Unfamiliarity with ship’s reaction to rudder commands</td>
</tr>
</tbody>
</table>

**Design proposition**

An exercise that requires closed loop control cognitive coupling but is per nature an open loop control coupling (feedback delayed more than 2 sec), can only be effective if the closed loop control is established again. According to Mehl, Mesenholl & Nachreiner (1993) there are two solutions to delayed feedback: Training of the people using the system or technical support that enables dynamic effect projections of user actions. In the latter case, the display present in the instructor station showing how the relevant factors affect the ship’s heading could be integrated in the Conning Display.

**Batch 1, scenario 2, person 2**

**Error**

When traversing the entrance to the harbour the ship starts to turn starboard and collides with the starboard quay.

**Performance-shaping factors**

The error here was that I did not know that it was possible to give rudder commands that are that big. I did not know that. I only told him [the helmsman] “hard to starboard” and because it [the command] was given in that way he naturally turns the rudder hard to starboard. I did not think that the ship could do rudder commands that are that great. (81-86)

When approaching the harbor opening, the ship is affected by the current, drifting it southwards towards the southern mole of the harbor. Probably intending to ensure that the ship does not drift into the southern mole, the captain gives the command to correct the ship’s movement. He executes this by saying “hard to starboard”. Unfortunately “hard to starboard” is a relative command that only makes sense in relation to the ships general rudder span. Probably used to smaller rudder spans, the
captain expects the ship to turn fewer degrees than it actually does as the helmsman initiates the turn. After the typical rudder command-effect delay mentioned in the previous error, the ship starts to turn towards the northern quay very fast. Adding to the already high rate of turn towards starboard is the breach of breakpoint water, leaving only the stern affected by the current, further boosting the rate of turn towards starboard.

The captain resorted to behavior that was successful in other domains. For him the fact that "hard to [direction]" is a command that is relative to the ship’s rudder span was not obvious at that moment. He had probably used the command a lot on other, known ships and associated it with a particular effect on the ship’s heading and thus did not think about the relative nature of the command. Thus, he transferred behavior appropriate on ships with a smaller rudder span to a ship with a much bigger rudder span as it seemed to be internally compatible but turned out to be incompatible (internal incompatibility).

The other cognitive couplings in this error sequence are essentially the same as in the previously mentioned event. The task is multidimensional (speed, current, rudder angle etc., all have to be put into relation for the turn that is to be made), compensatory (no pursuit display showing future course and actual course in one integrated view), simultaneous (instrument checks and rudder commands can be executed freely, in an order determined by the captain), monitive as the captain has to monitor how the rudder commands effect the ship’s heading and open loop as the captain only gives a single verbal rudder command to the helmsman for initiating the turn.

Probably having given that command a lot in the past, the captain did not notice cues that could have made him realize the potential danger in the way the command was given. This is probably due to the fact that the feedback in terms of an effect on the rate of turn of the ship does not show until after the rudder command has already been given due to the delay in ship reaction. In that stage, it is already hard to counter as the “counter rudder” also takes some time to take effect as evident from the following citing:

[…And it takes quite a while before the rudder gets back to straight ahead. That was probably the point that ruined it in the end (88-89).]

Thus, the way a rudder command was posed leads to the error for which the captain could not find a solution besides giving counter rudder which was executed to late (typical for information and goal overload).

Table 12. Cognitive Couplings Batch 1, scenario 2, person 2

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitory</td>
<td>Monitors Effect of Rudder commands/displays</td>
</tr>
<tr>
<td>Type of Control</td>
<td>(+) Closed-loop</td>
<td>Gives a single rudder command to helmsman</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, Current, ship</td>
</tr>
</tbody>
</table>

Table 2. Cognitive Couplings Batch 1, scenario 2, person 2
<table>
<thead>
<tr>
<th>Necessary Operation</th>
<th>(+) Simultaneously</th>
<th>Can choose freely when to execute actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Presentation</td>
<td>(+) Compensatory</td>
<td>Difference between actual state and desired state is not displayed</td>
</tr>
<tr>
<td>Internal Compatibility</td>
<td>(+) Internal Incompatibility</td>
<td>Relative rudder command with wrong rudder type in mind given by the captain to the helmsman</td>
</tr>
</tbody>
</table>

**Possible solution/design proposition**

It is generally hard to tell what the crew should have done in this situation. It seems as if the situation they found themselves in could have been prevented if appropriate feedback on rudder command outcome would have been given by the system without delay (as mentioned in the previous error). Then the crew could have executed counter rudder directly after giving the wrong rudder command, thereby countering the effect. Another solution would be to clearly display the rudder span for the captain making it implicitly clear what an outcome relative rudder commands would have.

**Batch 1, scenario 2, person 3**

**Error**
The crew leaves the designated waterway.

**Performance-shaping factors**

Yes, and that was already the current [affecting] but that was actually okay for us as we wanted to turn relatively early around there. But it turned out that the current there was too strong for the maneuver (405–407).

You know the speed of the current, but I personally find it quite hard to conclude from that when I have to start turning so that I can get around there in a good way and not drift away. (435–437)

The captain stated that he did not expect the current to affect the turn of the ship to such an extent as was the case in this scenario. When setting this into relation with the cognitive coupling dimensionality this means that the current must be taken into account not as an influencing constant but as an influencing variable that has different effects on the ship depending on at which angle it hits the ship. The current simply affects the ship differently depending on how the ship is located towards it, so when the ship turns the effect of the current will be quite different from one turn degree to another.

The turn was initiated too late, indicating that the captain was under information and goal overload, which resulted in the correct action being initiated too late. The following citation illustrates this point.

Yes, because if we start to turn here, then the current will push us, further and further. That means that our entire curve will become prolonged by the amount of speed that the current
vector adds. And that is something that I never would have expected to be that strong. If I do a
normal course change then I of course expect that the ship will complete the turn in this and that
time and in this and with this and that radius. [...] And if the current adds up to everything then
everything gets incredibly prolonged. [...] I thought that we would make it with that rudder
command. But it didn’t work. Then, you have given a stronger rudder command but it is already
too late. (409-417)

Table 13. Batch 1, scenario 2, person 3

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitory</td>
<td>Monitors ship’s reaction to rudder commands (the ship</td>
</tr>
<tr>
<td></td>
<td></td>
<td>starting to drift unexpectedly e.g.)</td>
</tr>
<tr>
<td>Type of Control</td>
<td>(+) Closed-loop</td>
<td>Has to react dynamically to changing scenario demand (change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>of route and current, e.g.)</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, ship inertia,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>current changes while turning</td>
</tr>
<tr>
<td>Necessary Operation</td>
<td>(+) Simultaneously</td>
<td>Can choose freely when to do what</td>
</tr>
<tr>
<td>Information Presentation</td>
<td>(+) Compensatory</td>
<td>Difference between actual state and desired state is not</td>
</tr>
<tr>
<td></td>
<td></td>
<td>displayed</td>
</tr>
<tr>
<td>Internal Compatibility</td>
<td>(+) Internal Incompatibility</td>
<td>Mapping between rudder commands and ship reaction is unclear</td>
</tr>
</tbody>
</table>

Possible Solution
Although the drift is displayed by the log, it does not give direct information on the extent future
course changes will be affected by the current. The correct action would have been to give a higher
rudder command (30° instead of 10°). Thus, some kind of feedback on future ship reactions should
be made available by the technical environment.

Batch 2, scenario 2, person 4

Error
When casting off, the ship scrapes against the quay with the stern part of the ship. During this
exercise the helmsman wore the eye-tracking glasses.

Performance shaping factors
As the interview had to be conducted with the helmsman and not with the captain, the interview
does not give any conclusions about why the error occurred as the captain gave the orders and the
helmsman did not always know why, did not know the intentions behind the commands. However, conclusions about the cognitive couplings of the captain can be derived from the combination of audio and CCTV-video recording.

Figure 14. The view on the distance between ship and quay of the rear part of the ship is hidden from view by the bridge itself (white circle). The visible distance between the foremost part of the ship and the quay is denoted by the shaded white area.

What at the first glance may seem like a minor incident only leading to a few scrape marks, turns out to be much more problematic. According to the instructor the ship’s rudder can be damaged when the rear part of the ship hits the quay. As the rudder is one of the main instruments necessary for manoeuvring this can immobilize the entire ship until the damage is repaired.

