Suitability of Tangible and Touch for Ship Navigation

Candidate name: Tore Hånde

University College of Southeast Norway
Faculty of Technology and Maritime Sciences

MASTER THESIS

May, 2016

Abstract

Tangible User Interfaces and Touch interfaces have become increasingly popular as ways of providing direct coupling between the user and the interface, but how suited are these input methods for ship navigation? This thesis investigates the possible differences in workload, user experience and visual gaze when using tangible and touch controls on a tabletop display for the purpose of ship navigation. An experimental study using within-subject design was conducted. 21 experienced navigators participated by navigating a ship in a simulator environment, using tangible and touch controls. To empirically test the two controls, three methods of data collection were implemented. 1) NASA R-TLX measuring subjective workload. 2) User Experience Questionnaire (UEQ) measuring User Quality, User Design and Attractiveness. 3) Video recordings to establish where the participants were looking (Visual Gaze) during navigation. The data was analyzed as paired t-tests. Findings indicate that the tangible controls are more suitable for ship navigation than the touch controls. Workload was perceived as higher with the touch controls. User experience was rated higher with the use of tangible controls, and it was found that the participants spent far more time looking at the interface during navigation with touch control compared to tangible control, which required less visual attention.

Keywords: Tangible User Interface, Touch Interface, Ship Navigation, Tangible controls, Touch controls, Tabletop display navigation.

Acknowledgements

I would like to take this opportunity to thank my supervisor Professor Kjell Ivar Øvergård for the guidance and encouragement he has provided me throughout the writing process of this thesis. I have been lucky to have a supervisor that cares about the maritime industry, research and statistical analysis as much as he does. In short, I could not have wished for a better supervisor.

A special thanks is also given to Bjørn Fjærli, who sat patiently explaining the necessary programs used for data coding and analyzing, while also being busy with his own work. He was also kind enough to dispose his office to me, so I could focus on coding in quiet surroundings.

I also wish to extend my gratitude to the students finishing their Maritime Management program at the University College of Southeast Norway (HSN) for sharing their knowledge with me, the discussions and for motivating me to work harder. Among these students, a special thanks is in order for Kristin Skyrud for proofreading my thesis.

Lastly, I want to once again thank both Kjell Ivar Øvergård and Bjørn Fjærli for allowing me to participate in the MACS project, which has provided me with a sneak peak into technology that might shape the future of ship bridge interfaces.

Tønsberg, May 2016

Tore Hånde

Table of Contents

Figures and Tables

Introduction

The ship's crew consists of people with different professions and ranks. These are essential to the safe voyage and operation of the ship. On the bridge, the captain or an officer of the watch is tasked with navigating the ship - hereafter referred to as the Navigator. A ship's bridge contains several information sources, such as available routes, ECDIS, Radar, propulsion systems, weather conditions and communication systems. These sources inform the Navigator about how the systems are working, the journey and the current surrounding environment (Schager, 2008). As a result of this, the navigator is progressively becoming a manager - who has to manage a combination of systems of varying complexity (Bowditch, 2002). Thus, it is important that the Navigator is able to stay updated on all of these sources while simultaneously navigating the ship.

When we think about navigation systems and equipment onboard vessels, we tend to think control panels, levers and graphical user interfaces. However, technology has evolved towards different interaction approaches; such as touch and tangible user interfaces. Our daily lives are filled with reminders of the former with the increasing number of tablets and smart phones. Vessels in the Maritime industry have also been increasingly equipped with touch solutions, with everything from simple touch functions for stopping/starting smaller systems, to bridge map interfaces and established systems such as ECDIS. But these interfaces have not yet replaced some of the most standard forms of navigation, which is still done with conventional thruster control systems, consisting of levers and buttons. Since navigation can be compared with driving, in the sense that distraction can have serious consequences for both people and material - it is important to create interfaces that do not require unnecessary visual attention (Bjelland, Hoff, Bjørkli, C.A. & Øvergård, 2007). Both touch and tangible could be likely candidates as input methods to offer new ways of interacting with information and controlling of the ship. Therefore, a new approach to bridge navigation, using tangible and touch controls on tabletop displays, is currently being developed with both input methods capable of accomplishing the same tasks (Völker, Nakajima, Thoresen, Itoh, Øvergård, & Borchers, 2013a; Völker, Nakajima, Thoresen, Itoh, Øvergård, & Borchers, 2013b; Völker, Corsten, Hamdan, Øvergård & Borchers, 2014; Völker, Cherek, Thar, Karrer, Thoresen, Øvergård & Borchers, 2015; Völker, Øvergård, Wacharamanotham & Borchers, 2015; Thoresen, Øvergård & Hancke, 2015).

There is limited empirical evidence on this subject, as tabletop displays with tangible and touch controls have not been considered for vessel navigation. Hence, the aim of this thesis is to explore how Navigators respond to the two different input methods of touch and tangibles, and how suitable these are for vessel navigation.

Background

Traditionally, simple input devices such as mouse and keyboard have been the norm, where the mouse is used as a way of interacting with the graphical icons displayed on a screen. This solution, usually referred to as Graphical User Interface (GUI), describes your typical Personal Computer (PC) running Windows or Macintosh software. As time moves forward, new ideas have resulted in new interfaces that offer different ways of interacting with information. One of these interfaces is termed Tangible User Interface.

Tangible User Interfaces. In the mid 1990s one such interface was introduced as what we now know as Tangible User Interfaces (TUI). The TUI was originally introduced as Graspable User Interfaces, by Fitzmaurice, Ishii, and Buxton (1995), in the form of "bricks". In their article, they proposed a concept where elements from the virtual user interfaces take physical forms, essentially allowing control of virtual objects through the physical input devices (Fitzmaurice et al., 1995). In other words, an idea that the physical (tangible) representations of information serves as direct control mechanics of the digital information (Ishii, 2008). Since its introduction, several studies have explored the TUIs capabilities.

One of the strengths of the TUI is that it takes advantage of the complex skills we develop in our daily life while manipulating the physical environment by giving physical form to digital information (Ishii, 2008). Instead of relying on a mouse to perform multiple actions with the use of pointing and clicking, physical input devices can be specialized for a particular task, offering direct manipulation. Fitzmaurice and Buxton (1997) argue that these specialized physical input devices provide performance advantages over other interfaces. Since the devices can be constructed to match the skills obtained by interacting with similar devices in the real world, this might reduce the time needed to learn the interface compared to other generic interfaces.

The TUI device can be distinguished from the traditional GUI mouse, due to its functions being dedicated to specific purposes, while the mouse needs to both navigate and click to get the desired effect. Fitzmaurice et al. (1995) referred to this as space-multiplexed (e.g. tangible) and time-multiplexed (e.g. mouse). In this sense, time-multiplexed input is seen as the controlling of different functions at different times, while space-multiplexed input refers to the controlling of physical objects dedicated to one function, within a space. By comparing space-multiplexed input against time-multiplexed input, Fitzmaurice & Buxton (1997) found specialized space-multiplexed input devices to outperform time-multiplexed. Because of these qualities, different TUIs have been made for different purposes, such as the MIXI TUI, which is a used to perform live electric music (Pedersen & Hornbæk, 2009).

Touch Interfaces. Another interface that has become increasingly popular these last few years is the touch interface. In this thesis, touch interfaces are defined as an interface relying on touch maneuvers to control functions. Here, hand gestures or the finger is used as the input device to interact with functions on a display, such as widgets, rather than with the use of the traditional mouse. Touch has been very prevalent in the form of smartphones and tablets and continues to be integrated into other devices. The reason for the increasing amount of touch offered solutions could be explained by the increasing improvements to hardware and software, allowing for a more responsive experience. Other factors, such as the small form factor and the robustness as compared to a computer mouse (Albinsson & Zhai, 2003) could also help explain the trend. However, it is important to note that touch interfaces are not without its limitations. Wigdor, Forlines, Baudish, Barnwell & Shen (2007) noted that the use of the finger as the input device could be problematic, since the finger is larger than a pixel on the display, commonly referred to as the fat finger problem. This could result accuracy difficulties during input. Similarly, the finger and arm can obscure the target before touching the display, as well as other parts of the screen (Wigdor et al.'s, 2007; Albinsson & Zhai, 2003). Arguably, the most obvious and important limitation of touch interfaces is that it requires the user to look at it to perform actions - one cannot use the touch screen without looking at it (Bjelland et al., 2007)

Comparing TUI with Touch Interfaces

One of the most apparent differences between TUI and Touch interfaces is their input method. With tangibles, we have a feel of something physical, which is weighty and responsive, providing the user with haptic/tactile feedback. While with Touch, we rely on hand gestures, or touches. Essentially, TUI enables the interaction with physical objects, translated into virtual objects, while touch interfaces are direct interaction with the virtual objects. As previously mentioned, TUI requires specialized physical controls, and as such can prove to be costly to produce. Virtual touch controls (widgets) however, can be modified or changed within the software, without the need to produce something physical, and does therefore not represent the same costs. Another important difference between the two is the size of the controllers. Both touch and TUI controls can be created to different sizes, but the touch "widgets" could suffer from being smaller, since the ability to target the functions could prove to be hard, as noted earlier. With TUI, the controls could be "built up", essentially using the available space more effectively. However, the size of tangible controls also means more space is taken up in general compared to touch, since it combines the physical and virtual worlds.

