University of
South-Eastern Norway

Exploring Skill Acquisition and Intrinsic Motivation in Head Mounted Virtual Reality Simulator Training

Candidate name: Sunniva Veie

University of South-Eastern Norway
Faculty of Technology, Natural Sciences and Maritime Sciences

MASTER THESIS

May 2018

Abstract

Simulators have long been utilized in maritime training, from relatively low-fidelity simulators of the 1950s to the full-mission simulators currently in use. Technological development has steadily increased the effectiveness of training simulators, and the last ten years focus has shifted toward implementation of head mounted display virtual reality (HMD VR) systems in training. While a number of studies comparing different display types have been initiated, limited research has investigated the effectiveness of HMD VR simulator systems in training. The purpose of this study was to contribute to this research gap by investigating the effect of HMD VR on elements of skill acquisition and intrinsic motivation.

An experimental study comparing task performance and subjective intrinsic motivation in HMD VR and desktop setup was conducted to validate the effectiveness of the system. The study consisted of two experimental groups, a HMD VR group (n=25) and a desktop group (n=25), both completing the same experimental tasks. Results indicated previous gaming experience to affect performance and intrinsic motivation to be higher for participants in the higher-level immersion HMD VR group, but contrary to hypothesis, results indicated task performance to be better in desktop than in HMD VR.

Acknowledgments

I would like to express my gratitude to my main supervisor Dr. Salman Nazir, for his constant support and guidance. He has given me the opportunity to greatly expand my knowledge and experience within TARG and has shown me respect and trust in all work we have conducted together. Additionally, I would like to thank my co-supervisor Dr. Steven Mallam for great support and motivation throughout the process, especially during data analysis. Jørgen Ernstsen, for sharing his knowledge about statistical analysis, the rest of TARG and the InnoTraining project team for allowing me to take part in an exciting project inspiring my thesis.

The participants of the experimental study enabled collection of valuable data for my analysis, and I would like to extend my appreciation for the time they dedicated contributing to my work. Finally, I would like to thank my family, especially my sister and role model, for supporting my endeavours and believing in me.

Abbreviations

Abbreviation	Explanation
HMD	Head Mounted Display
IMI	Intrinsic Motivation Inventory
InnoTraining	Innovating Maritime Training Simulators Using Virtual and Augmented Reality
KDI	Kongsberg Digital
KPI	Key Performance Indicators
SDT	Self-Determination Theory
SME	Subject-Matter Expert
TARG	Training and Assessment Research Group
USN	University of South-Eastern Norway
VR	Virtual Reality

Table of Contents

1. Introduction

1.1. Research Background

Humans today are highly dependent on technology. Nearly all daily activities are somehow aided by electronics. Technological solutions like smartphones, tablets and computers are continuously utilized in aspects like communication, monetary transactions and transportation. As technology becomes ever smarter, there seems to be no limit to how these electronic aids can be used in furthering process optimization. The combination of human dependency on technology and the urgency to acquire and utilize new technological solutions, result in a society where utilization of technology is continuously growing (Gartner, 2016).

After the so-called “information technology revolution” of the 1970-80s, technological innovation and its influence on society increased tenfold (Forester, 1990). This was the beginning of the “Information Age”, where a new social structure built on technological aids and solutions became part of social processes (M. Castells, 2009). The introduction of the World Wide Web in the early 1990s, enabled a worldwide network of connections within and across societies (History, 2010). This facilitated information-sharing across great distances, affecting further technology development and its influence on societies worldwide.

Many now state that the world has left the Information Age and ventured into the so-called “Experience Age” (Jenkins, 2017; Newman, 2016; Wadhera, 2016). An era where people are no longer satisfied with merely attaining knowledge and information. Now desiring to experience everything the world has to offer (Newman, 2016). Development of technology mirrors this change, by continuously offering new solutions designed to aid action rather than simply attainment of information (Jenkins, 2017).

New technology is introduced in most domains, resulting in changes in procedures and equipment used. The use of computers and other technical devices in learning exercises has become normal procedure for both individuals and institutions (Venkatesh, Thong, & Xu, 2012). As utilization increases, investigations into the effectiveness of such systems have been initiated. Ai-Lim Lee, Wong and Fung (2010) indicated positive effects of technological aids in education and training, revealing usefulness of such solutions.

1.1.1. Simulators and training

Simulation has been utilized in several contexts, but is stated to originate from simple game-simulations of battles utilized in development of warfighting strategies (Vincenzi, Mouloua, Wise, & Mouloua, 2008). The Hindu game “Chaturanga” was used in similar contexts, as well as the game “Go” from Eastern culture and the Western game of chess (Vincenzi et al., 2008). In addition, Allen (1987) stated that in the 17th century Prussians trained for warfare with military chess and toy soldiers.

Similar types of war games were introduced to the United States military in the 1880s through Major William R. Livermores book “American Kriegsspiel” (Vincenzi et al., 2008). In 1910 a “kiwi bird” flight simulator was used in the Wright Brothers training school, where students could practice flying in a stand-still simulator where motion was added through a motor-driven system (Vincenzi et al., 2008). Flight simulators were also utilized by the United States military during World War 2, for training aviation warfare.

In the 1930s the first digital simulator was created, and from this point on different variations and advancements of simulations has been utilized in various fields (Holst, 1982; Miles, Pop, Watt, Lawrence, & John, 2012; Rahm et al., 2016). A digital simulator can be explained as an electronic device displaying a simulated scenario, imitating features of real-life tasks and procedures (Merchant, Goetz, Cifuentes, Keeney-Kennicutt, & Davis, 2014).

Simulators are diverse in terms of fidelity, the degree to which the simulator matches its real-life counterpart scenario (Hontvedt, 2015). This may affect the user’s perception of immersion in the virtual environment. High-level immersion will provide an enhanced sense of presence for the user, likely affecting his/her experience and engagement (Alshaer, Regenbrecht, & O'Hare, 2017).

1.1.2. Maritime simulation

Training in the maritime domain consists of a mixture of theoretical and practical elements, were practical ship training are essential for achieving necessary certifications (Hontvedt, 2015; IMO, 2011). Initially, a maritime occupation could be achieved without much formal education and was therefore often initiated at a young age with practical training onboard a ship. After the introduction of the International Convention on Standards of Training, Certification and Watchkeeping amended in 1995 (STCW), ensuring a standard for maritime training and attainment of certificate of competence, training practices became more procedural and controlled (IMO, 2011).

Nowadays, maritime training is mainly conducted in classrooms and in simulated scenarios. Training in a simulated scenario allows for learning complex tasks in a virtual environment similar to reality, providing practical experiences and knowledge (Ai-Lim Lee, Wong, & Fung, 2010). Table A-II/1 and A-II/2 of the STCW code indicates requirements for nautical bachelor programs, demonstrating how simulators are used for specific learning objectives towards acquiring certificates (IMO, 2011). Castells et. al. (2015) stated how researchers have found simulators to be powerful training tools, enhancing training effectiveness and learning outcome, increasingly utilized in maritime training.

The enhanced usability has led to constantly growing simulator diversity. In relation to the STCW code, DNV GL presents four classes of maritime simulators in their Standard for Certifications No. 2.14: Class A (full mission), Class B (multi-task), Class C (limited task) and Class S (special tasks) (DNV GL, 2011). The standard from DNV GL also provides a more detailed division of maritime simulators, consisting of ten subsections based on function (DNV GL, 2011). Thus, a broad specter of simulators is utilized in different aspects of maritime training and education.

1.1.3. HMD virtual reality

Investigations into the possibility of a wearable system displaying virtual environments was initiated decades ago. The first head mounted virtual reality (HMD VR) system was developed already back in the 1960s, but the usability of this system and those in close succession were limited (Payatagool, 2008). As technology and computer power developed, increasing demands for commercial video gaming systems pushed advancements of high-fidelity HMD VR simulators (Berg & Vance, 2017).

The last ten years HMD VR simulation has gradually entered the gaming market, and many are now testing the usability of these systems in aspects of training (Buttussi & Chittaro, 2018; Kolo, 2017). In a HDM VR simulator, the virtual environment is displayed through lenses in a headset worn by the user. The user experiences the virtual scenario as though it surrounds him/her. Through tracking of head movements, the device alters the viewpoint of the user to correspond with the direction of the head (Richardson, Powers, & Bousquet, 2011). This raises perceived simulation immersion, through realistic experiences.

With the increasing usability of such systems, many are realizing the applicability of these immersive solutions in areas like education and training (Chessa, Maiello, Borsari, & Bex, 2016; Mineev, 2017; Niehorster, Li, & Lappe, 2017). Studies have revealed that high immersion VR may enhance learning outcome by increasing engagement of students in learning processes, by providing experiences more interesting and motivating (Mineev, 2017). Thus, positive effects of utilizing HMD VR in training and education have been identified, highlighting the possibility of practicing scenarios either impractically positioned in time and space, dangerous or unethical to conduct in real-life settings (Jensen & Konradsen, 2017).

1.1.4. Innovation based project inspiring master thesis

The Innovating Maritime Training Simulators using Virtual and Augmented Reality (InnoTraining) project is aiming for development of a HMD VR simulator, which will enhance training outcome and make simulator training more accessible. The project is led by the Training and Assessment Research Group (TARG), part of the Department of Maritime Operations at the University of South-Eastern Norway (USN), in cooperation with Kongsberg Digital (KDI). The author is a member of TARG and developed this study, in collaboration with the rest of the project team, as a contribution to the InnoTraining project.

InnoTraining is an undergoing project, and validation of usability of the system under development may be highly useful for further advancement. To enable investigation of the usability and effectiveness of HMD VR simulator training, an experiment comparing navigational training using two different setups was conducted. HMD VR and desktop setup were utilized. Two researchers developed and ran the experiment, the author and Anders E. Edwinson (Researcher B) (Edwinson, 2018).

