HØGSKOLEN
I ÅLESUND
Aalesund University College

Master's degree thesis

IP501909 MSc thesis, discipline oriented master

Impact of eye-trackers on maritime trainer-trainee experience

Siyang Zheng

Number of pages including this page: 71

Aalesund, 29.05.2014

HØGSKOLEN I ÅLESUND
Aalesund University College

Mandatory statement

Each student is responsible for complying with rules and regulations that relate to examinations and to academic work in general. The purpose of the mandatory statement is to make students aware of their responsibility and the consequences of cheating. **Failure to complete the statement does not excuse students from their responsibility.**

Please complete the mandatory statement by placing a mark ***in each box*** *for statements 1-6 below.*		
1.	**I/we herby declare that my/our paper/assignment is my/our own work, and that I/we have not used other sources or received other help than is mentioned in the paper/assignment.**	☒
2.	**I/we herby declare that this paper** 1. Has not been used in any other exam at another department/university/university college 2. Is not referring to the work of others without acknowledgement 3. Is not referring to my/our previous work without acknowledgement 4. Has acknowledged all sources of literature in the text and in the list of references 5. Is not a copy, duplicate or transcript of other work	Mark each box: 1. ☒ 2. ☒ 3. ☒ 4. ☒ 5. ☒
3.	I am/we are aware **that any breach of the above will be considered as cheating, and may result in annulment of the examination and exclusion from all universities and university colleges in Norway for up to one year, according to the Act relating to Norwegian Universities and University Colleges, section 4-7 and 4-8 and Examination regulations paragraph 31.**	☒
4.	**I am/we are aware that all papers/assignments may be checked for plagiarism by a software assisted plagiarism check**	☒
5.	**I am/we are aware that Aalesund University college will handle all cases of suspected cheating according to prevailing guidelines.**	☒
6.	**I/we are aware of the University College's rules and regulation for using sources paragraph 30.**	☒

Publication agreement

ECTS credits: 30

Supervisor: Sashidharan Komandur

Agreement on electronic publication of master thesis

Author(s) have copyright to the thesis, including the exclusive right to publish the document (The Copyright Act §2).
All theses fulfilling the requirements will be registered and published in Brage HiÅ, with the approval of the author(s).
Theses with a confidentiality agreement will not be published.

I/we hereby give Aalesund University College the right to, free of charge, make the thesis available for electronic publication: ☒**yes** ☐**no**

Is there an agreement of confidentiality? ☐**yes** ☒**no**
(A supplementary confidentiality agreement must be filled in)
- If yes: **Can the thesis be online published when the period of confidentiality is expired?** ☐**yes** ☐**no**

Date: 29.05.2014

MASTER THESIS 2014

FOR

STUD.TECHN. SIYANG ZHENG

IMPACT OF EYE-TRACKERS ON MARITIME TRAINER-TRAINEE EXPERIENCE

Station keeping operation is always a critical performance for offshore vessels, since these specialized vessels need a relatively fixed platform with respect to the ocean floor to finish operations successfully. Currently, the most popular way to keep the vessel stationary is to apply Dynamic Positioning(DP) system. This system provides an automatic compensational feedback, by means of active thrust, to withstand the environmental forces impacting the ship. But when the DP system fails, the bridge operators must be able to supersede the DP system and keep the vessel stationary by manual operation. This is a skill that all the DP operators need to be trained in and be able to perform well. The existing training methods are long and expensive. The training is very elaborate because the operator needs to learn how to work and react when the ship loses position due to external forces. Based on the feedback from DP trainers, the manual maneuvering is a highly visual-dependent operation. Eye-tracking is a promising technology that could change the current training method by providing extra information to both trainers and trainees.

A new training method for manual maneuvering operation with Eye-trackers and an experimental evaluation are designed in this thesis. The Eye-trackers are used in two phases: one for helping trainers' intervention, and one for providing eye-tracking video for debriefing activity. The Eye-trackers improve the quality of trainers' intervention and that between training level and eye-movement patterns. The author has tested the new designed training method by series of Case-Control experiments in the Sandøy bridge simulator of Høgskolen i Ålesund.

The thesis work shall include the following:

- Pre-study Eye-trackers applications and current DP training problems
    - Study relation between eye movement patterns
    - DP training centre around the world and applied technology current situation
    - Eye-tracker application in other training settings
    - Study Sandøy simulator training problems, analyze and extract the current problems
- Experimental design
    - Design the experimental arrangement, scenarios, subjects groups, variables and checklists
    - Find proper experiment subjects, apply help from operation experts
- Conduct the experiment
- Data and result analysis, discussion
- Analyze how much the Eye-tracking affects the quality of intervention by trainers
- Utilize Eye-tracking data for training debriefing and assess quality of debrief

The thesis should be written as a research report with summary, conclusion, literature references, table of contents, etc. During preparation of the text, the candidate should make efforts to create a well arranged and well written report. To ease the evaluation of the thesis, it is important to cross-reference text, tables and figures. For evaluation of the work a thorough discussion of results is needed. Discussion of research method, validation and generalization of results is also appreciated.

In addition to the thesis, a research paper for publication shall be prepared.

Three weeks after start of the thesis work, a pre-study have to be delivered. The pre-study have to include:

- Research method to be used
- Literature and sources to be studied
- A list of work tasks to be performed
- An A3 sheet illustrating the work to be handed in.

A template for the A3 sheet is available on the Fronter website under MSc-thesis. This sheet should also be updated when the Master's thesis is submitted.

The thesis shall be submitted as two paper versions. One electronic version is also requested on a CD or a DVD, preferably as a pdf-file.

Supervision at AAUC: Sashidharan Komandur,

Contact at: Internal project

Researcher. Sashidharan Komandur	Prof. Karl Henning Halse
Supervisor	Program coordinator

Finish: 1. June 2013 Signature candidate: _______________________

PREFACE

This thesis is submitted in partial fulfillment of the requirements for MSc degree in Ship design at Department of Advanced Marine Operations (AMO), Aalesund University College. It contains work done from January to May 2014. My supervisor on this project has been Dr. Sashidharan Komandur, at the maritime human factors laboratory. The thesis has been made solely by me; part of the text, however, is based on the research of others, and they have been duly referenced.

The broad area this thesis falls into is maritime human factors. This thesis project focuses on advanced maritime training, one of the three research tracks at maritime human factors laboratory. This thesis is a part of internal research project at Aalesund University College, which is known for its advanced simulator training facilities. The funding for this project came from MARKOM 2020.

I would firstly thank my supervisor and Aalesund University College, Sashidharan Komandur. He gave me a lot of trust and flexibility on the project. Thanks to him, I was able to deal with such a challenging project, to have such amazing research experience and to collect my data so quickly and complete the project in time.

Moreover, I would extend my gratitude to the experienced maritime course specialists Arnt Håkon Barmen and Tron Richard Resnes for their professional help. Without them, I could not gain enough maritime training knowledge for my project and I could not finish my data collection. Also, many more people are really appreciated for their help on my data collection. They are the students of Nautical science program at Aalesund University College. Thanks to their enthusiasm and participation. Thanks to my colleagues at maritime human factors lab Girtz strazdins and Yushan Pan for their motivation and supports.

Last but not least, I would like to thank my family, friends for their support.

ABSTRACT

The purpose of this study is to investigate the effect of using eye-trackers on manual manoeuvring training in bridge simulator, which is a crucial part of Dynamic positioning training.

In the first part, results of short survey of dynamic positioning, manual manoeuvring training and eye-tracking technology are presented. These shows that growing demand from oil and gas industry for qualified DP operators but the training itself is long and expensive. On the other hand, increasingly domains of application using eye-tracker as an assistive tool for training. Therefore it would be significant to design a new DP training method, specifically manual manoeuvring with assistant of eye-trackers.

The thesis then designs a case-control experiment to identify the difference between normal training method and new training method which applies eye-trackers to change trainers' intervention and debriefing activity. The detail of experimental design includes participants selection, scenario design, training procedures, outcome measurement, experiment time arrangement, variables and hypotheses definition and experiment execution. The results indicates that eye-tracker is not a necessary tool when experts intervene and guide the novices familiarize the operation. But it improves the efficiency of debriefing with great extent. The results also indicate that a right visual strategies of DP operators can improve the performance of manual manoeuvring operation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TERMINOLOGY

Fixations moments when the eyes are relatively stable, taking in or understanding information. It represents the amount of processing being applied to objects.

Saccades Rapid eye movements changing the fovea to a new location of interest.

Abbreviations

DP Dynamic Positioning

PSV Platform Supply vessel

AHV Anchor handling vessel

AOI Area of Interest

AMO Advanced Marine Operation

DTU Display/Transmit Unit

1 INTRODUCTION

1.1 Problem description

Station keeping is always a critical performance for offshore vessels, since these specialized vessels need a relatively fixed platform with respect to the ocean floor to finish marine operation successfully, such as offshore drilling and cargo transmission. Currently, the most popular way to maintain the vessel stationary is to apply Dynamic Positioning (DP) system to. This system provides an automatic compensational feedback, by using its own propellers, thrusters, position reference sensors, combined with wind sensors, motion sensors and gyrocompasses, to withstand the external environmental forces that impact the ship. The DP system increases the vessels' station keeping performance dramatically. It reduces human operators involvement in the station keeping operation with accurately automating the thrust adjustment. But the DP systems can possibly break down due to various reasons, including Global Positioning System (GPS) and other sensor failure. The bridge operators need to be trained to be able to supersede the DP system and keep the vessel station within certain boundaries by manual operation. This is a skill that all the DP operators need a lot of training in and have to be able to perform well in critical situations.

The current training method are long and expensive. The main reason is that the training is very elaborate because the operators need to be familiar with vessel's properties such as weight, thrust forces and vessel feedback to the thrust forces, and learn how to work and react properly when the ship loses position due to extreme external forces. In the meantime, the maritime industry is increasingly demanding new trained DP operators to support the offshore industry production.

Hence, it is important to find new technical assistive tools to increase the efficiency and improve user experience of dynamic positioning training.

1.2 Motivation

The Eye-tracking technology has been used in wide range of domains. The tool can be used as *diagnostic* and *interactive*. As a diagnostic tool the eye-tracker provides objective and quantitative data of users' visual attention. As an interactive tool, it serves as a powerful input device which can be used by visually mediated applications [1]. For example, eye-trackers are used in research application, market research, medical research, usability testing, human-computer interaction, and also for training purpose [2]. After the development of portable eye-tracker glasses with real-time eye-tracking video, the eye-trackers are extended to training usage which improved user experience with perspectives of feed forward training, trainers' intervention and after-activity debriefing [3].

According to DP operation experts in Aalesund University College, the manual maneuvering is a highly visual dependent operation. The experts are very experienced getting rich information through visual interfaces. Experts experience shows that, most of the novices spend too much time looking at the monitor screens or less preferable references when they should be looking outside the bridge instead to receive inefficient visual information. Also a study on laparoscopic surgery training emphasizes that experts and novices have significant difference in eye movement strategies during the surgery and it is a new way of assessing skills [4]. In the context of dynamic positioning training, the novices are supposed to collect situational information only by visual attention and get used to looking at the right places at the right time on the bridge. This requires a high level of competence and attention from the operators.

This thesis analyzes Eye-trackers as an efficient way to improve the training quality and efficiency.

1.3 Research questions

The usage of eye-tracker technology in this master thesis are in two aspects. The first application is to use the eye-tracker and real-time video to assist trainers in terms of providing intervention to trainees. The second application is to use the eye-tracking video to assist debriefing activity. The experimental design and data analysis are focusing on whether eye-tracking technology improved the manual maneuvering training in these two aspects. And eye-movement pattern of participants during the operation will be analyzed to verify the correlation with performances.

1.4 Objectives

This master thesis will design new training procedures for manual maneuvering training with Eye-trackers applied and the training method will be tested by a case control experiment. It includes scenario design, subjects’ group selection, definition of hypothesis and variables, outcome measurement and data analysis. The Eye-trackers will be employed in two phases. The first part of experiment is expert’s intervention and novices familiarization. In this phase, the purpose is to improve trainers' intervention quality with help of eye-trackers. The second part is to use the eye-tracking video as assistant tool in debriefing activities. The experiment measurements are designed to verify the impact of eye-trackers through quantitative and qualitative data.

1.5 Thesis out line

The thesis was organized in the following parts:

Preliminary study and representation. This includes background information of eye-tracker technology and application, DP operation and manual maneuvering, simulator training. It also includes the evaluation of method before conducting the experiment.
Experiment: This includes the experiment management, experiment variable and hypothesis and experiment execution.

Discussion: This includes the findings after conducting the experiment, and hypotheses testing.

Conclusion: This includes concluding remarks and description of future work.

Reference and Appendices: This includes the reference used in this thesis, the documents used in experiment, and summary of post-processed experiment data.

2 BACKGROUND AND THEORETICAL BASIS

This chapter will introduce the background of the thesis. It mainly contains the background of Eye-tracking technology, Eye-tracker application in training and background of current DP training approach.

2.1 Dynamic positioning training

With implementation of dynamic positioning system the operators do not need to react to the wind, current and wave for correcting the vessel's course or position. The system perform the correction automatically. Although a DP system is able to fulfill all the operational requirements, the training is still necessary. First of all, the DP operators should learn to work with the system, what are the different operational modes, and they should learn how to manually steer the vessel with controllers or joysticks. This is extremely important skill requiring a lot of precision. Moreover, it is important for the operators to learn how to take over the system when facing emergency. Such as GPS system failure extreme weather conditions.

Figure 1. Control system on bridge chair

Figure 1 shows a typical Rolls-Royce chair arm system. It include joysticks that control Azimuth thruster, bow/stern tunnel thrusters and port/starboard main propellers. Additionally, there is a multi-functional joystick on the outside of right arm. It is a combination of motion control of a vessel, which means to directly control a vessel in surge, sway and yaw motion have to be controlled. Different configurations of thrusters are matched with different chair arm control system.

2.1.1 Dynamic positioning operation

DP operation is defined as: "*A means of holding a vessel in relatively fixed position with respect to the ocean floor, without using anchors accomplished by two or more propulsive devices controlled by inputs from sonic instruments on the sea bottom and on the vessel, by gyrocompass, by satellite navigation or by other means*" [5].

Advantages

- No tugboats needed

- Offshore set-up is quick
- Power saving
- Highly accurate

Operational mode:

DP system has several different types of applications. There are several operational modes [6]:

- Manual/joystick mode: the operator has full control over the vessel;
- auto-heading mode: the system maintains required heading automatically;
- auto-position mode: the system maintains required position automatically;
- auto area position mode: the system maintains automatically within a specified area, while using minimum power;
- auto track mode: the vessel steers automatically along a pre-defined course;
- auto pilot mode: the vessel follows a specified track described by a set of way-points;
- follow target mode: the vessel follows a constantly moving target such as a remotely operated vehicle.

2.1.2 DP applications

In the offshore industry, the following application of DP system can be found on different vessel types:

- Platform supply vessel (PSV): for supplying offshore platforms.
- Anchor handling vessel (AHV): an offshore rig should be anchored as long as it reaches the desired position. This should be done accurately by an anchor handling vessel.
- Floating crane operation: the crane vessel should keep the correct heading and position.

2.1.3 Manual maneuvering

As discussed before, typical DP failure mode is one important courses of dynamic positioning training. Whenever the DP system is not able to work due to failure of GPS or other crucial part of the system, the operators need to maintain vessel's position by manually maneuvering it using controllers on the bridge arms. In Figure 1, there are four joysticks used for manual maneuvering. Among them, the joystick on the outside of the right chair is a multi-functional, which can control motions in surge, sway and yaw directions. The other three joysticks are separately controlling stern, bow tunnel thrusters, Azimuth thruster and port and starboard main propellers. When the vessel require fully manual, all the joysticks can be used to maintain the vessel's station.

Visual information sources:

There are mainly two major sources that provides information about external forces and vessel movement: monitors displaying sensor data in the bridge , and the bridge windows revealing relevant vessel motion relative to the rig or other references.

2.1.4 Simulator training

A simulator is device that re-creates carrier objects and corresponding scenarios by combining physical factors and visual solutions. Examples: flight and driving simulators. The simulator is built on simulation technology. It imitates operations of a real-world process [7].

Simulators are extensively used in the training of civilian and military personnel [8]. The simulators are rational alternative to real-world training due to lower and mitigated failure consequences. As a result the trainees will learn valuable lessons in a safe virtual environment and lifelike scenarios.

Bridge simulator:

The Sandøy bridge simulator can be described as a 'virtual simulation', where actual players use simulated systems in a synthetic environment. The simulator is designed for dynamic positioning training and anchor handling training. It consists of two control stations: anchor handling and vessel steering. In this experiment, only the steering station will be used.

2.1.5 DP training courses

There are numbers of DP simulator training centers around the world. In Norway, the major DP training centers are built by Kongsberg or Rolls-Royce and Offshore Simulator Centre [24]. The training processes in these centers are very similar which include basic knowledge, operation and operation under critical situations.

As an example, the training courses in Kongsberg DP center are [25]:

- DP basic operator course
- DP operation task
- Typical DP failure mode

This thesis will focus on the last process, the DP failure mode more specifically the manual maneuvering mode, to design new training procedures that change intervention and debriefing activities with using eye-trackers and verify the impact of user experience and training quality.

2.1.6 Experiential learning debriefing

Debriefing is an activity used in many types of studies. It is a process for receiving explanations of summaries and suggestions of studies or investigations after participation is complete. Trainers' intervention and after-training debriefing can be both considered as debriefing.