Recent studies comparing TUI and touch interfaces have found that TUIs have increased motor cognitive benefits (Antle & Wang, 2013) and are found to be easier to acquire and manipulate (Tuddenham, Kirk, & Izadi, 2010). Similarly, it has been found that TUIs results in faster completion of task, as well as being easier to learn compared to touch interfaces (Lucchi, Jermann, Zufferey, & Dillenbourg, 2010).

For example, Voelker et al. (2015) created an experiment based on a split-plot research design following Øvergård et al.´s (2007) study on turnable knobs, and presented participants with two tangible controls and two touch widget controls to perform a simple rotational task. The tangibles consisted of a knob and a puck, while the touch were widgets in the form of one touch (requiring input from one finger on the touch-screen) and two touch (require input from two fingers on the touch-screen). The movement time, from when the rotation started until the trial was completed, was used as measurement, as well as accuracy pertaining to overshoots measured by the cursor exiting target area. The results indicated that the tangibles were faster overall, with an average of 20%, and produced less overshooting compared to one-touch widget (Voelker et al., 2015).

However, not all studies have found the TUI to perform better than the touch interface. Hancock, Hilliges, Collins, Baur & Carpendale (2009) found the touch interface to be preferred for 2D rotation and translation task, but noted that tangible was more precise, and preferred for 3D objects. Similarly, Kratz, Westermann, Rohs, and Essl (2011) found the touch interface to be superior for phone interaction compared to the tangible counterpart (Capwidgets). What both of these studies have in common is that the tangible controls used were "eyes on", meaning that the digital information and tangible device was on the same screen, possibly obscuring the view. Following up on the latter, Voelker et al. (2015) compared tangibles with virtual touch widgets in the conditions of eyes-off and eyes on tasks, and found that the tangibles were still faster and less error prone than touch in both conditions. Additionally, it was found that the touch widgets lost their performance in the eyes-off condition.

From previous research, it is indicated that the tangible is the overall favored input method compared to touch. Because of this, it is expected that the TUI will outperform the touch interface for this experiment.

In order to compare the tangible and touch interfaces, the use of a tabletop display with support for both interfaces was acquired. The TUI condition will be using physical objects placed on the tabletop to control, while touch condition consists of virtual widgets shown on the tabletop display. Both interfaces involve hands-on interaction within the space of the tabletop. The tangible controls in this experiment are passive and do not rely on external components, such as cameras to detect movement.

From this point, TUI devices will be referred to as tangible controls, while touch widgets will be referred to as touch controls.

Aim of this thesis

The purpose of this thesis is to empirically test the two conditions of touch control and tangible control for vessel navigation with the use of subjective workload, user experience questionnaire and visual gaze (observation).

Hypothesis. Three main hypotheses was formed for the thesis:

H1: Tangible controls will have lower mental workload than touch controls

The lack of tactile feedback has been found to have a significantly negative effect on subjective workload (Noy, Lemoine, Klachan, & Burns, 2004). Since touch controls lack the tactile feedback available with tangible controls, it is expected that the touch control will produce higher mental workload.

H2: Tangible controls will score higher on the User Experience Questionnaire.

Since previous studies have found tangible user interfaces easier to learn (Lucchi et al., 2010) and faster to acquire and manipulate (Tuddenham et al., 2010) compared to touch, combined with the familiarity the controls bring, it is expected that the participants will rate user experience higher with the tangible controls.

H3: More time will be used looking down at the interface while using touch controls.

Due to the lack of haptic/tactile feedback when using touch controls (widgets), it is expected that the participants will more frequently look down on the tabletop when navigating the vessel.

Methodology

About the Project

This present project was conducted and carried out by a professor, a researcher and a master student at University College of Southeast Norway (HSN). The research was under the MACS (Maritime Control Systems for the future) project and was based on cooperation between HSN and Kongsberg. The MACS project is based on a goal is to create new interfaces for use in maritime control systems. In cooperation with professor Kjell Ivar Øvergård and research assistant Bjørn Fjærli, I was allowed to attend the gathering and coding of the data, and to base my thesis on the process and findings. The data collection for the experiment was conducted in the autumn of 2015, spanning over 2 months. The coding of the data was finished December 2015. During the autumn, the task that established itself as the most time demanding was the coding of the video recordings taken during the experiment.

The tangible controls used in the experiment are based on research on tangibles that are passive, but are constantly detected by the capacitive touch screen without the need to use touch to initiate this connection, and without the need for batteries to operate (Voelker et al., 2013). Lilaas AS, a partner in the MACS project, manufactured the tangible controls. They are in this experiment used in combination with a multi-touch tabletop that supports the input of both touch and tangible controls.

Participants

For the experiment, a total of 21 ship officers and captains were tested in the two conditions of touch controls and tangible controls in a within-subject design. The participants were experienced navigators who underwent a simulator course at a maritime training center in Vestfold, Norway. Age was not recorded, but the participants were from the end of their twenties to the end of their fifties. Both genders were present in the experiment, though it was a clear domination of male participants. Written consent was obtained from the participants prior to starting navigation with the two interfaces.

Data collection

The hypotheses will be answered by performing an experiment using these three elements for data collection RAW NASA TLX (R-TLX; Hart & Staveland, 1988), User Experience Questionnaire (UEQ; Laugwitz, Held, & Schrepp, 2008), Visual Gaze - meaning where the participants look during navigation training.

NASA TLX. To measure the perceived workload, the NASA TLX (The National Aeronautics and Space Administration Task Load Index) (Hart & Staveland, 1988) was chosen. This is one of the most popular tools for measuring subjective mental workload. The NASA TLX is a multi-dimensional workload measuring scale, which estimate the workload of operators while performing a task, or immediately after task completion (Hart, 2006). The scale is referred to as multi-dimensional because it uses six dimensions to assess the subjective workload of operators. These consist of mental demand, physical demand, temporal demand, performance, effort and frustration. The first three dimensions (Mental, physical, and temporal demand) relate directly to the demand the operator felt during the task, while the last three (performance, effort and frustration) relate to the interaction between the operator and the task (Felton, Williams, Vanderheiden, & Radwin, 2012). These dimension/subscales are rated by using bipolar scales, consisting of twenty steps, which the participants mark between the ranges of zero to twenty (see Appendix A).

The TLX used in this experiment was the RAW TLX (RTLX), which is a common modification to the original. In the original TLX the usual way of presenting results is by using overall workload score, calculated based on the weighted average of the subscales (Felton, Williams, Vanderheiden, & Radwin, 2012). The RTLX eliminates the weighing of the scales and instead use the average of each subscale and/or estimate overall workload based on the average of each subscale (Hart, 2006). In comparisons between the RTLX and the original TLX no differences were found concerning sensitivity (Hart, 2006), which indicate that they are equally valid forms to evaluate subjective workload.

UEQ. The questionnaire used in this experiment was based on the User Experience Questionnaire (UEQ) available at: http://www.ueq-online.org/. The original UEQ consist of 26 questions and 6 scales responding to novelty, attractiveness, Perspicuity, Dependability and Efficiency (Rauschenberger, Schrepp, Cota, Olshner, & Thomascheweski, 2013). In this experiment a modified version of the UEQ was administered, containing 12 questions, with 2 of the questions responding to each of the 6 scales. A previous user test was conducted to test the correlations/hitrate between the questions, resulting in the use of two questions corresponding to each scale in this experiment. These scales were then grouped into Attractiveness, User Quality and Design Quality. The questions used can be seen in Appendix A. Boring/Exiting and Interesting/Not interesting was grouped together into

Stimulation, Conventional/inventive and Leading edge/Usual into Novelty, Bad/Good and Attractive/Unattractive into Attractiveness, Easy/Complicated and Confusing/Clear into Perspicuity, Secure/Not Secure and Does not meet expectations/Meets expectations into Dependability, Inefficient/Efficient and Practical/Impractical into Efficiency (*see figure 1 for context*). The idea behind the UEQ is to offer a quick assessment tool for the end users covering a comprehensive impression of user experience (Laugwitz, Held, & Schrepp, 2008).

