ASSESSING THE EFFECTIVENESS OF MULTI-TOUCH INTERFACES FOR DP OPERATION

F. B. Bjørneseth[1,2], *M. D. Dunlop*[2] and E. Hornecker[2]

[1] Rolls-Royce Marine AS, Common Control Platform, Parkgata 3, 6003 Aalesund, Norway,
Email: froy.bjorneseth@rolls-royce.com

[2] Computer and Information Sciences, University of Strathclyde, Glasgow G1 1XH, Scotland, UK,
Email: mark.dunlop@cis.strath.ac.uk, eva.hornecker@cis.strath.ac.uk

ABSTRACT

Navigating a vessel using dynamic positioning (DP) systems close to offshore installations is a challenge. The operator's only possibility of manipulating the system is through its interface, which can be categorized as the physical appearance of the equipment and the visualization of the system. Are there possibilities of interaction between the operator and the system that can reduce strain and cognitive load during DP operations? Can parts of the system (e.g. displays) be physically brought closer to the user to enhance the feeling of control when operating the system? Can these changes make DP operations more efficient and safe? These questions inspired this research project, which investigates the use of multi-touch and hand gestures known from consumer products to directly manipulate the visualization of a vessel in the 3D scene of a DP system. Usability methodologies and evaluation techniques that are widely used in consumer market research were used to investigate how these interaction techniques, which are new to the maritime domain, could make interaction with the DP system more efficient and transparent both during standard and safety-critical operations. After investigating which gestures felt natural to use by running user tests with a paper prototype, the gestures were implemented into a Rolls-Royce DP system and tested in a static environment. The results showed that the test participants performed significantly faster using direct gesture manipulation compared to using traditional button/menu interaction. To support the results from these tests, further tests were carried out. The purpose is to investigate how gestures are performed in a moving environment, using a motion platform to simulate rough sea conditions. The key results and lessons learned from a collection of four user experiments, together with a discussion of the choice of evaluation techniques will be discussed in this paper.

Keywords: Usability Evaluation Techniques, HMI, Gestures, Multi-Touch, Safety Critical, Dynamic Positioning

1. INTRODUCTION

The maritime environment is deeply rooted in traditions and has the last years experienced an interesting and user-challenging technological development from suppliers of maritime equipment. The automation systems are continuously growing more advanced and the mariners have to keep up with technology. The demand of increased computer and technology related knowledge can for some feel overwhelming, while for others it feels natural and a part of everyday life. The division is often, but not exclusively, age related with the younger generation of mariners feeling more comfortable with technology than the older generation [Paul and Stegbauer, 2005]. The increasingly advanced automation systems controlling modern vessels lead to increasingly advanced and complex user interfaces. Furthermore, a typical operator must interact with many different systems, often with different interface styles, during an operation. On Dynamic Positioning (DP) vessels, which is the key focus of our work, the operator position can become stressful as (s)he must interact with at least three different systems – each with its own graphical user interface (GUI) and display. In addition, the operator must lead the radio communication, have an eye on the propulsion system and maintain "constant observational awareness" of the environment around the vessel. This can be a challenge both mentally and physically and the cognitive load can increase if presented with too much information [Lazet and Schuffel, 1977]. The physical strain also affects the operator if the equipment is poorly ergonomically placed [Galliers et al., 1999]. Depending on the ship owner, the ship yard and the suppliers of equipment, the composition of the equipment in the operator station can vary considerably and is often ergonomically sub-optimal.

Human Machine Interface (HMI) work has a long history in maritime settings, but is often given low priority due to perceived increased development time and economic pressures. The economic aspects play an important role in a vessel's lifecycle and issues concerning HMI and usability are in many cases not a part of the discussion until late in the cycle when it is often too late and expensive to make vital changes to obtain an optimal solution [Sillitoe et al., 2009]. Today's trend seems to move towards a more noticeable awareness around HMI issues, but is still not always properly accounted for. An overall increased mental load when using a system is both tiring and leaves less mental capacity for handling safety-critical events. Such events are not prominent in every-day operation, but when they occur a high mental load can reduce the operator's experience to

This is a late draft of the paper – please see Proc. International Conference on Human Performance At Sea (HPAS), Glasgow 2010 (ISBN 978-0-947649-73-9) for final version.

the level of a novice [Redmill and Rajan, 1997]. Poorly fitted equipment combined with low usability causes a long–term problem for the operators. Unlike personal consumer equipment which can often be easily replaced if the consumer is unhappy with the interface or usability, equipment installed on vessels typically lasts many years and will not be replaced before its operating time has ended. The overall aim of maritime-HMI research is to lower the operator's cognitive load and make the workflow more efficient by introducing interaction techniques known from other HMI domains, such as mobile technologies and personal computers, while also assessing them by using traditional usability methodologies. In safety-critical situations a lower cognitive load will require less attention on how to operate the system and enable more focus on the actual operation.

Within our work we are interested in multi-touch interaction – a form of interaction that was popularised by Apple on the iPhone range but which has existed in research labs since the early 1980s [Lee et al., 1985]. This interaction style seems to have a great potential for bringing the interface closer to the user. Our on-going research investigates multi-touch interaction on DP-systems. In particular we have investigated if it is possible to carry out the tasks faster and more safely when operating the Rolls-Royce Icon DP system using multi-touch interaction. Our hypothesis is that the user interface will be brought closer to the operator by enhancing the operator's possibilities for directly interacting with the interface of the maritime software application by using multi-touch gestures. This ties the advanced maritime interfaces together with its increasing resemblance to modern technological consumer products where multi-touch has introduced a new dimension of interaction techniques.

In this paper we discuss the methodologies used: an iteration of creating prototypes and assessing their usability through user studies. This is supported by an observational study to get insight in the DP operator's real-life situation. First we will give an overview of background and technologies used, then describe the topic concerning prototyping on different levels and last describe our studies. The studies consist of one observation study and four different iterations of user studies. The observation study was carried out to gain more knowledge on how DP operations are carried out in a real-life situation. This gave a good support for further studies. The initial study is based on the paper prototype where the aim is to investigate which gestures feel natural to use when operating a DP system. The second study is based on the results from the first, but where the aim is to investigate the efficiency of using multi-touch gestures vs. traditional buttons/menus when operating the DP system. The two last iterations concerns a pilot study where the aim is to investigate how motion affects task performance when doing tasks using multi-touch and a main study. The main study investigates operating the DP system in a moving/non-moving environment while comparing the usage of gestures vs. buttons/menus when operating. For each study the motivation for the methods chosen will be outlined together with the key results and lessons learned.