During the course of the exercise the helmsman, wearing the eye-tracking camera, attempted to convince the captain not to turn too strong towards starboard as this would, in his opinion, result in the ship not having enough manoeuvring space for the its rear part. The captain however, thought otherwise, as evident from this transcribed excerpt of the CCTV simulator video recording (“H” denotes the helmsman and “C” the captain):

21:18 CCTV

H: I don’t want to turn around yet as our ass [rear part] has to get away from the quay. That’s why I don’t want to give rudder right away. It will happen fast later anyway.

C: But it will go away from the quay anyway [4 sec]. Shall we do it your way or mine?

H: Yes, then tell us what to do.

C: But before that we have to agree.

H: You can decide.

C: Very well, then give 30° starboard and kick ahead.

The helmsman did not want to turn before the rear part of the ship had enough manoeuvring space but the captain overrode this objection. As indicated by this citation, the captain thought that the
ship would not touch the quay during the turn. Additionally, the crew does not get any audio feedback in terms of collision noise when the rear part of the ship scrapes against the quay as evident from the audio recording of the eye-tracking video and the CCTV camera on bridge.

One factor contributing to the captain’s decision to turn anyway could be the fact that the crew did not have much experience with casting off as this was the first exercise where this manoeuvre was required. The way in which the ship's stern can break out could thus be regarded as lacking internal compatibility; the captain does not think that the ship will break out in a safety critical way when turning. This becomes evident from the previously citation in which the captain answers “But it goes away from the quay anyway” to the helmsman’s objection, indicating that she does not think that the ship will collide with the quay during the turn.

Another stressing factor is the fact that the task-relevant actual system state is not displayed in the system (compensatory information display). The desired state consists of the ship not colliding with the quay after rudder commands are given during the cast off manoeuvre. With the current simulator layout the relation of actual and desired state can only be deduced by combining the information of the change in distance to the quay and the change in orientation towards the direction in which the crew wants to cast off and imagining how they will be affected by rudder commands. Therefore the information presentation essentially leads to the compensatory task of deducing this relation. This is further complicated by the fact that the crew cannot completely discern the distance to the quay from the outside view when casting off as this view is obscured by the bridge itself (Figure 14). Fortunately, the crew is able to see the distance between the rear of the ship and the quay through an onboard side view camera (See CCTV video recording 11 minutes 40 seconds “camera is on stern”).

Obviously, the task monitoring of this camera can be seen as problematic. It is positioned at the starboard side of the ship, while the area of interest (white circle, Figure 14) is on the portside. This results in the uncomfortable fact that the person monitoring these two information sources has to shift his/her gaze constantly. While this is not problematic in itself, the fact that looking at one information source makes it impossible to monitor the other, is. Thus, the captain has to look outside, put on memory what the outside looks like, and turn towards the display showing the camera view.

This monitive cognitive coupling of observing both the window and camera leads to an information gap. This gap stems from the fact that the window and the side view camera cannot be monitored at the same time. But for the captain both views are important. He/she needs the side view to discern when the distance to the quay is great enough to initiate a turn away from it. The window view is needed in order to see the ship’s heading in relation to the other ships and objects in the vicinity. The captain has to decide what view to look at excluding other information sources from peripheral vision. This could have contributed to the fact that the captain turned inappropriately.

The task is also multidimensional with the dimensions consisting of all the factors that have to be taken into account when projecting the outcome of a safe rudder command: orientation towards the other ships and obstacles (window view), distance to the quay (side view), velocity, acceleration and steering and thrust commands all have to be considered in order to figure out what the future position will turn out to be.
The type of control that is executed is of open loop and consist of calling the instructor for distance to the quay (the instructor simulates deck personnel that is called). All other actions are carried out by the helmsman.

Table 14. Batch 2, scenario 2, person 4

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitory</td>
<td>Has to monitor distance to the quay and distance to the ships ahead</td>
</tr>
<tr>
<td>Type of Control</td>
<td>(-) open loop</td>
<td>Calling instructor in order to get to know distance to the quay in regular intervals</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, ship inertia, distance to quay</td>
</tr>
<tr>
<td>Necessary Operation</td>
<td>(+) Simultaneously</td>
<td>Can choose freely when to do what</td>
</tr>
<tr>
<td>Information Presentation</td>
<td>(+) Compensatory</td>
<td>Difference between actual state and desired state is not displayed</td>
</tr>
<tr>
<td>Internal Compatibility</td>
<td>(+) Internal incompatibility</td>
<td>Captain does not expect the ship’s stern to “break out” a</td>
</tr>
</tbody>
</table>

The captain reacts to the dissonance of being confronted with potential objections to a proposed course with a fixation, the critique is regarded as unfounded although it turns out not to be.

**Design proposition**

As the crew lacks the experience of casting off, they should have information on how rudder commands will rotate the ship, more specifically they should have information providing insight into the issue of whether they will make the stern collide with the quay or not. This could for example be done with a simple graphical display of the turn radius (on the Conning Display e.g.) with a numerical value behind it, indicating that turns away from land can only be initiated when e.g. a distance of at least 4 meters to land is established. The captain would then know that turning right away can damage the ship and initiate discussions on alternative approaches to the problem.

The captain would also perhaps have acted otherwise if the 2 information displays, camera and outside view, could have been monitored in a better way. This could be done by e.g. locating another display showing side view on the portside of the bridge. The captain could then in one glance see how the distance to the quay changes and what the ships orientation towards other ships is. The previously mentioned “gap” would probably be smaller.

Another solution could be to slightly modify the solution proposed in the exercises from batch 1. If the rotation of the whole ship could be added over the projected course then collisions would be visible directly before the command cannot be reversed due to inertia.
Batch 2, scenario 2, person 6

Error 1
While casting off, the crew does not keep enough distance between the stern and the quay, resulting in the stern scraping against the quay during the turn.

Error 2
When the crew approaches the quay where they should and want to berth the crew fastens the bow hawser too soon resulting in the ship hitting the quay due to the rubber band effect.

Error 3
The crew tries to berth again and collides with the quay again due to the aforementioned rubber band effect.

Performance-shaping factors
Several errors occurred in the course of this exercise.

Error 1:

Figure 15. Batch 2, scenario 2, person 6: The distance between ship and quay (shaded area) is hard to tell as a part of it is hidden from view by the bridge itself (white circle)

I don’t think that we heard it. Okay, yes we thought that we simply give rudder towards the quay and try to get the stern away from the quay. What our problem consisted of was simply that we chose an angle that was too small. We did not swing away completely, that is 40, 50 °, like we should have (254-257).

The crew decided to cast off in bow direction. In order to accomplish this, they swung away the bow so they would not collide with the ship berthing straight ahead of them. After completing this, the crew initiated the right action, they gave a rudder command to the left (portside) direction towards the quay in order to get maneuvering space for the stern part of the ship. However, when the turn was initiated, the back (stern) part of the ship scraped against the quay anyway, which in real-life situations can result in critical damage to the rudder. Thinking that they had enough maneuvering space, the crew simply did not notice that the distance between stern and quay was too small.
In the interview the captain stated that he thought that the error occurred because the angle at which the bow was pointing away from the quay was too small. Talking to the instructor during the exercises revealed that he considered the lack of space between the stern part of the ship and the quay to be the cause of the error with the correct action being the initiation of counter rudder in the direction of the quay in order to create maneuvering space. This correct action was carried out by the crew however, it was not forceful enough, in the sense that the ship did not turn sufficiently towards the quay to free enough space for the stern.

Much like in batch 2 person 1 scenario 2, the crew did not have much experience with casting off as this was the first exercise where this maneuver was required. The way in which the ship’s stern may “break out” could thus be regarded as lacking internal compatibility, the captain does not think that the ship will break out in a safety critical way when turning but rather that the turn will be within safe limits.

More interesting is the information presentation as this crew lacks one source of information that Batch 2, scenario 2, person 4 had available, the side view camera (as evident from the eye-tracking recording). This means that they have no possibility to actually see the distance to the quay as the view on this section is obstructed by the bridge layout. The captain relied on another method, he called the instructor at regular intervals to get to know the exact distance to the quay.

However, this information from the instructor is only a snapshot of the actual distance, making it hard to infer the rate of change of the distance to the quay. And this inferred rate still has to be set into relation with the multiple dimensions, exactly as in Batch 2, scenario 2, person 4 (position towards other ships, acceleration, etc). The necessary operations are again simultaneous as the captain can choose freely when to initiate which actions.