Ernesto Yturralde, experiential trainer and researcher, explains: *"In the field of experiential learning methodology, the debriefing is a semi-structured process by which the facilitator, once a certain activity is accomplished, makes a series of progressive questions in this session, with an adequate sequence that let the participants reflect what happened, giving important insights with the aim of that project towards the future, linking the challenge with the actions and the future."* Debriefing is an effective process in training, in which the professional facilitator or trainer could be able to improve trainee's learning quality by giving suggestion based on past performance.

2.2 Eye-tracking technology

As one of the most significant features of the human face, eyes and their movement plays an crucial role in receiving visual information from real world and expressing one's desires, needs, cognitive processes, emotional states, and interpersonal relations. Gaze estimation or tracking are important for many domains of research including human attention analysis, human cognitive state analysis, gaze-based interactive user interfaces, gaze contingent graphical displays, and human factors [9].

2.2.1 Technical background

Eye-tracking is a process of capturing the gaze point or the motion of an eye relative to the head and an eye-tracker is a device for measuring pupil position and eye movement.

Eye-trackers types:

The main function of eye-trackers is measuring rotations of the eye in different ways. There are three categories of eye-tracking technology: measurement of the movement of a special eye-attaching lens; optical tracking without direct contact to the eye; and measurement of electric potentials using electrodes placed around the eyes [2].

Technologies:

The mostly obtained method to measure eye movement is to apply video-oculography system. An overview of the principal and components of Video-oculography system, are shown in Figure 2. The system obtain information of user's eye movement and image data from cameras. The eye location in the image is normally recognized by using corneal reflections(CR). Calibration is needed before using eye-trackers. It includes reorganization of eye position and movement and image of scenario, sometimes the head pose estimation, the gaze position of the user are able to estimate.

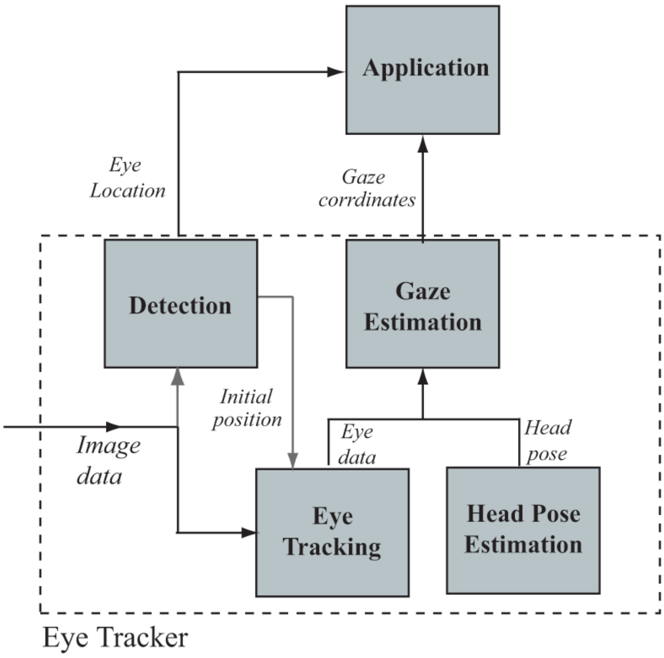

Figure 2. Eye-tracking techonology [9]

2.2.2 Fixation _ eye-movement metrics

The main measurements used in eye-tracking research are fixations and saccades.

Fixation: moments when the eyes are relatively stable, taking in or understanding information. It represents the amount of processing being applied to objects [10].

Number of fixations overall: higher number of fixations overall indicate less efficient search or sub-optimal layout of the interface [11].

Fixations per area of interest: More fixations on a particular area indicate that it is more noticeable, or more important, to the viewer, relative to other areas [12].

Percentage of participants fixating an area of interest: If a low proportion of participants is fixating an area that is important to the task, it may need to be highlighted or moved [13].

2.2.3 Eye –tracking history

The first Eye-tracker was built in late 1800s. It was difficult to build from technical point of views and mostly mechanical, therefore not comfortable for the users [2]. Starting from year 1950, researchers developed numbers of technologies for eye-trackers. The most common of them are:

- Lenses system with mirrors in the 1950s to 1970s. This technology has high precision allowing very detailed eye-movement record.
- Electromagnetic coil system. It measures the electromagnetic induction in silicon contact lenses placed on an anaesthetized eye. It is been considered as the method with most precision. But this method was proved to alter the saccades of user who wear them.
- Electro-oculography system. It measures the electromagnetic variation if the dipole of the musculature of eyeball moves. This system only measured horizontal movements and suffers from interference of electromagnetic noise of surrounding muscles.
- The Dual Purkinje System. This was an expensive system with costly maintenance, had a small visual field of recording. It was extremely precise and accurate. Compare to electro-oculography system, it did not need to place something directly onto the users eyes.

2.2.4 Manufacturers and customers

The Eye-trackers' manufacturers and customers were combined before the 1980s. Today, a single role is divided into two groups: the manufacturers and researchers [2]. There are several Eye-tracker producers nowadays. The products are suitable for different customer groups due to different product features and technologies. The major customers groups and corresponding manufacturers are discussed in following part.

- The academic researcher group. It is the oldest customer group and the largest. They utilize eye-trackers in almost all the disciplines of science which emphasizing precise timing, accuracy, precision, and high sampling frequency in data. The main manufacturers that provide for this demanding group are SR research, SensoMotoric Instrument, Applied System Laboratory (ASL) and Tobii Technology.
- Another large and new group of customer is the media and advertisement consultants. This group often needs the eye-tracker for simple usage. They normally want to rent the equipment instead of buy it. The main features they are looking for is fancy presentation of the data, such as heat maps. The Tobii technology has dominated this customer group.

- Human factors researchers are making up a small group and existed for a long time. They need the eye-tracker be able to using in field or scenarios. ASL, SMI, Smart-Eye and Seeing Machine occupied most of the customers.

2.2.5 Using eye-trackers with training purpose

In addition to the applications mentioned above, eye-trackers are widely used in training purpose.

Figure 3. ASL eye-tracking glasses

For example, ASL mobile glasses (Figure 3) eye-tracker has been used in to the athletics and sports training and, including golfing, volleyball, basketball, skiing, race car driving and soccer (Figure 4). The researchers often change the training procedures with respect to feed forward, crisis intervention and after-training debriefing.

Figure 4. ASL application in driving and football training

2.3 Related work

A master student Sathiya Kumar Renganayagalu graduated in 2013 with a master thesis 'Eye-trackers as assistive technologies in maritime training' [23]. His research focus was on improving quality of familiarization and intervention in dynamic positioning training. Figure 5 shows the experimental setup used in his study. In contrast to Renganayagalu's thesis, this work focuses on ' manual maneuvering training' instead of 'DP task training', and adds eye-tracker application to the 'debriefing' activity.

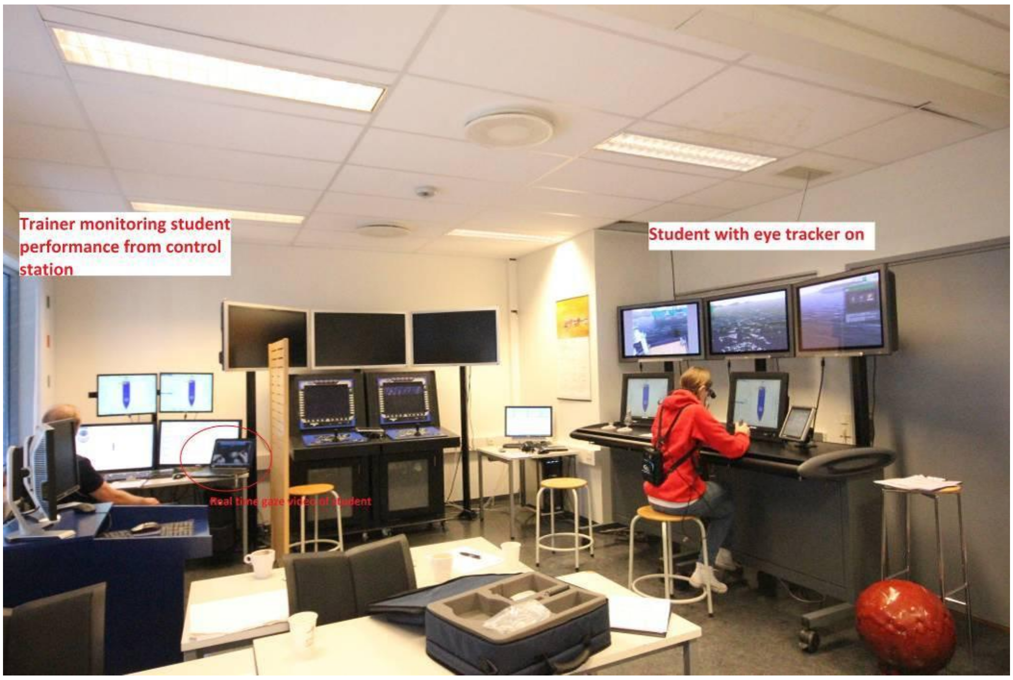

Figure 5. Experimental setup of DP training, from previous study [23]

3 METHODS

This chapter presents the related methodologies in this thesis.

3.1 Case-control experiment

A case-control study is one kind of observation study in which the subjects are not confronting a random situation. In this type of study, two existing groups have different outcomes due to controlled incomes. The comparison between the case group and control group can provide a qualitative or quantitative indication of whether or how much the factors influence the studied process [14]. The key feature of case-control experiment is having one controllable variable while keeping all other variables equal between the case and control groups.

3.2 Outcome measurements

One of the objectives of this experiment was to evaluate the performances of novices. This section described the methods used in the experiment to collect data of participants' performances.

3.2.1 Trajectory, footprints of vessel's motion

The footprints of vessel's motion are generated by the simulator itself. It represents the vessel's exact position on the global coordinate and recorded by the virtual Global Position System. The footprints are marked as a serious of positions on the GPS screen at a frequency of 1 position per 3 seconds. In addition, the real time vessel's heading in the global coordinate system is displayed on the same screen. Both vessel's heading and position footprints are important to evaluate the performance level of manual maneuvering. A small amplitude of fluctuation of heading or vessel's movement means more stable operation that performed by operator.

3.2.2 X bar and R chart

Many quality features can be expressed by numerical measurement. A measurement feature is called a variable [15]. Control charts for variable are one of the widely used primary tools in quality analysis and control. When dealing with a quality variable, it is usually necessary to monitor both the mean value of the variable and its variance. Control of the process average level is usually done with the control chart for means, called X bar control chart. Process variance can be monitored with a control chart for the range, called R control chart. The R control chart is more widely used[15].

Statistical basis of the charts

Assuming $x_1, x_2 ... x_n$ is a sample of size *n*, the average of this sample is:

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n}$$

Suppose there are totally *m* samples to be analyzed, each sample contains *n* data. The sample size is normally 4, 5 or 6. These small sample sizes result from the composition of reasonable subgroups [15]. From the samples, one can get $\overline{x}_1, \overline{x}_2 ... \overline{x}_m$ as the average of each sample. Then the process average is expressed as:

$$\overline{\overline{X}} = \frac{\overline{X}_1 + \overline{X}_2 + \ldots + \overline{X}_m}{m}$$

Therefore, $\overline{\overline{x}}$ would be the center line on the $\overline{x}$ chart.

Then use the range method to calculate the control limits. For sample $x_1, x_2 \ldots x_n$, the range of it is the difference between the largest and the smallest data, which is:

$$R = x_{\max} - x_{\min}$$

Hence the ranges of *m* samples $R\ ,R\ \ldots R_m$ can be calculated. The average range is:

$$\overline{R} = \frac{R_1 + R_2 + \ldots + R_m}{m}$$

Based on the *R* bar, the center line, upper control limits and lower control limit of $\overline{x}$ chart can be calculated:

UCL = $\overline{\overline{X}} + A_2 \overline{R}$

Center line = $\overline{\overline{X}}$

LCL = $\overline{\overline{X}} - A_2 \overline{R}$

The constant A_2 is different for various sample sizes [15].

After the center line, upper control limits and lower control limit of a X bar chart were decided, put the mean value of each sample $\overline{X}$ on the chart.

Process variability can be inspected by using values for the sample range *R*. the center line and control limits of the *R* chart are expressed as:

UCL = $D_4 \overline{R}$

Center line = $\overline{R}$

LCL = $D_3 \overline{R}$

The constants D_3 and D_4 are different for various values of sample size *n* [15].

After the center line, upper control limits and lower control limit of a R chart were decided, put the data range of each sample R on the chart.

3.2.3 Check sheets

The check sheet is a document for collecting data in real time and in the place where data is produced. The data collected in the check sheet can be qualitative or quantitative [16].

Check sheet for defect types: If a process or operation has been identified as a training for improvement, it is important to know the types of defects that happened in the process. By identify the types of defects, one can remove reason of the defects or reduce the frequency [17].

Check sheet for defect causes: If a process or operation has been identified as a training for improvement, the check sheets may focus on the identification of the causes of the defects [17].

In this experiment, the check sheets will be focused on identifying types of mistakes that the novices made by definition of several different evaluation categories. Also some effort will be made to identify what kind of causes lead to failures.

The check sheets are designed for the experts of the operation who are able to evaluate the performance of the novices during the training based on their own experience and observation during the experiments. Both an expert and the author participated in designing the check sheets. It was separated into four categories, 'Controllers', 'Visual attention', 'Reference', and 'Footprint' (see Appendix 2).

3.2.4 Questionnaire

The feedbacks from the novices are of great value when verifying the users' perception in this thesis. Therefore a proper self-report questionnaire is important. Questionnaire is an tool used in human-subjects experiment to collect information from respondents. The results collected from questionnaires can be statistically analyzed if the questionnaires are well designed [26].

The questions are differentiated into open-ended and closed-ended questions. An open-ended question asks the subjects to formulate his/her own answer; a closed-ended question has the subject pick an answer from a given number of options.

The closed-ended questions in the questionnaires are either a scale or index. There are four types of response scale are distinguished [18]:

- Dichotomous: The answer has two options.
- Nominal-polytomous: The answer has more than two unordered options.
- Ordinary-polytomous: The answer has more than two ordered options.
- Bounded continuous: The answer is presented with a continuous scale.

In this experiment, a bounded continuous questionnaire will be used to collect novices' personal feedback which will be a referable data to verify whether the application of eye-trackers impact the user experience of manual maneuvering training. The self-report questionnaire was separated into two parts, part 1 'pre-experiment' and part 2 'after-experiment' (see Appendix 1).

Drawbacks of questionnaires:

Although the questionnaires are inexpensive, convenient, and easy to analyze, the questionnaire can often have a lot of problem. For instance, unlike interviews, the researchers conducting the research may never ensure whether the subjects understood the question in the questionnaires. Also, because the questions are so specific to what the researchers are asking, the received information can be minimal.

3.2.5 Fixation numbers in AOIs

As discussed before, more fixations on a particular area indicate that it is more noticeable, or more important, to the viewer compared to other areas. The areas studied in this experiment are the rig, the container and the screens. The fixation numbers in each area of interest will be compared with other variables.

Area of interests:

Figure 6. AOIs set-up

There were three areas studied and set up as the area of interests, which are rig, container and the monitor screen. Each of the object gave different visual information of vessel position and motion to the DP operators.

4 EXPERIMENTAL DESIGN

This chapter defines the overall approach of the experiment management. It will highlight the scenario definition.

4.1 Experiment organizers

The experiment organizers include project leaders of Bachelor of Nautical Science program, Arnt Håkon Barmen and Tron Richard Resnes representing the experts, and the author of the thesis. The experts helped the author to recruit experiment participants and provided professional suggestions during scenario development. They also joined the experiment as trainers, a critical part of the experiment.

4.2 Participants

Non-probability sampling: As a method of sampling, non-probability sampling is not often used. Sampling mostly represents random/probability sampling, which is to select individuals randomly and use them to represents characteristic of the whole population. The non-probability sampling method does not meet these features and it cannot be used as the method to represent general population. However, researchers also claimed that the application of probability sampling for representation are suitable for large scale studies while non-probability sampling are usable for in-depth research where the main focus is often to understand more complex processes [19]. If the study is only focusing on a specific case, there is need to use the probability sampling method [20]. The purpose of this thesis is to study the training process for manual maneuvering operation and the only involved population is sailors on board. Therefore it is reasonable to apply non-probability sampling.

The subjects of case and control groups should come from the same population as control groups [21]. In this experiment, the subjects consist of nautical science students, both from first grade and second grade. The reason why they were chosen is their background knowledge on ship bridges and maritime operations. It allows them to quickly understand the operation with a short introduction to the scenario. Nevertheless, they have almost no experience on bridge control operation, therefore they can still be treated as novices to the manual maneuvering operation. Another reason of choosing the nautical science students is their high passion and motivation to join the activities in the anchor handling bridge simulator. Their high passion of participation ensured them to treating the training course seriously and none of the participants biased by money. This choice also saved funding of the project.

There were totally fourteen students joined the experiment. Seven of them became participants in case group and another seven student comprised the control group.

4.3 Scenario

The scenario of this experiment includes description of initial position of the vessel, exterior environmental forces, operator's position, and training procedure.

A lot of parameters and experiment procedures have to be decided during scenario analysis. Many of the parameters are qualitative by nature and cannot be described quantitatively. Therefore, it is important to have several trials before starting the experiment. The trials were held in February 2014. There were totally 4 test runs and 7 participants took part in them. The goal of the test runs was to clarify parameters of the scenario, such as operation details, vessel position, experiment procedures, and expert check sheets.

4.3.1 Vessel position

The objective of manual maneuvering is to keep the vessel's station close to the offshore platform. The distance and angle between the vessel and the platform should be in a specific range to make sure the platform is visible and the vessel will not collide with the platform in case of a small and sudden ship motion. Once the vessel's position and heading are decided, it should be constant number for all the subjects.