2. MARITIME SAFETY AND TECHNOLOGY

The advanced technology onboard vessels today leads to an increased amount of rules and regulations set by IMO and other large classification authorities such a Lloyd's Register that have to be complied with. Safety is the first priority and preventing accidents onboard has full focus. Accidents do however still occur and often they are connected to what is referred to as human error. The reasons behind a human error related accident can be widespread and are not always directly connected to personal fault as a result of inattention or unregulated behaviour.

2.1 HUMAN ERROR

Whenever maritime safety is discussed, it is not long before somebody produces a statistic showing that most accidents at sea are caused by human error. The statement is usually made in a tone of resignation, as though accidents are unavoidable [IMO, 1997]. Furthermore, accident reports show no improvement in the number of injuries and lives lost at sea since 1995 [Lloyd's Register, 2007]. Although shipping has increased implying a better safety record, there is still considerable room for improvement with sea safety improving considerably slower than, say, the airline and car industries over the same period. With this in mind we can ask; is it possible to improve safety by introducing technologies and interaction techniques better known from the consumer market into the maritime domain? Can taking advantage of the user's previous knowledge [Mills, 1998] of personal and mobile electronic equipment be an advantage?

Reason [Reason, 1990] discerns active failure (of front-end actors, e.g. operators) and latent failure. Latent failure originates from preceding actions, involves working conditions and load, competing demands, and is caused by designers, developers, decision-makers and managers. Latent failure is the type of failure that is frequently seen onboard vessels today. Active failure involves the human in the process and the operator can in some cases be blamed. There are two main approaches to handle the problem of human error [Song, 2009]. One approach would include increasing the number of well trained crew members. The second approach would be to look for ways to improve the working environment of the human onboard ships. The last is the more long –term solution which solves the actual problem.

2.2 HUMAN COMPUTER INTERACTION ON MARITIME EQUIPMENT

HMI on maritime equipment has not always been, and is still not always a priority in the maritime realm. The economic aspects play an important role even though the majority of accidents onboard vessels are attributed largely to human errors. The errors are often due to misunderstandings during stressful situations, and not system failure [Mills, 2005]. Poor design is often blamed, and there has been a trade-off between the usability of the maritime equipment and issues such as the safety-critical aspect, and also robustness. There will however, to some extent always be a compromise between the design, technical issues and maritime directives. Modern technology does become cheaper and there has been legislation that pushes safety onboard vessels forward [Mills, 2000]. The maritime industry is conservative about novel technologies due to safety issues, but with time, the industry will most likely adopt new innovations supported by research that will enhance safety onboard.

When developing equipment and graphical user interfaces for the maritime environment, *'knowing the user'* is of paramount importance in good design [Faulkner, 2000]. This underlies the different methods used to obtain knowledge about the situation where the product is to be used. These methods can however often be poor substitutes to real life experience [Mills, 2005]. The best designers of maritime equipment are most likely the mariners themselves, who have experience and know what requirements the equipment must be capable of handling. A contradiction is when new equipment for maritime environment is to be designed. The user knows what goal(s) to reach, but not how to get there or which tools to use. To depend solely on the user's information, can in many cases be inefficient and time-consuming due to predisposed opinions and habits. One of the products that underwent a rapid development the past few years is dynamic positioning systems. This is a product that demands performance on all areas from low-level control systems to top-end graphical user interfaces and input devices.

2.3 DYNAMIC POSITIONING SYSTEMS

To keep the vessel in a fixed position close to offshore installations without using anchors, a system was developed that automatically compensated natural forces such as waves, wind and current. This is called a Dynamic Positioning system (DP) and its technology has developed from the first simple systems in the 1960's to today's advanced systems covering single, double and triple redundancy according to the operation's safety critical level.

A Dynamic Positioning system (DP) can be defined as: *A computer controlled system to automatically maintain a ship's position and heading by using her own propellers and thrusters.*

A DP system [Bray, 2003] can be seen as a complete system that includes operator stations, position reference sensors, gyro compasses (detects true north by using an electrically powered fast spinning wheel and friction forces, in order to exploit the rotation of the earth), and a range of different sensors that give feedback to the operator about the ship's position and the forces that influence its direction.

2.4 MULTI-TOUCH INTERACTION

Multi-touch is a human computer interaction technique together with the hardware that implements it. It allows the user to interact with the computer without the conventional input devices (mouse, keyboard). Multi-touch consists of a touch-display that can recognize more than one point of touch and there is a range of different technologies that implements it. Two of these technologies, optical and capacitive sensing have been utilized in this research project.

Interacting directly with an application's interface has in the last few years been proposed as the new way of interacting with computers in the future. Multi-touch has been commercialised by Apple, and young user groups are already well acquainted with the world of gestures and directly touching the surface to reach their aim of interaction through the use of handheld gaming platforms such as Nintendo DS and mobile phones. Although Apple was first to popularize it, multi-touch and bi-manual interaction have been a topic since Jeff Han spread interest with his first public presentation of multi-touch interaction in 2006. This demonstrated the principle of Frustrated Total Internal Reflection [Han, 2005], a low-cost multi-touch sensing technique. The interaction with both GUI and software seemed easy and natural, with flowing movements and simple gestures. The demonstration utilized a large rear-projected display in front of the user, like a workbench. This inspired the idea of implementing multi-touch/bi-manual interaction into maritime equipment, hence a DP system, due to the direct control of the interaction techniques. This can possibly enhance the DP operator's feeling of control when using a DP system.