1.1.5. Research purpose

The ability to acquire skills through simulator training has been investigated by many, with varying results, but little research has been conducted regarding HMD VR (Bertrand, Bhargava, Madathil, Gramopadhye, & Babu, 2017; Buttussi & Chittaro, 2018; Webster, 2015). Motivation for learning may greatly affect the training outcome, and studies have indicated that training in virtual environments can improve motivation (Ai-Lim Lee et al., 2010). However, research confirming the effect of immersion on motivation is limited.

The main objective of the author’s study was therefore to analyse whether HMD VR yielded better results than desktop, related to performance and intrinsic motivation for learning. Investigating how training setup affected different aspects of skill acquisition and motivation, like task performance, attentiveness and enjoyment. As skill acquisition may include diverse aspects depending on context, the term was in this study defined related to what was measured in the experiment. Investigating the outcome of skill acquisition, instead of the entire process (Langan-Fox, Armstrong, Salvin, & Anglim, 2002). Thus, skill acquisition in was in this context expressed by the following formula:

$$\textit{Skill acquisition} = \textit{Task performance}$$

1.2. Research Question

The following research question was developed to guide research toward the objective goals and contribute further information and data on aspects lacking examination in previous HMD VR research:

> *"What are the effects of HMD VR simulator training on task performance in skill acquisition and intrinsic motivation for learning, compared to traditional desktop simulator training?"*

Three hypotheses were created during the initial stages of the study, to aid investigation of the usability of HMD VR as a training tool (see Figure 1 for visualization):

H1: There will be a higher level of performance in navigational searching tasks in HMD VR than in desktop.

H2: The individual difference *prior gaming experience* will affect task performance of the participants.

H3: The participant's intrinsic motivation for learning will be higher in HMD VR than in desktop setup.

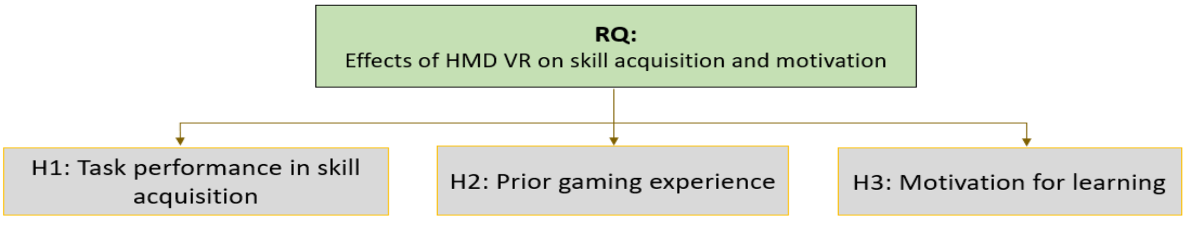

Figure 1: Visualization of the research problem, with related hypothesis

1.3. Significance to the field

Although simulators have been applied to maritime training for many years, the research on the pedagogical outcome and learning efficiency of these systems are rather limited (Sellberg, 2017). The accelerating development of VR technology increases the need for validation of usability and effectiveness. Currently, few researchers have investigated the effectiveness and validity of HMD VR simulator systems for maritime training, as no maritime HMD VR system is commercially available yet. This study will contribute information on the usability of the HMD VR systems in maritime training, which may aid validation of the system, further development and optimization.

Lacking research on training in HMD VR and validation of this system as a training instrument, results in a call for further study of these subjects. Several studies have investigated the effectiveness of other training systems like desktop, but very few look at the effectiveness of these VR systems (Ai-Lim Lee et al., 2010). Additionally, limited research has compared navigation in HDM VR and desktop (Buttussi & Chittaro, 2018). Thus, this study will contribute indication of the effectiveness of HMD VR as a training tool, increasing the knowledge within this field.

1.4. Definitions

One concept may have several different definitions. To clarify the intended meaning of the main aspects addressed in this study, Table 1 presents an overview of definitions.

Table 1: Definitions utilized in the thesis

Aspects	***Definition***	***References***
Avatar	A representation of the user in a virtual world, performing the commands of the user. Providing the user with a virtual body.	(Alshaer et al., 2017)
First- person gaming	A game setting allowing the user to perceive the virtual scenario via an avatar. Virtually the user is the avatar in the game.	(Murias, Kwok, Castillejo, Liu, & Iaria, 2016)
Games	Systems based on clear rules with variable outcomes. Users influence outcome, as it is of emotional interest. Here: video games (computer-based, utilizing video output to display a gaming scenario)	(Garris, Ahlers, & Driskell, 2016; Merchant et al., 2014)

Gamification	A broad term used to explain enhancing experience and user engagement through the use of video game elements in systems unrelated to gaming.	(Kim & Ahn, 2017)
Head mounted display	Relatively small simulators, which may be worn on the user's head, as a headset or goggles. Provides high-level immersion and user experience.	(Jensen & Konradsen, 2017)
Simulator	Here: a digital device used to display the simulation, often used as a tool in training or other demonstrations were the simulation is preferable to real-life.	(M. I. Castells et al., 2015; Merchant et al., 2014)
Simulator sickness	Here: physical discomfort experienced during or after being exposed to a virtual environment in a simulator.	(Chessa et al., 2016)
Virtual reality	A computer-generated world allowing sensory insight into the virtual environment through human-machine interaction. Offering realistic, immersive experiences engaging the user in the simulated scenario.	(Cibulka, Komulainen, Mirtaheri, Nazir, & Manca, 2016; Moore, 1995)
VR Simulators	Simulators have different levels of technological advancement, affecting virtual presentation. VR simulators provide high immersion simulation of VR.	(Mineev, 2017)

1.5. Ethical considerations

During the initial stages of the study, after the experimental design was cleared with the supervisor, a notification form for the research experiment was submitted to the Norwegian Centre for Research Data (NSD) (project number 57860). The experiment was accepted after the necessary revisions were made by the researchers.

Data was collected after informed consent forms were filled out, ensuring the complete awareness of the participants as to the experimental procedure. They were informed, both through the informed consent form and oral reminders, of the possibility of withdrawal and the guarantee of anonymization. The informed consent form stated that personal information, measurements, video- and audio recording would be securely stored at USN campus, with access granted only to listed researchers and members of TARG.

Actions were taken to avoid simulator sickness among the participants. Abiding by the ISO guidelines for cybersickness, participants could choose to end the experiment at any time if they experienced any physical discomfort (ISO, 1991a, 1991b). The exposure time in the simulators was kept to a maximum of fifteen minutes as several researchers recommends exposure time to be limited to 30 minutes (Deb, Carruth, Sween, Strawderman, & Garrison, 2017; Mania, Troscianko, Hawker, & Chalmers, 2003).

A simulator sickness questionnaire was implemented as part of the experimental procedure, but results showed generally low levels of simulator sickness experienced. This may be due to the positive effects of individual control provided in the virtual environment, which may limit unease and discomfort of the individual (Howard, 2017). The navigation method utilized in HMD VR was selected after a pre-testing conducted by the researchers revealing the method to provide high personal control and comfort, and low chance of inducing simulator sickness.

2. Literature Review

A comprehensive literature review was conducted to analyse previous research of aspects related to the research topic of this study. Four main topics were investigated to develop a fundament for the study of the effect of HMD VR simulator training on skill acquisition and motivation (see Figure 2). Several keywords utilized in searching for literature were identified: *Virtual reality simulator, head mounted display, skill acquisition, performance, motivation, training and education* etc. The keywords were used in different research engines, including Scopus, Oria and Google Scholar. The literature was collected continuously throughout the study process, to identify new developments and findings.

Figure 2 demonstrates the main topics of the literature review and relation to the hypotheses.

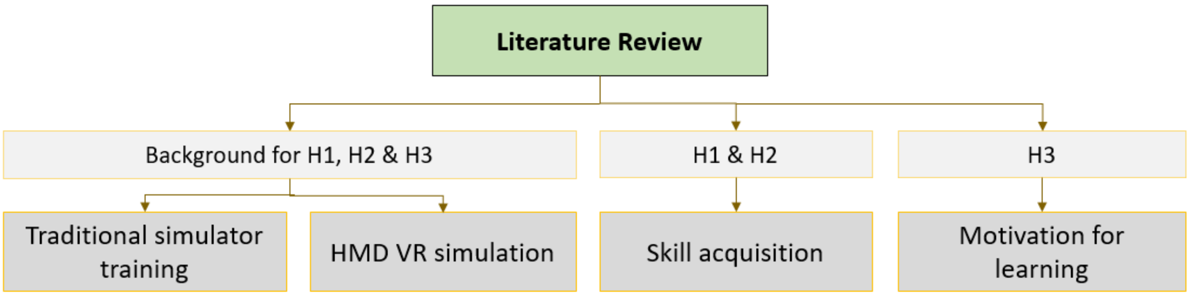

Figure 2: Literature review divisions, with related hypotheses

The first section (2.1.) of the review examines previous research connected to simulators utilized in training in general, and simulators currently utilized in maritime training. Section 2.2. addresses commercially available HMD VR systems, and their usability in different domains. In addition, this section discusses future prospects of such systems in maritime training. In section 2.3. studies of skill acquisition, learning and performance will be reviewed. Here studies of skill acquisition in maritime simulation, and the effects of immersion on performance are discussed. The final section (2.4.) investigates aspects of motivation for learning, with special focus on intrinsic motivation. Furthermore, it addresses how enhanced immersion may affect the motivation of trainees, by increasing user engagement.

2.1. Simulator Training

Digital simulators provide a virtual environment where users can repeatedly interact with a simulated scenario, without the emergency of failure or the cost of extensive and distant environments (Jensen & Konradsen, 2017). Merchant et. al. (2014) described how digital simulators have been utilized in training and education since their introduction in the 1980s. The early developments of training simulators were technologically limited with low learning output, but the expanding processing power of computer systems gradually increased the usability of simulators in higher education (Merchant et al., 2014).