It is also noteworthy that the captain does not interact actively with the technical environment except when calling the instructor. Everything else is accomplished by means of communicating with the rest of the crew, mainly by giving commands. Thus, the captain exerts open-loop control over the ship and monitors the instruments as these commands are carried out.

The cognitive system is relatively resistant against dissonance once a choice of action has been made. The captain probably fixated on his command when he initiated the turn especially since feedback on the collision is missing (no “crash sound”) and the appropriate maneuver was executed.

Table 15. Cognitive Couplings Batch 2, scenario 2, person 6

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitor</td>
<td>Has to monitor distance to the quay and distance to the ships ahead</td>
</tr>
<tr>
<td>Type of Control</td>
<td>(-) open-loop</td>
<td>Giving rudder commands to the helmsman</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, ship inertia,</td>
</tr>
</tbody>
</table>
Error 2 and Error 3:

Yes, we definitely did not think that the ship would react that extremely. Because we weren’t that fast, I can’t remember how fast (309-310).

When the crew approaches the quay and berths, the bow of the ship is immediately drawn towards the quay and collides with it. According to the instructor, this occurs due to the fact that the crew gave the command to fasten the bow hawser at too high a speed. As can be seen in the interview citing, the captain simply did not think, that the ship’s speed was too high for berthing. It is not unusual to have some speed on the ship when berthing, especially if handling a small ship. The crew then uses the hawsers in order to lose the rest of their speed and to get closer to the quay.

The captain deems the speed be low enough for berthing (approximately 0.7 knots) and gives the command to fasten the bow hawser. As the hawser tauten, the ship’s heading changes very fast towards the landside which results in the ship hitting the quay.

As evident from the citing he/she did not know what this presumed safe speed would lead to such an extreme reaction when berthing. According to the instructor a solution would be to give a rudder command in the opposite direction of the quay, so that the rate of turn towards the quay is countered. But, due to ship inertia, this command has to be given before fastening the hawsers which means that the captain must know how the ship will react before berthing at that speed.

This was not the case in this exercise as the captain did not know that the ship would react as extremely as it did. This leads to the conclusion that the accident could not be prevented once the hawsers were fastened.

No, [there is] nothing [you can do]. It pulled us like a rubber band and it happened so fast that we were already at the quay as we noticed it (320-321).

Thus, a design solution must aim at communicating the safe speed to berth at by displaying this information. This conclusion is further underlined by the fact that the crew made this error twice during the exercise.

The task is mainly monitory as the captain has to *monitor* the ship’s displays in order to discern when to berth. The berthing task is *multidimensional* as it consists of (at least) 4 dimensions that have to be
set into relation to imagine the ship's reaction namely the ship's speed, its orientation towards the quay (parallel or orthogonal), the distance to the quay and the rubber band effect by the hawser. The operations are simultaneous, the captain can choose freely when to execute which commands and has to coordinate these with each other.

The berthing command is given under open-loop control by giving the instructor a verbal command. As open-loop control is executed once and does not consist of a continuous regulation of a parameter. Therefore this one command must be very precise. Thus, the system should provide adequate support for finding this “right” command which carries implications for the next cognitive coupling, the information presentation.

The information presentation in this exercise is compensatory. The captain has to find out the safe speed to berth at himself as the actual system state (that should signal safe speed) is not visible in the external world. As a result he berths at a low speed which he considers as safe and is surprised by the extreme result that is incompatible to the ways he thought the ship would react (internal incompatibility). In the phase after the ship had fixated the hawsers, the captain did not know what to do (information and goal overload).

Table 16. Cognitive Couplings Batch 2, scenario 2, person 6

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitory</td>
<td>Has to monitor distance to the quay and distance to the ships ahead</td>
</tr>
<tr>
<td>Type of Control</td>
<td>(-) open-loop</td>
<td>Watching the ship as the rudder commands are being executed by the helmsman</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, ship inertia, distance to quay</td>
</tr>
<tr>
<td>Necessary Operation</td>
<td>(+) Simultaneously</td>
<td>Can choose freely when to do what</td>
</tr>
<tr>
<td>Information Presentation</td>
<td>(+) Compensatory</td>
<td>Difference between actual state and desired state is not displayed</td>
</tr>
<tr>
<td>Internal Compatibility</td>
<td>(--) Internal compatibility</td>
<td>Ship’s reaction to the rubber band effect is not in accordance with captain’s expectations</td>
</tr>
</tbody>
</table>

**Design proposition**
As mentioned before the Conning Display should provide the safe speed to berth at. Thus, the compensatory task of finding out what speed to berth at is transformed into a pursuit task which consists of adjusting the actual state to the displayed desired state. By doing this one of the basic
needs of open-loop control (that the speed adjustment command is precise) is satisfied and the ship will react in a way that is in accordance with the captain's expectations and thus internal compatible.

8.2.3. Conning Display “on”
The errors during which the Conning Display was in the condition “on” will be presented.

Batch 1, scenario 4, person 1

Error

Figure 16. Error Batch 1, scenario 4, person 1

Yes, [we] crashed into the mole [...] (12-13)

While attempting to pass through the harbor opening, the crew collides with the northern mole.

Performance-shaping factors
The crew sat their heading towards the northern part of the mole, thinking that the current would drift them southwards into the middle part of the harbor opening. During the course of the exercise they thus were approaching the opening too far north, as they estimate of the magnitude of the current was too pessimistic (253-255).

Because we actually were a bit too far north for the entrance. We were basically too pessimistic concerning the drift. (253-255)

While in the direct vicinity of the northern mole, the crew checked the side view and concluded that they would probably make it anyway. This assumption was proven true until they hit the breakwater point.

Yes you can see it with the help of the camera system that you drift towards starboard. I look on it all the time and say: “that could still work out!”.

Yes, as I said, the bow was carburized [...]. (92-95)
The breakwater point is a point in the water, where one part of the ship (stem) is positioned in the current and the other part (bow) is located in waters without current. The ship consequently starts to turn, in this case towards starboard in direction of the northern mole. The rate of turn is increased if the ship is going at low speed as the current hitting the stern part of the ship has longer time to affect it, thus accelerating the rate of turn.

Exactly, well we weren’t as central as we originally had intended. Well okay, we wanted to pass it quite narrowly and as slow as possible. (247-248)

The statement points out the fact that the crew was going with low speed on purpose when hitting the breakwater point (247, 248). Consequently these circumstances forced the ship to turn towards the mole, a situation the crew was not expecting and in which the captain perceived himself as being too passive (108-110).

No, no I think we did not react fast enough, we did not hit the thruster down, forward. We were too passive in this situation, we let it drift too much. (108-110)

**Possible Solution**

The situation could have been mastered if the crew had initiated a kick ahead in time. This maneuver involves hitting full power with the thruster and giving a rudder command in the direction that is demanded. This ensures that the ship briefiy gets the speed necessary to be maneuverable. For this crew, kick ahead would have meant to accelerate the movement towards the mole for a short period of time, until the rudder command has effect. The behavior necessary, initiating kick ahead, could be seen as incompatible to the behavior the crew is used to from other domains of life. The problem could is easier to understand if transferred to a new domain. Who would for example accelerate a car if heading towards a wall in order to avoid it? Normally, one would hit the brakes. Therefore accelerating is something that does not come directly to mind, as it is a behavior that is remote from what one would usually do. This also becomes evident from the following interview statement (101-106).

In that case you first have to give full speed to get the appropriate pressure on the rudder blade. And that’s simple, if no water is pushed, then there is no rate of turn, no doubt. But if you head like that, towards a wall, then you think logically, like in the car, that you should perhaps reverse the speed. And naturally that does not work on a ship (101-106)

Thus, the behaviour that would be correct in this context simply suffers *internal incompatibility* in the maritime context. At the same time the captain has to handle *multiple dimensions* (current, speed, rudder angle). The task is *compensatory* in the sense that the captain does not have access to direct information of current state (heading into the mole) and desired state (safe conduct). The necessary operations are *simultaneous* as the captain can choose freely when to initiate actions and the captain has to handle a *monitive* task in watching the side view camera and the other parameters in order to discern when to initiate the kick ahead.