Position:

The vessel's position is decided by the position of container hanging by the rig crane. Considering that the rig in the scenario is fixed, and the container stays at the same position relative to the rig through all the experiment, it is reliable to use the container as the reference of initial vessel's position.

Heading:

The vessel's heading is described in the global coordinates. Since the rig is fixed, the vessel heading degree would decide how much portion of the rig is in operator's visual range. The vision of the rig is important for understanding the relevant position between vessel and rig during the manual maneuvering operation. Hence this is an important factor in scenario design. After empirical several test, the vessel heading is decided as 315 degrees, which means the starboard side of ship is 45 degrees from rig side. The side view of the scene is depicted in Figure 7, top view in Figure 8, and the field of view of the operator in Figure 9.

Figure 7. Overview of the scenario configuration

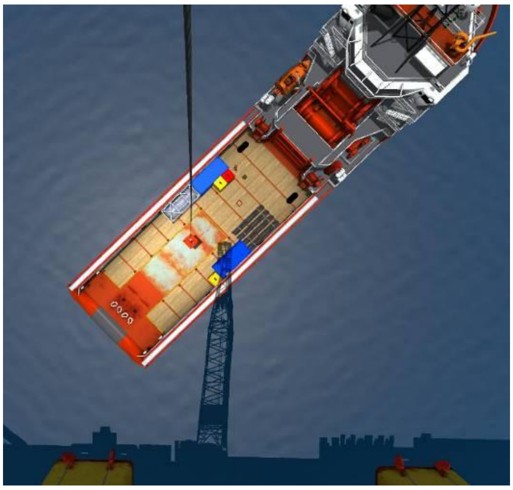

Figure 8. Top view from tip of the rig's crane

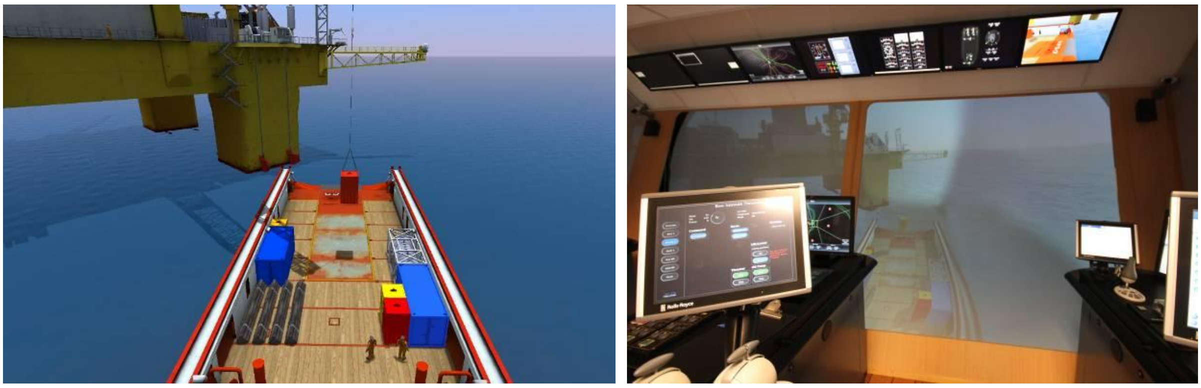

Figure 9. Operators field of view in the bridge simulator

4.3.2 Environmental forces

The instructor can modify the exterior environmental forces. These include wind speed and direction, waves height and current speed and direction. It is important to have proper exterior forces because these are the main factors that impact the vessel's motion and operation difficulty in test scenarios. After several trials the following environmental force parameters were selected as optimal to provide challenging ye reasonable training experience for the novices.

Current: 0.5 knot from east

Wind: 12 knots from north

Wave height: 2 meters

The subjects could get understanding of the environmental forces by either checking the environmental force censor or observing the vessel's motion by looking through the bridge windows.

4.4 Training procedures

4.4.1 Introduction of operation and control system

For any types of training, the trainees must have a basic understanding of the activity they are participating in. In the context of manual maneuvering training, the novices should be familiar with the purpose of the operation and used equipment on the bridge. In the introduction, experts are asked to perform a manual maneuvering to the trainees while explain the function of the involved controllers on the bridge arms.

4.4.2 Operation practice

All the subject had a 20 minutes practicing time to be familiar with the operational system. They were assisted by an expert who was standing next to them and intervening during the practice, see Figure 10.

Figure 10. Practicing with expert

Case and control groups:

The participants in the case group will wear the ASL glasses eye-tracker during the training. The expert will give feedback assisted by the ASL Eye-tracker’s real-time data.

For the participants in the control group, the expert will use the traditional method to give intervention to subjects without Eye-tracker assistant.

Since the expert have eye-tracking real time video as the assistive tool, the interventions that expert provided to participants in case group were expected to involve more suggestion on visual preferences.

4.4.3 Manual maneuvering task

All the subjects had 3 trials without the expert's intervention. After each trial, the performance was evaluated by expert through the check sheets. There should be three check sheets evaluation finished for each participants. From experts' experience, the key point to maintain position under manual operation is always paying attention the relative distance between the platform and vessel. It is easy for novices to lose attention.

Failure criteria

Before the operation starts, the vessel position was maintained by the DP system in the simulator and the container was hanging up by the rig's crane in the middle of the deck. When the operator turns off the DP system the ship will lose the initial position due to environmental forces. The operator has to use thruster controllers to manually maintain the position. If the operator does not react to the ship motion or operated improperly, the ship will get further away from the initial position due to lack of compensation or over-compensation to the environmental forces. Since the entire operation is a dynamic process, it is necessary to have a clear judging criterion to evaluate whether the operator can properly maneuver the vessel. The task is successful only if the containers stays above the deck area for the whole duration of the operation. The container position is checked by the expert by observing it's shadow in the visuals – as the sun is set to be directly above the container, the shadow represents a perfect projection of the container on the deck plane.

4.4.4 Debriefing and test trial

An after-training debriefing were performed for all the subjects based on their performance in the third trial. The assisting tool that expert used in debriefing is different from case and control groups.

Case group

The expert gave debriefing with the assistance of ASL eye-tracking video of the third trial. It is recorded by the ASL system and shown on a laptop.

Control group

The expert will give debriefing with a video from a scene camera of the third trial.

After the participant had feedback from expert in debriefing activity, they were asked to have one more trial their performance were evaluated with same tool as the previous three trials.

4.4.5 Summary

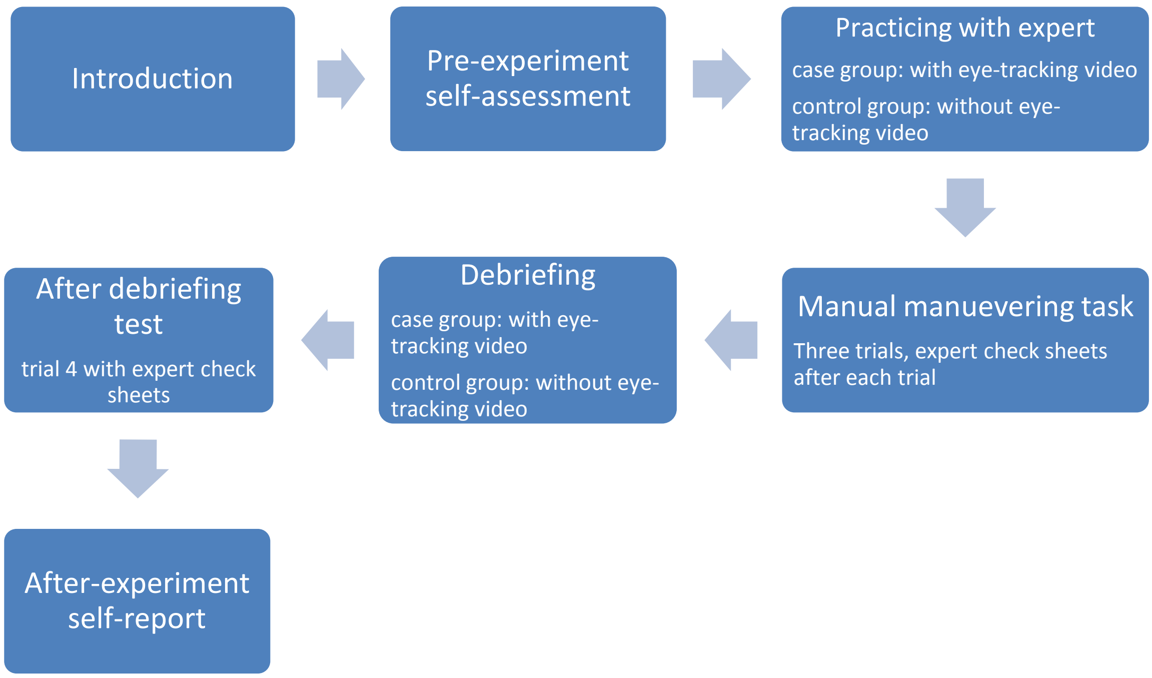

Figure 11. Summary of the training procedures

Figure 11 is the summary of training procedures. The different treatment for case and control groups are also included.

4.5 Experiment Measurement

The experiment measurements are consist of 5 parts.

1. Expert's check sheets
2. Participant's self-report
3. Footprint of the vessel's position and heading, generated by the simulator system
4. Number of fixation extracted from eye-tracking video

The first two represent subjective data while the latter two are objective measurements. The experiment focused on whether and how the eye-tracking technology improves the manual maneuvering training. Therefore the measurements are designed to evaluate the novices' performance in each trial from three different perspectives: expert's feedback, trainees understanding of his/her own performance, and quantitative data showing objective results.

Expert's check sheets

After each trial, the expert was asked to use check sheets to evaluate the participant's performance. For all the four trials of each participant, including three trials in manual maneuvering task and one trial after debriefing, the performances were evaluated by check sheets.

Participant's self-report

The self-report questionnaire were separated into two parts. The first part was finished after the introduction. The second part was finished after the last trial.

Footprint of the vessel's movement

The footprint of the vessel's movement was displayed on a screen in simulator room. It is a standard screen on the bridge showing the real-time vessel position and heading in the global coordinate system. The footprint of each trial is a data of major importance showing the novices' performance because it represents the vessel's movement and rotation relative to the initial position and heading. However, because the Sandøy simulator was not designed and built for research purpose, the footprint and vessel heading cannot be extracted directly from the system by software or other approach currently. Hence a video camera is used to record the information from the screen, see Figure 11. The camera was positioned to have the best view range screen images capture, see Figure 12. It is position was maintained constant during the whole experiment, for all the trials. Figure 13 shows a detailed description of the camera set-up.

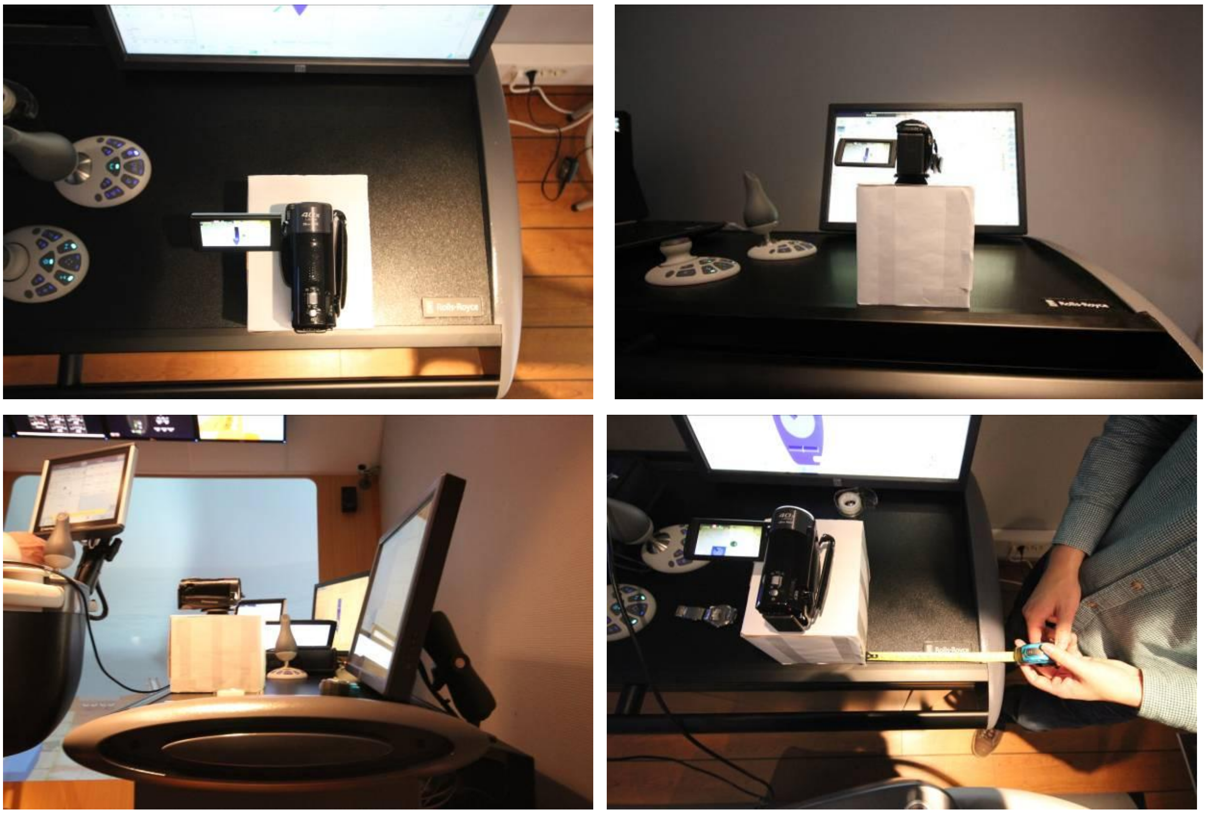

Figure 12. Set-up of the camera that captures the position and heading footprint screen

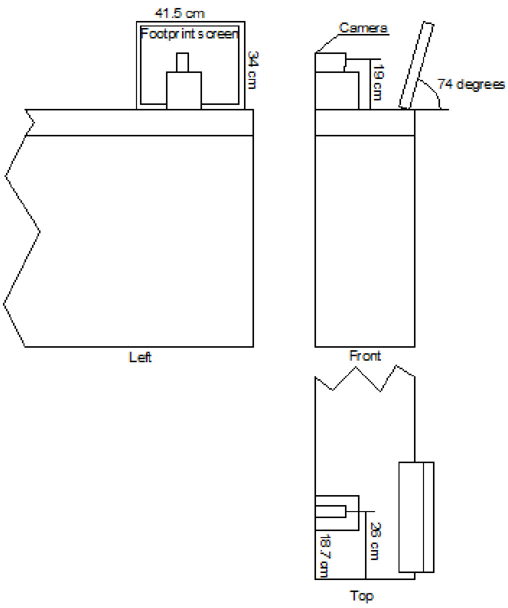

Figure 13. Detailed description of the footprint capture camera setup

Video analysis:

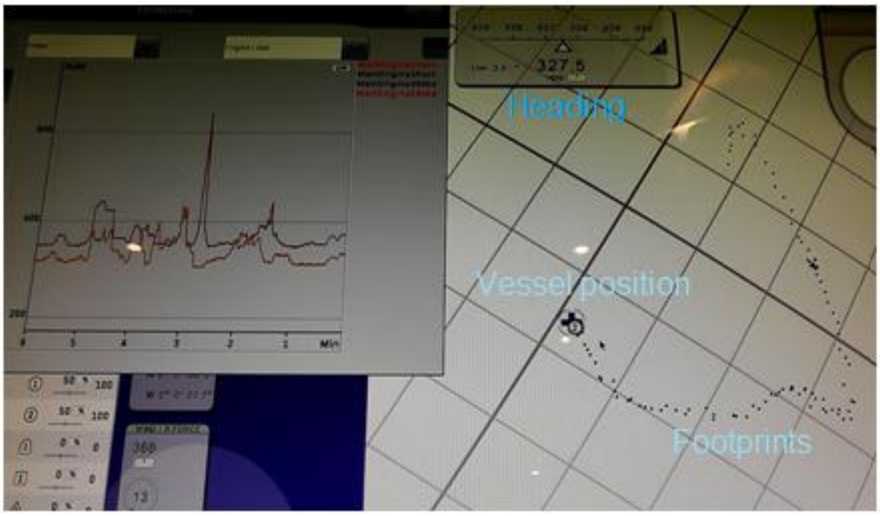

Figure 14. Screen shot of footprint video

Figure 14 shows an example of a screen shot of footprint video. From the footprint screen, the data of heading, vessel position of every moments can be read. The grids on the screen was scale of 5 meters by 5 meters. Since the author extracted the footprints manually. It is impossible to keep a high accuracy for entire video. Therefore the accuracy for reading the footprints was 1/10 of the grid size, which was 0.5 meters.

It is clear that the chosen accuracy was not accurate enough to predict the precise movement of vessel. But for this experiment, the data that expected from footprint was a range of vessel's movement. It normally varies from 10 to 20 meters. The accuracy of 0.5 meters was enough for this expectation.

After a footprints record were manually recorded, it needs further treatment because the footprints on the screen was the point that *20* meters left from centre line of vessel.

Vessel position calculation:

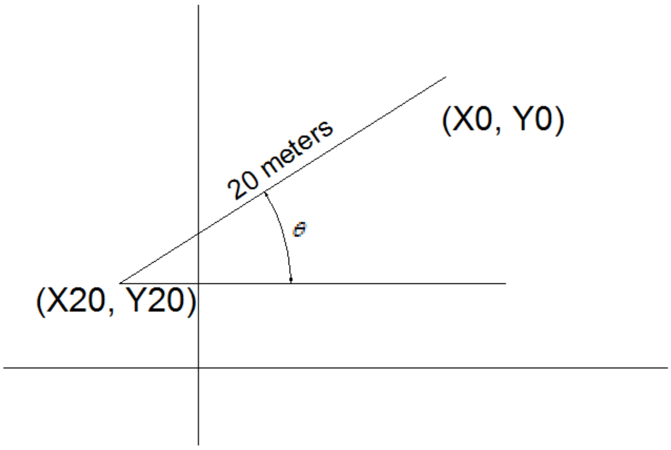

Figure 15. Footprints calculation

Figure 15 shows a draft of the relation between footprint and vessel position.