The effectiveness of training simulators has previously been investigated within a wide range of domains and industries. Cibulka et. al. (2016) investigated simulators used in the process industry, providing an insight into the output of training and the cost of acquiring such systems. Additional research has been performed on training simulators in the medical domain (Barnes, Burns, Nesbitt, Hawkins, & Horgan, 2015; Lee & Lee, 2018), and simulator use in military training (Soares, Corazza, Piovesan, Azevedo, & Vasconcellos, 2016; Webster, 2015). Studies of simulator use in other fields, like safety training (Buttussi & Chittaro, 2018; Li, Liang, Quigley, Zhao, & Yu, 2017; Nazir, Sorensen, Øvergård, & Manca, 2015) and education (Ai-Lim Lee et al., 2010; Merchant et al., 2014), has investigated the potential of utilizing simulators in training every-day activities and procedures.

2.1.1. Simulators in maritime training

Non-digital training simulators appeared in the maritime domain in the 1950s (Sellberg, 2017). The price of the early maritime simulators was considerable. So expensive it was cheaper to hire and maintain a training vessel, even for longer periods, than acquiring these simulators (Barsan, 2009; M. I. Castells et al., 2015). Therefore, simulators were rarely implemented in maritime training. However, as more advanced and affordable solutions entered the market and the introduction of IMOs international regulations for personnel standards in the 1990s changed training demands, implementation of simulators in maritime training increased (Barsan, 2009).

An explanation for this technological acceptance in maritime training may be the realization of simulator effectiveness in providing individual experience of tasks related to future work responsibilities (Vincenzi et al., 2008). Castells et. al. (2015) described simulation as a powerful training tool, imitating real-life practices of any shipping system and providing students with the opportunity to achieve more effective training outcomes. In simulators, the trainee is exposed to scenarios with diverse complexities, depending on simulator fidelity (Basak, Unver, Moss, Watts, & Gaioso, 2016). Additionally, benefits of these training systems compared to other training methods like onboard- or classroom training have been identified (Tvedt, Oltedal, Batalden, & Oliveira, 2018).

Even though the costs of simulators have been reduced and the usability increased, there is constant room for improvement (Tvedt et al., 2018). Berg and Vance (2017) stated that the goal of virtual reality is to provide an experience of physically being inside the virtual world. The last ten years much focus has been on enhancing virtual reality technology, creating a commercially accessible tool for high-fidelity experience of virtual environments. With this focus, new simulated solutions are regularly being introduced to increase the virtual experience (Chessa et al., 2016; Niehorster et al., 2017).

Table 2 presents a selective overview of articles identified related to previous research conducted related to traditional simulator training.

Table 2: Selected studies investigating the utilization of simulators in training

Subject	Reference	Scope
History of simulator use in training	(Merchant et al., 2014)	Investigating effectiveness of VR-based instruction on students' learning outcomes in higher education
Simulators in training	(Nazir et al., 2015)	Analysing effect of training methods on industrial operators' situation awareness
	(Barnes et al., 2015)	Looking at the effectiveness of home surgical VR simulation training
	(Soares et al., 2016)	Validating VR simulation for military reaction time evaluation
	(Cibulka et al., 2016)	Investigating VR training simulators in the process industry
Simulators in maritime domain	(Hontvedt, 2015)	Examining performance of work tasks in a full-mission ship simulator
	(M. I. Castells et al., 2015)	Improving deck officers' competences using simulators
	(Sellberg, 2017)	Review of simulators used in bridge operations training and assessment
	(Tvedt et al., 2018)	Investigating simulated way-finding training for maritime vessels

2.2. Immersive HMD VR Simulation

In 1960 Morton Heilig created the first HMD VR simulator, which did not provide motion tracking but simply a non-interactive head mounted film display (Burdea & Coiffet, 2003; Heiling, 1960). One year later the first motion-tracking HMD VR simulator was created, but it was not until 1968 that such a system was connected to a computer providing simple computer-generated simulations (Goran, 2016; Society, 2017). Ivan Sutherland (1965) published an essay called "The Ultimate Display'', where he presented what is considered by many to be the original vision of immersive virtual reality simulation (Berg & Vance, 2017). Sutherland described the ultimate VR display as a "room" where everything, the very existence of matter, is computer-controlled. Much remains for undetectable simulation to become reality, but HMD VR is a step in the right direction.

When the technology company Oculus Rift introduced their consumer-priced HMD VR system in 2013, a new generation of head mounted simulator technology introduced solutions to the prior issue of simulator quality and availability (Jensen & Konradsen, 2017). Before this, technological limitations restricted the applicability of such systems in domains like training and education (Vincenzi et al., 2008). With Oculus' head mounted commercial system entering the market, the quality of further developments of this type of simulators grew substantially (Jensen & Konradsen, 2017).

2.2.1. Immersive properties of HMD VR

The famous magician and master of illusion Harry Houdini once stated: ''What the eyes see and the ears hear, the mind believes.'' (Berg & Vance, 2017, p. 2)

Previous studies comparing different training platforms and simulator fidelity have indicated that HMD VR enhances user performance and learning outcome through higher level immersion (Bertrand et al., 2017; Buttussi & Chittaro, 2018; Janßen, Tummel, Richert, & Isenhardt, 2016). Immersion was defined by Murray (1997) as a state where one is isolated from the real world through total consumption in an alternative reality (Janßen et al., 2016). In a training scenario immersion may provide an improved overview of situations, by expanding the field of view of the user in the virtual environment and enhancing sensory understanding of the scenario. The user may perceive what he/she sees as reality if the level of immersion is high.

HMD VR systems provide head- and body tracking, enhancing the user's experience of immersion (Jensen & Konradsen, 2017). Head tracking presents vestibular information to the user through head rotations, which Richardson et. al. (2011) found enabled spatial orientation in the virtual scenario through coordination of movement and balance.

2.2.2. HMD VR implementation

During the 1990s HMD VR simulator development shot speed after Virtuality Group presented several arcade machines and headsets (Goran, 2016; Society, 2017). Hence, these simulator devices were initially developed for and introduced to the gaming industry, but were gradually implemented into other domains as usability increased (Berg & Vance, 2017).

Studies based on different experiments with simulators have presented HMD VR simulators as the next generation training devices (Bertrand et al., 2017; Jensen & Konradsen, 2017; Li et al., 2017). As these simulators enhance the accessibility of physical practice, they may prove highly useful in learning activities reliant on practical experience for optimal training outcome (Alshaer et al., 2017; Marks, Estevez, & Connor, 2014). These types of simulators will be accessible to universities as the procurement costs are relatively low and the ease of use in different scenarios are high (Janßen et al., 2016). If the system in addition provides realistic simulations, the effect of training should be equal to that of traditional real-life training (Mania et al., 2003).

The HMD VR systems commercially available today have been tested and used in training research, with varying results (Buttussi & Chittaro, 2018; Li et al., 2017; Mallaro, Rahimian, O'Neal, Plumert, & Kearney, 2017). These systems are currently being implemented in mathematics training in Norwegian primary schools, with promising results (Jørgenrud, 2018; Vissgren, 2017). Additionally, examples of using HDM VR systems in enterprise and retail have been identified. In USA Walmart is implementing HMD VR simulators in all of their 200 training centers, for training employees in management, customer service and similar tasks (Kolo, 2017). Denis Dorozhkin et. al. (2017) investigated the usability of these commercial systems in operating room fire safety training, with promising results, but stated that such solutions have yet to be implemented in educational medical training.

Although the application of commercial VR systems in training has been investigated to some extent, fully virtual HMD-based training modules are essentially non-existent. In domains like aerospace, automotive and military, HMD VR have been utilized as supplements to other simulators in several training facilities (Berg & Vance, 2017). Thus, HMD VR is so far mostly used to extend the field of view of other simulators, as stand-alone HMD VR systems for training and education have yet to be developed.

Similar obstacles exist in the maritime domain where traditional simulators have been utilized in training for many years, but the development of stand-alone HMD VR systems for maritime training is not yet completed. Increasing amounts of research is currently being conducted toward the utilization of HMD VR in maritime training, but little proof of its usability is yet to be identified. Speculators in maritime technology are expecting maritime HMD VR training solutions to arrive sometime during 2018 (Wingrove, 2018). As these solutions have not appeared thus far, researchers investigating HMD VR application in maritime training may be limited to analysing commercially available gaming solutions.

Table 3 presents an overview of selected research on immersive HMD VR simulators is.

Table 1: Selected studies contributing information on HMD VR simulator training

Subject	**Reference**	**Scope**
Comparison of HMD VR and other training platforms	(Buttussi & Chittaro, 2018)	Investigating different effects of virtual display on presence and learning
	(Bertrand et al., 2017)	Exploring effects of presentation method and simulation fidelity
	(Mallaro et al., 2017)	Comparing HMD VR and large-screen displays for an interactive simulator
Properties of HMD VR	(Jensen & Konradsen, 2017)	Reviewing the use of HMD VR in education and training
	(Richardson et al., 2011)	Investigating whether gaming experience affected performance in HMD VR
HMD VR training	(Alhalabi, 2016)	Virtual reality systems effect on students' achievements in engineering education
	(Berg & Vance, 2017)	Analysing the use of virtual reality in product design and manufacturing
	(Li et al., 2017)	Testing simulated earthquake safety training

2.3. Skill Acquisition

Skill acquisition is defined in several ways, depending on context. The term is often explained by a division into three main categories based on Ackerman (1988) theory of skill acquisition: Cognitive skill acquisition, psychomotor/autonomous skill acquisition, affective skill acquisition (Jensen & Konradsen, 2017; Shahriari-Rad, Cox, & Woolford, 2017).

Cognitive skill acquisition is related to the ability to understand and handle intellectual tasks through the utilization of individual knowledge and information (VanLehn, 1996). *Psychomotor/ autonomous* skill acquisition, however, concerns the ability to learn and manage the bodily manoeuvring needed to perform a task (Jensen & Konradsen, 2017). *Affective/associative* skill acquisition is related to the ability to control or manipulate own attitude and emotions in different situations and tasks, to positively influence choices of action and obtain the desired outcome (Jensen & Konradsen, 2017; Kraiger, Ford, & Salas, 1993).

Additionally, skill acquisition is often related to the process of learning, which entails gathering information on an aspect and utilizing this information to execute actions or processes (Garris et al., 2016). The similarity of these two terms is prominent in research, where learning also has been stated to result in three main outcomes: cognitive, skill-based and affective (Kraiger et al., 1993).