Due to the high levels of stress the cognitive system struggles to establish equilibrium and thus appears to hesitate in initiating the kick down and actually initiates it too late, when the collision was already a fact (*information and goal overload*). No control is exerted on the system as the open-loop
control of initiating the kick ahead was done when the error was already had occurred. This is evident in both video and Interview data (Figure 17, statement 118).

Oh my, that’s it. Now I put the [thruster] lever on the table, but I think that the collision has occurred already. (118)

![Figure 17](image.png)

*Figure 17. The captain initiates the kick down too late, when the crash is already audible*

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitory</td>
<td>Monitors distance to the mole</td>
</tr>
<tr>
<td>Type of Control</td>
<td>------------</td>
<td>Open-loop initiation of kick ahead done after error occurred</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, Current, ship inertia, etc</td>
</tr>
<tr>
<td>Necessary Operation</td>
<td>(+) Simultaneously</td>
<td>Can choose freely when to do what</td>
</tr>
<tr>
<td>Information Presentation</td>
<td>(+) Compensatory</td>
<td>Difference between actual state and desired state is not displayed</td>
</tr>
<tr>
<td>Internal Compatibility</td>
<td>(+) Internal Incompatibility</td>
<td>Car analogy proves counterproductive</td>
</tr>
</tbody>
</table>

**Design proposition**

You know your rate of turn, we did have a straight ahead course, we knew that we would drift towards the side [...] and from that we could have drawn the conclusions, that you have to recover the maneuverability of the ship by [inducing] the appropriate revolutions on the propeller. That did not work out in this case. (124-129)
The captain knows that the ship is turning in an unintended way and also knows that he has to initiate the kick down in order to regain control of the already turning ship. But he hesitates and executes the kick down when the ship has already collided with the mole. The Conning Display could break the indecisiveness of the captain by displaying clear advises for action in this situation. The relevant information that should be displayed in this situation is the difference in current affecting the parts of the ship and the speed the ship is running at. All of this information is already present in the Conning Display. What remains to be done is to establish the critical link the different parts of information. So, if a ship runs at low speed and enter waters where one part of the ship is affected by current and the other isn't then an arrow connecting this information could start to flash and the word “kick down?” could appear on the screen. The captain then immediately knows what action to take and that it is legitimate in this context (although the action per se may seem unintuitive).

Another solution would be to integrate the instructor view of the course over ground and the course over water in the Conning Display main view. Then the crew would have seen right away that their course would lead to a collision further ahead in time.

The design solution has to be located somewhere in the vicinity of the thrusters or the side view display as these are the captain's main elements of interest during this error sequence. Therefore, if a design solution is to be implemented, it should be placed somewhere in the vicinity of these instruments to ensure that the captain does not lose time and information by searching all over the bridge. Especially since, in this error sequence the time frame for selecting and executing an appropriate action is extremely small (captain's estimate about 10-20 seconds, 276-277). The following pictures should illustrate these constraints(Figure 18 and 19).

*Figure 18. Side view camera highlighted in red*
The Conning Display is not visible when the captain looks through the side view camera (Figure 18) but the more when he intends to initiate the kick ahead with the thruster (Figure 19). But he initiates the kick ahead too late due to the previously mentioned factors and cannot avoid the collision.

The information for when to initiate the kick down would be visible on the Conning Display when looking at the thrusters, thus making it available for action selection in this gaze pattern. The captain could notice it by just altering his gaze slightly upward.

**Batch 1, scenario 1, person 2**

**Error**

You already knew at that point that the ship would not make it. You knew [...] that it was too late. Because [...] you had experienced how slowly it turns here [referring to first turn]. Therefore it was actually obvious from my point of view. You had to turn much earlier and we should have cut the corner directly, then it would not have happened (39-43).
Figure 20. Crew (white circle) leaving the designated waterway. The point where the prescribed track is left is highlighted (black circle).

During the second turn in starboard direction the ship (Figure 19) leaves the designated track in eastern direction leading to a position outside of the waterway.

Performance-shaping factors

It’s this weighting, you know? (126-127)

After the first turn the crew is placed south of the waterway separation line (central yellow line, Figure 20). This positioning is an optimal starting point for the initiation of the next turn towards starboard according to the instructor. Unfortunately for the crew, the ship starts to drift slightly southward (white arrow, Figure 22). As they probably thought that their new position would not let them get “around the corner” (39-43), they initiated a corrective action, steering the ship in portside direction. This led to the situation as depicted in Figure 22. The ship (black circle) is still positioned somewhat south of the central radar line but with a turn rate of 12° in port direction. This portside turn was initiated so the crew would not leave the waterway when turning starboard by cutting the corner (white circle).
The ship (black circle) started drifting southwards (white arrow). The crew initiates a turn towards port side to ensure that they don't leave the marked out waterway when they turn starboard, “around the corner” (white circle).

As the ship has a portside turn when the turn in starboard direction was initiated and due to the rudder-effect delay, the ship reacted very slowly to the rudder command to starboard. The rate of turn the ship has to portside is reversed, leading to a straight ahead heading, without rate of turn in any direction. Realizing this, the crew initiates another rudder command (figure 23). The ship then starts to turn but it is however too late and the ship leaves the designated waterway and turning outside of it.

The crew (black circle) executes another starboard rudder command as the command given before simply stopped the portside rate of turn without resulting in a starboard rate of turn.
The following citing gives implications on why the error occurred:

Suddenly you have to master this task and for that you have to pass through such a big part [of the map]. That's actually a real challenge as you don't really know, how extremely the ship will react. (4-7)

This strongly resembles the citing from batch 1, scenario 1, person 2, the captain simply does not know exactly how the ship will react to rudder commands. Although it was known that the ship they were manoeuvring is large and thus inert, the extent to which the inertia slows the ship's reactions was unknown. The captain stated that he underestimated the inertia. Thus the link between rudder commands and the resulting ship reaction and future position does not seem to be evident for him. The task is *compensatory* in the sense that the captain does not have access to direct information of actual and desired state. The necessary operations are again *simultaneous*, the captain may choose order of actions freely. *Open loop* control is exerted as the captain gives one time rudder commands which the helmsman executes. The task is mainly *monitive* as the captain monitors the rudder commands being executed via the displays and the outside.

You know the instruments but you do not know how extremely the ship will react. (65)

As the captain does not know how the ship will react to instrument manipulation he must observe these reactions by monitoring the displays and the outside view. The statement also indicates that the functional coupling between instruments and ship generally has low *internal compatibility* as the ship behavior after manipulations is unclear. Essentially, this error sequence can be classified with the same cognitive coupling as described in Batch 1, scenario 1, person 1, although this time, the Conning Display is switched on.

It is not surprising that the captain did not know when to react as he was suffering from *information and goal overload* as a characteristic typical for interactions that are classified with high demanding couplings.

And you were actually indecisive all the time if you really should start turning or if you should wait. And then the current affected and from that point on you knew that you actually won't make it any more. Because after that, things happened fast. (47-50)

Table 18. Cognitive Couplings Batch 1, scenario 1, person 2

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitory</td>
<td>Monitors ship’s reaction to rudder commands</td>
</tr>
<tr>
<td>Type of Control</td>
<td>(-) Open loop</td>
<td>Open loop control is exerted as the captain gives one time rudder commands which the helmsman executes</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, ship inertia, etc</td>
</tr>
</tbody>
</table>
**Necessary Operation**

<table>
<thead>
<tr>
<th>(+) Simultaneously</th>
<th>Can choose freely when to do what</th>
</tr>
</thead>
</table>

**Information Presentation**

<table>
<thead>
<tr>
<th>(+) Compensatory</th>
<th>Difference between actual state and desired state is not displayed</th>
</tr>
</thead>
</table>

**Internal Compatibility**

<table>
<thead>
<tr>
<th>(+) Internal Incompatibility</th>
<th>Mapping between rudder commands and ship reaction is unclear</th>
</tr>
</thead>
</table>

---

**Design proposition**

*See Batch 1, scenario 1, person 1*

**Batch 1, scenario 2, person 3**

**Error**

No collision occurs. However, the crew gets very close to one, entering water that is not intended for use according to the scenario instructions.

After passing through the opening of the harbor the ship starts to turn towards starboard and heads directly in direction of the northern quay. The crew manages to stop the vessel very close to the quay.