Assuming the position of vessel is (X_0, Y_0) and the position of footprint on the footprint screen is (X_{20}, Y_{20}). Known that the initial position of footprint is on the origin point of Axis, and initial heading of vessel was 45 degree. Then the calculation of vessel position is:

$$X_0 = X_{20} + 20 \times \cos\theta - 20 \times \cos 45$$

$$Y_0 = Y_{20} + 20 \times \sin\theta - 20 \times \sin 45$$

An example of vessel position treatment of a footprint video is in Appendix 3. There were four video were manually recorded and transferred for each participant.

Number of fixations extracted from eye-tracking video

After the experiment finished and eye-tracking video recorded, the author used the software 'ASL Results Plus'. This software is used for post-processing the eye-tracking video recorded from fields.

AOIs adjustment:

As described in section 3.2.5, the areas of interests (AOI) were defined into three different part. They were 'Container', 'Rig', and 'Screen' (see Figure 5). But the position of defined AOIs in the scene need to be adjusted to overlap the original selected object through the entire video because the head position of the user was not fixed and the defined AOIs are not moving with the corresponding object automatically. Therefore, the position of AOIs were adjusted with a certain frequency. Theoretically, the more frequent adjustment made, the better accuracy of results can be calculated. The author used 20 frames as the frequency of adjustment, which is approximately 0.67 second.

4.6 Experiment risk

There were some unavoidable risks that could affect and decrease the validity of the experiment. All the identified risks would be analyzed and evaluated as much as possible, for instance the back ground and experience on the manual operation of the subjects, inconstant interventions from experts and different studying speed between subjects. Although the experiment design were expected to be valid for the purpose of research, and it was designed with concept of highly constant. There are still very big possibility that the design was not proper enough and the data measured are not valid as expected because of the various reasons such as human subjects which cause a lot of uncontrolled variables.

4.7 Experiment time arrangement

Before the formal experiment, several pre-experiments were set up to test all the details of the experiment. As shown in Table 1, most of the experiment was performed between March 24th and March 28th. Two more participants joined the experiment on April 10th. All the experiment were performed in the timeframe from 10:00 to 16:00. Each of the participants spent approximately 1.5 hours to finish the experiment with half an hour preparation, resulting in 2.0 hour total for each participant.

Table 1_Time arrangement

Date	24.3.2014	25.3.2014	26.3.2014	27.3.2014	28.3.2014	10.4.2014
Participation	Participant 1	Participant 3	Participant 5	Participant 8	Participant 11	Participant 13
	Participant 2	Participant 4	Participant 6	Participant 9	Participant 12	Participant 14
			Participant 7	Participant 10		

5 EXPERIMENT - THE VARIABLES AND HYPOTHESES

For verifying the impact of eye-tracker in manual maneuvering training, an experiment was designed and conducted. Hypotheses and variables are greatly important for experimental design. The subjects involved in this experiment are first year and second year students from bachelor of Nautical science program. All the participants have basic understanding of Dynamic Positioning system from relevant courses. All the data from each participant was recorded by eye-trackers, video cameras, self-reports, and expert's check sheets.

5.1 Experiment definition

The goal of the experiment is to verify whether there is a significant difference, in terms of trainer-trainee experience and trainees' performances, between manual maneuvering training with and without help from eye-trackers. The core concept of experimental design and hypothesis definition is case-control study. Hence it is important to define proper measurable variables to identify individual performance and feedback for the experiment.

- **Objects of study**: As described before, the object of the thesis was to clarify the change in training efficiency after apply eye-tracker technology. The selected subjects were the first and second grades students from nautical science program, Høgskolen i Ålesund. All these students are very similar to the real trainees that could join the DP training because they already had the Dynamic Positioning course in their education ensuring basic knowledge of dynamic positioning. These conditions make them well-suited and them become ideal experimental subjects and simplified the experimental design because it was unnecessary to provide introduction to dynamic positioning.
- **Purpose**: The purpose of this experiment is to evaluate the performance of the subjects based on variables extracted from outcome measures and compare case group with control group.
- **Quality focus:** The primary effect under this experiment is the novices' operational quality of manual maneuvering operation.
- **Perspective:** The perspective is from the author's point of view.
- **Context:** The experiment was conducted in the bridge simulator. The study was a case-control experiment. The author is interested in the impact of eye-tracker to trainer's intervention and the debriefing effectiveness.

5.2 Planning

Base on the experiment definition in chapter 5, the experiment need to be properly planed in order to control the experiment.

5.2.1 Context selection

As described in Chapter 4, the experiment scenario was selected as manual maneuvering operation and will be executed in Sandøy bridge simulator in Aalesund University College.

The experiment is designed to apply eye-trackers in manual maneuvering training and to compare the changes to the existing training method. Manual maneuvering training has been available for some years already but the methods are different between trainers and training centers. To document the detail of training method with check sheets and compare with other measurements, including vessel footprints and novices' fixation analysis, would provide trainers, trainees, and the author a new perspective of the operation.

5.2.2 Variable description

Following the criteria of independent variable [22], none variables were selected in this experiment. All the variables are divided into four categories.

1. Participants' personnel quality assessment
   - Self report questionnaires part 1[Se]
2. Evaluation of novices' performance:
   - Time duration[Time], recorded from the footprint video and scene video
   - Change of time duration between fourth trial and mean of three trials[TC]
   - Check sheets evaluation[Che]
   - Change of check sheet evaluation between fourth trial and mean of three trials[CC]
   - Distance[Dis], extracted from footprint screen video
   - Heading[He], extracted from footprint screen video
3. Novices' visual attention:
   - Fixation numbers on AOIs[Fix], post-processed by software 'ASL Results Plus'
4. Novices' perception of the training:
   - Participants self report feedback[Fe]

5.2.3 Hypotheses

The null hypotheses and alternative hypotheses were listed below. There are totally 8 hypotheses in the experiment, see Table 2.

Table 2_Hypotheses definition

ID	Description	Representation
$H1_0$	In the first three trials, there will be similar mean value of time duration between case and control group	$\text{Time[case]} = \text{Time[control]}$
$H1_1$	In the first three trials, there will be significant difference in mean value of time duration	$\text{Time[case]} > \text{Time[control]}$
$H2_0$	There will be similar change of time duration in 4th trial between case and control group	$\text{TC[case]} = \text{TC[control]}$
$H2_1$	There will be significant difference in change of time duration in 4th trial	$\text{TC[case]} > \text{TC[control]}$
$H3_0$	In the first three trials, there will be similar scores of check sheets between case and control group	$\text{Che[case]} = \text{Che[control]}$
$H3_1$	In the first three trials, , there will be significant difference in scores of check sheets	$\text{Che[case]} > \text{Che[control]}$
$H4_0$	There will be similar change of check sheets score in 4th trial between case and control group	$\text{CC[case]} = \text{CC[control]}$
$H4_1$	There will be significant difference in change of check sheets score in 4th trial	$\text{CC[case]} > \text{CC[control]}$
$H5_0$	There will be similar 'distance to initial position' between case and control group	$\text{Dis[case]} = \text{Dis[control]}$
$H5_1$	There will be significant difference in 'distance to initial position' between case and control group	$\text{Dis[case]} \neq \text{Dis[control]}$
$H6_0$	There will be similar 'vessel's heading change' between case and control group	$\text{He[case]} = \text{He[control]}$
$H6_1$	There will be significant difference in 'vessel's heading change' between case and control group	$\text{He[case]} \neq \text{He[control]}$
$H7_0$	There will be similar fixation numbers in each AOI after debriefing between case and control group	$\text{Fix[case]} = \text{Fix[control]}$
$H7_1$	There will be significant difference in fixation numbers in each AOI after debriefing between case and control group	$\text{Fix[case]} \neq \text{Fix[control]}$
$H8_0$	There will be similar participants' feedback between case and control group	$\text{Fe[case]} = \text{Fe[control]}$
$H8_1$	There will be significant difference in participants' feedback between case and control group	$\text{Fe[case]} > \text{Fe[control]}$

6 EXPERIMENT - EXECUTION

After the experiment is designed an planned, it needs to be strictly put into effect in order to collect the data that was planned for analysis. This chapter introduced several important points in experiment operation.

6.1 Preparation

The date of conducting the experiment was decided based on trainers' availability. The participants were notified in advance so that they can arrange their personnel schedule. Since most of the participants came from same grade and class, the sequence of their participation was selected by the students themselves.

Different tools and material were needed in the experiment.

Tools:

- ASL eye-tracker, Tobii eye-tracker
- Two video cameras. One was used to record the vessel's footprint generated by the simulator. Another one was used to record the entire view of the simulator
- Sandøy bridge simulator. Getting the scenario ready, including position of container, vessel's position and heading, environmental forces

Forms:

- Expert's check sheets, four copies for each participant.
- Participants self report questionnaires
- Consent form
- Experiment diary for the author

Preparation:

- Fully charged batteries of eye-trackers and cameras
- A room with white board for eye-tracker's calibration

6.2 Execution

All the experiments were done within one and half hours. Each participant provided author raw data as expected. Both experts and participants followed the exact the same sequences of procedures, described in section 4. 4. The following raw data was collected from each participant:

- Four videos recording the vessel's footprints of each trial,
- Four videos recording the simulator overview of each trial,
- Four copies of check sheets, filled by the experts to evaluate each trial,
- Four eye-tracking videos recording participants' eye-tracking data of each trial;
- One self report questionnaires including pre-experiment self-assessment and after-experiment self-report.

6.3 Data validity

All the raw data and forms from fourteen subjects were successfully collected. However, due to technical reasons, data of three participants were discarded because the data was invalid. This section discussed the possibly invalid data that could lead to mistakes and incomplete analysis.

Self-report questionnaires

The questionnaires in the experiment were only design for evaluate the participants personnel quality and after experiment feedback. It is also a great tool to collect the background of the participants, such as the attended courses and onboard working experiences. These information were important for selecting participants and results analysis. It is a defect of this experiment that did not design the questionnaires to collect it. The direct influence to the experiment was that one of the participants had onboard working and manual maneuvering experiences. Therefore the data from this participant was unavailable.

Experts' intervention

In the practicing part of the experiment, the expert intervened and guided the novices familiarized with operation. The important point in this part was to capture how and if the interventions from experts due to whether using the eye-tracking real-time video as assistant tool. It was expected to verify the effect that the interventions from experts were different from case group to control group. However during the experiment, the experts did not use the eye-tracking real-time video as much as expected. Therefore it was only slightly different between case and control group in terms of experts' intervention. The reason: the experts were focusing a lot on guiding the novice on how to use the system to control the vessel. The expert said:

"I was focusing much more on telling the students to operate the controllers properly to compensate different kinds of vessel motion."

Due to this mistake in experiment, the data from manual maneuvering task was invalid to provide evidence for hypothesis.

Eye-tracking video

The ASL eye-tracking glasses were selected as the primary tool with real-time video. However, it did not perform stable during the experiment and just part of the eye-tracking data is available for further analysis. The possible reason could be signal interference in the simulator environment that affected eye-tracking data wireless transmission. Such behavior was not experienced in any of the previous test runs in the simulator facilities. Therefore it was an unexpected issue that affected the data validity. After the data collection, only eye-tracking data of participants 5, 10, and 13 were intact enough for further analysis.

Vessel's footprint

As described in section 4.5, the vessel's footprint during operation were recorded by a camera. The set-up of the camera's position was decided to have the best view range to the footprint screen and was expected to be constant in every trials. However, due to mistakes made by the author, three participants' video were not intact. It caused the footprints of these three participants incomplete.

7 RESULTS

7.1 Descriptive statistics

This chapter presents the data after post-processing. For showing the results of the experiment, several graphical statistics were posted for data visualization and distribution. All the participants finished experiment but due to some errors in data collection, the outcomes of participant 3, 6, and 9 are not complete. The usable portion of their data is listed in Appendix 4.

7.1.1 Pre-experiment self-report

Manual manoeuvring operation requires also some personnel characteristics which help the operators to make the right decisions calmly in critical situations. Therefore before starting the experiment, all the participants were asked 7 questions to fill the self-assessment sheet, see Figure 16. The answers of each question are bounded continuous, which means the answer was presented as a continuous scale. The answers in this questionnaire were presented in 6 scales.

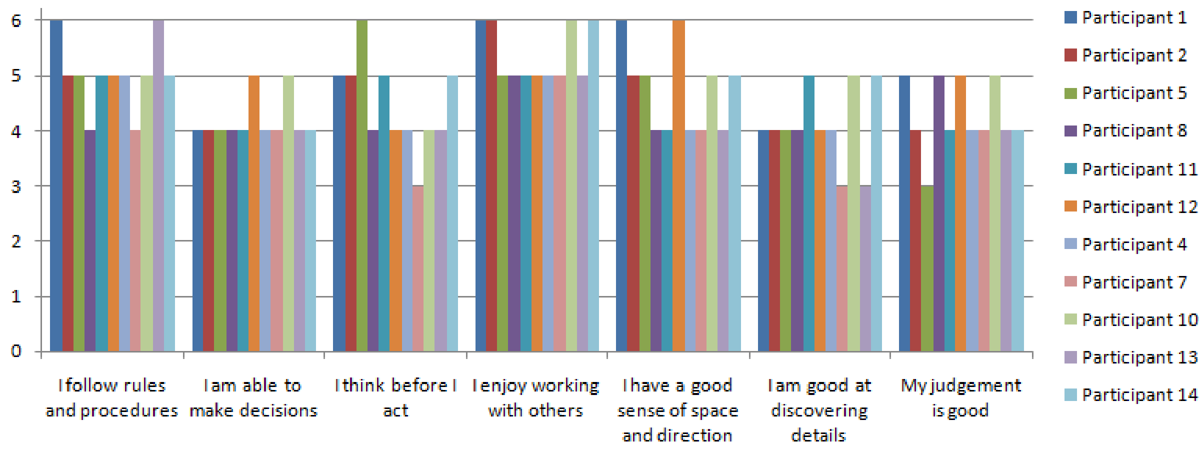

Figure 16. Self-report questionnaires part 1: pre-experiment phase

Most of the participants think themselves willing to follow the rules and procedures when it is need in their normal life. Only participant 7 and 8 marked *4* for this question. Almost all the participants chose *4* for decisions making ability. The third question '*I think before I act*' showed that only participant 7 gave the lowest scale. All the participants express they enjoy to work with others. The rest three topics are great important personal qualities that needed in manual manoeuvring operation. About sense of space and direction, participants 1 and 12 gave themselves 6, full scale in this category, while participants 2, 5,10, 14 were 5 and other five participants showed less confidence on this ability. about detail discovering, participants 7 and 13 had the lowest scale with only *3*. Participants 10, 11 and 14 have the most confident with scale of *5* and rest of the participants gave themselves *4* out of *6*. The last ability was personnel judgement. Participants 1, 8, 10, and 12 gave *5*, which was higher scale than others. Participant 5 was least sure on this ability and marked *3* on this category. The rest of the participants chose *4* on this ability.

7.1.2 Time duration of each trials

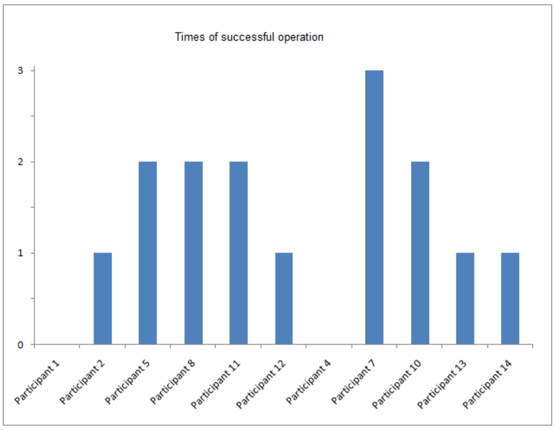

Figure 17. Times of successful operation

Figure 17 showed the times of successful operation of eleven participants. It contains four trials including the trial after debriefing. Participants 1 and 4 had not success to maintain position for one time. Participant 7 had the best performance which was 3 totally times. Participants 2, 12,13 and 14 had one times successful operations while participants 5,8, 10 and 11 had 2 times successful operations.

Time duration was the most direct indicator of the performance of the operation. The time duration of all the four trials in case and control groups is shown in Figure 18 and Figure 19 respectively. The maximum time duration is 300 seconds and also represents the participant successfully maintained the vessel's position in this trial. Any time duration less than 300 seconds means that the participant failed to maintain position in this trial and crossed the safety boundaries at corresponding time moment.