2.3.1. Performance

Langan-Fox, Armtrong, Salvin & Anglim (2002) researched similar aspects in their study of processes in skill acquisition. Like Ackerman, these researchers presented a three-stage process of the term (Langan-Fox et al., 2002). However, they investigated the inclusion of performance as an important factor in their three stages of skill acquisition. In analysing how individual abilities, like cognitive ability, affected the performance throughout the learning process, the researchers identified performance as an outcome of the skills acquisition process. Along this process performance became more automatic, less cognitive demanding and of higher quality.

Even though the stages of Langan-Fox et. al. (2002) are slightly different from the categories of Ackerman (1988), these performance-based stages have definite similarities to his skill acquisition categories:

In the first stage, like in the cognitive phase, the individual is unfamiliar with the task and dependent on declarative knowledge. Performance at this stage is therefore highly cognitively demanding, and controlled information processing is required to complete tasks. The second stage is less reliant on cognitive processes, as the individual is more familiar with the task.

Thus, performance is more autonomous and the individual can focus on achieving good results. This stage is loosely tied to the affective/associative category, where other aspects than only the task at hand are the focal point if the individual. In the third and last stage, the individual is no longer consciously controlling actions. The task is completed autonomously, like procedure. Thus, performance is less demanding over time as skills are acquired and become automated. The desired end-result of this process is stated by Langan-Fox et. al. (2002) to be skilled performance.

2.3.2. Goals and previous gaming experience

Several aspects, other than cognitive abilities, have been theorized to influence performance. For instance, Langan-Fox et. al. (2002) found a correlation between task-specific goals and performance level. Wood, Mento and Locke (1982) stated that years of research constitute how level of performance is positively influenced by the implementation of task-specific and challenging goals. The goal should adequately challenge the individual's abilities and raise goal commitment, to increase performance by motivating the individual to succeed (Langan-Fox et al., 2002).

Another aspect indicated to have a great effect on performance in a virtual scenario, is previous gaming experience (Murias et al., 2016). The effect of individual differences on skill acquisition in virtual environments was investigated by Münzer and Zadeh (2016) identifying that a person familiar with video gaming systems and manoeuvring in virtual worlds had fewer problems adapting to the simulator equipment and scenario. Research has also revealed that perceptual and cognitive abilities related to learning are correlated with previous experience with video games (Richardson et al., 2011). Additionally, Fabroyir & Teng (2018) found that video gaming experience more specifically influenced virtual navigation behaviours, and Tvedt et. al. (2018) stated such individual differences to be even more influential to training outcome than treatment. This may indicate that prior gaming experience may be a highly influential factor of VR training outcome.

2.3.3. Measuring performance

There are several methods for measuring performance, highly dependent on the task performed and the desired outcome of the measurement. Lundwall, Sgro and Fanger (2018) measured reaction time by counting number of correct interpretations of visual cues received, and Huber (1985) measured number of moves require to exit a maze (score). Grubel et. al.

(2017) measured task performance in VR navigation by measuring the accuracy of angle and distance from goal and time to task completion.

Additional examples of methods utilized in performance analysis include measurement of quality or outcome based on subjective measures and outcome data such as mortality rate (Savjani, Haseeb, & Reay, 2018), measuring speed and task completion time (Bailey, Konstan, & Carlis, 2000), and investigating aspects of memory through number of correct answers (Cummings & Quimby, 2018). Thus, performance is often measured dependent on different methods of data collection, and are generally investigated through measurements related to the objective of the analysis.

2.3.4. Effects of immersion

Researchers investigating different aspects of virtual reality, seem to be labouring under the same idea of immersion positively affecting learning outcomes (Jensen & Konradsen, 2017). Researchers have analysed this effect, in different level immersion simulators, on aspects influencing performance. Reiners et al. (2014) found that high-level immersion resulted in more attentive and goal-oriented individuals, concentrating on tasks provided. Loup et. al. (2016) identified those using HMD in learning to be more engaged, and Alhalabi (2016) found that participants voluntarily extended the duration of training sessions when learning was conducted in high immersion systems.

In a study comparing HMD VR and classroom lectures, Rasheed, Onkar, and Narula (2015) found that the high immersion simulator alternative provided better results for spatial awareness. Ray and Deb (2016) revealed that individuals receiving instructions via HMD VR over a period, performed better than those receiving the same instructions in classrooms. Studies have also investigated the effect of immersion and visual complexity in relation to scanning task, identifying that high immersion provided better results when training for target identification (Ragan et al., 2015)

Table 4 presents a selective overview of literature identified related to skill acquisition and performance.

Table 2: Selected studies contributing knowledge on elements of skill acquisition and performance in virtual reality

Subject	**Reference**	**Scope**
Skill acquisition & Learning	(Ackerman, 1988)	Investigating individual differences during skill acquisition
	(Jensen & Konradsen, 2017)	Reviewing the use of HMD VR in education and training
	(Kraiger et al., 1993)	Researching cognitive, skill-based, and affective theories of learning outcomes
	(Garris et al., 2016)	Reviewing games, motivation, and learning and connections between them
Performance	(Langan-Fox et al., 2002)	Investigating performance as an outcome of the skill acquisition process
Previous gaming experience	(Murias et al., 2016)	Looking at the effects of video game use on performance in VR navigation
	(Fabroyir & Teng, 2018)	Analysing navigation in virtual environments using HMD
Effects of immersion	(Reiners, Wood, & Gregory, 2014)	An experimental study on immersion in VR for education and training
	(Ragan et al., 2015)	Investigating effects of immersion on VR training effectiveness
	(Loup, Serna, Iksal, & George, 2016)	Looking at effect of immersion and persistence on engagement in learning

2.4. Motivation for learning

Another aspect given much attention throughout the years in research of education and learning effectiveness is motivation (Hanafi, Said, Wahab, & Samsuddin, 2017; Hidi & Harackiewicz, 2000). Many have attempted to define the concept and discuss its influence on performance. Guay et al. (2010) defined motivation as the reasons behind all behaviour. Pintrich & Schrauben (1992) and Wolters (1998) explained motivation as an individual’s decision to engage in an activity, and persist toward accomplishing good results in set activity (Garris et al., 2016). The descriptions may alter, but most researchers present motivation similar to Gredler, Broussard and Garrisons (2004) definition of motivation as a human trait that initiates or repress actions (Lai, 2011).

The concept of motivation is often divided into two main sections: Intrinsic motivation and extrinsic motivation (see a graphic explanation of motivation in Figure 3)

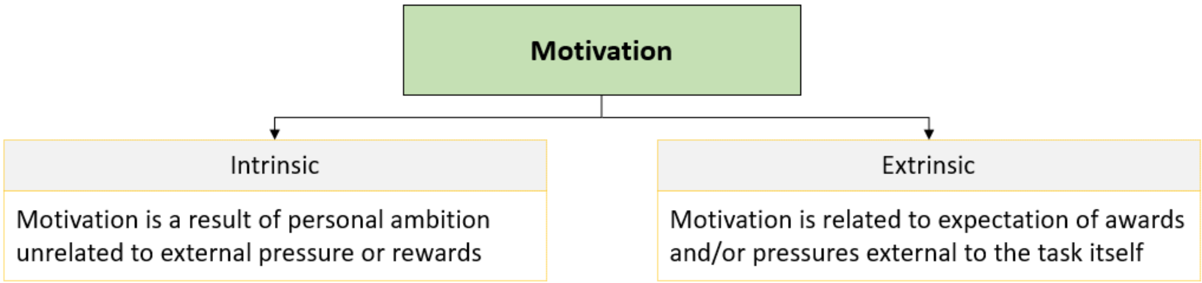

Figure 3: Demonstration subdivision of motivation, based on the theory of (Deci & Ryan, 1985)

Intrinsic motivation is often related to the intrinsic value perceived in conducting a task, as it is led by personal interest, enjoyment and pleasure (Guay, 2010). In contrast, extrinsic motivation is conceived from rewards or other external consequences of performance (Guay, 2010). As this study mainly focuses on the measurement and analysis of intrinsic motivation in HMD simulation training, extrinsic motivation will not be awarded extensive attention in this literature review.

Research has indicated intrinsic motivation to be more desirable in education than extrinsic, as it has been suggested to result in better learning outcomes (Lai, 2011). This may be because intrinsically motivated individuals are motivated to learn and master new skills because personal of interest or engagement, which may lead to an extension of motivation over time and is therefore more desirable than those motivated by external and more short-term factors like rewards (Sansone, Harackiewicz, & Sansone, 2000). Garris et. al. (2016) found that individuals with high-level motivation will be eager to learn, focused on the task and interested in performing well. They also identified motivated learners to be more self-determined, persistent in their actions and driven towards success.

2.4.1. Self-determination theory

In their study of intrinsic motivation in human behaviour Deci and Ryan (1985) investigated aspects of motivation in relation to the self-determination theory. This theory identifies three central psychological needs essential to the attainment of intrinsic motivation: Competence, autonomy, and relatedness. Looking at these three needs in an educational context, a student will obtain perceived *competence* when given the opportunity to succeed, *autonomy* if the success is achieved by individual endeavours, and *relatedness* when barriers between themselves and authoritarian figures are removed (Monteiro, Mata, & Peixoto, 2015; Ryan & Deci, 2000).

When these three psychological needs are satisfied, a high level of self-motivation and comfort will manifest in the student, enhancing task performance (Monteiro et al., 2015; Ryan & Deci, 2000). Thus, according to the self-determination theory, the level of intrinsic motivation should increase as a result of the individual's experience of achievement and self-control of outcome (Eccles & Wigfield, 2002).

2.4.2. Measuring motivation

As motivation is a concept concerning a diverse array of subjective aspects, there have been identified challenges in measuring motivation (Lai, 2011). Much of motivational research is related to children, focusing on aspects which could also be relative to adults but making the measurements more understandable for children. Turner (1995) investigated assessment of motivation, focusing on children's motivation for learning. He confirmed that most aspects of cognitive motivation, like achievement and interest, is not clearly measurable because of the difficulty of direct observation (Turner, 1995). Additionally, he identified that most measurements of motivation were conducted through self-report measures, which are often individually biased and result in generalized output rather than task specific.