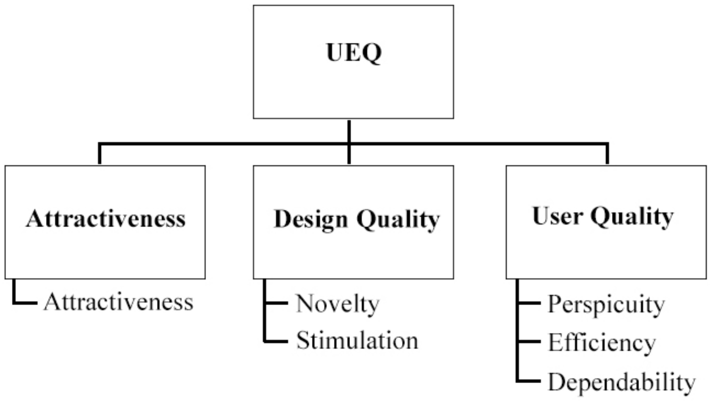


Figure 1. The User Experience Questionnaire scales as used in this thesis

VISUAL GAZE. To collect data about the participant's point of focus during the experiment, a controlled observation was implemented. Nachmias and Nachmias (2008) stated *"Controlled observations methods are characterized by clear and explicit decisions made as to what, how and when to observe"* (p. 196). This was made possible by employing video recording to the project. The reasoning behind this was to explore where each participant were looking during navigation of the ship, which was termed *visual gaze*. The participants were recorded when performing navigation with both tangible and touch controls.

The digital video files were then later scored into quantitative data using behavioral observation software Noldus Observer XT. Each participant was measured in regards to how much time they spent on the visual gaze points; Conning display, ECDIS display, tabletop display (interface), outside and elsewhere. This was made possible by assigning a key to the behavior, which the Noldus Observer registered and assigned a timestamp. The duration of the gaze points was then calculated between two different points. Since the recordings were done in the daytime with a medium lit room, the points of focus were identified relatively securely. Two cameras were used *(see figure 1)*. The first camera was placed to the left side of the participants, getting a view of the participants and the tabletop interaction. The second camera was placed over the middle 55" ship surrounding display, getting a front view of the participants.

Research Design

The research design was an experimental study using within-subject design, which allows for control of individual differences - hence within-subject design offers higher statistical power than between subject designs (Charness, Gneezy, & Kuhn, 2012). This does however introduce different components, such as order effects, where skills and experience obtained from a previous treatment may influence the next. To counter this, the order of which interface the participants would start with was randomized.

The experiment used random assignment of the subjects to conditions. However, it could be argued that it is not fully randomized in regards to sampling, because the subjects were connected to a specific simulator facility, which was deemed necessary since the availability of experienced navigators is limited.

Variables. The input method participants used for the experiment were the independent within-subject variable. The two different types of input methods used were Tangible and Touch control. The first dependent variable was perceived workload during the two conditions. The second dependent variable was the UEQ. The third dependent variable was the visual GAZE measurement.

Data analysis. For the coding of video recordings, the behavioral observation software Noldus Observer XT was used. The statistical analysis of the collected data from NASA R-TLX, UEQ and Visual Gaze was done in IBM SPSS Statistics version 23.

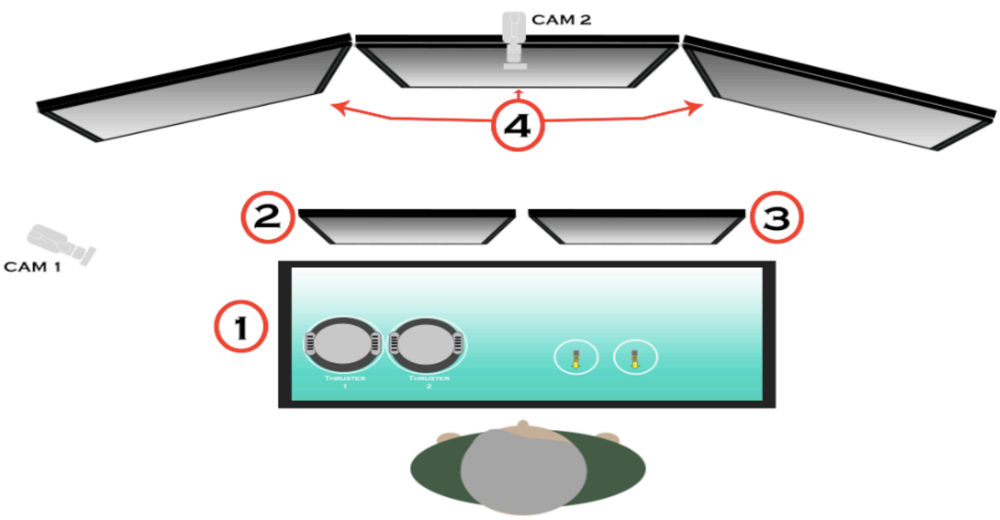


Figure 2. Simulator Overview

Simulator setup

The experiment was conducted by the use of a bridge simulator, stationed at Vestfold Innovation Park, in cooperation with Kongsberg Maritime. The simulator was running K-Sim Software, and consisted of six displays; purposely built to test the two input methods.

Two 32" LCD monitors were placed in front of the participants for ECDIS and Conning respectively. The ECDIS (Electronic Chart Display & Information System) display presented the map of the area, coupled with general route information (3). The Conning display presented information such as speed, depth, and thruster information (2). Three 55" LCD monitors were used to run the simulator bridge view, showing a 3D representation of the ship surroundings (4), essentially replicating what the navigators would be able to see when navigating the vessel in real time.

The solution to combining both touch and tangible controls in the same experiment was a 54" touch input screen (tabletop display), measuring 108 x 63.5cm, situated horizontally in front of the participants (1). The tabletop display served as the input platform for navigation with both conditions of touch and tangible controls, and measured about 95cm from the ground. This allowed the experiment to be eyes-off, as the actual simulator view, and vessel information was not on the same display as the controls. Thus, the two input methods were not obstructed by other information, allowing Navigators to see all the five screens containing the necessary information while navigating the ship.

In the tangible control condition, tangibles were placed directly onto the touch input surface, where they coupled/registered to a virtual widget on the tabletop display. The tangible controls used resembled levers available onboard vessels in design. The thrusters were controlled by *rotating* the main body of the tangibles for angle, and the smaller wheels at each end of the lever for speed. See *figure 3* for context.

In the touch control condition, "widgets" were used as a means of manipulating the available thrusters (Propulsion and steering). Manipulating these widgets by touch would allow participants to adjust the angle by following the white circle in a circular motion, and increasing speed by moving the yellow circle slider up and down. See *figure 4* for context.

A total of four thrusters were available for each condition. In tangible control, two bow-thrusters[1] were controlled with one tangible lever, while the remaining three thrusters[2] (azimuth) was controlled by three separate tangible levers. In the touch condition each thruster had its own widget.

---

[1] Bow thruster: A type of propeller built into the bow of the vessel, with the purpose of providing better maneuverability.
[2] Azimuth thruster: A type of marine propeller whose axis can be rotated 360 degrees to any horizontal angle, hence eliminating the need for a rudder.

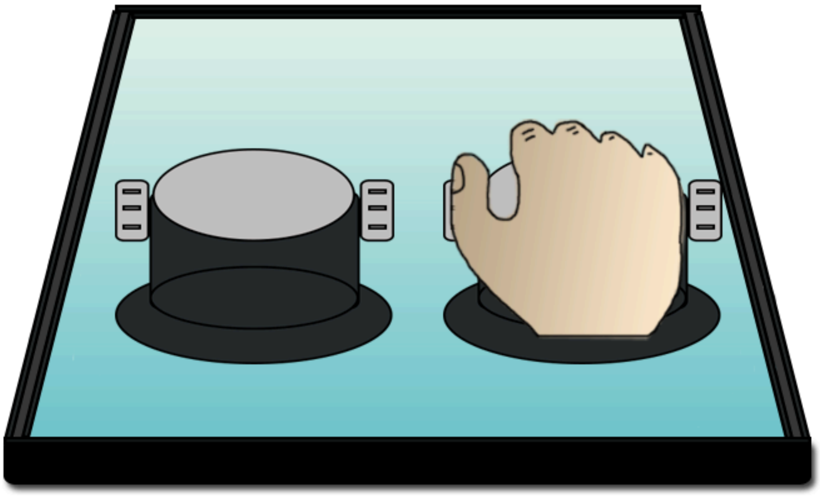


Figure 3. Tangible Controls for Navigation

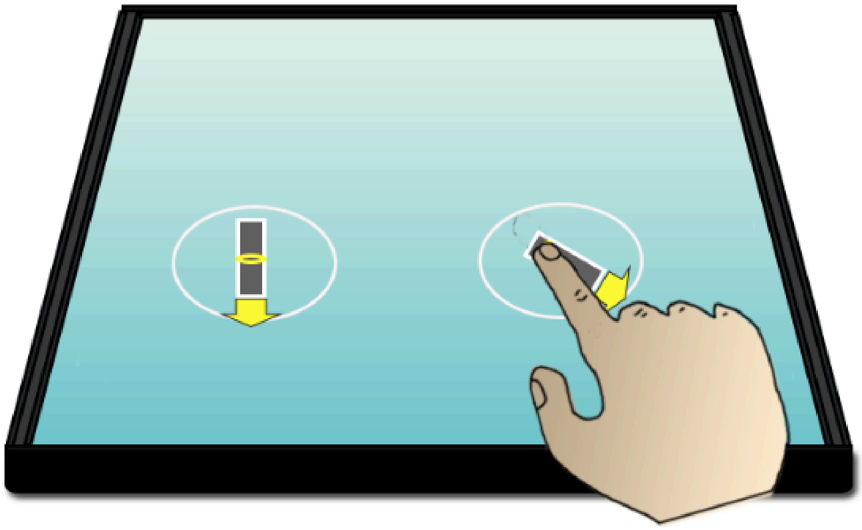


Figure 4. Touch Controls for Navigation.