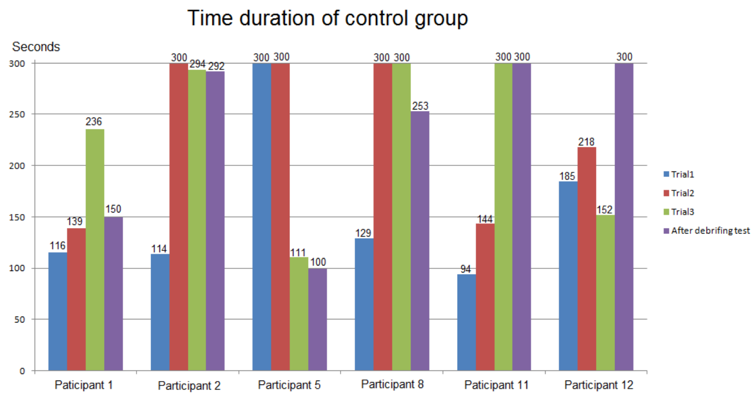

Figure 18. time duration of control group

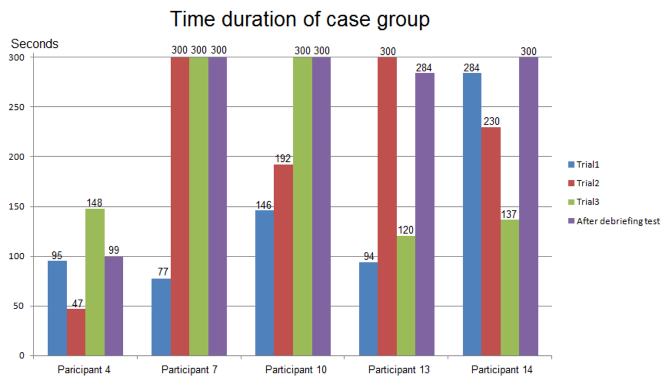

Figure 19. time duration of case group

Figure 18 presents the time duration of all the participants in control group. Participant 1 had the lowest performance overall and there were some improvement from trial 1 to trial 3. But the '*after debriefing test*' was worse again. Participants 2, 8, 11 and 12 also showed improvement after the first trial and all of them were able to maintain the vessel's position for 5 minutes for at least once. The data of participant 5 showed that the task was well performed in the first two trials the attention decreased in the rest of two trials including '*after debriefing test*'.

Figure 19 is showing the data summary of case group. Participant 4 was not able to perform well in all the four trials and it was hard to see improvement overall. The data of participant 14 showed some decrease through first three trials and a big improvement in the last one. All the other three participants experienced improvement.

7.1.3 Expert check sheets

The expert's evaluation for each participant's trials was made based on experts' experience and own observation during experiment. It contains a combination of subjective evaluation and objective data. The scores from expert's check sheets of four trials are shown in Figures 20 and Figure 21, both in case and control groups. Each check sheets had four sub-topics and sixteen questions in total. All the questions in check sheets were scaled into 5. Participants 1 to 12 were instructed by expert 1. Participant 13 and 14 were instructed by expert 2. Two experts could have different standard to evaluate the performance.

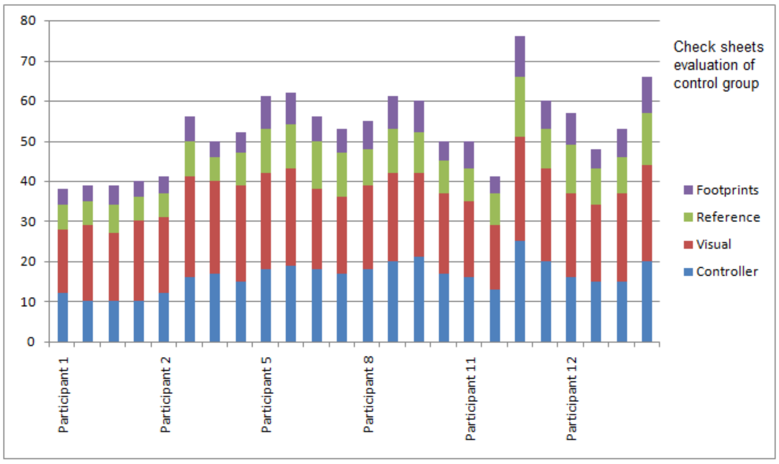

Figure 20. check sheet evaluation of control group

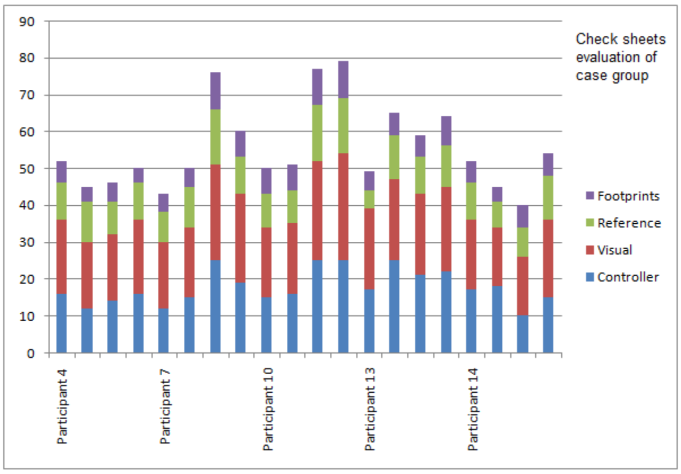

Figure 21. check sheet evaluation of case group

This data shows how the experts felt about the quality of the operation. All the evaluation was done directly after each trial based on experts own perception. Some of the performance were extremely good, such as participants 7, 10, and 11. Some of their trials received score 95% or more. It can also be seen that participant 1 and 4 had lower score compare to other participants, while showing same results in terms of time durations.

7.1.4 Pre-analysis of time duration and check sheets

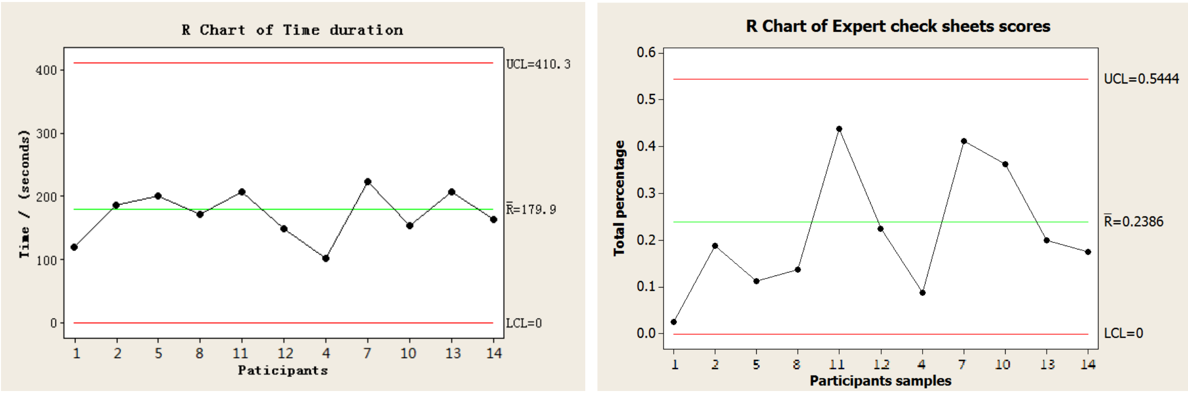

Figure 22. R charts of time duration and check sheets

Figure 22 showed the R charts of time duration and results from check sheets. The value in the left figure which represented each participant was the range of time durations. Similarly, in the right figure showed the range of scores in percentage. Figure 18 provided an assessment of stability of each participant and the entire operation. The time durations of each participant are very stable compare to scores from check sheets. It is clear that besides participant 1 and 4, other participants' range of time duration were around 180 seconds, which represents during their operations, they had more significant improvement than participant 1 and 4. From the perception of the experts (R chart of expert check sheets scores), the entire operation were also within the control limit although it fluctuated much more than the time durations. However, the range of score changes also showed that participants 1 and 4 had less improvement than others.

7.1.5 After-experiment self-report

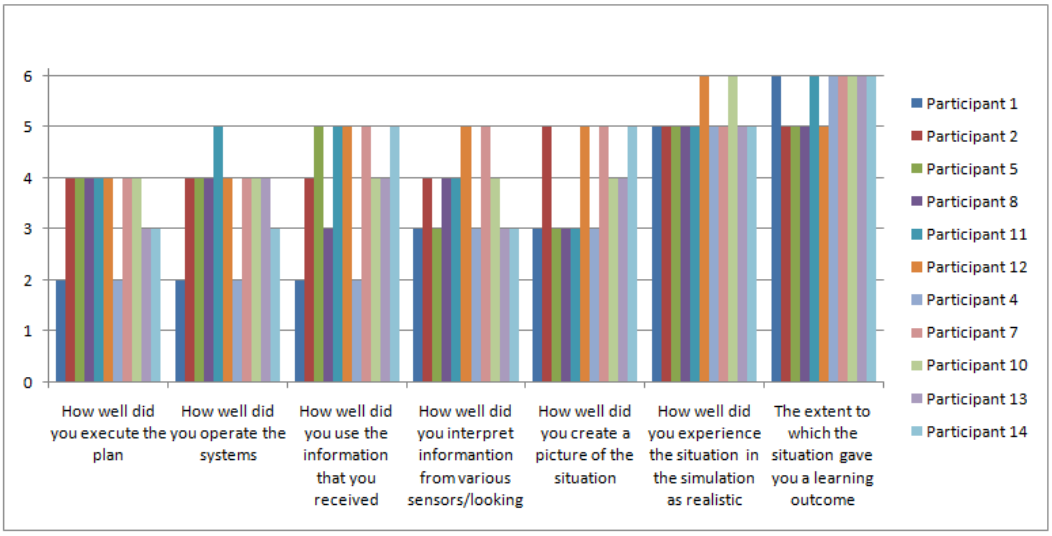

Figure 23. After-experiment self-report

Figure 23 is the self-report by the participants after finish the training. From the first two items, participants 1 and 4 were not satisfied with their performance in the experiment. All the other participants had almost same number for these two items. Item 3 through 5 represent the information processing. These three items showed big difference among participants, however participants 1 and 4 had less points compare to others. The last item is the participants' own perception of learning outcomes. All the students in case group gave it full scale and four students in control group gave five out of six for this item.

7.1.6 Footprint of vessel movement

The footprint of each trial are manually processed from the footprint screen video. Each of the footprints contains the following information, plotted over time:

- Displacement on global X, Y axis, 3D plot
- Global X displacement
- Global Y displacement
- Distance from initial position
- Vessel's heading change from initial angle

Figure 24 show an example of the 1st trial of the participant 12, all the plot were extracted from Appendix 3.

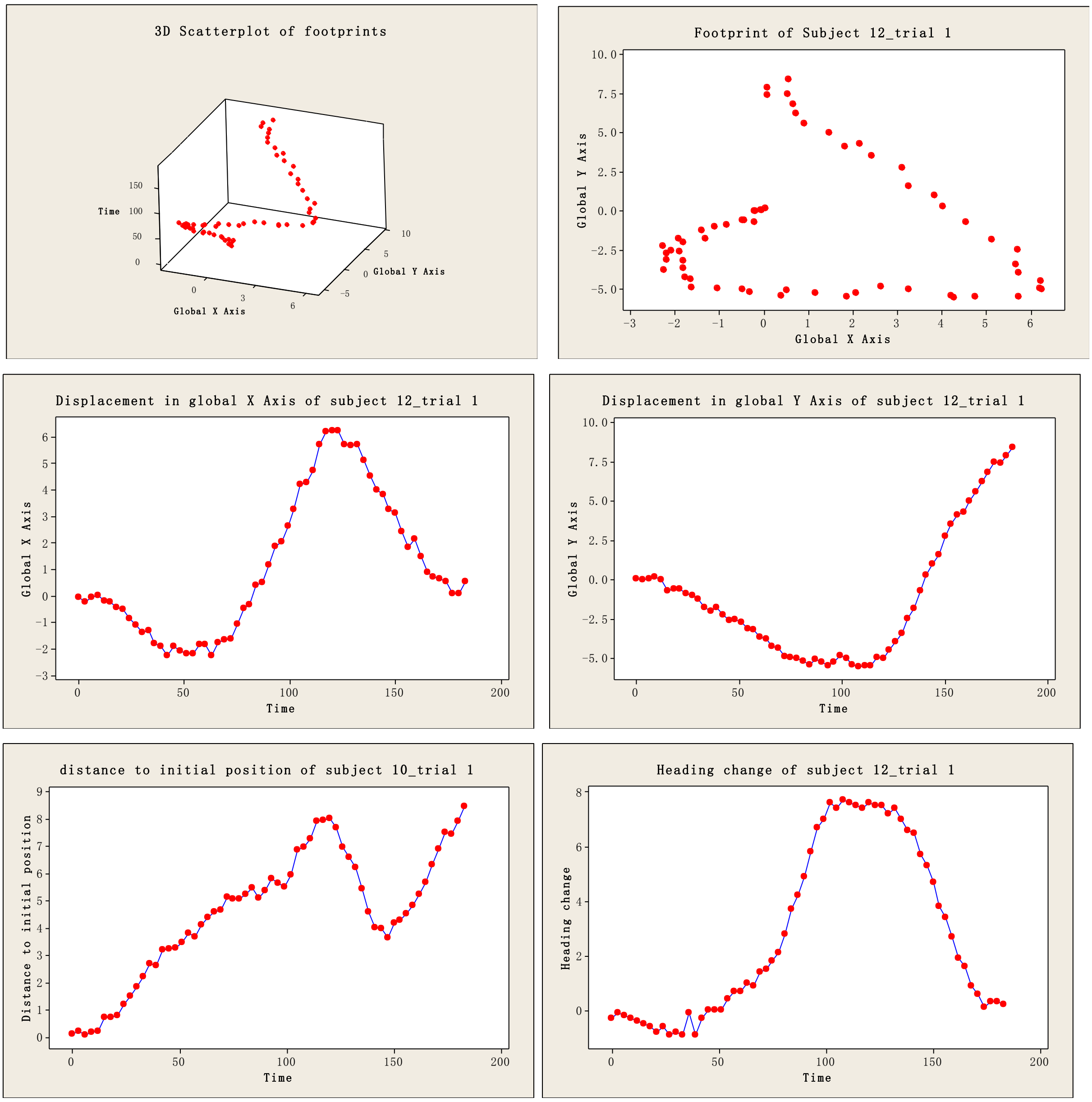

Figure 24. Plots of participant 12, trial 1st

According to the curves of the footprints, some features can be generated including *mean value*, *standard deviation*, *maximum and minimum values and data range* [see Appendix_1].

All these features of each trials and participants can be transferred into R charts, which is a standard way to analyze the data variability. But only the data range of *distance to initial position* and *heading change* of each trial would be analyzed, because the remaining three charts contain redundant data: the position and heading displacement. And in normal situations, trainers suggested to evaluate the trainees' performance based on these two features.

Distance to initial position:

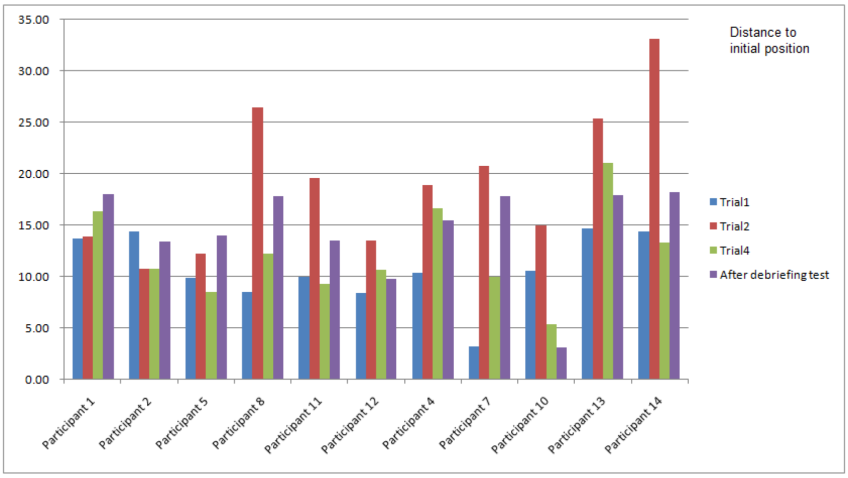

Figure 25. Distance to initial position

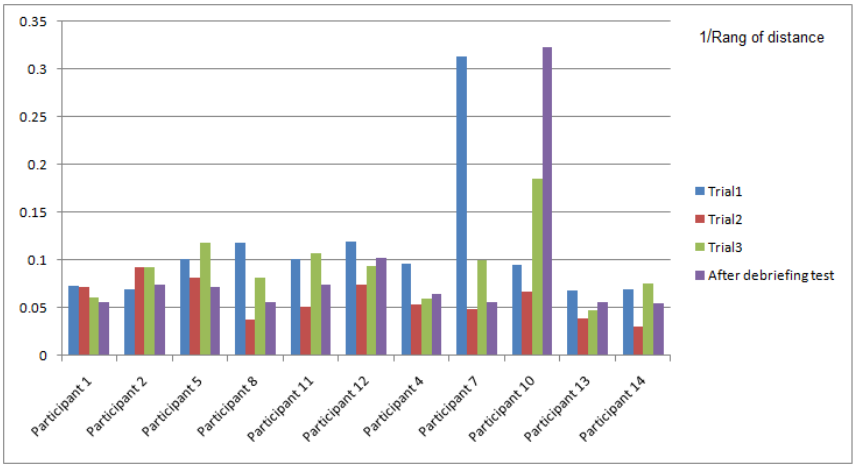

Figure 26. Reciprocal of distance to initial position

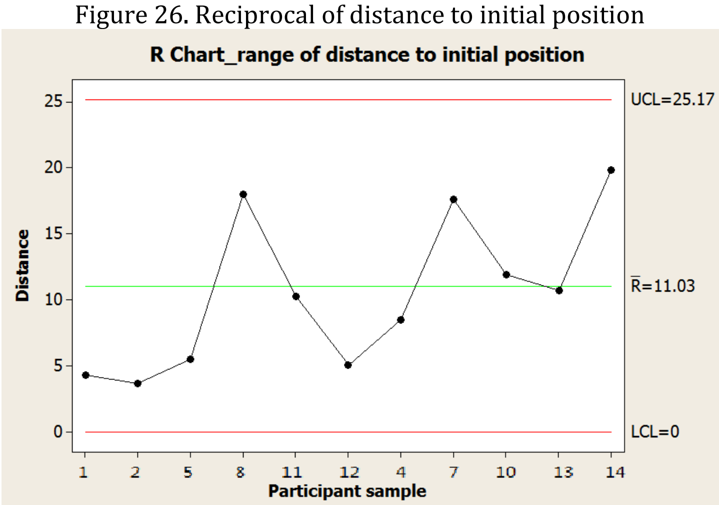

Figure 27. R chart of distance to initial position

The entire data of distance to initial position were posted in Figure 25. Some of the trials exceeded the criterion with small displacement such as the first trial of participant 7 of which the time duration was 147 seconds and the distance was only 3.2 meters. Some of the trials exceeded the criterion with big displacement such as second trial of participant 14, the time duration was 230 seconds and the distance was over 30 meters.