The difficulties of measuring motivation are prominent, but research have resulted in several measurement instruments. Most of these instruments are created for measuring motivation among children, but some for adults as well. In studies of children's motivation, instruments like the Children's Academic Intrinsic Motivation Inventory created by Gottfried (1986), and the Instrumental Competence Scale for Children by Lange & MacKinnon (1987) have been published and accredited (Lai, 2011). An instrument much used in motivational research among older individuals, like students, is the Intrinsic Motivation Inventory (IMI) questionnaire (Ocampo et al., 2017). This questionnaire is developed from the beforementioned Self-Determination Theory and measures intrinsic motivation through seven subcategories (Monteiro et al., 2015).

The IMI has previously been utilized in several studies of intrinsic motivation in different areas like sports (Fonseca & Brito, 2001; Gutiérrez, Ruiz, & López, 2010; McAuley, Duncan, & Tammen, 1989), reading (Grolnick & Ryan, 1987), computer-related activities (Deci, Eghrari, Patrick, & Leone, 1994; Hanafi et al., 2017), recovery training (Levac, Driscoll, Galvez, Mercado, & O'Neil, 2017), and training and education (Filak & Sheldon, 2010; Monteiro et al., 2015; Ocampo et al., 2017; Wang, 2017). In addition, the usability of the instrument has been examined, resulting in indications of high validity (Leng, Ali, Baki, & Mahmud, 2010; McAuley et al., 1989)

2.4.3. Effects of immersion

Several aspects can increase or reduce motivation for learning, and immersion may be especially influential on motivation in simulation training. Kanfer (1991) and Tannenbaum and Yukl (1992) stated in their studies that the level of effort, intensity and persistence exerted by an individual will affect motivation for training. Thus, motivation may be influenced by level of individual engagement. The level of immersion of a training simulator may affect the motivation of the user by providing scenarios relatable to reality, thus increasing training effectiveness (Jensen & Konradsen, 2017). It may also encourage more active and self-controlled student learning, by providing the student with an interactive training platform (Ai-Lim Lee et al., 2010; Janßen et al., 2016).

In addition to affecting personal engagement, perceived immersion may influence individual comfort and eagerness to utilize a system. These aspects were studied by Howard (2017), investigating the application of the Uncanny Valley Theory (UVT) in simulator training. The UVT indicate that personal comfort in an experience is related to level of perceived reality. The theory suggests that highly realistic or non-realistic experiences are perceived as pleasant, while slightly realistic experiences may cause unease (Howard, 2017). The researcher found the theory to be applicable to simulators, and that a simulator with either high- or low-level realism would yield optimal results. Thus, indicating that level of realism influences perceived immersion and comfort of the user, affecting learning and motivation.

A feature often incorporated into immersive simulator scenarios used in training is gamification. By incorporating elements and strategies from video games, the training instructions are perceived as more interesting and the scenario more engaging (Brazil et al., 2018; Kim & Ahn, 2017). Garris et. al. (2016) identified several studies which identified the incorporation of gaming elements as a driver of motivated individuals and improved learning outcome.

An overview of selected articles identified related to motivation for learning are presented in Table 5.

Table 3: Selected studies contributing knowledge on motivation, focusing on intrinsic motivation

Subject	Reference	Scope
Motivation	(Hidi & Harackiewicz, 2000)	Looking at methods for motivating the academically unmotivated
	(Lai, 2011)	Reviewing of studies of motivation
	(Garris et al., 2016)	Reviewing games, motivation, and learning and connections between them
Intrinsic motivation	(Guay, 2010)	Investigating intrinsic motivation for elementary school children
	(Ryan & Deci, 2000)	Analysing among other SDT and the facilitation of intrinsic motivation
SDT & IMI	(Deci & Ryan, 1985)	Looking at intrinsic motivation and self-determination in human behaviour
	(Ryan & Deci, 2000)	Investigating the SDT and facilitation of elements of intrinsic motivation
	(Monteiro et al., 2015)	Using IMI to investigate motivation in first language and mathematics learning
Immersion and motivation	(Hanafi et al., 2017)	Trying to improve students' motivation using virtual learning environments
	(Jensen & Konradsen, 2017)	Investigating the use of HMD VR in education and training
	(Wang, 2017)	Testing the effect of VR on learning motivation and academic performance

2.5. Summary

Simulators have been utilized in training and education for many years, with increasing effectiveness. Studies have investigated simulator application in training, revealing usability in many domains. The initial simulators for maritime training had limited effectiveness, but as technology enhanced with new simulator solutions they were increasingly utilized in maritime training and education. Limited research has been conducted on the usefulness of simulators in maritime training, but the investigations conducted have mainly yielded positive results.

Commercially available HMD VR simulators are with increasing frequency being utilized in training, but to this day limited systems specifically designed for training have been revealed. Resulting from the lack of HMD VR systems developed specifically for training and education, maritime training included, limited research has been conducted on the effectiveness of these systems. Studies have indicated a positive effect of immersion on learning outcome, revealing the affordances of such systems.

Several factors may influence task performance in a skill acquisition process. Clear, specific goals and cognitive ability may impact learning outcome. In a virtual context, previous gaming experience have been indicated to affect performance because of system familiarity. The effect of immersion on performance have been investigated, most results indicating a positive correlation between level of immersion and performance.

Intrinsic motivation is highly individual, affected by among other personal interest and engagement. Even though much research has been conducted on motivation, there still exists ambiguity as to how best to measure it. Several methods were identified, most consisting of subjective measurements. Although the research on the effect of immersion on motivation is rather limited, existing studies indicate immersion to positively affect intrinsic motivation. The concept should therefore be further investigated to reveal the effect of high-level immersion simulation on motivation for training.

3. Research Methodology

3.1. Background

This study was initiated and developed as a result of the author's participation in the Innovating Maritime Training Simulators using Virtual and Augmented Reality (InnoTraining) project. The goal of this project, established in 2017 and awarded 13 million NOK by the Research Council of Norway, is to develop a HMD VR simulator with high usability and range regarding training scenarios and assessment.

Two researchers, the author and Researcher B, created and conducted an experimental study in a virtual environment. Both researchers collected data throughout the experiment but focusing on a different aspect of the experiment, contributed data to two master theses. The author's thesis focusing on comparing aspects of HDM VR training to desktop training, while the thesis of Researcher B concentrated on assessment of quality of experience (Edwinson, 2018). Thus, the experiment was conducted as a cooperation, resulting in two different master's theses with two sets of research questions and scope.

Figure 4 presents a timeline of the thesis project, from main development initiation.

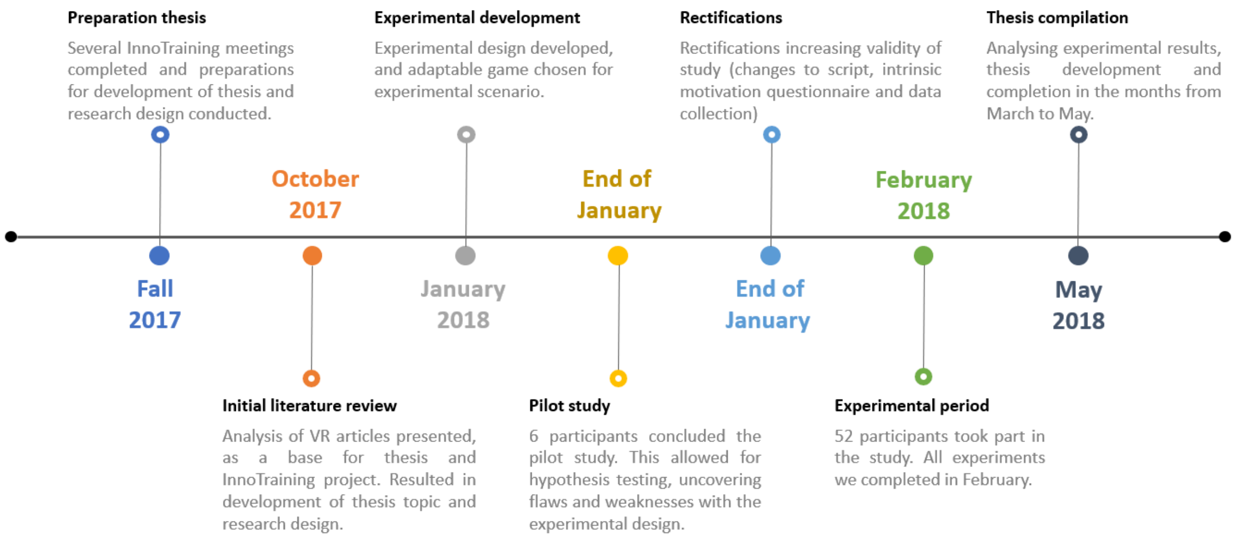

Figure 4: Timeline of the thesis project. Includes main events like pilot study (Mallam et al., 2018) experimental period and completion.

3.2. Experimental Design

A mixed method focusing on quantitative research was developed to investigate the effect on HMD VR on skill acquisition and motivation for learning. This research method is characterised as a mix between qualitative and quantitative, because qualitative interviews were included as an additional source of data complementing the quantitative measurements (Small, 2011). An experimental between-subjects factorial design was selected to obtain data on the experience of users (Psychology World, 1998a). Participants were divided into two groups, a control group conducting the experiment on a desktop setup and an experimental group completing it in a HMD VR setup. This design investigates the results from the two groups, allowing comparisons between them (see Figure 5 for visualization).