Multi-touch and bi-manual interaction has through several studies shown to be more efficient than

Figure 1: PSV Havila Foresight

Figure 2: DP Operation onboard PSV

traditional input techniques. One of the initial studies of two-handed input was presented by Buxton and Myers [Buxton and Myers, 1986]. They concluded that the two-handed method tested outperformed the one-handed technique, which was most commonly used in 1986 at the time (and still is today). What appears interesting is the fact that poor design can make interaction with two hands worse than with one [Kabbash et al., 1994]. It is unclear whether occlusion and reaching over the tabletop can counteract the benefits of such interaction [Forelines et al., 2007]. This will increase the need of well- designed GUI's especially in a maritime environment where safety is of utter importance.

The majority of DP systems available on the market do not have advanced 3D graphics, including manipulation of the camera in the 3D scene, implemented. The Rolls-Royce DP system is however based on a 3D engine, which makes new types of user- interaction possible, together with a correct scaling of all visualization. With use of 3D, multi-touch and gestures, the original three degrees of freedom can be extended to six. This means that the user will be able to control the camera in the 3D scene by using gestures in three additional DOFs [Hancock et al., 2007], which are referred to as pitch, roll and heave. This can lead to the user feeling closer to the system and more in control.

3. OBSERVATIONAL STUDIES

To observe and report are techniques widely used for both social research and usability related research. In social research the observation is often part of an ethnographic study where the researcher immerses him/herself in the environment observed for months or even years [Bryman, 2008]. For usability studies and gathering knowledge around processes connected to carrying out specific tasks, smaller studies combined with interviews of users are more beneficial and commonly used. As mentioned earlier, "knowing the user" is important, but it is often difficult for users to express their views and put these in the context of wider HMI work. The benefit of being an outsider observing the users is that the observer might question issues that the user may never have thought about. This gives a wider angle to finding the right solutions while still grounding it in the end users actual use of the systems and his/her environment.

In our work the observations were anchored in the guidelines given by Jordan and Henderson concerning interaction analysis of videodata [Jordan and Henderson, 1995] and Alan Bryman's work on social research methods [Bryman, 2008]. Our observational study was a non-participant observation, where the observer did not take part in any tasks or daily routines. This is one of the best-known methods of research in the social sciences and primarily associated with qualitative research. It entails a relatively prolonged immersion of the observer in a social setting in which he or she seeks to observe the behaviour of members of that setting [Bryman, 2008].

In this case the bridge crew of an offshore vessel is in focus and the main goal of the observation is to gather knowledge on how platform supply DP operations and related activities are carried out. This can also be called an overt "micro-ethnography" [Bryman, 2008], where being overt reflects the fact that the researcher is not "under-cover" pretending to be a part of the crew. It is often normal to have a key informant that initially gives the observer access to the group and also key information. In this case the key informant was the Chief Officer who invited and informed the observer throughout the observation. To gather as much information as possible regarding issues related to being a mariner and working offshore semi-structured interviews were carried out in addition to observations. This is an interview technique that encourages the natural flow of a conversation instead of a fixed setup with the interviewer asking questions and noting down or recording the answers [Bryman, 2008]. In this case the interview guide, which held the topics of the interviews, was memorized and incorporated into normal everyday conversation. Semi-structured interviews often give longer and more supplementary answers. Through the whole observation study, concurrent field notes were written. Field notes play an important role when the study is to be analysed and similar sections are coded/organized and given labels to component parts that seem to be of potential theoretical significance [Bryman, 2008]. In addition to the procedures around how to carry out the observation, the guidelines given by Jordan and

Henderson [Jordan and Henderson, 1995] were utilized to plan what to look for, which questions to ask and how to structure the video recordings of the operations. A detailed observation study, selectively making use of video recording, was carried out to investigate how the DP operator operates the DP system in its authentic environment, and to find out which tasks are more frequent during the different operations. In addition, the situation around the operator's workplace onboard was analysed and it was also investigated if there were any specific movement patterns between the different equipment situated on the bridge. The study was conducted over a period of three days in early February 2010 onboard the Platform Supply Vessel (PSV) Havila Foresight. See Figure 1.

The participants of this observation study were the crew onboard Havila Foresight and two representative from Rolls-Royce Marine AS. The vessel's work tasks for the three day period, was to deliver drilling equipment, food and different liquids contained in the vessel's tanks below deck to four platforms in the Norwegian sector of The North Sea. On supply vessels the sailing schedule and tasks are determined just before loading the vessel with cargo, this meant that our schedule could not be finalised until arrival in harbour. After three days, four DP operations were observed and the crew was also observed when steaming to the oilfield, between the platforms and on return to shore. Due to similarity in operations only one DP operation was video recorded. See Figure 2. In total 7 observations were conducted (with one night-time DP operation unobserved). The participants observed were the captain, the first officer, two second officers and one midshipman. For the semi-structured interviews the captain and the first officer participated. This felt natural due to that they were the highest ranked officers onboard and also the spokesmen for the rest of the crew. Following a lightweight interview script, the semi-structured interviews were carried out in the shape of a normal conversation, where the captain and the first officer were asked questions while they were on duty on the bridge. The questions were asked during free periods between operations.

The observations were divided into four categories in addition to the semi-structured interviews. The first category concerned observing the crew on the bridge while steaming towards a goal (i.e. platform) and the second category concerned observing the operator during a DP operation. The third category concerned observing the crew on the bridge when steaming between oilrigs and the fourth category concerned observing the crew while returning from the oilfield to shore. Each category was supported with a set of questions in line with the guidelines given by Jordan and Henderson [Jordan and Henderson, 1995]. The questions concerned briefly: who was situated on the bridge, communication and movement patterns on the bridge, and also any usability issues with the equipment onboard. During the DP operation the official start and end to the operation was investigated, if there were any repetitive patterns, communication between the operators and also territorial issues. In addition the interaction between the operators and their abilities to work together was observed. The semi structured interviews consisted of questions revolving around the operator's daily routines when on watch, if any incidents had occurred and how they solved the issues.