Procedure

Data collection was done at a maritime training center in Vestfold, Norway. Navigators taking courses at the training center were asked to be a part the experiment and volunteered to do so. After signing the consent form the participants got a short introduction to the simulator and the tangible and touch controls. This was done to help familiarize the participants with the different input methods, and depending on their experience with similar controls, the time it took for the participants to be comfortable enough to start the sessions varied, by an estimated average of 2-3 minutes. The scenario the participants were requested to navigate involved a starboard turn and a port turn in a canal in Rotterdam. The participants performed two sessions of navigation, one with touch control and the other with tangible control, with each lasting approx. 10-12minutes.

During the experiment the participants were videotaped. Immediately after completing one condition, the participants were administered the print NASA TLX and a User Experience Questionnaire, which they were instructed to fill out.

Validity and Reliability

Validity is concerned with whether a measure of a concept actually measures that concept (Bryman, 2012). Cronbach & Meehl (1955) argued four types of validation; Predictive validity, concurrent validity, content validity and construct validity. Reliability, on the other hand, refers to the consistency of measures of a concept, and is characterized by factors such as stability, internal reliability and inter-rater/inter-observer reliability (Bryman, 2012).

For this experiment, the participants were divided into blocks of four, with randomization between the blocks and within the blocks. Similarly, as to which interface (control input) the participants started with was randomized between tangible and touch control.

Concerning the measuring instruments, the NASA TLX instrument has been used as a standard for measuring subjective workload for over 30 years. Since that time it has been translated to several languages, subjected to many evaluations relating to reliability, sensitivity and utility, and has been compared to other workload measuring methods (Hart, 2006). NASA TLX was found to be equally acceptable as a subjective workload measuring instrument in regards to sensitivity as to the Overall Workload (OW) scale, Modified Cooper-

Harper scale (MCH) scale, and the Subjective Workload Assessment Technique (SWAT) (Hill, Lavecchia, Byers, Bittner, Zaklade, & Christ, 1992). In a similar study NASA TLX showed high convergent validity and concurrent validity to other instruments (Rubio, Diaz, Martín, & Puente, 2004). As the inventor of NASA TLX recently said: *"The years of research that proceeded subscale selection and the weighted averaging approach resulted in a tool that has proven to be reasonably easy to use and reliably sensitive to experimentally important manipulations over the past 20 years"* (Hart, 2006, p.1). In this experiment the printed English version was administered (see Appendix A).

Studies on the User Experience Questionnaire (UEQ) has indicated acceptable levels of reliability and construct validity (Laugwitz, Held, & Schrepp, 2008). The UEQ questions were randomized to reduce the occurrence of the participants ticking the same number throughout the questionnaire. Instead of every question ranging from positive to negative, some were negative to positive.

With both the NASA TLX and the UEQ the data was manually added into SPSS and then double-checked by two individuals to verify that the inputs were correct.

Ethical Considerations

Because the experiment involved video surveillance, approval by NSD (Norwegian Centre for Research Data) was required. The application was approved (project number 44167). Prior to starting the sessions involving navigation with tangible and touch controls, the participant signed a consent form. The participants were informed about video surveillance, data collection methods, and that the data collection is confidential and anonymous. No discomfort was reported to have been experienced from the participants during or after the experiment.

Results

This experiment examined the effects that the two different input methods of touch and tangible had on participants, in regards to perceived workload, user experience and where the attention was during navigation (visual gaze). In all three cases paired t-tests were used on the three dependent variables, with CI 95%. Cohen's *d* was calculated for the paired samples by dividing the paired samples test mean with the standard deviation.

$$d = \frac{Mean}{SD}$$

Cohen defined the effect sizes range from small (0.2) to medium (0.5) to large (0.8) (Cohen, 1992). Due to video recording errors, one participant was not counted in the analysis on the visual gaze.

Inter-Rater Reliability

The Visual Gaze data was based on video recordings and manual input of codes. Two video files recorded from two different angles of each participant, and imported into Noldus Observer XT. The video recordings were analyzed in 0.5x speed. After the data was coded an inter-rater reliability test was performed, showing Cohen's Kappa = 0.85, 95% CI [.80, .92], and an average percentage agreement ranging from 87% to 96%. The value of Cohen's Kappa can be interpreted as a perfect agreement if 1, and chance agreement if 0. The scale used to assess this agreement states that if < 0: Less than chance agreement, 0.01-0.20: Slight agreement, 0.21-0.40: Fair agreement, 0.41-0.60: Moderate agreement, 0.61-0.80: Substantial agreement, 0.81-0.99: Almost perfect agreement (Viera & Garrett, 2005). The inter-rater reliability was 0.85 and therefore sufficient and deemed substantial and more than acceptable.

.

NASA R-TLX

Subjective workload measured by NASA R-TLX with touch and tangible controls.

Table 1

NASA R-TLX Results

Dimensions of NASA RTLX	Mean, (SD) Touch	Mean, (SD) Tangible	Mean Difference between groups [95% *CI*]	Effect size Cohen's *d*
Mental Demand	11.00 (4.43)	9,09 (4.44)	-1.91 [-3.84, 0.02]	0.437
Physical Demand	6.73 (3.97)	5.73 (3.88)	-1.00 [-3.06, 1.06]	0.215
Temporal Demand	7.18 (3.25)	5.73 (3.21)	-1.45 [-2.51, -0.40]	0.609
Performance	15.09 (4.63)	15.50 (3.75)	0.41 [-2.02, 2.84]	0.074
Effort	10.68 (4.57)	8.86 (4.06)	-1.81 [-3.85, 0.21]	0.397
Frustration	6.09 (3.69)	4.00 (3.79)	-2.09 [-3.82, -0.36]	0.535
Overall workload	47.59 (15.15)	38.90 (15.7)	8.68 [-15.85, -1,51]	0.536

Note. CI = confidence interval. Graphical representations for each dimension can be seen in Appendix C.

The average overall workload for touch controls was higher than with the tangible controls (Mean difference = 8.68, t_{21} = 2.517, p = 0.020, d = 0.536), indicating a medium effect size between the two conditions, with the tangible solution having lower overall workload. For every dimension except Performance, score was higher for touch controls than that of tangible controls. The Performance dimension was similar among touch controls and tangible controls and produced the lowest effect size. The larges difference could be observed within the temporal dimension (Mean difference = -1.45, t_{21} = 2.861, p = 0.009, d=0.609), indicating a medium effect size between the conditions, with tangible controls having lower workload. These findings are in accordance with hypothesis 1.

User Experience Questionnaire

User Experience as measured by the UEQ, presented with the use of the scales Attractiveness, Design Quality and User Quality.

Table 2

User Experience Questionnaire Results

UEQ SUM	Mean, (SD) Touch	Mean, (SD) Tangible	Mean Difference between groups [95% *CI*]	Effect size Cohen's *d*
Attractiveness	11.00 (2.27)	12.23 (1.34)	-1.23 [-2.14, -0.31]	0.59
Design Quality	24.19 (3.59)	24.29 (2.45)	-0.095 [-1.17, 0.98]	0.04
User Quality	28.00 (5.20)	32.68 (6.07)	-4.68 [-7.74,-1.62]	0.68

Note. CI = confidence interval. Graphical representation can be seen in Appendix D.

With the UEQ, the answers were collected and gathered into Attractiveness, Design Quality and User Quality. In Attractiveness, participants scored higher with tangible controls than with touch controls (Mean difference = -1.23, t_{21}= -2.783, p= 0.01, d = 0.59), which indicates a medium effect size. Similarly, with User Quality, participants scored higher with tangible controls than with touch controls (Mean difference = -4.68, t_{21} = -3.184, p = 0.004, d = 0.68), also indicating a medium effect size. However, in regards to Design Quality no noteworthy difference was found between the tangible and touch controls, with a miniscule effect size (Mean difference = -0.95, t_{20} = -0.185, p = 0.855, d = 0.04).

Because of these findings, hypothesis 2 can be seen as partially rejected, as not all aspects of user experience was found to be better for the tangible controls as compared to the touch controls

Visual Gaze

To make a more realistic representation of the participants gaze points during the navigation with both tangible and touch controls, the data was converted into percentages. This was done so that it is possible to see how much time (in percentage) each participant spent on each gaze point for the two conditions.

Table 3

Visual Gaze Results

.