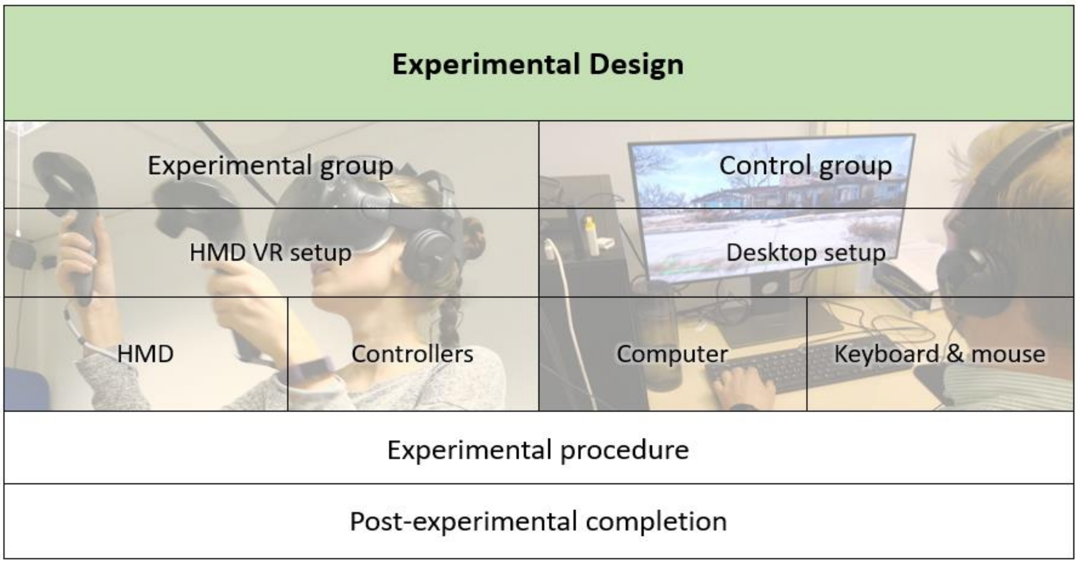

Figure 5: Between-subjects design of the experimental study

In this study, two independent variables were investigated: Training condition (HMD VR and desktop), and gaming experience (high, medium and low). The dependent variables were skill acquisition (task performance) and intrinsic motivation, investigating how training method influenced performance and intrinsic motivation. See Figure 6 for a visual explanation of variables and relationships.

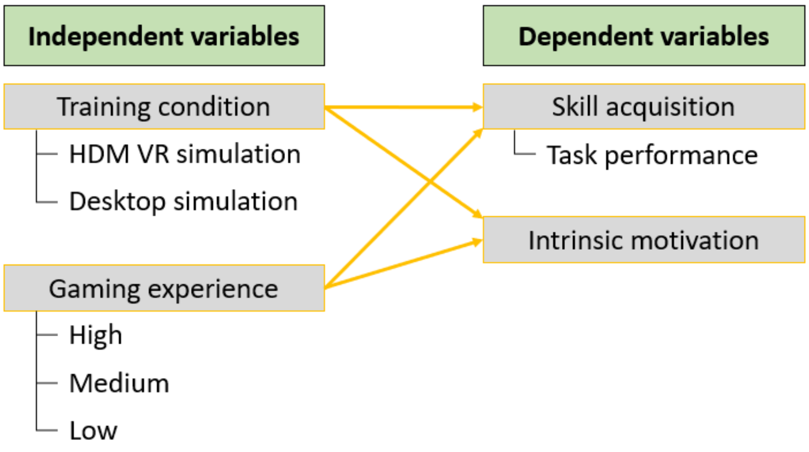

Figure 6: Variables investigated in the experimental study

The independent variables were investigated through the following intervention: Navigation and localization tasks in a small maze (more information on the intervention in section 3.6.). The dependent variables, performance and intrinsic motivation, were measured through both objective and subjective measurements defined by the researcher. Data on the KPIs, time to completion and number of figurines collected, were registered enabling assessment of task performance. The KPI measurements also provided information of improvement in task performance throughout the experiment, indicating level of skill acquisition (Langan-Fox et al., 2002). The dependent variable intrinsic motivation was measured by individual subjective perception of the experience. These subjective measurements were conducted through an intrinsic motivation questionnaire based on the Intrinsic Motivation Inventory (IMI), and an exit interview probing participant's experience (Monteiro et al., 2015).

3.2.1. Design of exercise

Familiarization exercises were included to provide practice on using the controllers and system. The tasks in the familiarization ranged from basic tasks like navigation across a street and inside houses, target practice (shooting tin cans), and localisation and collection of items. Studies have shown that performance is enhanced by preparing for a given task with similar exercises (Elliott, Hayes, & Bennett, 2012). Thus, by enabling training similar skills as needed in the main task the participants were better equipped to perform well in the main task.

Figure 7 indicates the main areas and tasks in the familiarization.

Figure 7: Main areas of tasks in the familiarization. Numbering indicates the order of tasks (1=Beginning to 6= Last part of scripted familiarization)

The first image depicts the starting point of the familiarization, where the participants entered the virtual scenario. The second and third picture visualize the first task of shooting down tin cans from shelves. This task was included to accustom the participants to the controls. Picture four shows a room where several figurines, later used in the main task, were positioned. The participants were taught how to find and collect these three figurines by using the controllers. The two last pictures, five and six, depicts the last part of the scripted familiarization (see Appendix 8.2. for the experimental script).

The main task was designed to test navigation and localisation skills of the participants, by placing them in a maze scenario where the goal was to find and collect as many hidden figurines (see figurine to the right) as possible within five minutes. The maze was small, but by utilizing similar interior throughout the maze the challenge of navigation was enhanced. Ten figurines were hidden at random locations in the maze, eight inside cupboards or cabinets and two out in the open positioned in shelves (see Figure 8).

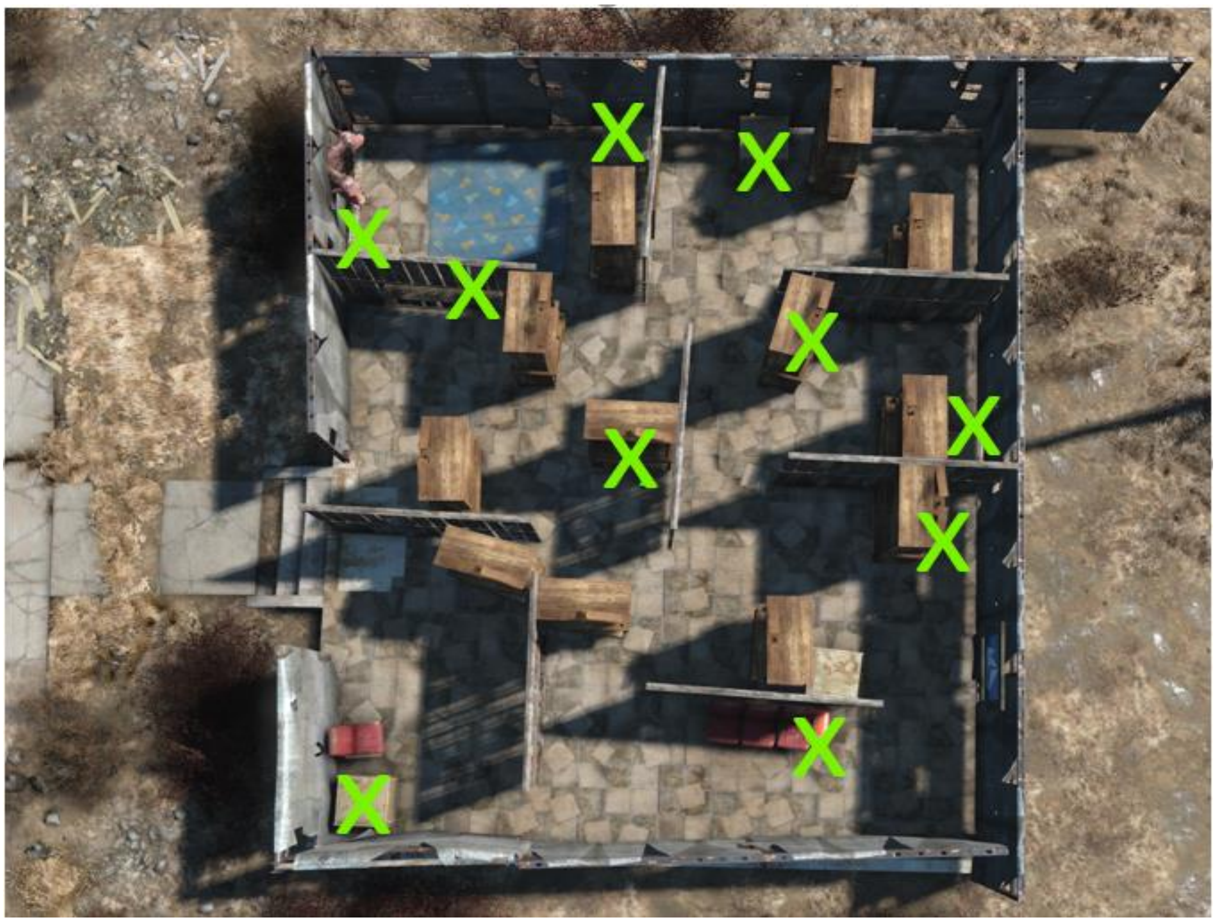

Figure 8: The maze seen from above. The markings (x) in the picture indicates the positions of the figurines in the maze.

To minimize the chance of participants experiencing simulator sickness as a result of moving around, a teleportation method for manoeuvring in the virtual world was selected. This could have been performed by using the touch-pads on the controllers, but a few in-house tests with this method uncovered high possibility of inducing simulator sickness. By utilizing the teleportation method for moving around, the chance of inducing simulator sickness was minimized (Meng-Lin & Voicu, 2017). The HMD VR also provided the use of intuitive interfaces of head- and hand movement, which have been proven to have a positive effect on the precision of the movements and higher comfortability (Bowman, Johnson, & Hodges, 2001).

3.3. Setting

This study took place in the TARG lab at the University College of South-Eastern Norway, located in Horten, Norway. This lab was set up by the TARG team in association with the InnoTraining project and tailored to the needs of the study by the researchers. See Figure 9 for a picture of the lab.

Figure 9: Setting of the experiment in the TARG lab

3.4. Participants

52 participants volunteered as participants for this study. Two participants from the control group were unable to complete the experiment due to simulator sickness, resulting in a total of 50 participants. These were divided into either the experimental group (EG, n=25) or the control group (CG, n=25) upon arrival, to avoid bias in group assignment. To ensure even distribution in the two groups, the participant following those unable to complete were placed in the same group. See Table 6 for participant data.