**Performance-shaping factors**

While passing through the current the ship drifts very close to the southern mole. The ship does not collide with the mole, however it starts to turn as breakwater point is passed. While passing the breakwater point the ship is located in waters with differing currents making the ship slowly turn to starboard and thus changes the heading in direction of the northern quay. The breakpoint effect is even stronger as the ships speed is low. The captain decided to purposely enter the harbor at low speed so they would not have to face the problem of not being able to stop in time once they are in the harbor (95-104).

I had to think of not being too fast all the time as I was afraid that I would not be able to stop. Therefore the speed was a bit too low and as the current affected us from behind and started to turn us, I simply tried to do things that would counter the rate of turn, but did not really know what would help. I simply tried a lot of things and kind of said “try some bow thruster” or “do this...”. And that did not work out. And as I saw that, that we were heading towards here [the quay], then it was a collective decision to completely reverse the engines and to just try to avoid hitting it (95-104).

As several rudder commands were given right before the ship entered the current, the captain was expecting the ship to turn to starboard and could therefore not attribute the rate of turn to the breakpoint phenomenon. If he would have done this, he would have reacted right away with the appropriate commands. But instead he was surprised by the development of things. This is illustrated by the following interview statement (104-109):
But on that I have to say, that I did not expect that to happen. […]And that could be my error, in hindsight, when I think about it. I head towards it and do not take into account that I turn towards starboard, that is I did not notice it right away. I basically thought that it belonged that way, the current affect me and then I’m in [the harbor]. If I had anticipated it, I would have noticed the slightest rate of turn tendency and I had countered it accordingly. (104-109)

The captain did also not know how to react in a correct way as he basically tried everything he could think of (97-100).

[…]I simply tried to do things that could counter the rate of turn but, to be honest I did not know what would work out best. (99-100)

The captain was sure that the rate of turn was due to previous rudder commands. He mentally established a relation between this rudder command and the rate of turn as indicated by the outside view and the rate of turn display. This relation did not reflect the true nature of the increased rate of turn that actually resulted from the breaching of breakpoint water. The captain combined the wrong parameters, the rudder command and the rate of turn, as the actual link was breakpoint current and rate of turn. So he had a multidimensional view of the situation that was initially constructed from the “wrong” parameters and thus led him to believe that the rate of turn would not be a problem. The fact that he attributed the ship’s reaction to be associated with the previous rudder command and not with the breakpoint current indicates that the situation was suffering from internal incompatibility, a the ship’s reaction was incorrectly attributed to the ships steering system.

The captain only monitored the ship during this exercise without giving commands at all. He could freely determine what to do when (simultaneous operations). The information display was compensatory in the sense that he could not see how the actual state differed from desired state. Had he known that the ship would turn more than implied by the rudder commands, he would immediately have known that he would have to counteract this dangerous rate of turn. As the captain was not giving any commands he did not exert any control over the system (no control).

Table 19. Cognitive Couplings Batch 1, scenario 2, person 3

<table>
<thead>
<tr>
<th>Cognitive coupling</th>
<th>Demand</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Task</td>
<td>(+) Monitory</td>
<td>Monitors ship’s reaction to rudder commands</td>
</tr>
<tr>
<td>Type of Control</td>
<td>-----------</td>
<td>No control exerted</td>
</tr>
<tr>
<td>Dimensionality</td>
<td>(+) Multi-dimensional</td>
<td>Speed, Acceleration, ship inertia, etc</td>
</tr>
<tr>
<td>Necessary Operation</td>
<td>(+) Simultaneously</td>
<td>Can choose freely when to do what</td>
</tr>
<tr>
<td>Information Presentation</td>
<td>(+) Compensatory</td>
<td>Difference between actual state and desired state is not displayed</td>
</tr>
</tbody>
</table>
Internal Compatibility | (+++) Internal Incompatibility | Mapping between ship reaction and rudder commands was established in the wrong way

Possible Solution
The correct action would have been to execute a kick ahead as indicated in the following quotation (51-56).

If the ship passes the current here where it is the strongest [...] and if the bow is already in the harbor and the stern outside of it, then the ship starts to turn. And if you counter it with the appropriate rudder command and a kick ahead moving the stern through there quickly, then the exercise is quite easy. I got to know that from the other students that did it in that way[...].(51-56)

Design proposition
A clear display of breakpoint situations, when the ship is going at low speed and the current only affects one part of the ship. The captain said, that he noticed the breakpoint by feeling. Therefore, a design proposition must encourage the new user to use the kick down immediately by presenting this possible action in a highlighted way.

8.3. Putting it together
If we look at the all the different errors committed from the participants, a clear pattern emerges. All errors regardless of whether the Conning Display was switched on or off, occurred under similar cognitive couplings. The necessary operations can be chosen freely (simultaneous operations cognitive coupling mode). The captain has to monitor several instruments on the bridge as the interrelation of task-relevant multiple dimensions that are displayed by the instruments must be imagined in order to infer the how actual and desired states stand in relation to each other (information display leading to compensatory actions). As the students also lack experiences necessary to make the correct connections between the task-relevant dimensions, they are confronted with situations that do not match their internal picture of what would happen when the controls on the bridge are manipulated. Thus, they are confronted with situations that are internal incompatible and thus stressful.

The most promising approach would be to address the issue of information presentation. If the actual and desired state can be displayed clearly, changing the display from compensatory to pursuit, all the other cognitive couplings will shift into a less stressful mode as well. The dimensionality of the task is reduced as the combination of the relevant dimensions is displayed by the external system and therefore does not have to be inferred by the individuals on the bridge. Thus, only one dimension, the relation of the desired state and actual state must be processed. This should also make the task more internally compatible to the individual's expectations of ship behavior after instrument manipulation, thus clarifying the mapping between instrument and function in the context of the task at hand.

8.3.1. Design proposition
When asking the most successful participant (he and his crew did not commit any errors) in the maneuvering exercises if he planned ahead during the exercise, he stated that he was constantly
thinking 3 minutes ahead into the future when giving rudder commands. This gives us a hint about what the current state consists of in the maritime context. This is completely different from e.g. car driving where the car reacts directly when the steering wheel is turned. Imagine driving a car on the motorway when every steering command would take minutes to have effect! And even more disturbingly, you might not get any cues at all for the first minute if the car ever will start to turn. Thus, paradoxically as it may sound, the actual state of a ship extends into the future in the sense that it incorporates what your momentary orientation towards a desired state is. This orientation is not only influenced by the delay from rudder commands to take effect but also from other ship intrinsic factors like e.g. speed, acceleration. Additionally, ship-external factors affect it as well, for instance current and wind.

What should be done is to project the ship into the future, taking all ship-external factors (current, wind, etc) and the ship intrinsic factors (rudder command inertia, momentary acceleration, position etc) into consideration. The basic components for creating this projective function are already “present” on the bridge. The ECDIS, electronic chart and information display, shows all the external information (except for wind) that is relevant for the scenario, like current, water depths and ship position. The ship’s delay in terms of rudder commands to take effect, can be extracted out of the manuals present on the bridge and be set into the formula proposed by Sheridan (2002). Thus, what the Conning Display would do is essentially to provide the functionality that was requested by the personnel in the simulator facility: to display the relevant information in a refined form.

This design solution would help in preventing errors by providing direct feedback on the ship’s current state and would thus aid the students to identify wrong actions and to counter them right away. If the captain in the exercise Batch 1, person 2, scenario 2 (MB2TA2) would have had direct feedback on what happens 3 minutes into the future after giving the wrong (relative) rudder command he could have reversed it immediately. All the exercises in which the current posed difficulties would become easier as the user would immediately see how it will influence future ship position. Thus, the students would become aware of wrong actions associated with low internal compatibility.

Therefore, this “projector” should be added to the functional requirements as identified by the interviews. The projective function would probably improve the performance of crews in every scenario (see following sections) except for one crew. This is the crew MB3 whose ship collided with the quay due to berthing at too high speed. It is impossible for the predictive function to know when the captain wants to berth as this has to do with intentions of the crew. And if the crew has already executed the berthing procedure at too high a speed, it is too late to show the negative effects as they cannot be prevented any more. However, the reasoning that the current and the desired state have to be displayed clearly in the system holds as well. If the Conning Display would display the information of what the safe speed to berth at is for this ship type, then the captain could adjust his speed thereafter. Thus the functional requirement “display of safe speed to berth at” should be added to the requirements identified through the interviews.