Throughout the observations and analysis from the PSV Havila Foresight, a picture of a well-organised and formal vessel emerged. They carried out the tasks given with ease and followed procedures precisely, which is necessary on vessels working in safety-critical environments. However the personal relations between the crew members reflected an informal organisation that respected the ranking of an officer, but had an informal and cheerful tone between each other. They had an overall good working environment. The observations gave a good base of knowledge on how platform supply DP operations at sea were carried out in real life. For platform supply vessels the majority of time is spent on steaming to, from and between oil platforms and also waiting to get access within the 500 meter safety zone around the platforms. The discoveries made during this observation were that the pace onboard was much lower than anticipated. This can of course vary between different types of DP vessels, but what was anticipated in this case was a more hectic scenery on the bridge with lots of equipment interaction. The level of stress does increase if weather conditions are bad, but in general for platform supply DP vessels the pace is comfortable and slack time onboard is often used to browse the internet, check the weather reports and fishing. The most frequently used equipment on the bridge during steaming to or from a destination was the logbook, the coffee machine and the captain's chairs on the front bridge. There were always at least two officers on the bridge, with one always being on watch. They swapped from being on watch and doing other tasks, such as filling in entries into the logbook.

During a DP operation the DP operator stations were naturally the most frequently used equipment. During DP operation, the officer in command of the DP system maintained the view out the aft windows and aft deck the majority of the time. The operator's good overview of the aft deck and the actions happening on deck during operation give an advantage to ensure that safety on deck is maintained.

The observation provided detailed knowledge of the routines onboard and which tasks were more important than others. The interaction with the system had peak time when the operator closed in on the oilrig. The main interaction technique was using the input devices, such as the joystick and the heading wheel. They occasionally glanced on the belonging displays sometimes followed by quick interactions with them. A problem that was highlighted by the operators was a button that could be hit accidentally. This caused a change of state without the operator being aware. This could possibly cause dangerous situations. In addition to the above, it was also interesting to observe and understand the communication patterns between lower and higher ranked officers, between the vessel and the oilrigs and also between the vessel and shore base.

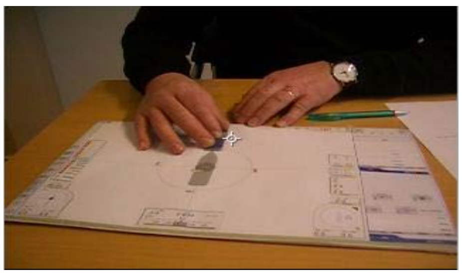


Figure 3: Paper Prototype

Figure 4: 1st Gen Prototype

Figure 5: 2nd Gen Prototype

The benefits of collecting observation data such as described above are that it provides a much more detailed understanding of the processes onboard a vessel. This will provide better knowledge when developing equipment and will save both time and money when the knowledge gained can prevent stepping into the most obvious pitfalls. The most beneficial time to do an observation study is in the early stages of the research or development process. An important preparation was to read the related literature mentioned above to gather information about what to look for and which questions to ask.

The limitations of this observation study are that only one vessel has been observed and that the observation was very time consuming. It did however give a valuable insight in what life at sea onboard offshore vessels is like and how the procedures concerning the different operations are carried out.

4. PAPER PROTOTYPING

Prototype development is a well- known technique for testing concepts and designs [Dix et al, 1997]. There are several different levels of prototyping varying from lo-fi (low fidelity) prototypes made in low-cost and easy accessible material to working prototypes made of hardware and software. The lo-fi prototype is often the first one created to test the basic functionality and to study which direction to follow before investing heavily in development.

A good initial study can save resources and prevent obvious errors during product development. The close to full functioning prototype is created in the last stages of the development and demands a larger amount of resources. Each prototype goes through an iteration of usability studies to discover errors and faulty design decisions. This is called iterative design where the design can be modified and redesigned to correct any false assumptions that were revealed in the testing. This initial prototype utilize the throw-away approach, due to that the results from the testing is used for next iterations, but the prototype itself is discarded and is not to be used as the final product [Dix et al, 1997].

The aim of our initial prototype study was to investigate which gestures would feel natural to use when operating a touch-screen DP-system [Bjørneseth, 2008]. To find out which gestures a selection of test participants were given a set of tasks to carry out using their hands directly on a lo-fi prototype. The participants were eight Rolls-Royce employees with experience from developing DP systems and tuning /installing DP systems onboard vessels. There was not given any guidance on how to proceed through the exercises or what gestures to use. This was due to the desire to investigate if it was possible to find common suggestions for movement/gesture for each task across participants.

The interface presented to the test participants, was a simple rectangle shaped piece of cardboard where a printout of the DP system's 3D interface was glued on. See Figure 3. The prototype display was placed in a desk-like position in front of the operator, adjusted to support usage of both hands. In the centre of the interface a grey boat was visible. This was displayed in a grey colour following the colour scheme the DP interface uses when the vessel is in a transitional between two positions. The vessel has a blue colour when it is in position and is not moving. To represent the blue vessel indicating the vessel in position, a small boat cut out from cardboard was used. On top of the small vessel a blue print-out from the authentic system was glued on top of it. The users moved this cardboard vessel when conducting the tasks given. A camera was used to record the movements on the surface of the prototype. The participants were given the same nine tasks, but in a randomized order. Before the tasks were carried out, the participants were encouraged to move the vessel in any way they found natural, regardless using one or two hands or touching the prototype display with more than one point. The tasks given were to move the vessel in all linear directions and to change the vessel's heading by rotation. Last the participants were asked to suggest method on how they would zoom in the 3D scene, pitch and roll the vessel. The last minutes were spent on a post-task walkthrough in addition to a general discussion regarding which gestures were preferred. In this experiment no quantitative data was collected and there were no hypotheses or experimental variables. This was due to it being a small experiment where the aim was not to compare different interfaces, but to investigate the possibilities within an interface.

The usability methods used to obtain the results needed from this study was to utilise the lo-fi prototype to do the simple tasks with a small collection of participants with knowledge about DP systems and maritime processes. Their knowledge was utilized to get a wider picture of why the hand gestures suggested could be usable in a DP system. To make the most out of the small experiment, the post-task walkthrough supported the results with the participants' thoughts on the different gestures selected. Video and audio recordings were useful tools to review the data and as a support details were noted down throughout the experiment. The combination of the above gave results worth building a new study on to investigate the impact of hand gesture interaction further. The outcome provided four hand gestures that the users felt were natural to use when operating the touch-screen DP system by directly manipulating the vessel in the system's 3D scene. These gestures created the basis for developing new and more advanced prototypes, with the gestures implemented. This made it possible to do user studies to investigate the pros and cons of using gestural interaction in maritime graphical user interfaces. The limitations of paper prototyping are because of their simplicity that paper prototypes do not support the evaluation of fine design detail. Due to the use of paper and a human operator, this form of prototype can not be reliably used to simulate system response times [Retting, 1994].