Table 4: Participants demographics (percentage of total), including demographics of the two experimental groups

Characteristics	**Participants**	***Experimental group***	***Control group***
N	(50)	*(25)*	*(25)*
Participant age (years)			
Mean	29,16	*29.48*	*28.84*
SD	10,81	*11.16*	*10.66*
Range	20-69	*20-69*	*21-61*
Participant gender			
Male	35 (70%)	*18 (72%)*	*17 (68%)*
Female	15 (30%)	*7 (28%)*	*8 (32%)*
Participant gaming experience *			
High	20 (40%)	*9 (36%)*	*11 (44%)*
Medium	14 (28%)	*7 (28%)*	*6 (24%)*
Low	16 (32%)	*9 (36%)*	*8 (32%)*

Table 7 explains the three levels of previous gaming experience. These three levels were defined, similarly as Murias et. al. (2016) defined previous video game usage in their study, by measuring participants previous experience with video gaming on a 7-point Likert scale (1=no experience to 7=playing several times a week).

*Table 7: * Level of gaming experience explained*

Level of gaming experience	**Explanation**
High	Playing once a week or more
Medium	Playing once a month or once every other month
Low	Playing never

The sample selection was conducted based on a convenience sample, as participants were selected dependent on availability in the experimental period (see timetable in Figure 4 for the experimental period) and the majority of the sample was students and faculty at USN (Bui, 2014). The inclusion criteria for the study were: Age above 18 and no previous epileptic seizures. Additionally, the snowball sampling method was utilized by prompting all participants to promote the experiment to friends and colleagues, to enable information of the experiment to expand at high speed throughout the population (Everitt, 2002). This method is often used to obtain access to unattainable samples of populations, enabling connections with new participants through already established connections (Cohen & Arieli, 2011). In this study, the snowball sampling method was used to increase the effectiveness of promotion, and to utilize the power of "word of mouth" (Whitler, 2014; Wyner, 2014).

Additional promotion of the experiment was conducted through a promotion campaign composed of posters positioned around the university, flyers, a stand including HMD VR demonstrations with the Oculus Rift, and promotion in lectures (see Appendix 8.1. for poster/flyer). Participation in the experiment was booked through an customized booking website. In compensation for their time, each session lasting approximately 90 minutes, participants were offered a 100 NOK voucher for the university canteen.

3.5. Materials

3.5.1. Equipment

The virtual world utilized in this study was a modified version of the action role-play video game Fallout 4. This game was created by Bethesda Game Studios and downloaded by the two researchers through the online gaming platform Steam (Bethesda Game Studios, 2018). As this game allows for modifications, the researchers were able to redesign parts of the gaming scenario to fit the experiment.

The gaming scenario was the same for both experimental groups, but the equipment used was different. The experimental group was conducting the experiment in a HTC Vive head mounted display, with HTC Vive controllers (see Figure 10). The headset provided a 360-degree head tracking, leaving the user with an all-round view of the scenario (HTC Corporation, 2018). Vive controllers with 360-degree tracking and response were used for manoeuvring and interacting with objects in the virtual environment, visible to the user as the virtual avatar's hands. AKG K518 headphones provided sound effects and background noise from the gaming scenario.

Figure 10: HMD VR setup

Moving around the virtual environment was conducted through teleportation with the controllers or real life bodily movements tracked by the equipment. Thus, bending down would transfer into the virtual environment, as the equipment tracking provided virtual mirroring of physical movements. By clicking different buttons on the controllers, the participants interacted with elements in the virtual environment like picking up objects or shooting a gun. A computer screen displayed the direct vision of the HMD, to enable the researchers to follow the experience of the participants.

The control group attended the experiment on a desktop setup (see Figure 11). Here the scenarios were displayed on a 27-inch screen. By using the computer keyboard and a mouse, the participants were able to navigate through the scenario and interact with objects in the virtual environment. Sound effects and background noise were provided through the same AKG K518 headphones as worn by the experimental group.

Figure 11: Desktop setup

In both the HMD VR and desktop setup, the scenario was presented through first-person view. Thus, the player perceived the world from the view of the virtual avatar. This view was chosen to increase the similarity to the head mounted display. By using similar views in both setups, the variability affecting the results were minimized. This increased the probability of results being due to the effect of immersion level of HMD VR and desktop.

3.5.2. Measurement instruments

Four types of measurement instruments were used in this experiment: Performance measurements, questionnaires, interview, and physiological measurements. As the physiological measurements were mainly collected for the InnoTraining project, they will not be given much attention here. In total, five questionnaires were included in the experiment, but only two of these provided data to this thesis. These two questionnaires measured demographic data and aspects of intrinsic motivation. Other data collected throughout the experiment provided data for InnoTraining and Researcher B's thesis (Edwinson, 2018). Thus, the measurement instruments utilized to collect data for this study was performance measurements, two questionnaires and the exit interview.

Performance measurements

The performance of the participants was measured through collection of data on KPIs selected: time to task completion and number of figurines collected. These measurements were manually registered by the researcher during and after each round. The data was entered into an excel sheet, to provide an overview of the results and enable comparison of participants. In addition to the manual recordings, screen recordings of all three rounds were administered by the researchers. Thus, visual recordings of performance in the tasks were obtained, providing information on virtual actions.

Questionnaires

Five questionnaires were utilized in the experiment, to optimize the data output. These were: introduction questionnaire, subconscious learning questionnaire, quality of experience questionnaire, intrinsic motivation questionnaire, and simulator sickness questionnaire. Of these, only the introduction questionnaire and the intrinsic motivation questionnaire provided output for this study, and therefore only these two will be discussed further.

An *introduction questionnaire* was used to obtain information on the participants previous gaming experience and their health (See Appendix 8.3.). This questionnaire was a researcher-made instrument created to collect demographic data and information about individual differences predicted to influence performance, like gaming- and simulator experience.

An *intrinsic motivation questionnaire*, based on the Intrinsic Motivation Inventory (IMI) multidimensional measurement device, was developed by the researcher and used to measure intrinsic motivation after task completion. The IMI is a standardized measurement instrument grounded in the Self-Determination Theory, measuring seven different subcategories of intrinsic motivation (Monteiro et al., 2015). The subcategories are: Interest/Enjoyment, Perceived competence, Effort/Importance, Pressure/Tension, Perceived choice, Value/Usefulness and Relatedness.

The questionnaire used in this study was composed of twelve questions, related to six of the seven levels of intrinsic motivation developed by the creators of the IMI: Interest/Enjoyment – items 2 and 8, Perceived Competence – items 1 and 7, Effort/Importance – items 3 and 9, Pressure/Tension – items 6 and 11, Perceived choice – items 10 and 13, Value/ Usefulness – items 14 and 16. All items were scored on a 7-point Likert scale ranging from 1 (“Not at all true”) to 7 (“Very true”)” (See Appendix 8.4 for full questionnaire). Thus, all levels were included, except for the Relatedness level. This level was excluded because of its focus on relations between humans in team exercises, which were not the topic of this study’s experiment (Monteiro et al., 2015).

In addition to the intrinsic motivation queries, the questionnaire included six questions related to aspects of performance and experience. These were included to receive additional information on the experience of the participant. Thus, this questionnaire collected data on elements of intrinsic motivation and additional data on other aspects.

Interview

A short exit interview, developed by the researcher, was administered at the end of each experiment to obtain further information on the experience of the participant and their view of the HMD VR simulator as a training tool. The interview consisted of six main questions related to participation, task strategy, equipment used and perceived future prospects of VR simulation (see Appendix 8.5. for complete interview guide). The interview was developed with support from four subject matter experts (SME's) from TARG, and was audio recorded.

3.6. Procedure

An experimental procedure of approximately 90 minutes was executed for all participants attending. Figure 12 visualizes the overall procedure, with the main experimental sections: Experiment initiation, Intervention and Completion.

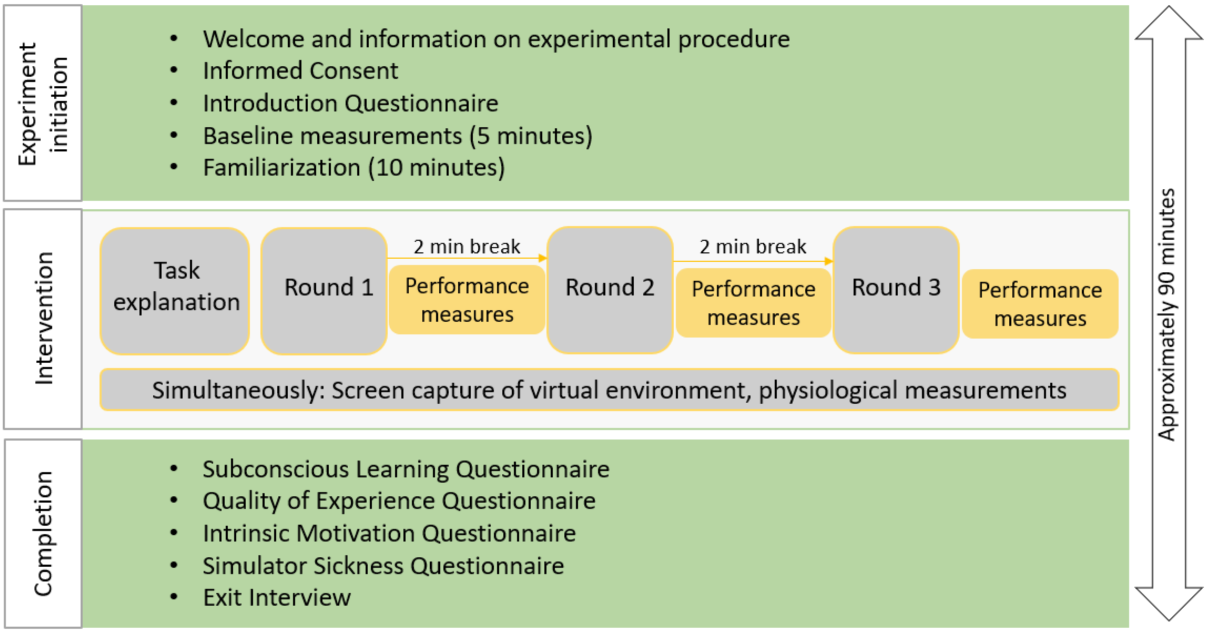

Figure 12: Experimental procedure, intervention includes a demonstration of experimental rounds

The experiment initiation period included all aspects completed before the main experimental task. Intervention included the main components of the main experimental tasks. The completion section entailed main activities completed post-testing.