Figure 26 is the reciprocal of distance to initial position. Since the numbers of this feature are expected to be as small as possible, therefore the reciprocal of this feature was more intuitionistic for evaluating the performance, which means the bigger number in the reciprocal chart is, the better performance that the participant had. The best performance were participants 7 and 10, respectively 3.2 meters and 3.1 meters. The data of participant 7 was contradictory to the time duration, which in the operation was only 146 seconds. By checking the video of participant 7, trial 1st, it showed that the vessel lost position because of exceeding the transverse displacement criterion, which means the container touched the ship side. It was much easier to exceed transverse criterion than longitudinal criterion. And also because of the relative position between vessel and container was decided by both vessel's heading and coordinate. The contradiction between distance to initial position and time duration can be explained.

The variability of distance to initial position were plotted in Figure 27. Statistically, the entire operation was with the upper and lower control limit and that means it was in control.

Heading change:

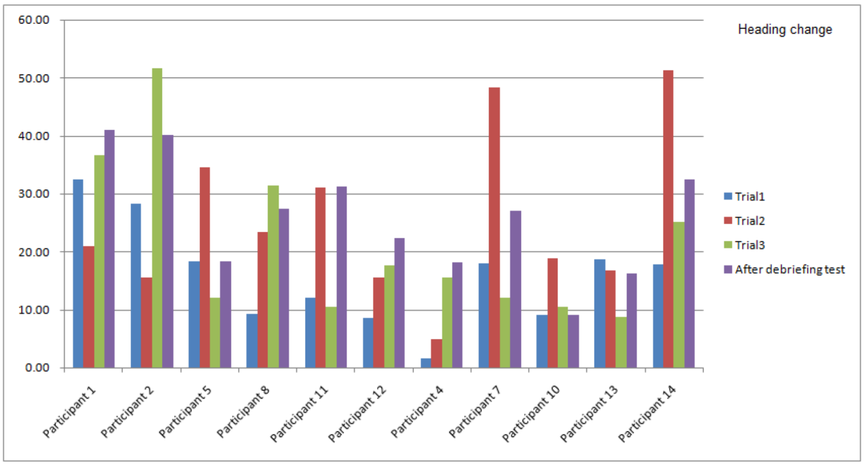

Figure 28. Heading change

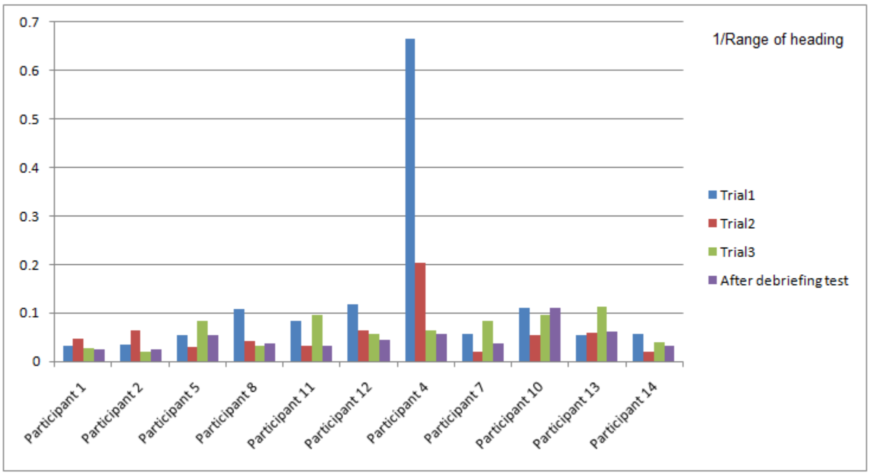

Figure 29. Reciprocal of heading change

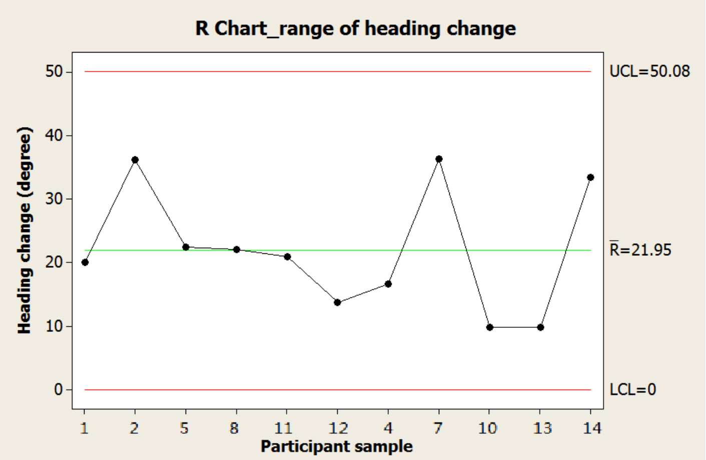

Figure 30. R chart of heading change

Similar with distance to initial position, the data of heading change were plotted in Figure 28, Figure 29 , and Figure 30, which are heading change, reciprocal of heading change and R chart.

7.1.7 Fixation numbers at Area of Interests

During the participants working on the operation, their eye-tracking videos were recorded by the ASL laptop and these videos can be post processed to get metrics such as fixation numbers/percentage at AIOs.

Based on the previous data including the time duration, expert's check sheets and features of footprints and also the eye-tracking video quality, participant 10, 13 showed some significant improvement from 3rd trial to 4th trial, which is the test trial after debriefing. Hence these 6 trials are selected to conduct more analysis in terms of eye-tracking videos including fixation on AIOs.

Fixation data of 3rd and 4th trial of participant 10:

The third trial of each participant is the last one in the third part of the experiment, where they have most practice by themselves compared to the first and second trial. Therefore the eye-movement pattern in the third trial was expected to be more stable and representative to their understanding of the operation. The fourth trial was the one after debriefing, where the participants had received feedback from experts again. The participants were expected to have better performance and visual focus.

There were four categories in each bar plot, including previously discussed rig, container and monitor screens, additionally the fixations out of these three areas, 'outside' area was also taken into consideration. Number of fixations, total average fixation duration on each AOI are the metrics used for analysis.

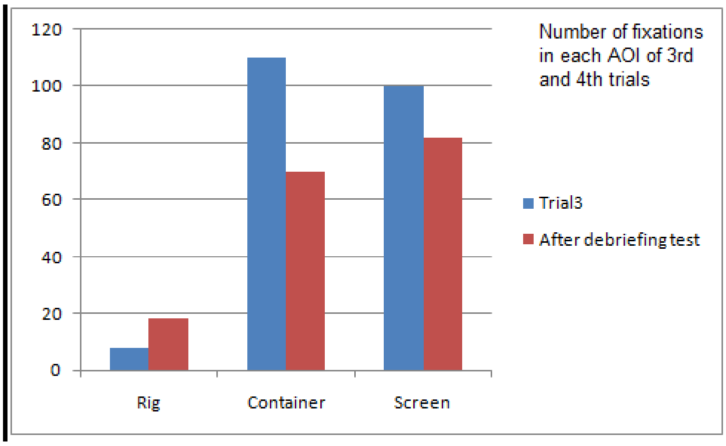

Figure 31. Number of fixations in each AOI of 3rd, 4th trial

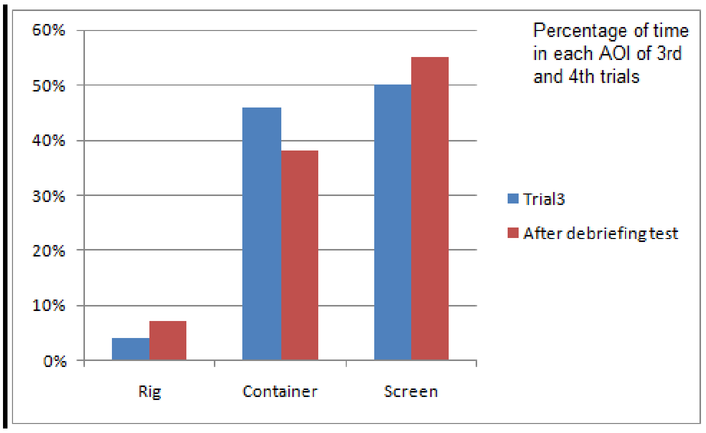

Figure 32. Percentage of time in each AOI of 3rd, 4th trial

From Figure 31 and Figure 32, there were 8 times of fixations on the *rig* in the third trial of participant 10 and the percentage was 4%. However in the fourth trial, there were 18 times of fixations and the percentage was 7%. The number of fixation and time percentage of container in the 3rd trial were 110 and 46%. In the 4th trial they were 70 and 38%. The number of fixation and time percentage of screen in the 3rd trial were 100 and 50%. In the 4th trial they were 82 and 55%.

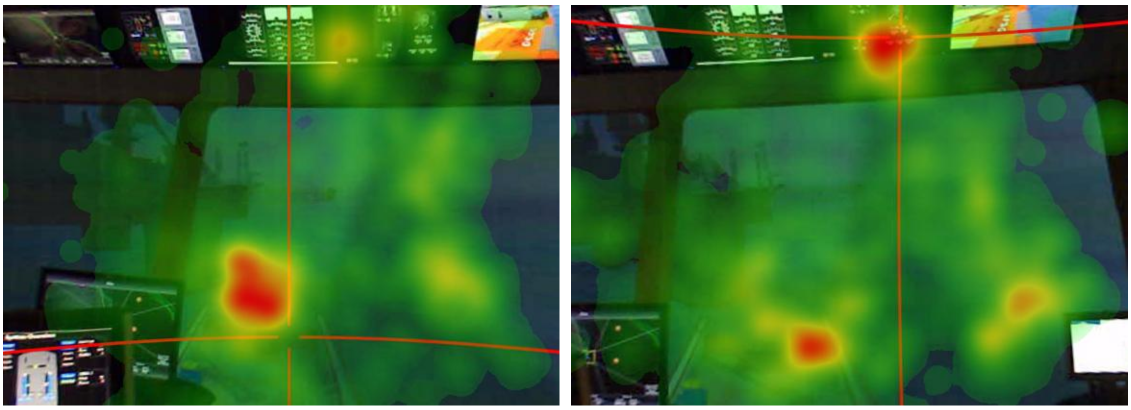

Figure 33. Heat map of 3rd, 4th trials

By visualizing the fixation data on participant's vision, Figure 33 showed the heat maps of 3rd and 4th trial. The color on a specific area represented the time duration that the participant looking at. Warmer the color was , longer the participant looked at. The comparison showed that in 4th trial, the participant spent less time looking at the container, more time on screen and rig than 3rd trial. It gave the same information as the number of fixation data.

Fixation data of 3rd and 4th trial of participant 13:

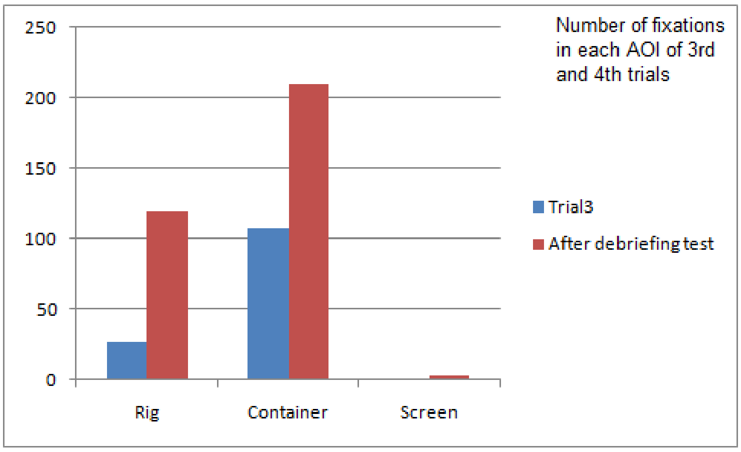

Figure 34. Number of fixations

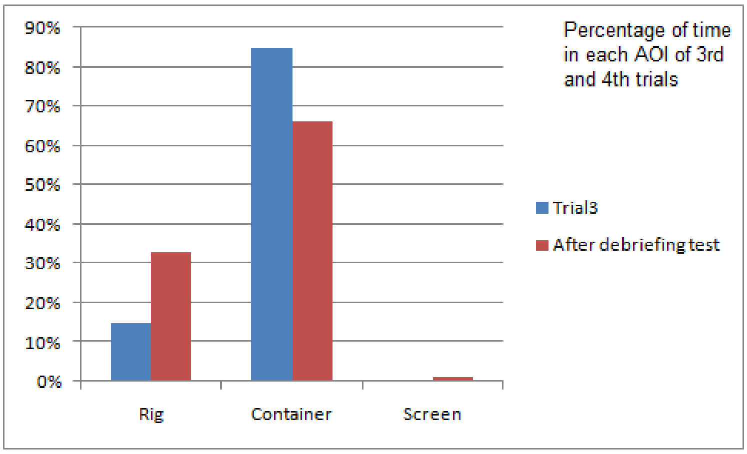

Figure 35. Percentage of time

From Figure 34 and Figure 35, there were 27 times of fixations on the *rig* in the third trial of participant 13 and the percentage was 15%. In the fourth trial, there were 120 times of fixations and the percentage was 33.5%.

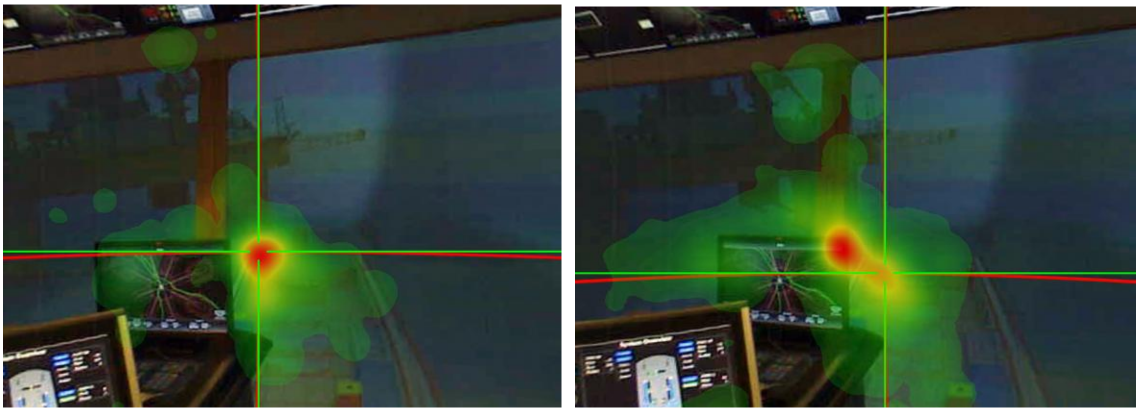

Figure 36. Heat maps of 3rd and 4th trial

From Figure 36, there were more green-colored area on the rig in the 4th trial. The heat maps showed same results as number of fixation data, which was that the participant spend more time looking at the rig.

7.2 Hypotheses testing

7.2.1 Time duration analysis

The data was collected by checking the operation time duration from recorded video of each trial. All the samples are independent, therefore the *t-test* method would be efficient to test the hypothesis 1, which involved two variables, Time[case] and Time[control].

Three trials:

Each participant is represented by the mean value of the first three trials.

Table 3. Mean value of the first three trials

Participants	1	2	5	8	11	12	4	7	10	13	14
Mean, time	163	236	237	243	179	185	96	225	212	171	217

The null hypothesis was that the time duration of case group and control group was same ($H_0 : \mu_x = \mu_y$). The alternative was that the time duration of two groups had significant difference and case group was higher than control group ($H_1 : \mu_x > \mu_y$). Based on the data in the table, the mean value of case group was $\bar{x} = 184$ control and the mean value of control group was $\bar{y} = 207$.

Calculate $t_0 = \dfrac{\bar{x} - \bar{y}}{S_p \sqrt{\dfrac{1}{n} + \dfrac{1}{m}}}$, where $S_p = \sqrt{\dfrac{(n-1)S_x^2 + (m-1)S_y^2}{n+m-2}}$

And S_x^2 and S_y^2 are the individual sample of variances, corresponding to case group and control group.

$$S_x^2 = \frac{\sum_n x_1 + x_2 + ... + x_n}{n-1}, \quad S_y^2 = \frac{\sum_m x_1 + x_2 + ... + x_m}{m-1}$$

Criterion of hypothesis test: $H_1 : \mu_x > \mu_y$: reject H_0 if $t_0 > t_{0.025, n+m-2}$.

The $t_{0.025, n+m-2}$ value is decided by the sample size n and m which in this case are 5 and 6.

It can be found that S_x^2=2867 and S_y^2=1248, S_p =44 and t_0 =-0.86. $t_{0.025, n+m-2}$ can be found as 2.262 (from appendix of book *Experimentation in Software Engineering: An Introduction* [22]). It is not possible to reject the null hypothesis.

After debriefing test:

For testing the effect of different method in debriefing, the time durations of *fourth trial* were compared with the *mean value of first three trials*.