5. INITIAL SYSTEM PROTOTYPING

The purpose of system prototyping is to discover errors and design faults before the final product is released. As mentioned above, prototyping is a part of an iterative design process that can be described by the use of prototypes and artefacts that simulate or animate a selection of features of the intended system. There are three main approaches to prototyping which are described by Dix et al. [Dix et al, 1997] as the throw-away approach, the incremental approach and the evolutionary approach. In our work three different prototypes were created that were built on the throw-away approach – discarding the prototype after collecting the data needed and using the knowledge gained to build the next product. The initial study above was conducted using a lo-fi paper prototype and was followed by two generations of software based prototypes tested on different hardware platforms.

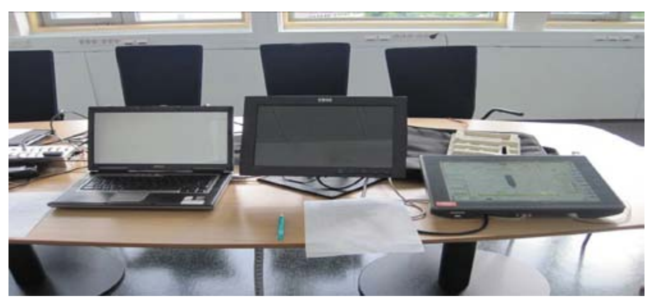
The first generation prototype was built by using the Rolls-Royce Icon DP system, a NextWindow multi-touch display using optical technology (See Figure 4). The standard DP system's graphical user interface (GUI) was extended in Java, while the NextWindow drivers were programmed in C++ and C#. This first generation prototype enabled a second iteration of user tests where the aim was to uncover if operating the DP system using multi-touch and direct interaction with the GUI's 3D scene could be faster and more efficient than using single touch and button/menu interaction with the GUI. The experiences obtained from the initial lo-fi prototype were built into this software prototype to be able to test the gestures found in a working environment. A user study was carried out following standard user laboratory study procedures that are widely used in interface design [Dix et al, 1997], having been adapted from psychology experimentation methods.

To fully exploit the advantages of prototyping, the natural steps between each generation of prototypes are usability testing and usability studies. The experiences obtained and the results gained from this provide the base for the next generation of prototypes. What is covered under the term usability study is a study that demands a well planned setup. This includes planning the study, doing individual sessions with each test participant, thoughts about the observer's role, the outcome of the study and which tools to use to obtain data and analyse the results. The usability testing concerns the separate test where the aim is to measure performance, accuracy, recall and subjective response. Usability studies can give a good insight into the user's response to the system and gives the possibility to weed out serious faults before the final decisions towards the product are made. For maritime equipment and software, the costs of replacing equipment with bad usability are so high that it is only done if the product represents a safety hazard. The process of developing controlled experiments that can provide robust results has been by Blandford, Cox and Cairns [Cairns and Cox, 2008]. The focus for controlled experiments is on quantitative data and it is important to select the appropriate population for the experiments. The participants recruited must be familiar with the tasks and have knowledge about the experiment's surrounding scenario. In this project the appropriate population for all user studies was participants who had knowledge of DP systems, but not extended experience. Through observing the test participants the results from this test were hoped to demonstrate the potential efficiency of using multi-touch.

The participants selected for our user studies were a mix of people with DP experience, students studying to be officers and DP operators on vessels and for the pilot study students with various backgrounds. This was because the system is safety-critical and from previous research it has been proven that under excessive stress the knowledge of an experienced operator is lowered to the level of a novice [Redmill and Rajan, 1997]. Before the user studies were carried out the ethical considerations were taken into account. This is important to maintain the participants' trust. This was done by making all participants sign consent forms and make them aware that they could leave the experiments at any time. None of the participants were in this case were particularly vulnerable (i.e. children), but some maintained their right to not have video clips or photos published.

To carry out a user study, the experimenter usually has a hypothesis. In order to test the hypothesis a set of dependent variables (what the experimenter will control) and independent variables (what will be measured) must be identified, with the value of the dependant variables depending on the independent variable. For small and simple studies there is normally only one dependent variable, but for larger studies this number can increase. For the last study done in this project three dependant variables were studied: average time spent on each task, average error rate on each task and reaction time to environmental distractions. In formal experiments it is also important to minimise the number of confounding variables that are varied unintentionally between the conditions during the experiment. Partly to address this, the studies were designed using a within-subject design where all participants repeated the same, or a very similar procedure, several times with different variations of the independent variable (experimental conditions). This approach can lead to learning effects, so the experiments were balanced with an even split of which experimental condition users would first encounter.

After the above setup was selected, a procedure describing the process of what the participants are supposed to do was fixed. This procedure ensures that all participants are treated the same and also makes it possible for others to replicate the experiments. To make the experiments more robust, pilot studies are recommended. For this particular project, three user studies and one pilot study supporting the last iteration of user studies were carried out. Post-experiment it can be desirable to collect some additional qualitative and quantitative data. This data is collected by conducting a post-task walkthrough and make the participants fill out questionnaires that can be quantitatively measured by using Likert-scaled questions combined with the participants' opinion on specific matters. Post-task walkthroughs and questionnaires were utilised for all the experiments in this project. To gather and safely keep the results of the experiments protocol analysis has several different methods. In this case paper and pencil in addition to video recording was used. By using the first generation prototype for a second iteration of user tests, it was possible to discover issues that concerned not only the gesture interaction, but also issues concerning the display technology. This emphasizes the advantages of doing prototyping and user- studies as an iterative design process. The drawbacks of prototyping are however the time spent on it together with not being able to test aspects such as safety and reliability. These features are often the most important, but will in a prototype be non-functional [Sommerville, 1992]. The



feedback from the test participants after finishing the

Figure 6: Static Lab Environment

user study and going through a post-task walkthrough led to the development of the second generation prototype which will be described in section 6.