GAZE Point %	Mean, (SD) Touch	Mean, (SD) Tangible	Mean Difference between groups [95% *CI*]	Effect size Cohen's *d*
Interface	0.3978 (0.118)	0.1673 (0.156)	0.23 [0.164, 0.297]	1.57
Outside	0.2971 (0.139)	0.3776 (0.204)	-0.08 [-0.135, -0.26]	0.67
ECDIS	0.2094 (0.135)	0.1674 (0.104)	0.04 [-0.007, 0.09]	0.38
CONNING	0.0943 (0.065)	0.2861 (0.176)	-0.19 [-0.266, -0.12]	1.17
Elsewhere	0.0013 (0.003)	0.0016 (0.005)	-0.0003 [-0.003, 0.002]	0.04

Note. CI = confidence interval. Graphical representation can be seen in Appendix D.

Finally, the visual gaze was analyzed. The percentage of gaze on the interface for touch controls was noticeably higher than with the tangible controls (Mean difference = 0.23, t_{20} = 7.215, p = 0.0000005, d = 1.57), showing a very large effect size, and indicating that the participants spent a much larger amount of time focusing on the interface in the touch conditions. Similarly, gaze on the ECDIS display for touch controls was higher than with the tangible controls (Mean difference = 0.04, t_{20} = 1.763, p = 0.09, d = 0.38), with a small to medium effect size, but did not produce any statistical difference.

In contrast, the participants spent more time gazing outside the ship while using tangible controls compared to touch controls (Mean difference = -0.08, t_{20} = -3.078, p = 0.006, d = 0.67, with a medium to large effect size. Similarly, gaze on the Conning display for tangible controls were substantially higher than on the touch controls (Mean difference = -0.19, t_{20} = -5.375, p = 0.00002, d =1.17), with a large effect size.

In both tangible and touch controls the amount of time spent elsewhere was similar for both conditions (Mean difference = -0.003, t_{20} = 0.189, p = 0.852, d = 0.04). Findings are in accordance with hypothesis 3.

Discussion

This thesis investigates the possible differences in workload, user experience and visual gaze for participants when using tangible and touch controls on a tabletop display for the purpose of ship navigation. By employing these three methods of collecting data, it was hoped that one of the two interfaces would produce less subjective workload among the participants and score higher on user experience. Similarly, for the visual gaze to provide insight as to where the participants focus their attention during navigation with the two interfaces, and hopefully indicate which of the interfaces require more attention. This could then be used for recommendation as to which of the interfaces are better suited for ship navigation.

Subjective Workload

Hypothesis 1 stated that tangible controls will have lower mental workload than the touch controls. Based on the results from the NASA RTLX, the subjective ratings showed that touch controls had a higher workload score than the tangible controls. The results were in accordance with hypothesis 1.

Although few studies have evaluated the subjective workload when comparing TUI against touch interfaces, one recent unpublished paper by Besancon, Issartel, Ammi and Isenberg (2016) found similar results using the same measuring instrument as in this thesis (NASA RTLX). Their findings indicate that TUI produce less workload than touch interfaces, with similar differences in the dimensions. Another experiment found tangible to produce less workload than its touch counterpart as information systems when driving (Hoff, Alsaker, & Bjørkli, 2002). These further help support our findings that tangibles produce less workload than touch. The results also revealed one particularly interesting dimension, which was the temporal dimension ($p = 0.009$). This was also the dimension that produced the largest difference, similar to what was found by Besancon et al. (2016). The dimension refers to how much pressure the participants felt, or how rushed the tasked was perceived. Tangibles have in previous studies been found easier to learn (Lucchi et al., 2010), and since participants got an equal introduction to both the tangible and touch controls, to get more familiar, it seems plausible that the increased pressure could also indicate more time is needed to learn touch controls combined with the need to look down at the interface to perform actions, as seen with visual gaze.

User Experience

Hypothesis 2 stated that tangible controls will score higher on the User Experience Questionnaire. Based on the results from the UEQ, the ratings showed that the tangible controls got higher scores than the touch controls on all scales, however, one scale did not yield statistical significance ($p > 0.05$). It should still be noted that attractiveness ($p < 0.05$) had a medium effect size ($d = 0.59$) and User Quality ($p < 0.05$) had a medium to large effect size ($d = 0.68$). Thus, the data generally supports hypothesis 2. However since there is no statistical significance in Design Quality, accompanied by non-existent effect size, Hypothesis 2 is partially rejected.

The results from the UEQ was rather unexpected, since both User Quality and Attractiveness scored higher with tangible controls compared to touch controls, but the Design Quality showed no particular difference. Design Quality consisted of the scales Stimulation (Boring/Exiting, Not interesting/interesting) and Novelty (Conventional/Inventive, Leading edge/usual). As seen in Appendix B, all four questions produced miniscule effect sizes, hence rated similar with both touch and tangible controls in the overall Design Quality scale. It seems as if the participants generally had a positive experience with the design of both the touch and tangible controllers, with questions all scoring around 6 on the 7-point scale. However, the overall results fit with a previous study by Lucchi et al. (2010), where participants found the TUI used in their experiment to be more easy to use compared to the touch interface. They also found that the participants became more stressed and irritated with the touch interface, but at the same time had more fun with the touch interface. The latter could help explain why the difference observed in our experiment within Design Quality was so small, seeing as the participants rated both inputs positive. Further, Widgor, et al. (2009) observed that users can experience lack of confidence with touch, due to the feedback uncertainty, accompanied by a increase in user frustration and confusion. Participants also said they preffered the tangible anternative when asked what input out of tangible, touch and mouse they would like to use again (Besancon et al., 2016). This general consensus supports our findings in that user experience is better with tangible controls.

Visual Gaze

Hypothesis 3 stated that more time will be used looking down at the interface while using touch controls. Based on the results from the video recordings (Visual Gaze), it was shown the participants spend far more time looking at the interface with the touch controls, showing a very large effect size ($d = 1.57$) and statistical significance ($p < 0.05$). The data clearly supports hypothesis 3, giving indications that touch controls require considerably more visual attention - hence, the interface might be dangerous to use in safety-critical work tasks which require information from other places than the interface (Fjærli & Øvergård, 2015)

Since touch interfaces require the user to look at the interface to perform the actions (Bjelland et al., 2007) the findings were expected. Unfortunately, limited research is available on the comparison between tangible and touch interfaces with focus being on where the users are looking. The closest to the controls described in this thesis is by Rümelin and Butz (2013), who found in their experiment that by using tangible controls (knob) on a touch screen helped provide blind interaction while driving, which was not possible with the other touch controls. Their experiment used three different touch controls, one tangible knob and one knob controlled car radio with visual display. Overall the control that required the least visual attention was the knob controlled radio, and not the tangible knob. However, I would argue the reason for this was that the tangible control still required touch input to operate, in addition to being eyes on. The knob controlled radio was eyes off and was controlled only by the use of the knob, with no need for additional input. However, both these controls showed that tactile feedback (knob) helped with keeping eyes on the road. Overall, this helps support our findings, in that the tangible controls require less visual attention.

In addition to this, studies on tangible and touch with focusing on visual attention is related to the surrounding entertainment system, or secondary task, and not the primary task, which is seen as the driving itself. Thus, the steering wheel in the car would have to be replaced with tangible or touch controls to give a more accurate comparison to our findings. This would likely increase the effect of the visual demand, which was observed quite clearly in our experiment, where the difference between the amount of time participants were looking at the interfaces were large (tangible 16.7% and touch controls 38.7%). This lends to suggest that the lack of Eyes-free tactile feedback with touch controls forced the participants to spend more time safeguarding their input. This could lead to crucial information being missed while navigation using only touch controls. This is also apparent since it caused the participants that used touch controls to spend less time observing the outside of the ship (29.7%). Meanwhile, with the tangible controls the outside of the ship gained the most attention (37.7%), suggesting that the tactile feedback provided the participants with the opportunity to focus their attention to what they saw as most important for a safe voyage.

However, the visual gaze also showed other interesting findings, such as that the ECDIS display, which gained more attention when navigating with touch controls (20.9%) than tangible controls (16.7%). The ECDIS screen contained route information and a general map of the area. Even though there is no statistical significance, it can be speculated that the reasoning behind this is due to the attention the interface required, which forced the participants to spend less time looking outside the ship and instead relied more on the map shown at the ECDIS screen. However, such over-reliance on the ECDIS could pose dangers, as navigators can be lulled into a false sense of security (Schager, 2008). With a similar difference to the interface, the Conning display in the tangible controls (28.6%) gained a lot more attention than with the touch controls (9.4%). The Conning display contained information such as: speed in knots, depth, heading, and information about each thruster such as the angle, set point and feedback. The cause of this could be that instead of safeguarding the input by looking down at the interface, as with touch, the use of the actual thruster information provided more detailed overall ship information, while simultaneously serving as visual safeguarding.