Table 4. Change of time after debriefing

Participants	1	2	5	8	11	12	4	7	10	13	14
Value	-13	56	-137	10	120	115	2	74	87	112	83

The null hypothesis was that the change of time durations of case group and control group was same ($H_0 : \mu_x = \mu_y$). The alternative hypothesis was that the debriefing in case group had better effect, which means the time duration would increase more than control group ($H_1 : \mu_x > \mu_y$). The testing method would be *t-test*.

Based on the data in the table, the mean value of case group was $\bar{x} = 71.6$ and the mean value of control group was $\bar{y} = 25.2$.

After calculation, it can be found that S_x^2=1712 and S_y^2=9200, S_p=77 and $t_0 = 1$. $t_{0.025,n+m-2}$ can be found as 2.262. It is not possible to reject the null hypothesis.

Although from statistical analysis it was not possible to give conclusion that the method in case group had better effect to participants' performance, however it is still clear that all the participants had improvement after debriefing in the mean time participants 1 and 5 in control group did not have any improvement.

7.2.2 Experts' check sheets analysis

The hypotheses made for experts' check sheets were similar to the time duration, which were designed for comparing first three trials and the difference between fourth trial and mean value of first three trials. The involved variables were Che[case] and Che[control].

Three trials:

The null hypothesis was '*there will be similar scores of expert's check sheets between case and control group' ($H_0 : \mu_x = \mu_y$).* The alternative hypothesis was that participants in case group had better scores ($H_1 : \mu_x > \mu_y$).

Similar as analysis of time duration, each participant was represented by the mean value of the three trials.

Table 5. Check sheets mean values

Participants	1	2	5	8	11	12	4	7	10	13	14
Value	38.7	49.0	59.7	58.7	55.7	52.7	47.7	56.3	59.3	57.7	45.7

Based on the data in the table, the mean value of case group was $\bar{x} = 53.34$ control and the mean value of control group was $\bar{y} = 52.42$.

After calculation, it can be found that S_x^2=38 and S_y^2=61, S_p=7 and $t_0 = 0.213$. $t_{0.025,n+m-2}$ can be found as 2.262. It is not possible to reject the null hypothesis.

After debriefing test:

With the same hypothesis testing method, the experts' evaluation was analyzed.

The null hypothesis was that '*the change of experts' evaluation of case group and control group was same'* ($H_0 : \mu_x = \mu_y$). The alternative hypothesis was that '*the debriefing in case group had better improvement than control group'* ($H_1 : \mu_x > \mu_y$).

Table 6. Check sheets improvement in 4th trial

Participants	1	2	5	8	11	12	4	7	10	13	14
Value	1.3	3.0	-6.7	-8.7	4.3	13.3	2.3	3.7	19.7	6.3	8.3

Based on the data in the table, the mean value of case group was $\bar{x} = 8.06$ control and the mean value of control group was $\bar{y} = 1.08$. After calculation, it can be found that S_x^2=48 and S_y^2=64, S_p=7.5 and $t_0 = 1.53$. $t_{0.025, n+m-2}$ can be found as 2.262. It is not possible to reject the null hypothesis.

The results from check sheets and time duration were not able to reject their null hypotheses. But the t_o values of time duration and check sheets, which were calculated for checking debriefing quality, illustrated similar results that the debriefing with help of eye-tracker had better effect than using the overview video, but not 95% available.

7.2.3 After-experiment self-report

The after-experiment self-report was consisted of the participants' own perception of their experiences, which in detail included performance, information treatment, and learning outcome. For compare each participants self assessment, only the last question of learning outcome was analyzed with *t-test* method. The involved variables were Fe[case] and Fe[control].

Table 7. Learning outcome

Participants	1	2	5	8	11	12	4	7	10	13	14
Value	6	5	5	5	6	5	5	6	6	6	6

As defined in section 5.2.3, the null hypothesis was '*There will be similar participants' feedback between case and control group* ($H_0 : \mu_x = \mu_y$)' and the alternative hypothesis was $H_1 : \mu_x > \mu_y$.

Based on the data in the table, the mean value of case group was $\bar{x} = 5.8$ and the mean value of control group was $\bar{y} = 5.3$.

After calculation, it can be found that S_x^2 =0.20 and S_y^2 =0.27, S_p =0.49 and t_0 = 1.58. $t_{0.025, n+m-2}$ can be found as 2.262. It is not possible to reject the null hypothesis.

Similar as the t_0 values of *after debriefing test*, 1 for time duration and 1.58 for check sheets, the null hypotheses could not be 95% rejected. The mean values of case and control group provided information that there are better improvement in case group. But the data collected in this experiment were not significant enough to reject the null hypotheses.

7.2.4 Summary and main findings

Table 8_Summary of hypotheses testing with t-test method

ID	$H1_1$	$H2_1$	$H3_1$	$H4_1$	$H8_1$
Mean, case	184	71.6	53.34	8.06	5.8
Mean, control	207	25.2	52.42	1.08	5.3
t_0 **value**	-0.84	1.00	0.21	1.53	1.58

Table 8 gave a summary of the important values in hypotheses testing for further discussions. for each tested hypothesis, the mean values of case and control group, the t_0 value were presented.

Table 9_Hypothesis tesing results

ID	Description	Results
$H1_1$	In the first three trials, there will be significant difference in mean value of time duration	Rejected
$H2_1$	There will be significant difference in change of time duration in the 4th trial	Rejected
$H3_1$	In the first three trials, , there will be significant difference in scores of check sheets	Rejected
$H4_1$	There will be significant difference in change of check sheets score in 4th trial	Rejected
$H8_1$	There will be significant difference in participants' feedback between case and control group	Rejected

Table 9 summarized the results for the tested hypotheses. There were 5 of 8 hypotheses were tested in chapter 7. All the alternative hypotheses were rejected. Although the current data were not able to show significant changes between new training methods with eye-trackers and normal training methods without eye-trackers. The results in table 8 still indicates that in this experiment the case group had better learning outcome than control group. The mean values of hypothesis 2 were 71.6 and 25.2 corresponding to case group and control group. Besides hypothesis one, the other hypotheses showed better performance in case group than control group. Other variables were discussed in the following chapter.

8 DISCUSSION

This chapter discussed the results in Chapter 8. The results were separately discussed to evaluate the effect from application of eye-trackers in operation practicing part and debriefing part.

8.1 Operation practicing

The subjects were first year and second year students from Nautical science program at Aalesund University College. The reason that they were chosen as the participants of experiment was that the background of these students was perfect for this experiment since their education was mainly about navigation and they have dynamic positioning course, which introduced basic knowledge. This advantage reduced the introduction work and moreover, some of the students even had real experience working on an anchor-handling vessel. Because most of the trainees that need DP training courses are experienced skippers, the nautical science students represented a test subject set very similar to usual trainees who join real DP training courses.

After the operation practicing with trainer's intervention, the performance of three trials was not significantly different between case and control groups. The average numbers of successful operation was 1.4 and 1.33 for case and control groups respectively. Both groups had a single participant who could not succeed in all four trials. The average time duration of case and control groups are respectively 184 seconds and 208 seconds. This variable even showed worse performance from participants in teh case group. The average score from check sheets evaluation was 53.34(case group) and 52.42(control group). Based on this data no evidence is found proving that practicing with eye-trackers had a positive influence to the participants. The case control study applied in this part of training was intended to compare the two different intervention method. But the insignificant results could indicate that the eye-tracker was not properly used, leading to the results that both case and control group had similar intervention quality during operation practicing. The experts feedback during the operation also gave similar information.

Based on the experts comments:

"I was focusing much more on telling the students to operate the controllers properly compensate different kinds of vessel motion."

Also the eye-tracking video captured from eye-tracker worn by experts shows that the experts did not pay enough attention on real time eye-tracking video when intervene in the case group.

8.2 Debriefing

As discussed in chapter 2, debriefing was an activity where trainers were able to improve trainees' learning quality by giving suggestion based on past performance. Eye-tracking videos have been widely used as an assistive tool for debriefing in many training domains such as sports training and driving training. Therefore part of the training method in the experiment was designed to test whether using eye-tracking video in debriefing activities of manual maneuvering training can improve the debriefing quality, comparing to traditional debriefing using normal scene video.

The quality of debriefing was measured by means of comparing the 4th trial(after debriefing) with the average performance of the first three trials. Three variables were available for comparison between case and control group. *Change of time duration*(TC[case], TC[control]), *change of check sheets scores*(CC[case], CC[control]) and *change of number of fixation*(Fix[case], Fix[control]). Although the hypothesis testing in chapter 7 did not prove significant difference between case and control group, the mean value of data still revealed some findings. These three variables show that, the debriefing in case group were more efficient than control group. The average number of CC[case] and CC[control] were 71.6 seconds and 25.2 seconds respectively. The case group had almost two times more improvement compare to control group in terms of operation time duration. In the same time, the experts also gave better scores to the case group compare to control group. The improvement in check sheets evaluation was 8.06 points and 1.08 points respectively.

It is clear that the participants in the case group had more information from experts that would help them perform better than the previous performances. Based on the eye-tracking video, the experts emphasized more on the best areas to focus and which objects to use as references, and which objects can be misleading reference points.

In Section 7.1.7, the number of fixation of participants 10 and 13 were presented. The comparison between 3rd trial and 4th trial showed huge difference in terms of fixation number and fixation percentage in different areas of interests.

Table 10_Fixation number of participant 10

Participant 10	number of fixation (percentage of time)		
	Rig	Container	Screen
3rd trial	8 (4%)	110 (46%)	100 (50%)
4th trial	18 (7%)	70 (38%)	82 (55%)

Table 11_Fixation number of participant 13

Participant 13	number of fixation (percentage of time)		
	Rig	Container	Screen
3rd trial	27 (15%)	108 (85%)	0 (0%)
4th trial	121 (34%)	210 (66%)	3 (2%)

Participant 10 and 13 were members in case group. The data in Their 4th trials had significant improvement compare to the first three trials. By looking into the eye-tracking data, both of them spent more fixation number and percentage of time looking at the rig and screens, just same as DP experts suggested.

These results provided evidence that besides being familiar with the control system and thruster power of a single vessel, the manual maneuvering operation needs the operators to collect situational information of surroundings by spending enough time looking at the right reference objects, which in this experiment was the rig.

8.3 After-experiment self-report

The goal of the self-report questionnaire was to collect trainees' perception of the experiment. It involved different questions such as self-assessment of the performances and learning outcome after finishing the experiment.

The questions '*how well did you execute the plan*' and '*how well did you operate the system*' reflect the participants' own evaluation of their performance. From the results in Chapter 7, there were no significant differences between case and control groups. It gave same conclusion as the results in Section 8.1 the practicing part in the experiment provided similar quality of feedback from the trainers. This resulted to similar time duration, check sheets evaluation and self-report assessment.

Questionnaire data of learning outcome were tested in section 7.2.3. The test rejected the hypothesis that *There will be significant difference in participants' feedback between case and control group*. But it was clear all of the participants in case group gave full scale for this question while only 2 participants in control group gave full scale and others were 5 out of 6.

9 CONCLUSIONS

The author of this thesis has conducted a case-control experiment that compared two training methods of manual maneuvering operation. All the participants went through the training procedures and eye-trackers were introduced to the participants of case group in the beginning of the experiment. The purpose of the experiment was to verify the impact of eye-trackers to manual maneuvering training, more specifically trainers' intervention and debriefing activities. All the participants were asked to learn the operation techniques and try to maneuver the vessel on their own. They had trainers' intervention and debriefing as input knowledge, which differed from case to control group. The participants' performance and own perception were collected by several methods, including expert's check sheets, video recording, and self-report questionnaires. After the experiment execution, collected data was post-processed and analyzed with the use of bar charts, R charts and t-test.

T-test method was used to test the hypotheses in Chapter 7. All the hypotheses were rejected due to low number of participants. Therefore, none of the hypotheses were statistically proved. Due to time limitation and students' schedules, it was not possible to have more participants for the experiment. As discussed in Section 4.7, each participant needs 2 hours to finish the entire experiment. It was maximum 3 participants joined the experiment in a single day. For getting enough sample size, the predicted time that needed is 10 more working days at minimum. Both the experts and the author did not have time to increase the sample size.

After the results analysis, three main conclusion can be made:

- The results indicate that the implementation of eye-tracker and the eye-tracking video were complementary way for debriefing activities of manual maneuvering training. Because of the advantage that eye-tracking video were able to visualize the participants' eye-movement during the operation, it was intuitional for experts and participants to understand and point out the error with visual attention.
- By compare eye-tracking data between 3rd trial and 4th trial which was the trial after debriefing, the results can be made that the participants understood to spend more visual attention on better reference object as experts suggested. And by looking at the right reference object, which in this case was the rig beside the vessel, they can operate better.
- The results and comments from expert also indicate the fact that the eye-tracker was not useful in the practicing part of the training. It was not used as expected since the expert focuses on guiding the participants to operate for most of time, which left little time for them to check the eye-tracking real time video and give intervention about visual attention.

10 FURTHER WORK

This chapter discusses the author's suggestion based on the experiment and conclusions.

10.1 Suggestion to the experiment

Based on the results of the experiment and the validity analysis, several suggestions were made for improve the experimental design for the future research.

- The design of participants' questionnaire should include the educational background and working experience. In this experiment, the author chose the participants as the students in Nautical Science. It was based on a clear reason that these students had basic knowledge on DP system. But their working experiences were not considered during the participants selection and the educational background was not specific enough. The current version of self-report questionnaire cannot provide enough information about the participants. To include the specific educational background and working experience would help the experiment a lot.
- The data transmission of eye-tracking video was interfered in the experiment. Since it had been never experienced before, the reason was speculated as the signal interference in simulator room, considering there were much more electrical devices compare to normal environment. Therefore, conducting similar experiment in any kinds of simulator, the researcher should use Display/transmit Unit (see Figure 37) of the eye-tracker components to record the video instead of wireless transmission.

Figure 37. ASL eye-tracker Display/Transmit Unit DTU

- If the vessel footprint are interested in the future projects, it is better to use an official software which is able to output the footprints of vessel instead of manually extract from video record. The author was trying to use the footprint of vessel to evaluate the performance of each trial. But the results of vessel movement were not able to draw any conclusion. Even though, the vessel movement is a valuable feature for DP operation and manual maneuvering. To develop a efficient method to analyze the footprint would be important.

10.2 Further work

The experiment results answered to the research question of this thesis and point out the probability of future work.

The effect of using eye-tracking video in debriefing activity was studied. It was proved that the eye-tracker is an efficient tool to improve quality debriefing activity of manual maneuvering training. So the application and conclusion could be used in future research or put into practice.

The results also provide evidence that the correct eye-movement strategy is important for a novice to have better performance in manual maneuvering operation. This conclusion can lead to an interactive design for training, which combines eye-tracker technology and alert system to give automatic feedback to trainees when they focus on less preferable object for too much time.

Eye-trackers' application in maritime training can be extended to other operations. Such as anchor handling operation and crane operation.

REFERENCES

[1]. Duchowski, A. (2000), "Eye-Based Interaction in Graphical Systems: Theory and Practice". *Siggraph 2000 Course 05 Notes.*

[2]. Holmqvist, K & Nystrom, M. (2011). *Eye tracking – a comprehensive guide to methods and measures.* Oxford University Press.

[3]. Applied Science Laboratories. *Sports training and research*. Retrieved November 29, from http://www.asleyetracking.com/Site/Applications/SportsandAthletics/tabid/263/Default.aspx.

[4]. Benjamin Law, M. Stella Atkins, A.E. Kirkpatrick, Alan J. Lomax, Christine L. Mackenzie. (2004). "Eye Gaze Patterns Differentiate Novice and Experts in a Virtual Laparoscopic Surgery Training Environment". *Eye tracking research & applications*, 41 - 48.

[5]. Jalitha Wills. (2007). *Dynamic positioning simulator interim report.*

[6]. Carlton, J.S. (1994). "Marine Propellers & Propulsion". Oxford, Butterworth Heinemann.

[7]. J. Banks, J. Carson, J. Nelson, D. Nicol. (2001). *Discrete-Event System simulation.* Prentice Hall, page 3.

[8]. B. Magerko & J. Laird. (2002). "Towards Building an Interactive, Scenario-based Training Simulator". *Proceedings of 11th conference on computer generated forces and behavioral representation.* 9 pages.

[9]. D. Hansen & Q. Ji. (2010). "In the eye of the beholder: A survey of models for eyes and gaze". *IEEE Trans. Pattern Anal. Mach. Intell* . 32: 478-500.

[10]. Poole, A. & Ball, L. (2005). "Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects". *Prospects", Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction.*

[11]. Goldberg, H. J, & Kotval, X. P. (1999). "Computer interface evaluation using eye movements: Methods and constructs". *International Journal of Industrial Ergonomics*.24:631-645.

[12]. Poole, A., Ball, L. J., & Phillips, P. (2004). "In search of salience: A response time and eye movement analysis of bookmark recognition". *People and Computers XVIII-Design for Life: Proceedings of HCI 2004*. London: Springer-Verlag Ltd.

[13]. Albert, W. (2002). "Do web users actually look at ads? A case study of banner ads and eye-tracking technology". *Proceedings of the Eleventh Annual Conference of the Usability Professionals' Association*.

[14]. Mann, C J. "Observational research methods. Research design II: cohort, cross sectional, and case-control studies".

[15]. Montgomery, Douglas, (2009). *Introduction to Statistical Quality Control*. Wiley.

[16]. Schultz, John. (2006). "Measuring Service Industry Performance: Some Basic Concepts". *International Society for Performance improvement.* p. 3.

[17]. Ishikawa, K. (1986). *Guide to Quality Control.* Tokyo: Asian Productivity Organization.