5.1 TESTING GESTURES VS. BUTTONS/MENUS IN A STATIC ENVIRONMENT

The second iteration of user studies was built on the results from the initial study with the lo-fi paper/cardboard prototype. This study was carried out using the 1st generation prototype that implemented the four gestures into the Rolls-Royce DP system's software and the NextWindow display. This study employed the same tasks as the first user study in order to investigate the difference in interaction time between the use of gestures versus buttons and menus in a static lab environment. See Figure 6. The experiment included one independent and one dependent variable and one hypothesis was selected for testing.

Eleven first year nautical students from Aalesund University College participated. They had knowledge of DP systems in general, but no practical DP experience. This would make it easier to recognize a trend when operating the system using the different methods, because experienced/expert users can be predisposed from other DP systems, which could distort the result of the experiment. The reason for choosing test subjects with little to no experience is also based on earlier research [Redmill and Rajan, 1997], which implied that the experience of a skilled user is reduced to the level of a novice under strain and extreme stress. A system that is easily understood and operated by a novice will also support the skilled user in a safety-critical and stressful situation. During the experiment two touch screen systems were used, one with multi-touch functionality and one with standard single touch functionality. This was connected to an authentic DP application where a Rolls-Royce Marine Controller was used to supply the GUI with data. Four of six available degrees of freedom (DOF) were tested; surge, sway, heave and pitch. The participants interacted with the vessel in two different conditions: button-based and using multi-touch. The tasks were identical for both conditions, but the methods used to interact were different. Initially the participants declared how well they knew Dynamic Positioning and operating DP systems. The experiment consisted of four parts: plenary session, introduction, series of tasks and a post task discussion. The students were briefed in plenary in a lecture room. All tasks were videotaped and the timestamp for each operation was recorded by the camera. All participants were given the same nine tasks to complete twice in each condition, in order to measure learning between first and second attempt. The tasks consisted of four tasks that changed the vessel's position and five tasks that oriented the camera in the 3D scene. After completing the tasks, the last minutes were spent on a post task discussion where questions concerning the understanding of the system, their overall impression of the interaction techniques presented and how they felt about operating a DP system using gestures were asked.

5.1(a) Findings and Methods used

To extract results from the experiment outlined above, the simple method used for the initial paper prototype study was extended and an experimental evaluation carried out to test the hypothesis. By using video recordings to time the different tasks, it was possible to measure the difference between the two presented methods and by using the timestamps for each task a statistical method was selected to analyse the data. Selecting the correct and most appropriate statistical test can be difficult. In this case a two-tailed paired t-test was selected due to the simple structure of the experiment with few variables and only one hypothesis. The outcome from the statistical tests gave an interesting and overall significant statistical result that supported the hypothesis. The difference was significant ($p<0.05$, two-tailed paired t-test ($p\approx 0.0013$)) and shows that direct multi-touch interaction performed faster overall. It was however not uniform across the tasks. Due to the within groups design on the study the transfer of learning was likely to occur. This was therefore measured and between the first and second attempt the users improved in both interactions techniques. Overall, when comparing the observational and numerical results, it is clear that using direct gesture interaction is faster. Furthermore, according to participants' comments, it is more intuitive than the traditional button/menu interaction, though this has not been scientifically proven. The participants suggested a better display surface with more resistance and also a system that is less sensitive to touch and included a rotation gesture. The reason for more resistance on the surface was that the glass overlay caused problems for participants with moist hands. Their fingers kept sticking to the surface and made it difficult to carry out a continuous gesture movement. A system less sensitive to touch was suggested because even a small movement caused the system to register a gesture or a touch point. There is a general optimism towards direct gesture interaction, provided the technology is improved and made as optimal as possible. This together with the detailed statistical data is out of the scope of this paper and will be separately discussed in a future research paper.

Reading the post task discussion before doing the statistical analysis very often gives a good pointer towards which results will be unveiled from the statistical calculations. This was also the case for this experiment. The participants felt the system was overall a good regardless of which interaction techniques used. After a user study has been carried out and reported, critique is an issue that is either welcomed or despised. Possible critique for this particular study could be the lack of error rate analysis. If this had been added as an additional hypothesis and variable the structure would increasingly be more advanced and possibly another statistical test should have been chosen, such as the much often used ANOVA test. The experiences from this study were the foundation of a new and extended study described in the next section.

6. REALISTIC PROTOTYPE TESTING

The second generation prototype was built on the same software base as the previous, Rolls-Royce Icon DP. However a new generation tablet-PC replaced the NextWindow touch display (see Figure 5). The Dell Latitude XT2 tablet has a 12.1" mulitouch screen, runs Windows 7 (the first mainstream OS supporting multi-touch interaction) and uses touch drivers from NTrig. The tablet computer's display surface feels better to touch and is less sensitive than the NextWindow glass overlay, solving the issues raised by users of the previous prototype. The second generation prototype was used in two different iterations of user studies where the aim was to investigate if and how movement would impact operating the system using multi-touch interaction versus buttons and menus.

6.1 TESTING GESTURES VS BUTTONS/MENUS IN A STATIC VS MOVING ENVIRONMENT USING 2ND GENERATION PROTOTYPE

6.1(a) Motion Platform Pilot Study

The purpose of this experiment was to investigate the differences between manipulating a computer displayed object using gestures or buttons in a static environment versus in a moving environment, using a tablet computer and a movement-platform to simulate sea movement. Eight students with various backgrounds from Aalesund University College participated. They utilised the 2nd generation prototype tablet computer with multi-touch functionality to carry out the experiment. In addition they were seated on a moving platform which moved according to settings which simulated different conditions, in this case rough sea.