General Discussion

It seems like the need for visual confirmation with the touch controls is the main reason the differences between the inputs were so large. This was also observed quite clearly with the visual gaze, where the touch controls gathered most attention. The findings indicate that the navigators had more time to observe other important information sources with the tangible controls. This is quite relevant for the consideration of the suitability of the two input methods for ship navigation. The navigator is already in an information rich environment, and the need for visual confirmation from the touch controls may also contribute to the increased subjective workload experienced, as well as the overall lower user experience score. Since, if workload is perceived as high, the operators are more likely to miss important information sources (Lehto & Buck, 2008). As the tangible controls provide tactile feedback, the need to look down at the interface is reduced, which allows the processing of arguably more important information. The tangible controls also resembled actual thruster controls used onboard vessels. Fitzmaurice and Buxton (1997) argued that specialized controls are visual and tactile reminders of associated tools. It is therefore assumed that the familiarity of the tangible controls also gives the interface an advantage. Bjelland et al. (2007) stated two major safety factors concerning usability of controls in cars; the visual attention the controls require and workload and distraction imposed on driver. Taking these into consideration with ship navigation, we see that the tangible controls performed far better than the touch controls, and therefore could be seen as safer and more suited for ship navigation.

Limitations

It should be noted that there are likely differences workload, user experience and visual gaze related to where the user interfaces are applied, as well as how the touch and tangible controls are designed. In a previous study on tangibles, different sizes were found to have an impact on user performance (Øvergård, Forstervold, Bjelland, & Hoff, 2007). Therefore the results cannot be directly translated into other solutions.

The sampling for the experiment was non-random, but there was however a random assignment to conditions.

The questionnaire containing the NASA TLX and UEQ (see Appendix A), measuring workload and user experience were administered after the participants had completed one of the input methods, thus they are vulnerable to the participants decay in memory. To minimize this effect, the participants were instructed to fill out the questionnaire immediately after the completion of navigation with each input method. The duration of navigation for the two input methods was relatively short, lasting approximately 10-12minutes, which also arguably contributed to minimizing the effect.

The UEQ was based on 12 questions (see Appendix A) instead of the original 26, as well as not being analyzed in the standard excel sheet available at http://www.ueq-online.org/. Therefore, our modified version could be subject to different results than the original. However, a previous user test conducted found the hitrate between the questions used in the modified version to be satisfactory, which resulted in the use of 12 questions in total. Since there are fewer questions, this also helps mitigate respondent fatigue (Bryman, 2012).

Future Research

Since our findings confirm the need for more visual attention with the touch interface, future research could be focus on finding ways to reduce this demand, such as employing touch interfaces that provide tactile feedback. This could for example be done with vibrations and auto tracking, where the "widget" is to appear where the finger is touching. Comparing tangible controls against touch controls providing tactile feedback for ship navigation could yield interesting and different results to the non-tactile one used in this experiment. It is, however, still believed that the tangible controls would perform better, due to the familiarity of the controls.

Further research on tangible user interfaces could also focus on different systems for ships, such as crane operation, to better understand which systems could be operated by TUI.

Similarly, since vessels often operate at high seas, it would have been interesting to see how tangible controllers would perform under these conditions.

This thesis explores the suitability of tangible and touch controls for ship navigation with the use of workload, user experience and video recordings. Additional research on TUI and touch interfaces in similar domains, such as the process industry and in relation to safety critical tasks could also help foster innovative ideas on how to apply such solutions to the maritime industry.

Conclusion

From the findings presented in this thesis, one might reasonably conclude that the tangible controls outperformed the touch controls for ship navigation. This thesis has shown that tangible controls for ship navigation produces less workload that the touch controls, generally score higher on user experience, and causes less need to look at the interface when navigating the ship. Thus it is concluded that TUI is more suited for ship navigation than that of touch interfaces.

References

Albinsson, P. A., & Zhai, S. (2003). High precision touch screen interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 105-112). DOI: 10.1145/642611.642631. CHI '03.

Antle, A. N., & Wang, S. (2013). Comparing Motor-Cognitive Strategies for Spatial Problem Solving with Tangible and Multi-touch Interfaces . *TEI '13 Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, (pp. 65-72). DOI: 10.1145/2460625.2460635

Bach, K. M., Jæger, M. G., Skov, M. B., & Thomassen, N. G. (2008). You Can Touch, but You Can't Look: Interacting with In-Vehicle Systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1139-1148), DOI: 10.1145/1357054.1357233, CHI '08.

Besancon, L., Issartel, P., Ammi, M., & Isenberg, T. (2016, April 12). Usability Comparison of Mouse, Touch and Tangible Inputs for 3D Data Manipulation . *Retrieved from Cornell University Library database (arXiv:1603.08735)* , n.a, arXiv:1603.08735v2.

Bjelland, H.V., Hoff, T., Bjørkli, C.A. & Øvergård, K.I (2007) A Case Study of Touch Based Interface for In-Car Audio Systems. *Design Journal*, 10(1), 24-34.

Bowditch, N. (2002). *The American Practical Navigator - An Epitome of Navigation* (Vol. 2002 Bicentennial Edition). Bethesda: National Imagery and MappingAgency.

Bryman, A. (2012). *Social Research Methods* (Vol. Fourth edition). New York: OXFORD University press .

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization , 81* (1), (pp. 1-8). DOI:10.1016/j.jebo.2011.08.009.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin, Vol 112(1)* , 155-159.

Cronbach, L. J., & Meehl, P. E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin, 52.* , 281-302.

Felton, E. A., Williams, J. C., Vanderheiden, G. C., & Radwin, R. G. (2012). Mental Workload during brain-computer interface training. . *Ergonomics 2012 May: 55(5)* , (pp. 526-537). DOI: 10.1080/00140139.2012.662526

Fjærli, B. A. B., & Øvergård, K. I. (2015). *Applicability of Touch Screens on Ship´s Bridges. MACS project report.* Horten, Norway: Kongsberg Maritime & Buskerud and Vestfold University College.

Fitzmaurice, G. W., & Buxton, W. (1997). An empirical evaluation of graspable user interfaces: towards specialized, space-multiplexed input. *CHI '97 Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, (pp. 43-50).

Fitzmaurice, G. W., Ishii, H., & Buxton, W. (1995). Bricks: Laying the Foundations for Graspable User Interfaces . *CHI'95 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 442-449).

Ishii, H. (2008). Tangible Bits: Beyond Pixels. *TEI '08 Proceedings of the 2nd international conference on Tangible and embedded interaction* , (xv-xxv). DOI: 10.1145/1347390.1347392.

Hancock, M., Hilliges, O., Collins , C., Baur, D., & Carpendale, S. (2009). Exploring tangible and direct touch interfaces for manipulating 2D and 3D information on a digital table. *ITS '09 Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, (pp. 77-84). DOI: 10.1145/1731903.1731921

Hart, H. G. (2006). NASA-TASK LOAD INDEX (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, (pp. 904-908). DOI: 10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Emperical and Theoretical Research. In P. A. Hancock, & N. Meshkati, *Human Mental Workload* (pp. 139-183). Holland: Elsevier Science.

Hill, S. G., Lavecchia, H. P., Byers, J. C., Bittner, A. C., Zaklade, A. L., & Christ, R. E. (1992). Comparison of Four Subjective Workload Rating Scales. *Human Factors: The Journal of the Human Factors and Ergonomics Society August 1992 vol. 34*, (pp. 429-439). DOI: 10.1177/001872089203400405

Hoff, T., Alsaker, M., & Bjørkli, C. A. (2002). Effects of Tangible User Interfaces (TUI) in In-Vehicle Information Systems. In: Alm,H.(ED):. *Proceedings of the Nordic Ergonomics Society's 34th Annual Congress on Humans in a Complex Environment, 2002.*

Kratz, S. G., Westermann, T., Rohs, M., & Essl, G. (2011). CapWidgets: Tangible widgets versus multi-touch controls on mobile devices. *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1351-1356). DOI:10.1145/1979742.1979773 Vancouver, BC, Canada: CHI EA '11.

Lucchi, A., Jermann, P., Zufferey, G., & Dillenbourg, P. (2010). An Empirical Evaluation of Touch and Tangible Interfaces for Tabletop Displays. *TEI '10 Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, (pp. 177-184). DOI: 10.1145/1709886.1709917

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger, *HCI and Usability for Education and Work,* (pp. 63-76). DOI: 10.1007/978-3-540-89350-9. Berlin: Springer-Verlag Berlin Heidelberg 2008 .

Lehto, M. R., & Buck, J. R. (2008). *Introduction to Human Factors and Ergonomics for Engineers* . New York: Lawrence Erlbaum Associates.

Nachmias, C. F., & Nachmias, D. (2008). *Research Methods in the Social Sciences* (Vol. 7th). New York: Worth Publishers.

Noy, Y. I., Lemoine, T. L., Klachan, C., & Burns, P. C. (2004). Task interruptability and duration as measures of visual distraction. *Applied Ergonomics, 35(3)*, (pp. 207-213).

Pedersen, E. W., & Hornbæk, K. (2009). mixiTUI: A Tangible Sequencer for Electronic Live Performances. *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, (pp. 223-230). ACM, New York, NY, USA,. DOI: 10.1145/1517664.1517713

Schager, B. (2008). *Human Error in the Maritime Industry: How to understand, detect and cope.* Halmstad: Vinnova and Bengt Schager.