Adding all the requirements listed under the interview result section and the ones proposed here, would lead to a display providing the information display functions necessary to prevent errors from happening (this will be discussed further in the discussion section). To facilitate the task of setting the
new Conning Display into relation with the committed errors, a short summary on how the proposed
design requirements will affect all groups mentioned in the result section, will follow.

**Batch 1**
Scenario 1 - Inertia can be taken into account due to projection enabling dynamic reactions to
dynamically changing situational demands, like current direction for instance.

Scenario 2 - This error would not occur as the effect of current breakpoint water on the ship’s rate of
turn would be made explicit before the breakpoint is breached, thus giving the captain and the crew
time to prepare and execute counter commands that will be in effect when the point is passed.
Additionally, rudder commands that are given in a relative way can be reversed directly as their
effect becomes visible immediately and not when it is too late to counteract.

**Batch 2**
Scenario 1 - This error would not occur as the effect of passing breakpoint water will be made clear
and the captain thus would probably choose another heading with which won’t lead to a collision.

Scenario 2a - The projection could display the whole ship and thus how the stern reacts to rudder
commands, in other words if the selected maneuvering actions will lead to a collision of stern with
quay.

Scenario 2b - As the safe speed to berth at will be displayed, the students could consider this when
they approach their berthing location of choice.

In summary, every error has to do with unintended effect of rudder commands. Thus, the change in
cognitive couplings will be related to the effect of rudder commands. I have not forgotten that I have
given other design recommendation as well in the previous analysis of the errors. They will remain in
that section as they denote “the waypoints” leading to the actual design proposal.
9. Qualitative method reflections

The qualitative methods will be evaluated with the validity and reliability concepts (e.g. Heiman, 2001) applied to the quantitative part of the study.

9.1. Observations

Observations were conducted in order to infer the cognitive processes underlying sequences that led to errors. Behavior was observed by means of video material that stemmed from several sources, these sources individual impact on the qualitative design should be pointed out.

Eye-tracking always requires calibration to the physiognomy of the participant. As the participant moves around freely and thus also may move the glasses unintentionally, the initial calibration of the eye-tracking instrument can be lost (e.g. when the participant scratches his head). This leads to the crosshair, indicating the gaze, being slightly removed from what the participant actually looks at. This can be very misleading if the researcher does not know what the participant usually looks at during exercises. The interviews tackled this issue by investigating when and to what extent the crosshair was maladjusted as some of the questions concerned what typical "gaze cycles" were, what instruments the participants usually checked during the exercise.

This can help when manually readjusting the crosshair to fit to actual gaze patterns. Additionally, it was also possible to discern how much "off" the crosshair was when the participant pointed at something in the simulator (When pointing at something, the person usually has to look at it).

9.2. Interviews

The interview questions were created in order to classify the interaction during scenarios in terms of cognitive couplings. As there is always a possibility that the participants do not remember the actual events in the simulator correctly, interview responses concerning their interaction could be misleading. This could be seen as compensated (Green et al. 1989) for by the observations conducted on the video material. Here coupling types and associated modes could be inferred from behavior when possible and compared with the findings from the interviews.

9.2.1. Catch 22

The interviews were conducted according to Dekker (2006), that is to repeat the researcher's understanding of what the interview partner had said during the interview at regular intervals (To clarify understanding when the researcher was not sure whether a correct understanding of the participant's answers had been reached).

As the researcher wanted to keep the participant as unbiased as possible this led to an uncomfortable catch 22. If the researcher repeated his understanding of the participant's account, it could lead the participant to be biased towards that account, leading him/her to confirm the researcher's account, regardless if it is flawed or downright wrong (For instance due to the fact that the participant has overconfidence in the competences and authority of the researcher (Milgram, 1963)). However, if the researcher would not verify his/her understanding, he/she would run the risk of establishing an incorrect understanding of the participant's account.
However, even understanding that was verified by the interviewee in order to “satisfy” the researcher, could be validated with the video recordings in order to clarify if the understanding is correct.

9.2.2. **Are the interview questions reflecting the theoretical framework?**

The questions were generated by the researcher. This was done according to the researcher’s understanding of the couplings, which could be incorrect. An approach addressing this issue would have been to let other researchers create questions in parallel and to compare these versions with the own one. Thus, the degree of consistency between the own understanding of the theory represented by the questions with that of another researcher could be assessed (investigator triangulation, Thurmond, 2001).

In a sense this was done, as some questions that were similar to those proposed by Klein as cited by Dekker (2006), which also concern systemic error causes, were considered to cover the same concepts and integrated in the question pool.

Additionally, other questions proposed by Klein as cited by Dekker (2006) were added, as they could provide insight into why errors occurred in a general sense, without explicit connections to the cognitive couplings. However, these questions could be associated with cognitive couplings in a second step after the analysis of the observations. The question of whether “the exercise or a sequence of the exercise went according to plan” was one of these general questions that provided much insight and that, in the course of the analysis, could be broken down into cognitive couplings (e.g. internal compatibility).

9.2.3. **Results reflection of researchers perspective**

The results of this study could reflect the perspective of the researcher having conducted the study. This could be especially critical in respect to classification of the cognitive coupling modes, a central concept in this study. An additional observer classifying the cognitive couplings could shed light on the question whether some coupling modes that were perceived as ambiguous (in the sense of *resolution*, see discussion) were interpreted correctly (investigator triangulation).

9.2.4. **Data analysis**

No explicit approach was used for combining the interview data with the observations. The interview data functioned like a heuristic in the sense that the researcher established some kind of understanding of what happened during an exercise and then observed the video material to see if this understanding had to be modified or was correct. This process was documented in the form that was presented under the result section.
10. Common method criticism

Certain critique affects both the qualitative and quantitative part of the study. This critique will be elaborated in the following sections.

10.1. Operationalization of performance

The operationalization of the variable performance could be regarded as suboptimal. In this study expert evaluation was chosen for the measurement of student performance as all of the instructors have had several years of seafaring experience, either as captain of ships or as pilots. Thus, they intuitively know what correct or incorrect behaviour consists of.

This approach leads to high dependence on instructor feedback for additional clarification. If there was an urge to further investigate interesting behaviour that, in the eye of a novice (me), seemed problematic for the course of the exercise, then further information could only be accessed by asking the instructors. Unfortunately time and resources are always limited as the instructors, besides monitoring simulator exercises, also have to give classes and are bound to other mandatory activities. Additionally, other researchers willing to replicate these findings could get different feedback on what errors consist of as the person filling the role as instructor may be different.

10.2. Other measurement approaches

Thus, other forms of measurement could be considered for future studies. The measurement of "time for an exercise to be completed" could serve as an indicator for how well the students performed compared to other students (everything else kept constant, that is). Although this ensures that criteria for performance evaluation other than the instructor experiences may be used, it is not clear in what way this would be "correct". Do crews that take longer for exercises really perform worse? It could be that they e.g. get to know the ship’s manoeuvring abilities better than other crews, thus having a better ground to stand on when working with that kind of ship in real-life situations.

Another approach could be to sensibilize the instructors to the resolution of the error analysis through briefing by researchers in advance. The instructors could thus, drawing in their experience, point out when exercises take longer than usual and give hints or guesses on why this occurred. This feedback could thus serve as valuable heuristic when further analysing the eye-tracking data for instance.

This was not done in this study as there was no clear idea about exactly how errors would be classified by the instructors. It was not know in advance what the instructors would deem to be an error and what for instance was regarded as "time consuming struggle in order to find a solution". This could also not be clarified in interviews with the instructors. They stated that the scenarios often lead to a broad variety of behaviour and thus several correct (but unexpected) solutions of scenarios may be observed for which no explicit criteria may be identified in advance.

10.3. Role division

Role division was enforced on the bridge, which would otherwise not have been part of the maneuvering exercises. The errors observed could thus have resulted from the students struggling with the role division and therefore been created by the experiment design itself. However, the
other crews that did not carry the eye-tracking device and were running the exercises in parallel, made comparable errors as evident from the instructor station recordings, suggesting that the behavior was not influenced by the application of roles.