The participants were presented a collection of photographs displayed in a standard photo viewer (Windows Picture and Fax viewer). They interacted with the displayed photos in four different conditions (interaction x environment). The tasks were identical, but the setting while interacting to achieve the task goal was different. The tasks were conducted in a non-moving and in a moving environment. In each environmental condition, the participants carried out the tasks using two different interaction methods, multi-touch interaction or the buttons and menus manipulate a picture in the photo viewer. The purpose of using gestures to manipulate a photo was to relate it to using similar gestures to manipulate a vessel in the 3D scene of a dynamic positioning system. This allowed us to investigate the pros and cons of using gestures in a moving environment. Between the task sets and post-task the test participants filled out NASA TLX questionnaires. NASA-TLX is a subjective workload assessment tool that allows users to perform subjective workload assessments on operator(s) working with various human-machine systems [Cairns and Cox, 2008]. The motion platform pilot gave insight in how to perform a larger study using the DP system and also gave an indicator towards the impact of movement and which technique was more efficient. The test-participants felt more comfortable to operate the interface using gesture interaction. In addition it gave insight into practical considerations such as the screen being slightly unstable, indicating that for the main study support of a device in shape of a lectern or similar is needed.

Doing a pilot study gives benefits so it is possible to pre-test conditions and gathers experience towards planning the main user study. The equipment can be tested and the researcher can also get acquainted with the environment where the study is to be carried out. In this particular pilot study a HSC ship simulator was utilized that demanded some training to operate. The next iteration using the second generation prototype was the follow-up study.

6.1(b) Using Gestures to Operate a DP System in a Static vs. a Moving Environment

The fourth and last iteration of user studies was carried out also using the 2nd generation prototype, but with the Rolls-Royce DP system running on the tablet computer. The aim of this experiment was to investigate the differences between using gestures or touch in a static environment versus a moving environment. This iteration used the authentic DP

Figure 7: Moving environment using HSC simulator

system, the HSC simulator with motion platform and a live visualisation where vessels were crossing at specified time intervals. See Figure 7. The study included 19 test participants with maritime background who encountered several conditions using gestures versus buttons and menus, and with and without movement of the motion platform. With buttons, in this case, the real context is using soft-keys on a display. A set of experimental parameters were prepared with independent and dependant variables and a set of four hypotheses.

Pre-experiment the participants filled out consent forms and gave information about themselves and their level of experience. The participants interacted with the vessel in four different conditions. Two Likert-scaled questionnaires and a comment-sheet were filled out, one between change of conditions and two (one Likert and one comment) after completing all conditions. The tasks were identical for all conditions, but the interaction style used to achieve the goal of the tasks was different. The instructions were given verbally in Norwegian read from a manuscript, so that it was the same for all participants. The tasks were conducted in a static and a moving condition where the participant carried out the tasks using multi-touch interaction manipulating the vessel in the 3D scene. The test participants used their hands to perform different gestures that changed the vessel's direction. The tasks were also performed in the traditional way using buttons and menus in a moving/non-moving environment. During the study the test participants also had to keep an eye out the window for crossing vessels. When they spotted a vessel in the visualisation, they notified the observer by saying: "Boat!". The reason for adding a distraction task to the experiment is connected to the discoveries made during the observation study. Here it was observed that the operator spent most of his/her time looking out the windows during operation to ensure safety on deck and around the vessel [Lumsden et al., 2008]. This study was thus a 2x2 study design resulting in the 4 conditions mentioned above. The conditions were tested in a within-group balanced study. The conditions are consistently tested with visualisation in the simulator, which means that the test subjects see a moving landscape when looking out of the bridge windows. For a within-subject design all users have to do four combinations of the conditions, with conditions in counterbalanced order to counteract learning effects.

6.2 Findings and Methods used

The experimental evaluation utilised the motion platform pilot study, hypotheses and experimental variables, within-group design, protocol analysis, questionnaires and a post-task walkthrough. To analyse the data the method used was adapted to suit the study design with more variables. The possible critique of the static environment experiments was addressed by investigating error rates and by selecting a more frequently used statistical test, repeated-measures ANOVA. The evaluation tested the hypotheses. For protocol analysis video recording was utilised to time the different tasks. While full analysis is still ongoing, initial results indicate that movement had little to no effect on the task performance. Under both conditions (gestures/buttons) the participants showed good awareness out the windows, but a difference between them appeared when performing tasks using gestures. During lookout using buttons, the participant lifted the fingers off the display, scanned the scenery and then shifted the attention down to the screen to finish the task. When using gestures, the participant kept moving the vessel while at the same time as scanning for crossing vessels. This gave a more flowing and dynamic interaction with the screen and supports that the interaction can be more efficient when using gestures in this type of scenario.

7. FUTURE METHODS

The study so far has focused on observing users and attempting to give a solid scientific foundation to attempt to prove the scientific hypotheses that multi-touch interaction with a DP system is more task appropriate than using the traditional interaction techniques mentioned. The results from the final tests do not appear to strongly support this, but detailed analysis still needs to be done to set any fixed conclusions. If Rolls-Royce Marine were to go forward with this, the next steps would be to build a fully operational system based on the results from the studies and to do think-aloud sessions with prospective users. The think-aloud technique involves participants thinking aloud as they are performing a set of specified tasks. Users are asked to say what they are looking at, thinking, doing, and feeling, as they go about their task [Lewis, 1982]. This provides a quick way of revealing problems people have with a working system and the possibility to correct them.

8. SUMMARY AND CONCLUSION

This paper has described the different steps of a research project where different methods were used to evaluate human computer interaction in a maritime environment, here a DP system. The project's aim was to investigate the possibilities of introducing multi-touch and direct manipulation of 3D objects in the DP system's GUI as an additional interaction technique. The methods used are well known from consumer-based research. Through several different iterations of prototyping followed by user studies it was possible to find answers to the presented hypotheses. The hypotheses were concerned with the differences between operating a DP system when using gestures versus traditional buttons and menus in a static versus moving environment. Three different prototypes were tested; where the initial was lo-fi and the first iteration of user tests produced the results needed to develop a second prototype where they could be implemented. The 1st generation prototype was software and hardware based and gave the possibility of investigating multi-touch interaction in a static environment through a second user study. This study resulted in interesting and significant results. The feedback from the test participants were taken into account and a 2nd generation prototype were developed using similar software adapted to fit new and better hardware. Two iterations of user studies were carried out using the 2nd generation prototype, one pilot study and one larger study. The larger study was the last study in this project that concerned operating a DP system using buttons/menus versus gestures in a static versus moving environment. In addition the participants were distracted in order to keep their focus on both the interface and the surrounding environment. Preliminary analysis of the data seems to indicate that the movement had little to no impact on performance and that using gestures during look-out lead to a more flowing interaction.