Rubio, S., Diaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods . *Applied Psychology Volume 53, issue 1*, (pp. 61-86). DOI: 10.1111/j.1464-0597.2004.00161.x

Rümelin, S., & Butz, A. (2013). How to make large touch screens usable while driving. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 48-55). DOI: 10.1145/2516540.2516557. AutomotiveUI '13 .

Rauschenberger, M., Schrepp, M., Cota, M. P., Olshner, S., & Thomascheweski, J. (2013). Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ).Example: Spanish Language Version. *International Journal of Artificial Intelligence and Interactive Multimedia, Vol. 2* , (pp. 39-45).

Tuddenham, P., Kirk, D., & Izadi, S. (2010). Graspables Revisited: Multi-Touch vs. Tangible Input for Tabletop Displays in Acquisition and Manipulation Tasks. *CHI '10*

Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, (pp. 2223-2232). DOI: 10.1145/1753326.1753662

Thoresen, C., Øvergård, K. I., & Hancke, U. (2015). Detection of Ungrounded Objects on Mutual Capacitance Touch Screens. In K. Chun, C. Schober, S. Tadigadapa, and J. Lee (eds.). *IEEE Sensors 2015 Proceedings,* (pp. 1542-1545). IEEE.

Viera, A. J., & Garrett, J. M. (2005). Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine* (37), (pp. 360-363).

Völker, S., Nakajima, K., Thoresen., Itoh, Y., Øvergård, K. I., Borchers, J. (2013a). PUCs: Detecting Transparent, Passive Untouched Capacitive Widgets on Unmodified Multi-touch Displays. In A. Quigley, G. Jacucci, M. S. Horn, M. A. Nacenta (eds), *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces, ITS'13*, co-located with UIST'13, October 6–9, 2013 (pp. 101-104), St. Andrews, UK. ACM 2013, ISBN 978-1-4503-2271-3. DOI: 10.1145/2512349.2512791

Völker, S., Nakajima, K., Thoresen., Itoh, Y., Øvergård, K. I., Borchers, J. (2013b). PUCs Demo: Detecting Transparent, Passive Untouched Capacitive Widgets. In A. Quigley, G. Jacucci, M. S. Horn, M. A. Nacenta (eds), Proceedings of the 2013 *ACM International Conference on Interactive Tabletops and Surfaces, ITS'13*, October 6–9, 2013 (pp. 325-328), St. Andrews, UK. ACM 2013, ISBN 978-1-4503-2271-3.

Völker, S., Corsten, C., Hamdan, N. A., Øvergård, K. I., & Borchers, J. (2014). An Interaction Model for Grasp-Aware Tangibles on Interactive Surfaces. In R. Dachselt, N. Graham, K. Hornbæk, M. Nacenta (eds.). *ITS 14 Proceedings of the Ninths ACM International Conference on Interactive Tabletops and Surfaces* (pp. 279-281). Dresden, Germany: Association for Computing Machinery (ACM). DOI: 10.1145/2669485.2669494

Völker, S., Cherek, C., Thar, J., Karrer, T., Thoresen, C. B., Øvergård, K. I., Borchers, J.

(2015). PERCs: Persistently Trackable Tangibles on Capacitive Multi-Touch Displays. In C. Latulipe, B. Hartmann, and T. Grossman (Eds.). *Proceedings* from *the 28th ACM Symposium on User Interface Software and Technology (UIST)* (pp. 351-356). Association for Computing Machinery. DOI: 10.1145/2807442.2807466

Völker, S., Øvergård, K. I., Wacharamanotham, C., & Borchers, J. (2015). Knobology Revisited: A Comparison of User Performance between Tangible and Virtual Rotary Knobs. In N. Nunes, E. Constanza, P. Olivier & J. Schöning (eds.). *Proceedings of the 2015 ACM International Conference on Interactive Tabletops and Surfaces*, (pp. 35-38). New York, NY: Association of Computing Machinery.

Widgor, D., Williams, S., Cronin, M., Levy, R., White, K., Mazeev, M., et al. (2009). Ripples: utilizing per-contact visualizations to improve user interaction with touch displays. *Proceedings of the 22nd annual ACM symposium on User interface software and technology* (pp. 3-12). DOI: 10.1145/1622176.1622180). UIST '09.

Wigdor, D., Forlines, C., Baudisch, P., Barnwell, J., & Shen, C. (2007). LucidTouch: A See-Through Mobile Device. *UIST '07 Proceedings of the 20th annual ACM symposium on User interface software and technology* (pp. 269-278). DOI: 10.1145/1294211.1294259

Øvergård, K. I., Forstervold, K. I., Bjelland, H. V., & Hoff, T. (2007). Knobology in use: an experimental evaluation of ergonomics recommendations. *Ergonomics: 50:5, (pp. 694-705).* DOI: 10.1080/00140130601168046

Appendix A: Questionnaire

The upper body contains the UEQ questions, while the lower is the NASA TLX questions

Please assess the user experience and workload now by scoring one “X” per line.

boring	○	○	○	○	○	○	○	exciting	1
not interesting	○	○	○	○	○	○	○	interesting	2
inventive	○	○	○	○	○	○	○	conventional	3
good	○	○	○	○	○	○	○	bad	4
complicated	○	○	○	○	○	○	○	easy	5
usual	○	○	○	○	○	○	○	leading edge	6
secure	○	○	○	○	○	○	○	not secure	7
meets expectations	○	○	○	○	○	○	○	does not meet expectations	8
inefficient	○	○	○	○	○	○	○	efficient	9
clear	○	○	○	○	○	○	○	confusing	10
impractical	○	○	○	○	○	○	○	practical	11
attractive	○	○	○	○	○	○	○	unattractive	12

Mental Demand – ***How mentally demanding was the task?***

Low High

Physical Demand – ***How physically demanding was the task***

Low High

Temporal Demand – ***How hurried or rushed was the pace of the task?***

Low High

Performance – ***How successful were you in accomplishing what you were asked to do?***

Low High

Effort – ***How hard did you have to work to accomplish your level of performance?***

Low High

Frustration – ***How insecure, discouraged, irritated, stressed, and annoyed were you?***

Low High

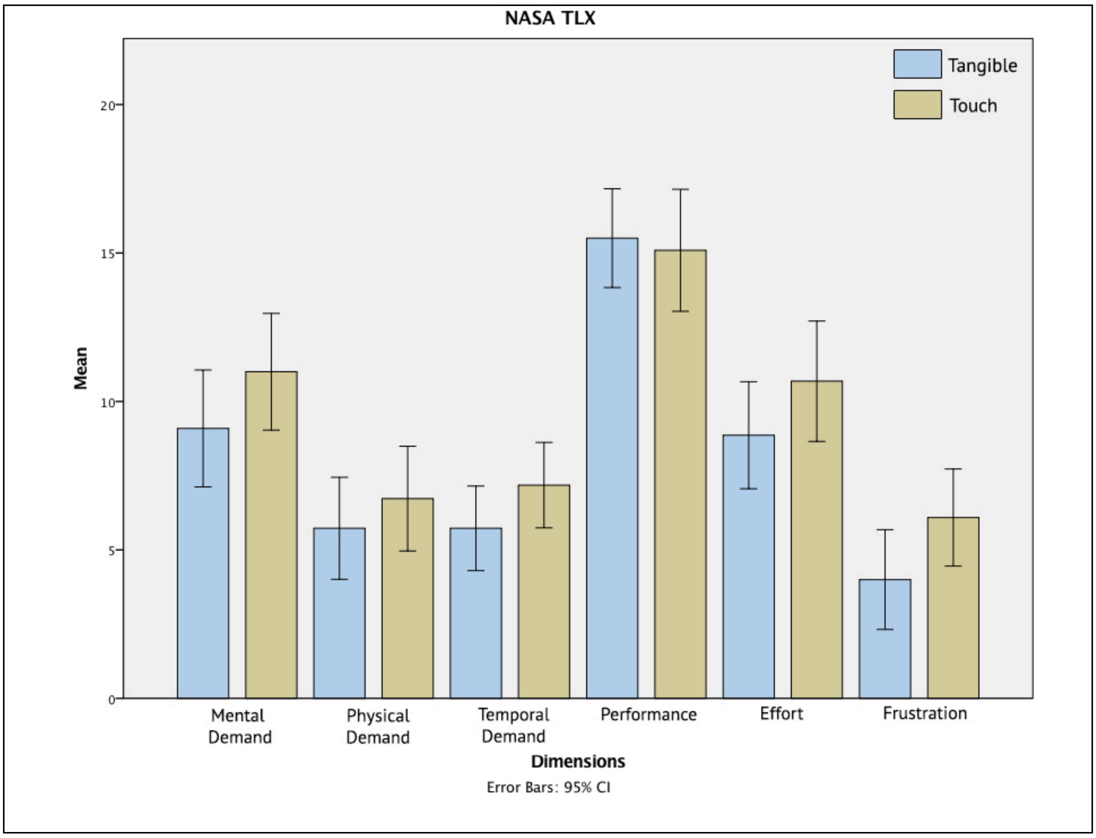
Appendix B: UEQ - Analyzing each question as paired t-test