10.4. Eye-tracking influencing performance
The eye-tracking instrument in itself could influence the performance of the crews. The person carrying the glasses did know why he/she carried them however, the mere awareness of the gaze being recorded might have been an performance-influencing factor. The captain could have missed important cues due to feelings of being uncomfortable for instance when looking at a person of the other gender. This possibility cannot be ruled out completely. However, the recorded crews made similar errors as the other crews conducting the same exercise in parallel (but not participating in the experiment). Thus, many errors similar to those committed by the eye-tracking crew were done by other crews as well. In future studies it would additionally be beneficial to compare the amount of errors committed by the eye-tracking crew in the control condition with the amount of errors committed by crews without eye-tracking (as discerned in the instructor screen capture video and from instructor feedback). If the amount of errors is not different, it could be argued that eye-tracking instrument not influencing the students’ performance. This was not done in this study due to limited resources.

10.5. Crew on the bridge - captain in focus
The focus of this study is on the captain and how he/she handles the manoeuvring task. The captain was the only one to participate in the interviews and to be recorded with the eye-tracking device. While the captain is not alone on the bridge he/she is the one responsible for the ship and always has the last word in debates. Therefore he/she should always have a very accurate picture of the situation on the bridge and thus probably is the most important person to focus on. Additionally, the performance of the rest of the crew was still observed as they could be seen and heard by the CCTV camera.
11. Discussion

Generally speaking, this study was conducted in order to see how/if the design layout of the Conning Display influenced student behavior. It was investigated if the Conning Display influenced student’s performance in a positive or a negative way. The quantitative between groups analysis revealed tendencies, for instance that the probability to commit errors for the first batch of exercises was 25 percentage points lower than without Conning Display. In the second batch no errors were made with the Conning Display and four without. This suggests that the Conning Display could enhance performance.

The within group comparison revealed that the errors were reduced by turning the display “on” from being “off”. However, the errors actually decreased when shifting the display from “on” to “off”. Therefore, the quantitative analysis was inconclusive and no decision on whether the Conning Display influences performance could be made, which shifts the focus on the qualitative analysis to shed light on this matter. However, it cannot be ruled out that student performance could have improved. Therefore, it was suggested to keep the functions stated as useful by the students in the interviews in the new layout of the display.

The qualitative analysis revealed that the errors committed all occurred under similar cognitive couplings types, all associated with comparatively high levels of stress which leads to an increase in errors (Sträter, 2005). This has certain design implications and reveals some of the complexity that the students have to cope with on their way to becoming certified mariners.

The design proposition in terms of functional requirements and in terms of the predictive function presented in the previous chapter of this study, aims at lowering the stress stemming from the cognitive couplings and will thus hopefully lead to safer crew behavior during simulator exercises.

11.1. Internal cognitive coupling and negative transfer

All the errors that occurred during simulator exercises were related to internal incompatibility. This seems fitting as the students have to face an unfamiliar new situation in which they have to command and manoeuvre a ship completely on their own. As they've had no previous experience with this besides a few introductory simulator exercises, they try to cope with the overwhelming flood of new information by setting it into relation with previous experiences for instance in other domains. Thus analogies for coping with dissonance at hand are established.

They are created through either positive or negative transfer (Sternberg, 2006). Negative transfers occur when “wrong” analogies are used, when the experiences made in one domain are applied in another domain but the result is erroneous behaviour. Consequently positive transfer occurs when the applied behaviour leads to the intended effect. Sternberg (2006) argued that negative transfer occurs when a transfer of behaviour from one problem to another is based on structural similarity without regard for the content of the problems.

Once the individual has found a suitable analogy a fixation on this analogy occurs. It will take very strong cues to break the decision of the individual to apply a specific behaviour as the behaviour has
proven successful in the past and thus is associated with the respective neural strength (in the terms of its activation pattern having become strong by means of repetitive activation).

Regardless of these conceptual reflections, on a practical level a system centred on safe performance of novices should always be designed in order to give room for this natural phenomenon of analogy establishment. Cognition is experience-driven (Sträter, 2005) and thus momentary cognition will always stand in some sort of relation with previous experience.

This is accomplished in the best way if the system can provide direct feedback to the user that makes it possible for him/her to see that the behaviour will not fulfil its intended purpose (that a negative transfer has occurred). The person can then immediately readjust his/her behaviour to undo the potential damage and will hopefully not resort to this behaviour the next time as he/she has noticed that its effects are negative. This is what the proposed design aims at doing. It provides direct feedback on user actions in relation to the desired state by means of projection. Therefore the design proposition presented in this study could be regarded as being designed to give room for the natural phenomenon of analogy establishment, thus leading to more resilient performance. Resilient performance means that fast feedback is provided to the students, which is necessary for preventing errors that would have occurred from unsafe acts (Hollnagel, 2005).

### 11.2. Dimensionality

As stated in the theory section, dimensionality is defined as the amount of independent parameters that have to be mentally combined to make sense of the task at hand. This leads to the fact that two types of dimensionalities can be distinguished, system dimensionality and process/interface dimensionality resembling the concept of system and interface complexity as proposed by Hollnagel (2005). While system dimensionality refers to the general dimensionality of a system, that is how many degrees of freedom it has, process/interface dimensionality denotes the dimensionality as mediated by the displays for a specific task at hand. Therefore question on the relation between these two concepts arises.

Is it always effective to reduce the interface dimensionality in order to highlight a task-specific interdependence of parameters or does the reduction maybe lead to the fact that operators spend less time setting this parameter in relation to other parameters, thus degrading performance? In other words what would lead to better performance, several displays showing the interdependence of parameters for certain tasks or the independent display of parameters leaving the human to mentally imagine their interrelationship? And if these two approaches have to be balanced against each other, what where is the optimal return/trade-off point?

Another question is how this interrelatedness of dimensions can be understood from a constructivist perspective. For the novice for instance it might be difficult to set the dimensions "affecting current", "ship speed" and "rudder angle" into relation to each other in order to discern if the resulting heading will be good or not. However, an expert mariner, having handled this ship type for several years, will know exactly how the ship will behave and can thus predict if the rudder command will have good or bad effect with relative ease. Thus, the stress put on an expert individual when imagining the interrelation of several dimensions seems to be less than for a novice.
Perhaps we can gain more insight into this phenomenon if we explore what this interrelation actually means. It could be regarded as indicating how the current state is oriented towards the desired state of a task by means of combination of independent parameters. For experts the way in which these dimensions are interrelated in relation to a desired state has been internalized, become skill through years of experience (Sternberg, 2006). In that way they have a feeling for the ship.

Perhaps seeing a bridge with the eyes and the mind of an expert would be like putting on an elaborate head up-display that enriches the already present information by highlighting information interdependence for at task at hand (Thereby aiding the establishment of “professional vision” Goodwin, 1994). With this enrichment, the difference between current and desired state is highlighted and would be easy to discern. Therefore, the information presentation that was originally compensatory has become pursuit by experience in the sense that the actual state and the desired state are clearly mediated by the external world for the expert with adequate professional vision.

Does this mean that the human can adapt to anything simply by virtue of experience, that the feared "killer argument" of more technically oriented system designers is actually true? The big issue here seems to be how much effort it takes for the human to force his/her way of thinking into a badly designed corset of technology. If the technology requires the human to react dynamically although his/her actions will only take effect after considerable delay, it would require more effort for the cognitive system to accomplish this task than if the system had provided feedback without delay. Or placed into the terminology of the cognitive couplings, if a safe interaction requires one mode of a cognitive coupling type but the novices in the domain show behavior that is clearly part of the other mode of the coupling type (for instance mostly doing active tasks when the demanded behavior is to be monitive) then one could argue that the cognitive distance between experts and novices is large (the modes of a coupling type are exclusive). If this distance is present in other cognitive coupling types as well, the cognitive distance between novices and experts is even larger, thus requiring a lot of training and experience to bridge it.

The layout of the Conning Display as proposed in this study aims at bridging this distance by supporting the students in being in the correct coupling modes as demanded by the task (by changing compensatory to pursuit, multidimensional to unidimensional, etc). Of course this raises the issue of automation overreliance and correspondingly automation surprises.

11.3. Design recommendations and training
Design recommendations always have to consider the irony of automation. If the students depend too heavily on the new design solution they perhaps would not be able to handle critical situations without it if it should malfunction (automation surprises). Additionally, they will almost certainly be on board vessels where the display and its functionality are not available which could impair their performance in real life.