After generating four different iterations of user studies and testing, it was possible to reveal issues that would have stayed hidden if decisions of selecting hardware and software were taken without utilising these methods of low-cost testing. In addition the results from extensive testing can be reused and used to create guidelines for similar types of problems to be addressed. The limitations of creating several prototypes are however that it is time consuming and often it is not possible to test all conditions to make it as authentic as possible. The other issues that limits the iterative design is that design decisions often made in the very beginning of the design process may be wrong. Dix et al. [Dix et al., 1997] state that when initial decisions are wrong, the design inertia can be so great as never to overcome an initial bad decision. In theory this means that an iterative design will discover changes that need to be made, but in practice there can be bad decisions within the basic design that are not unveiled and dealt with (finding a local minimum, but missing the global one). The other issue is that if a usability problem is diagnosed through testing, it is important to investigate the background of the problem and not only deal with the symptom. It is therefore important when working with an iterative design process to support the process with additional methods and thorough testing.

Overall for safety-critical environments such as the maritime sector, the process of investigating the product's surrounding environment and influencing factors is time consuming, thorough and expensive. However, this process is carried out once for each product area and by reusing the knowledge gained; user studies can more efficiently and often be carried out. That can save money and time by avoiding obvious pitfalls and faulty design decisions. In addition, it can give a more satisfactory product where the equipment is actually usability tested where the base for decisions made is grounded in reported research.

ACKNOWLEDGEMENTS

Many thanks to: Aalesund University College and the Nautical Institute, Mark and Eva for being ace supervisors, Rolls-Royce Marine AS for supporting my research and providing a great working environment, the Geek Ladies for great support and to the ones at home whom I could never be without.

REFERENCES

Bjørneseth, Frøy Birte et al. Dynamic Positioning Systems- Usability and Interaction Styles. NordiCHI'08. 43-52.

Bray, D. 2003. Dynamic Positioning, 2nd Edition.

Bryman, A., 2008, 'Social Research Methods', Oxford University Press

Buxton, W. and Myers, B. 1986. A Study in Two - Handed Input. ACM CHI Conference on Human Factors in Computing Systems. 321- 326.

Cairns, P. and Cox, A., 2008, 'Research Methods for Human Computer Interaction', Cambridge University Press.

Dix et al., 1997, 'Human Computer Interaction 2nd edition', Prentice Hall

Fairplay, LLoyd's Register. World Casuality Statistics. s.l. : LLoyd's Register Fairplay, 2007.

Faulkner, X., 2000. Usability Engineering, Basingstroke, Macmillan Press Ltd.

Forlines, C. et al. 2007. Direct- Touch vs. Mouse Input for Tabletop Displays. ACM CHI 2007 Procs. 647-656.

Galliers, J. et al. 1999. An impact analysis method for safety-critical user interface design. ACM Transactions on Human Computer Interaction (TOCHI), Volume 6, Issue 4. 341 – 369.

Han, Jefferson Y. 2005. Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection. UIST'05, October 23-27, 2005, Seattle, Washington USA.

Hancock, M., Carpendale, S., Cockburn, A., 2007. Shallow- Depth 3D Interaction: Design and Evaluation of One-, Two- and Three-Touch Techniques. CHI 2007 Proceedings, Novel Navigation, 1147 – 1156.

IMO, 1997, World Maritime Day, Optimum maritime safety demands a focus on people http://www.imo.org/Legal/mainframe.asp?topic_id=339&doc_id=889

Jordan, B. and Henderson, A., 1995, ’Interaction Analysis: Foundations and Practice’, Journal of the Learning Sciences, Vol. 4 , No. 1, pp. 39-103

Kabbash, P., Buxton, W., Sellen, A. 1994. Two-Handed Input in a Compund Task. CHI’94. ACM Human Factors in Computing Systems. 417 – 423.

Lazet, A. and Schuffel, H. 1977. Some applications of human engineering to wheelhouse design. This Journal, 30(7), 77 -86.

Lee, SK. et al. 1985. A Multi-Touch Three Dimensional Touch- Sensitive Tablet. CHI’85 Proceedings. 21 – 25.

Lewis, C. H. 1982. Using the "Thinking Aloud" Method In Cognitive Interface Design. Technical Report IBM RC-9265

Lumsden, J. et al., 2008, Evaluating the Appropriateness of Speech Input in Marine Applications: A Field Evaluation, ACM MobileHCI, pp. 343-346

Mills, S. 2005. Designing Usable Marine Interfaces: Some Issues and Constraints. The Journal of Navigation.58. 67 – 75.

Mills,S. 1998. Integrating information – a task-oriented approach. Interacting with computers 9. 225 - 240.

Mills,S. 2000. Safer Positioning of Electronic Fishing Aids. Cambridge University Press.Journal of Navigation, 53: 355-370.

Paul, G and Stegbauer C. 2005, ‘Is the Digital Divide Between Young and Elderly People Increasing?’, First Monday, Vol.10 No.10

Reason, J. Human Error. Cambridge : Cambridge University Press, 1990.

Redmill, F. and Rajan, J. 1997. Human Factors in Safety Critical Systems. Oxford: Butterworth Heinemann.

Rettig, M. 1994 ’Prototyping for Tiny Fingers’, Communications of the ACM, Vol. 37 Issue 4, pp 21 – 27

Sillitoe A.,Walker O., Earthy J. The Case for Addressing The Human Element in Design and Build. London : Royal Institution of Naval Architects, 2009. pp. 19-27.

Sommerville, I., 1992, ‘Software Engineering, 4th edition’, Addison-Wesley, Wokingham.

Song, Y. Analyzing Human Errror Triggered Accidents Onboard Ships Against Ergonomic Design Principles. London : The Royal Institution of Naval Architects, 2009. pp. 97-104.