[18]. Mellenbergh, G.J. (2008). *Tests and Questionnaires: Construction and administration.* The Netherlands: Johannes van Kessel Publishing.

[19]. Marchal, Martin N. (1996). "Sampling for Qualitative Research". *Family Practice.* 13: 522-526.

[20]. Lucas, Samuel R. (2014). "Beyond the Existence Proof. Ontological Conditions, Epistemological Implications, and In-depth Interview Research". *Quality & Quantity.* 48:387-408.

[21]. Rodrigues L & Kirkwood BR. (1990). "Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls". *International J of Epidemiology*.

[22]. Claes. W, Per. R, Martin. H, Magnus. O, Björn. R & Anders. W. (2000). *Experimentation in Software Engineering: An Introduction.*

[23]. Renganayagalu, Sathiya K. (2013). "Eye-trackers as assistive technologies in maritime training".

[24]. The Nautical Institute. (2014). "Nautical Institute accredited DP Training Centres, April 2014"

[25]. Kongsberg. *Bridge systems Dynamic Positioning courses*. Retrieved November 28. http://www.km.kongsberg.com/Internal/HUB0024.nsf/CourseGroups?ReadForm&Count=-1&Area=Bridge%20systems&Group=Dynamic%20Positioning

[26]. Mellenbergh, G.J. (2008). *Advising on Research Methods: A consultant's companion*. Johannes van Kessel Publishing.

APPENDIX

Appendix 1. Self-report questionnaires

Part 1

This part of the questionnaire deals with you as a person in general. Answer as best you can and be honest. The information you provide will be treated confidentially. Take a stand on these statements. Circle the option that best suits your needs (from 1 = "very poor" to 6 = "Votes very good")

1. I usually follow rules and procedures. 1 2 3 4 5 6
2. I usually have good ability to make decisions. 1 2 3 4 5 6
3. I'm thinking before I act. 1 2 3 4 5 6
4. I enjoy working with others. 1 2 3 4 5 6
5. I have a good sense of space / sense of direction. 1 2 3 4 5 6
6. I'm good at discovering details in totality. 1 2 3 4 5 6
7. My judgment is good 1 2 3 4 5 6

Part 2

Think back to the situation in the simulation and circle the answer that best fits (from 1 = "To a very small extent" to 6 = "To a very great extent").

How well did you manage to execute the plan? 1 2 3 4 5 6

To what extent did you manage to operate the systems that were available?
1 2 3 4 5 6

To what extent were you able to use the information that you received?
1 2 3 4 5 6

To what extent were you able to interpret information from various sensors / looking?
1 2 3 4 5 6

To what extent did you create a picture of the situation?
1 2 3 4 5 6

To what extent did you experience the situation in the simulation as realistic?
1 2 3 4 5 6

The extent to which the situation gave you a learning outcome?
1 2 3 4 5 6

Appendix 2. Experts' check sheet

1. Controllers

[Find Right controllers] Rate: 1 2 3 4 5

Does the DPO understand connection between controllers and propellers?

Tunnel thruster Bow: Yes ☐ No ☐

Tunnel thruster stern: Yes ☐ No ☐

Main propeller port: Yes ☐ No ☐

Main propeller stb: Yes ☐ No ☐

[Force direction] Rate: 1 2 3 4 5

Does the DPO understand the direction of force vector?

Tunnel thruster: Yes ☐ No ☐

Main propellers: Yes ☐ No ☐

[Reaction] Rate: 1 2 3 4 5

Does the DPO react correctly to the vessel movement?

Sway motion, push TT controllers to desired direction: Yes ☐ No ☐

Surge motion, push MP controllers to desired direction: Yes ☐ No ☐

[Smooth adjustment] Rate: 1 2 3 4 5

Does the DPO operate smoothly?

Note:

[Thrust feedback] Rate: 1 2 3 4 5

Does the DPO understand feedback delay of propellers?

Note:

2. Visual focus evaluation

[Screen] Rate: 1 2 3 4 5

Not focusing too much on the screen.

approximate time that checking screen: 10% 30% 50% 70% 90%

[Container] Rate: 1 2 3 4 5

Not focusing too much on the container.

approximate time that checking container: 10% 30% 50% 70% 90%

[Rig] Rate: 1 2 3 4 5

Does the DPO focus a lot on the rig?

approximate time that checking rig: 10% 30% 50% 70% 90%

[Distance] Rate: 1 2 3 4 5

Does the DPO always focus on the correlation between rig and vessel?

including distance and degree.

Note

[Vessel motion] Rate: 1 2 3 4 5

Is the DPO aware of vessel movement?

Note

[Concentration] Rate: 1 2 3 4 5

Is the DPO concentrating during whole operation?

Note

3. Reference position

According to the hanging container as the reference,

[Surge] Rate: 1 2 3 4 5

Is the DPO able to recover surge displacements?

Displacement range: 10m 20m 30m 40m

leave deck area Yes☐ No☐

[Sway] Rate: 1 2 3 4 5

Is the DPO able to recover sway displacements?

Displacement range: 5m 10m 15m 20m

ship sides collision Yes☐ No☐

[Yaw] Rate: 1 2 3 4 5

Is the DPO able to keep heading around 315 degree?

Rotation displacement range: 5° 10° 15° 20° 25°

4. GPS Foot print trajectory

[Displacement] Rate: 1 2 3 4 5

The radius of vessel foot prints during the operation.

Radius: 2.5m 5m 7.5m 10m 15m

[Speed] Rate: 1 2 3 4 5

The density of vessel foot prints during the operation

Note:

Appendix 3. Example of footprint treatment, participant 12, trial 1

time	seconds	X20(m)	Y20(m)	Heading	Angle	Heading change	X0(m)	Y0(m)	Distance to initial position
0:05	0	0	0	315	45	-0.3	-0.1	0.1	0.1
0:08	3	-0.2	0	314.9	45.1	-0.1	-0.2	0.0	0.2
0:11	6	0	0	314.8	45.2	-0.2	0.0	0.1	0.1
0:14	9	0.1	0.1	314.7	45.3	-0.3	0.0	0.2	0.2
0:17	12	-0.1	-0.1	314.6	45.4	-0.4	-0.2	0.0	0.2
0:20	15	-0.1	-0.8	314.5	45.5	-0.5	-0.2	-0.7	0.7
0:23	18	-0.3	-0.7	314.4	45.6	-0.6	-0.4	-0.6	0.7
0:26	21	-0.3	-0.8	314.2	45.8	-0.8	-0.5	-0.6	0.8
0:29	24	-0.7	-1	314.4	45.6	-0.6	-0.8	-0.9	1.2
0:32	27	-0.9	-1.2	314.1	45.9	-0.9	-1.1	-1.0	1.5
0:35	30	-1.2	-1.4	314.2	45.8	-0.8	-1.4	-1.2	1.8
0:38	33	-1.1	-2	314.1	45.9	-0.9	-1.3	-1.8	2.2
0:41	36	-1.8	-2	314.9	45.1	-0.1	-1.8	-2.0	2.7
0:44	39	-1.7	-2	314.1	45.9	-0.9	-1.9	-1.8	2.6
0:47	42	-2.2	-2.3	314.7	45.3	-0.3	-2.3	-2.2	3.2
0:50	45	-1.9	-2.6	315	45	0	-1.9	-2.6	3.2
0:53	48	-2.1	-2.5	315	45	0	-2.1	-2.5	3.3
0:56	51	-2.2	-2.7	315	45	0	-2.2	-2.7	3.5
0:59	54	-2.3	-3	315.4	44.6	0.4	-2.2	-3.1	3.8
1:02	57	-2	-3	315.7	44.3	0.7	-1.8	-3.2	3.7
1:05	60	-2	-3.5	315.7	44.3	0.7	-1.8	-3.7	4.1
1:08	63	-2.5	-3.5	316	44	1	-2.3	-3.7	4.4
1:11	66	-2	-4	315.9	44.1	0.9	-1.8	-4.2	4.6
1:14	69	-2	-4	316.4	43.6	1.4	-1.7	-4.3	4.7
1:17	72	-2	-4.5	316.5	43.5	1.5	-1.6	-4.9	5.1
1:20	75	-1.5	-4.5	316.8	43.2	1.8	-1.1	-4.9	5.1
1:23	78	-1	-4.5	317.1	42.9	2.1	-0.5	-5.0	5.0
1:26	81	-1	-4.5	317.8	42.2	2.8	-0.3	-5.2	5.2
1:29	84	-0.5	-4.5	318.7	41.3	3.7	0.4	-5.4	5.5
1:32	87	-0.5	-4	319.2	40.8	4.2	0.5	-5.1	5.1
1:35	90	0	-4	319.9	40.1	4.9	1.2	-5.3	5.4
1:38	93	0.5	-4	320.8	39.2	5.8	1.9	-5.5	5.8
1:41	96	0.5	-3.5	321.7	38.3	6.7	2.1	-5.2	5.6
1:44	99	1	-3	322	38	7	2.6	-4.8	5.5
1:47	102	1.5	-3	322.6	37.4	7.6	3.2	-5.0	6.0
1:50	105	2.5	-3.5	322.4	37.6	7.4	4.2	-5.4	6.9
1:53	108	2.5	-3.5	322.7	37.3	7.7	4.3	-5.5	7.0
1:56	111	3	-3.5	322.6	37.4	7.6	4.7	-5.5	7.3
1:59	114	4	-3.5	322.5	37.5	7.5	5.7	-5.5	7.9
2:02	117	4.5	-3	322.4	37.6	7.4	6.2	-4.9	7.9
2:05	120	4.5	-3	322.6	37.4	7.6	6.2	-5.0	8.0
2:08	123	4.5	-2.5	322.5	37.5	7.5	6.2	-4.5	7.7
2:11	126	4	-2	322.5	37.5	7.5	5.7	-4.0	7.0
2:14	129	4	-1.5	322.2	37.8	7.2	5.7	-3.4	6.6
2:17	132	4	-0.5	322.4	37.6	7.4	5.7	-2.4	6.2
2:20	135	3.5	0	322	38	7	5.1	-1.8	5.4
2:23	138	3	1	321.6	38.4	6.6	4.5	-0.7	4.6
2:26	141	2.5	2	321.5	38.5	6.5	4.0	0.3	4.0
2:29	144	2.5	2.5	320.7	39.3	5.7	3.8	1.0	4.0
2:32	147	2	3	320.3	39.7	5.3	3.2	1.6	3.6
2:35	150	2	4	319.7	40.3	4.7	3.1	2.8	4.2
2:38	153	1.5	4.5	318.8	41.2	3.8	2.4	3.5	4.3
2:41	156	1	5	318.4	41.6	3.4	1.8	4.1	4.5
2:44	159	1.5	5	317.7	42.3	2.7	2.2	4.3	4.8
2:47	162	1	5.5	316.9	43.1	1.9	1.5	5.0	5.2
2:50	165	0.5	6	316.6	43.4	1.6	0.9	5.6	5.7
2:53	168	0.5	6.5	315.9	44.1	0.9	0.7	6.3	6.3
2:56	171	0.5	7	315.6	44.4	0.6	0.6	6.9	6.9
2:59	174	0.5	7.5	315.1	44.9	0.1	0.5	7.5	7.5
3:02	177	0	7.5	315.3	44.7	0.3	0.1	7.4	7.4
3:05	180	0	8	315.3	44.7	0.3	0.1	7.9	7.9
3:08	183	0.5	8.5	315.2	44.8	0.2	0.6	8.5	8.5

Appendix 4. The features of footprint of all participants

Participants	Trials	Time(s)	Heading change			X displacement			Y displacement			Distance to initial		
			Range	Mean	St dev	Range	Mean	St dev	Range	Mean	St dev	Range	Mean	St dev
1	T1	116	32.50	1.14	7.24	9.37	1.40	3.09	11.66	−5.28	3.77	13.73	5.90	4.31
	T2	139	21.00	−7.46	5.76	8.79	−1.92	2.07	19.68	−5.98	5.37	13.87	6.97	4.64
	T3	236	36.60	−0.13	0.60	20.42	−0.13	5.32	25.77	−4.95	6.00	16.40	8.08	4.79
	T4	150	41.00	12.51	11.11	16.30	6.52	6.37	19.92	−5.33	4.25	18.02	9.65	6.01
2	T1	114	28.20	−3.91	6.33	6.27	0.31	1.73	14.74	−6.99	5.28	14.45	7.22	5.26
	T2	300	15.50	−1.87	3.23	8.24	0.78	1.71	11.08	−3.64	3.39	10.77	4.21	3.25
	T3	294	51.70	4.32	15.30	13.37	4.95	4.51	30.62	−2.07	9.44	10.80	10.16	5.85
	T4	292	40.10	−4.11	10.12	11.76	−1.53	2.98	19.80	−2.36	5.45	13.40	5.47	4.04
5	T1	300	18.40	−6.64	5.17	4.36	−1.35	1.25	14.67	−2.45	4.47	9.87	4.70	2.68
	T2	300	34.50	−1.97	9.72	10.50	−3.26	2.83	18.58	1.99	5.96	12.28	6.71	3.57
	T3	111	12.00	0.99	3.27	6.32	1.91	2.20	9.42	0.06	2.23	8.50	2.78	2.36
	T4	100	18.30	−7.19	5.77	12.70	−4.81	4.64	7.50	−3.69	2.35	14.00	6.31	4.89
8	T1	129	9.30	−3.57	3.23	7.90	−3.71	2.48	4.62	−2.52	1.45	8.50	4.53	2.79
	T2	300	23.40	−5.16	5.73	14.30	−1.80	2.70	27.20	−3.20	3.89	26.50	4.59	3.83
	T3	300	31.40	−1.39	8.64	6.97	−0.31	2.07	21.08	−0.10	6.63	12.23	5.97	3.50
	T4	253	27.30	−8.03	7.98	10.70	−5.11	3.77	15.15	−4.05	3.72	17.80	6.65	5.14
11	T1	94	12.00	−4.77	4.03	3.42	−0.78	0.76	10.04	−4.71	3.46	9.96	4.88	3.38
	T2	144	31.10	18.65	12.71	15.15	7.31	5.75	14.32	3.63	5.21	19.60	8.83	6.98
	T3	300	10.40	2.20	3.57	10.34	−1.69	1.90	5.65	−1.69	1.90	9.32	3.52	3.21
	T4	300	31.30	9.30	9.76	15.76	2.10	4.95	8.76	−0.57	2.80	13.50	4.92	3.56
12	T1	185	8.60	2.77	3.16	8.51	1.10	2.66	13.97	−1.20	4.05	8.40	4.52	2.33
	T2	218	15.50	7.06	5.06	13.11	4.69	4.47	15.00	1.67	4.41	13.50	6.35	4.86
	T3	152	17.70	−8.96	6.65	7.76	−4.00	2.75	8.86	−3.03	2.69	10.70	5.24	3.53
	T4	300	22.30	−8.13	7.20	6.67	−1.35	1.58	10.23	−2.25	2.62	9.79	3.41	2.14
4	T1	95	1.50	0.08	0.36	1.79	0.33	0.45	10.20	−3.72	3.46	10.40	3.74	3.38
	T2	47	4.90	0.74	1.40	16.71	3.64	5.50	16.13	2.32	5.17	18.90	6.26	5.93
	T3	148	15.50	1.64	5.87	2.43	0.99	0.55	25.24	−1.25	6.62	16.70	5.50	3.99
	T4	99	18.10	−6.88	5.94	6.00	−1.94	1.82	14.40	−5.06	4.60	15.50	5.43	4.93
7	T1	77	18.00	2.27	5.12	2.44	−0.88	0.81	3.00	−1.10	1.02	3.20	1.59	1.07
	T2	300	48.40	−14.77	14.47	22.11	−7.34	6.98	29.07	−2.63	7.79	20.80	11.87	5.00
	T3	300	12.10	0.51	2.45	9.30	−0.34	2.48	8.27	−3.31	2.47	10.00	4.25	2.28
	T4	300	27.00	−14.18	7.57	18.20	−9.41	6.19	13.14	−3.21	3.47	17.80	10.85	5.59
10	T1	146	9.10	−4.11	2.97	8.70	−3.50	2.72	7.01	−3.96	2.18	10.60	5.38	3.34
	T2	192	18.90	−7.37	5.69	15.30	−6.06	4.91	4.40	−2.21	1.20	15.00	6.65	4.79
	T3	300	10.40	−2.91	3.03	7.57	−1.31	2.35	4.69	−0.60	1.35	5.40	2.70	1.44
	T4	300	9.00	−3.26	2.04	3.88	−0.04	1.20	5.66	−0.27	1.43	3.10	1.70	0.80
13	T1	94	18.60	−7.27	6.70	3.43	−1.50	1.23	14.61	−5.32	4.99	14.67	5.58	5.08
	T2	300	16.70	−4.72	5.07	24.46	−6.35	6.60	15.06	0.78	4.19	25.39	7.95	6.20
	T3	120	8.80	−3.37	2.78	15.80	−2.72	4.42	21.77	−0.72	5.58	21.09	5.58	5.19
	T4	284	16.20	0.21	3.99	20.75	−0.06	4.68	17.18	−1.26	3.66	17.90	4.74	3.77
14	T1	284	17.80	7.08	5.75	13.82	2.34	2.91	20.62	1.73	4.40	14.42	4.95	3.40
	T2	230	51.20	9.92	15.63	27.17	1.66	6.62	27.66	2.78	6.07	33.13	5.84	7.54
	T3	137	25.10	6.46	7.32	11.19	−0.47	3.39	13.89	4.72	4.81	13.31	4.72	4.81
	T4	300	32.40	10.28	11.63	11.76	2.44	3.65	17.35	4.63	5.25	18.21	5.88	5.80