UEQ Rated as a 7 point scale	Mean, (SD) Touch	Mean, (SD) Tangible	Difference between groups [95% CI]	Effect size Cohen's *d*
Boring/Exciting	6.14 (1.125)	6.18 (0.733)	-0.04 [-0.53, 0.44]	0.041
Not Interesting/Interesting	6.23 (1.066)	6.27 (0.767)	-0.04 [-0.47, 0.38]	0.047
Conventional/Inventive	6.05 (1.359)	6.00 (1.049)	0.05 [-0.32, 0.41]	0.059
Bad/Good	5.41 (1.403)	5.91 (0.868)	-0.50 [-1.04, 0.04]	0.408
Complicated/Easy	4.09 (1.269)	4.73 (1.723)	-0.64 [-1.29, 0.01]	0.434
Usual/Leading Edge	5.73 (1.162)	5.68 (0.995)	0.04 [-0.35, 0.44]	0.05
Not Secure/Secure	4.32 (1.615)	5.05 (1.214)	-0.73 [-1.58, 0.13]	0.376
Does not meet expectations/Meets expectations	5.05 (1.362)	5.77 (1.152)	-0.73 [-1.29, -0.16]	0.568
Inefficient/Efficient	4.82 (1.296)	5.86 (1.082)	-1.04 [-1.68, -0.41]	0.73
Confusing/Clear	4.91 (1.269)	5.73 (1.241)	-0.82 [-1.54, -0.09]	0.503
Impractical/Practical	4.82 (1.332)	5.55 (1.224)	-0.73 [-1.29, -0.60]	0.568
Unattractive/Attractive	5.59 (1.333)	6.32 (0.646)	-0.73 [-1.39, -0.07]	0.489
Sum	63.14 (8.76)	69.86 (8.12)	-6.71 [-10.38, -3.04]	0.832

$p = .001$ $t = -3.816$ $df = 20$

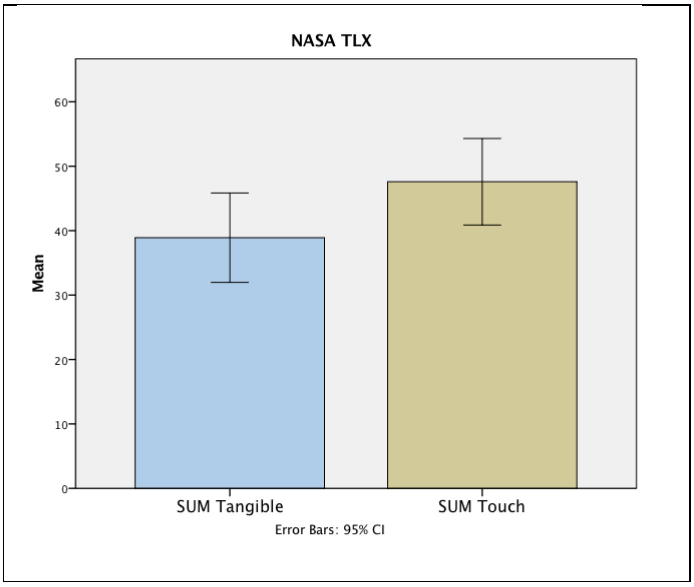
Appendix C: NASA R-TLX Graphs

All dimensions Graph



Note. Blue: Tangible controls, Brown: Touch Controls

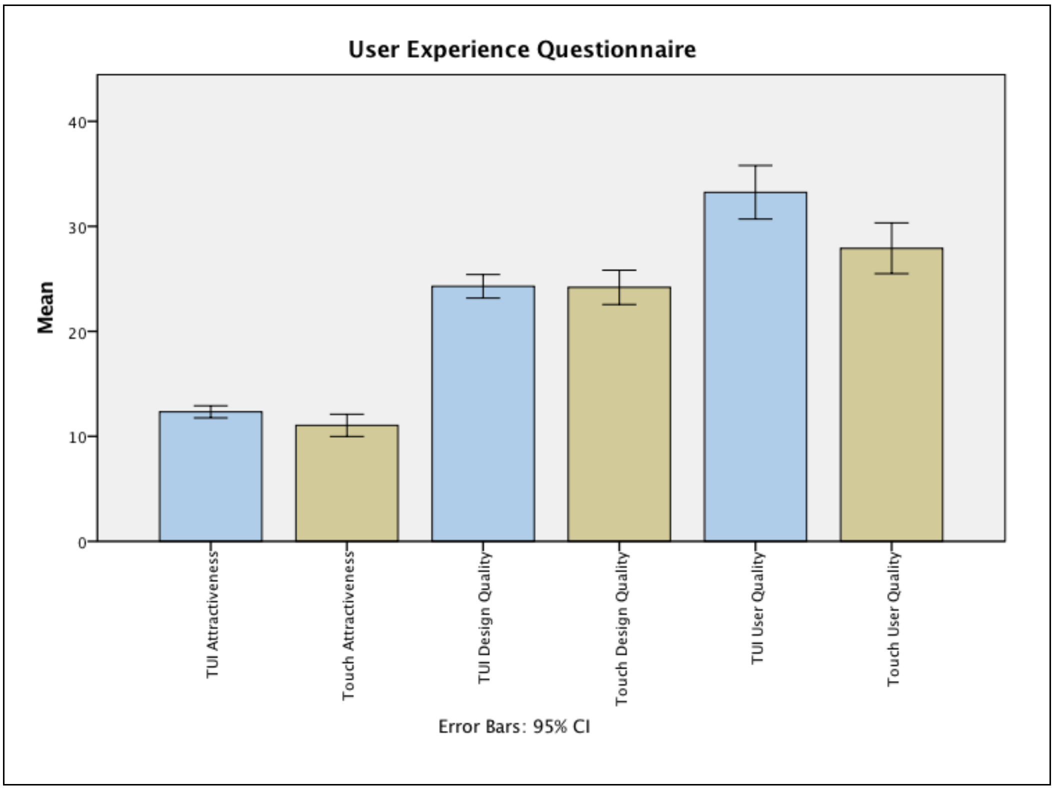
Overall workload comparison Graph



Note. Blue: Tangible controls, Brown: Touch Controls

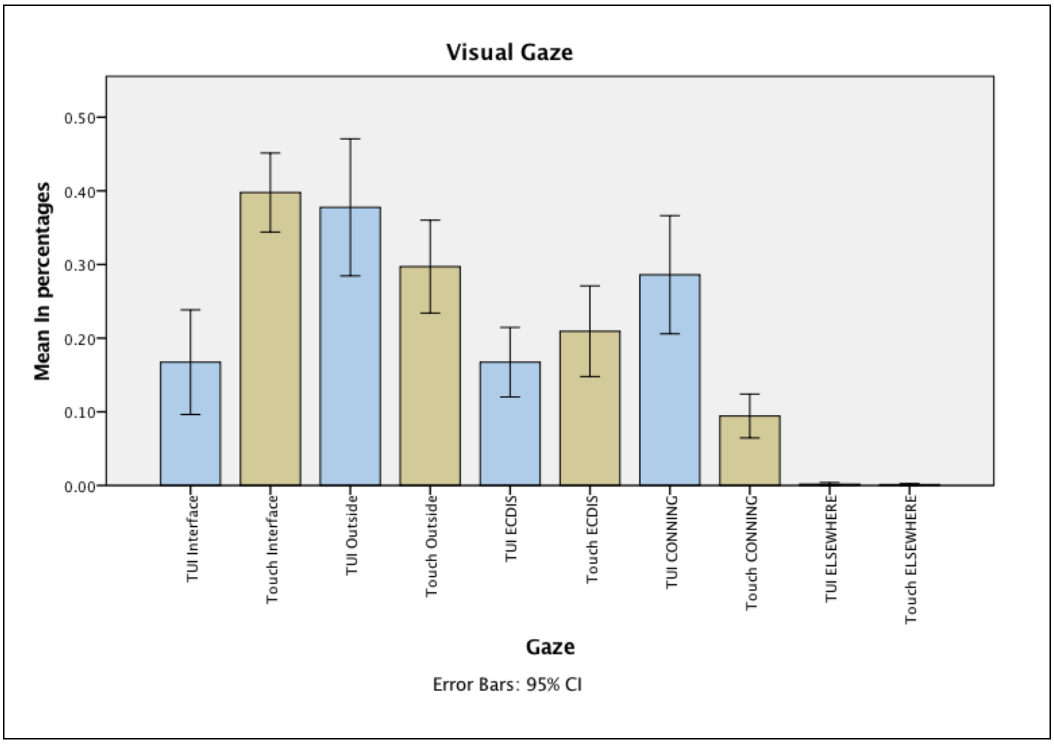
Appendix D: UEQ and Visual Gaze Graphs

User Experience Questionnaire Graph



Note. Blue: Tangible controls, Brown: Touch Controls

Visual Gaze Graph



Note. Blue: Tangible controls, Brown: Touch Controls