I argued that the layout of the Conning Display leads to more resilient performance. What this means is that the students, during conduct of the exercise, are able to discern what a safety-impairing action (unsafe act) is themselves. Essentially, this boils down into deciding what design philosophy to chose when addressing errors, preventing errors from happening altogether or letting act potentially leading to them occur but providing the tools necessary to recover from them. The latter is a resilient approach.
On a practical level the introduction of this (presumably) resilient layout in the simulator means that what for the instructor might seem as flawless performance could in truth be a continuous committing of errors that are recovered, thus giving the students more chances to experience how to recover from errors. Otherwise they would have to face problems like the crew in batch 1 person 1 scenario 4 that had to wait for the debriefing after they collided with an object. This “wasted” time could be spent much more efficiently in continuing the exercise in a resilient way, making it clearer and clearer what the “right” and “wrong” actions are under certain situational conditions.

Additionally, the crews have to manoeuvre several different ship types. As the use of different ship types shall demonstrate how differently they behave to rudder commands, the Conning Display could visualize how this different behaviour looks like by showing the preplotted course when rudder commands are given. Thus, students could perhaps develop a better understanding for how the ship behaves than from just using the regular instruments.

Of course these are only hypotheses about reality that must be validated in future experiments highlighting this study's iterative design layout. Ideally, the student's performance should not degrade too much when the display is shifted “off” during follow up studies as the students with the Conning Display have developed a better understanding for how different ship types react to rudder commands and therefore have developed a better understanding in what constitute safe actions in general (from the resilient conduction of the exercise as mentioned before).

11.4. Providing the “why”
What could have been done instead of this projective function could have been textual display like “these and these circumstances have kicked on, do not do [action]”. The design solution proposed here does not provide this kind of syntactical feedback. Instead, it lets the student commit unsafe acts but provides immediate feedback on why they are safety degrading. Thus, this display does not say "that" something is wrong, it lets the students discover what consequences actions can have and thus provides the “why” some action is wrong. The students can thus themselves initiate corrective action with the knowledge on why they do it, they know the intention behind their actions.

11.5. Resolution problem of the cognitive couplings and resolving mechanisms
It can sometimes be difficult to tell at what resolution the classification of the cognitive couplings should be conducted. The maneuvering task e.g., if considered under the cognitive coupling of type of task, cannot be properly divided into active and monitive as some elements of it are active (giving commands) and some elements are monitive (watching the command being executed). Thus, in the section leading to an error, the interaction may be classified as being both active and monitive. This could be addressed by letting an independent researcher classify the interaction of interest as well and comparing the results (investigator triangulation). Thus, a general feeling for the appropriate level of analysis could be established and perhaps, when enough data is available, be formalized. If this should not be viable, then other researchers, attempting to replicate classifications of video material would for instance not know on which “level” to start the analysis. The researcher would be exposed to the risks of getting stuck at a level of detail that is not “rewarding” in the sense that it does not generate useful data.

This problem of resolution also seems to apply to the resolving mechanisms. The behavior immediately preceding an accident often can be classified as a fixation on the factors leading to this
behavior. However, the immediately following dissonance of discovering that one was acting under wrong assumptions could classified as information and goal overload. Future research should address this issue in an explicit way.

11.6. Spin-off: Planning tool
Some of the crews that did actually manage to turn the ship in the third scenario of batch 2 were not doing this as they had originally intended. This is also a something that has troubled the instructors. During conversations some of them stated that it can sometimes be hard to discern if a crew actually carried out their exercise as planned when they did not commit any errors or if it was just by luck or on wrong assumptions. Even the fellow students conducting the exercise sometimes stated that they wondered if the crews that manage to get through the exercises without errors actually did it "according to how they originally intended" and if they would thus be able to repeat their results in another trial of the same exercise.

It would therefore perhaps be beneficial to provide a tool for the students where they can preplot their intended course through the exercise. These maps could be submitted to the instructor before the start of every exercise and be put in relation to how the actual outcome of the exercise was. This would on the one hand make it easier for the instructors to see if "correct" performance went according to plans and if not, how the students coped with the problems that arose during the exercise. As the briefings are done in presence of all of the participating students, other students could then get to know and benefit from strategies successful crews followed when the exercise did not go according to plans. Interview data suggests that this is in fact the case more often than not, especially since interview data revealed that the results of a planning phase are often regarded as a kind of blueprint. The situations encountered during simulator exercises are often simply too dynamic and complex to take into consideration beforehand.

It would certainly be beneficial to display these coping strategies when things did not go according to plans by displaying the intended way through the map and the actual course in the scenario. Thus successful crew behaviour may be distributed among all students, resulting in a greater repertoire of actions in critical situations.

11.7. Conclusions
This study has shown how the variability of human performance in naturalistic settings can be analysed out of a systemic perspective and integrated into a design proposal. Thus, in the long run a resilient and hopefully educationally sound tool, guiding the students towards the real-world domain of practice, can be created. Future research will have to clarify if this hypothesis will prove to be true.
References


Campbell, D. & Fiske D. (1959). Convergent and discriminate validation by the multtrait-multimethod matrix. Psychological Bulletin, 56 (2) (pp.299-302)


So-0681C1 Polaris Technical Manual Section 5c Radar-ARPA. Kongsberg.


Appendix A: Interview questions

• Did you use the Conning Display?
  o What functions did you use?
• How did the turning “off” or “on” of the Conning Display influence your performance?
• What were you seeing?
• What were you expecting to happen?
• What system parameters were relevant?
• Were these parameters interdependent?
• What instruments did you monitor?
• With what instruments did you stand in a more active relation?
• Were certain instrument inputs dependent on timely demands?
• What instrument inputs were of more of a one-time nature and less dependent on timely demands?
• Was the feedback of the instruments clear?
• Did alarms occur?
  o What kind of alarms?
• Did you have to find out yourself if you were in danger?
  o With the help of which instruments/displays?
• If you had to describe the situation to a fellow student, what would you have told him/her?
• What kinds of mistakes were likely at that point in time?
• Were you reminded of previous experience?
  o Did you know how to react from previous exercises?
• Did you have the necessary practical experience with the instruments in order to master this exercise?
• Have you undergone training for this situation?
• Were there any rules that applied clearly in this situation?
• Did you have to resort to other sources of knowledge in order to get to know what to do?
• What goals governed your actions at that time?
• Were these goals that had to be checked constantly?
• Were there any conflicting goals or did you have to compromise between goals?
• Were you under time pressure?
• Did you plan ahead mentally?
  o When?
• How did you think that you could influence the course of events?
• Did you think of or discuss alternate actions or was it clear what to do right away?
• Did these actions have to be done in parallel or sequentially?
• Was the outcome of these actions in accordance with your expectations?
• Did you have to reconsider your evaluation of the situation?

If the interviewee was participating in the watchkeeping exercises the following question was asked:

• What functions of the Conning Display were sensible?
Appendix B: Pilot study

As there were many unknown variables concerning the functioning of the technical equipment and the comprehensiveness of the interview questions a pilot study was conducted to shed light onto these issues.

A pilot study was conducted with one person to provide information on how the technical equipment (eye-tracking, screen recording) would function and how well the interview could be understood. The study was conducted in the simulator environment during the training exercises, with one person of the three crew members carrying the eye-tracking glasses. The crew was assigned by means of administrator selection.

The scenarios in the pilot study are different from the ones in the main experiment. The entire session was recorded (CCTV, instructor station screen capture and eye tracking data).

The battery time of the eye-tracking device also proved to be problematic as the batteries have to be changed about every hour. The exercises carried on longer than that, which means that the experimenter has to enter the bridge during exercise to change batteries and that the used battery has to be fully charged before the new battery is exhausted.

After the exercise an interview was conducted. In contrast to the later study, all of the participants were interviewed, not only the person carrying the eye tracking glasses. This was changed in the original study, as the information from the persons that were not captain were considered to not be relevant enough to justify the increased difficulty in finding interview slots that fitted everyone's schedules. This was especially problematic since the interviews could not be conducted after the recording session in the original study as watchkeeping exercises were recorded directly after the maneuvering exercises and because the researchers had to leave directly after the watchkeeping exercises.

The pilot study also revealed that some of the interview questions were hard to understand. Therefore these were revised. The problems with the interview questions mainly stemmed from the fact that they were based on Dekker (2006) and with the purpose to uncover reasons behind accidents. But if no accidents occurred during the session, then the interview questions should fill another purpose, to unveil the interaction that took place during the exercise that in turn may form the basis for comparisons in terms of cognitive couplings between exercises with accidents and exercises without. Additionally, some of the questions that should clarify the cognitive couplings overlapped with questions aiming to stemming from Dekker (2006).