3.6.1. Experiment initiation

After an introduction to the TARG lab, the participants were instructed to fill out an informed consent form to confirm their understanding of the study, data collection, anonymity and right of withdrawal. They were prompted by the researchers to pay special attention to the section instructing them not to disclose any detailed information of the tasks or procedure to any potential participant, to avoid affecting results of the study.

After signing the informed consent form, the participants were presented with an introduction questionnaire (see Appendix 8.3. for questionnaire). This questionnaire would reveal the level of previous gaming experience, which was predicted to affect their performance in the virtual scenario (Fabroyir & Teng, 2018; Münzer & Zadeh, 2016).

All the participants were equipped with sensors measuring heart rate and galvanic skin response. These measures were collected to identify physiological response to the tasks, but were collected for the purpose of the InnoTraining project and future data analyses outside the scope of this thesis. Before the experiment was initiated, a baseline was established through five-minute measurement with no distractions. Five-minute baseline was implemented based on previous research indicating it as sufficient (Filipovsk, Ducimetiere, & Safar, 1992; Honda et al., 2018).

As the volunteers in the study had diverse background knowledge, a familiarization period of ten minutes was initiated prior to the experiment. In the familiarization the participants practiced moving around and interacting with objects, based on instructions from the researchers. This allowed the participants to test the controls and develop an understanding of skills needed in the scenarios. By including this period, diversity effects were reduced and manoeuvring in the virtual environment became more natural for the user.

To obtain information on time perception in virtual environments, the researchers asked the participant: “How long do you feel you have been in the virtual world now?” As the participants were not given information on the length of this period on beforehand, the answer would indicate perceived time and immersion (Mineev, 2017).

3.6.2. Intervention

After a detailed description of the task, the experiment was initiated (See Appendix 8.2. for the experimental script). The participants were to complete three rounds in the maze, where the main task was to locate and collect as many of the ten hidden figurines as possible within five minutes. After each round a two-minute break was administered, the total number of figurines collected (score) and time was registered manually by the researcher.

As the time limit in each round was five minutes, this was registered every time except for when less time was used to find all ten figurines. If all figurines were collected in less than five minutes, the round was ended and time used registered. After each round was completed, the researchers asked the same question as after familiarization: “How long time do you feel you used this round?”

3.6.3. Completion

When three experimental rounds were completed, participants were asked to remove the equipment and relax, to obtain some distance from the virtual world. To collect subjective data on participants’ experience in the task, several questionnaires were administered. The control group were then instructed to go through the familiarization period and a round in the maze in the HMD VR setup. This was included to provide the HMD VR experience that was utilized in promoting the experiment. After questionnaires were completed and the control group had finished the trial in HMD VR, the physiological measurement equipment was turned off and removed. The very last aspect of the experiment was a short exit interview (see Appendix 8.3. for the script). This interview prompted further explanations and details of the participant's experience.

3.7. Data Analysis

As the participants were allocated into one of two groups prior to experiment initiation, data on two balanced groups with equal numbers of participants were obtained. When analysing the information collected in this study, both descriptive and inferential statistics were used. The data analysis of the quantitative data was conducted using the Statistical Package for the Social Sciences (SPSS Statistics 24) software. The two experimental groups were analysed, both independently and combined, and compared. Explanation of the data analysis procedure related to the three hypotheses is presented below.

3.7.1. Task performance

Task performance was measured through data collected related to KPIs throughout the three rounds of the experimental procedural task. To enable the expression of task performance as a function of time used and number of figurines collected each round, a total performance score per round for all participants was calculated. This score was estimated using the following function/formula:

$$Task\ performance\ score = \frac{Time\ used(seconds)}{Figurines\ collected}$$

Thus, task performance score demonstrated efficiency within the maze task, indicating the average time of collecting each figurine. Low score relates to high efficiency and performance, as a low performance score would indicate the participant finding and collecting the figurines fast.

Initially, independent samples t-tests were planned performed to compare the means of the two groups and investigate whether there existed a statistical significance between them related to performance (Frankfort-Nachmias, Nachmias, & DeWaard, 2015; Kent State University Libraries, 2018). However, violations of necessary assumptions resulted in the implementation of the non-parametric Mann-Whitney U test instead (Fay & Proschan, 2010).

3.7.2. Previous gaming experience

Previous gaming experience was measured on a 7-point Likert scale, from “never played before” to “plays every day for more than two hours”. These measurements were then transformed into three levels of prior gaming experience: 1= Never played before, 2= Plays one a month or once every other month, and 3= Plays once a week or more (Murias et al., 2016).

Descriptive statistics of performance score of each level of prior experience was investigated, comparing means and identifying trends in the data. Additionally, Spearman’s rho correlation test was conducted to investigate correlations between gaming experience and performance. Spearman was chosen over Pearson’s correlation, since the assumption of normal distribution in the data was not met. The correlation tests were initiated to investigate whether high level of experience correlated with high level performance.

3.7.3. Motivation for learning

Motivation for learning was measured through an intrinsic motivation questionnaire measuring six subscales of intrinsic motivation and an exit interview. Post data collection, data from the first two participants were excluded due to changes in the measurement instrument. These changes rendered the output from these participants substantially different than the rest. Thus, they were removed, reducing the total number of participants (n=48).

A reliability analysis of the six subscales was conducted to validate the internal consistency of the intrinsic motivation questionnaire. Independent samples t-tests were performed on the two experimental groups for each subcategory, to reveal statistically significant differences in the data. Additional qualitative data on motivation was collected from all participants through exit interviews. These interviews were recorded and processed by the researcher to contribute extra depth to the findings. Analysis of the interview data revealed patterns and themes indicating positive and negative aspects of the participants experience.

4. Results

This section will present an overview of the experimental results produced in the experimental study comparing HMD VR and desktop training, in relation to task performance in skill acquisition and motivation for learning. To enable clear presentation of the results related to the three hypotheses of the study, the results chapter is divided into three main sections: Task performance, effect of previous gaming experience on task performance, and intrinsic motivation. See Figure 13 for visual explanation and relation to hypotheses.

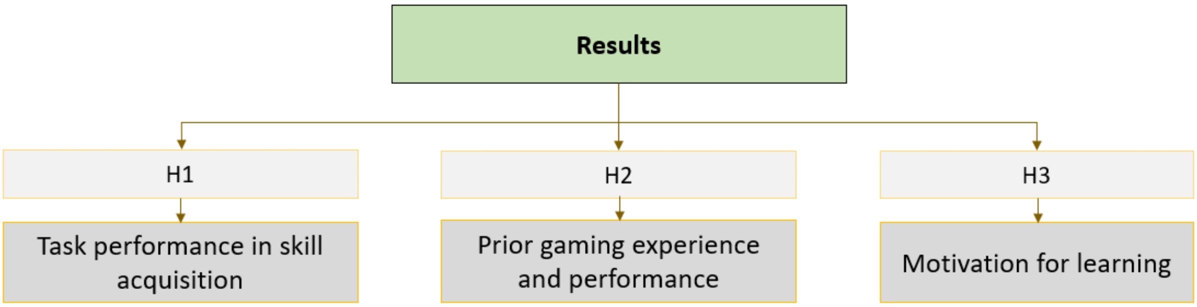

Figure 13: Division of results chapter, with related hypotheses

Section 4.1. presents the results from analysis of performance score, including a table of descriptive data and graphical visualisations of central tendencies of groups. Section 4.2. presents results from analysis of the effect of prior gaming experience on task performance. Additional graphs and tables are utilized to demonstrate data trends. The last section (4.3.) includes results from analysis of elements of intrinsic motivation, with a graph indicating central tendencies. An alpha level of .05 was used in all statistical testing, to reject the null hypotheses.

4.1. Task Performance

Hypothesis 1 proposed a significant difference between task performance in the two experimental groups and a higher performance level in HMD VR than in desktop. To investigate the accuracy of this hypothesis, comparisons of the performance in the two groups were initiated.

Investigating box plots of the data in each group and utilizing the outlier labelling rule, one clear outlier in the control group and one in the experimental group were removed (Hoaglin, Iglewicz, & Tukey, 1986). Thus, the total number of participants was reduced to (n=48). There were identified other outliers by analysing the box plots, but these were based on (g=1.5) as demarcation criteria for determining outliers. It has been stated that (g=2.2) is the optimal criteria for determining valid outliers (Hoaglin & Iglewicz, 1987). As SPSS only offers calculations with (g=1.5) and (g=3), outliers calculated with (g=3) were excluded.

Table 8 presents the measures of central tendencies related to the performance score of each round, after the two outliers identified were removed.

Table 8: Descriptive statistics for task performance score each round, divided by experimental setup

	Performance score round 1		**Performance score round 2**		**Performance score round 3**	
	HMD VR	Desktop	HMD VR	Desktop	HMD VR	Desktop
N	24.00	24.00	24.00	24.00	24.00	24.00
Mean	29.77	24.50	23.10	12.61	20.89	12.04
Std. Deviation	10.59	11.50	10.90	10.37	12.44	10.61
Variance	112.20	132.24	118.88	107.49	154.74	112.56

As low performance score translates to high task performance level (see section 3.7. for further explanation), Table 8 demonstrates overall higher level mean task performance in desktop setup than in HMD VR. Standard deviations and variances of the two groups are generally similar, but big. This indicates a wide spread in the data, and highly varying performance scores. Additionally, the percentage of participants obtaining top-score of ten figurines, were higher in desktop setup for each round. Figure 14 demonstrates mean performance scores of the two groups.

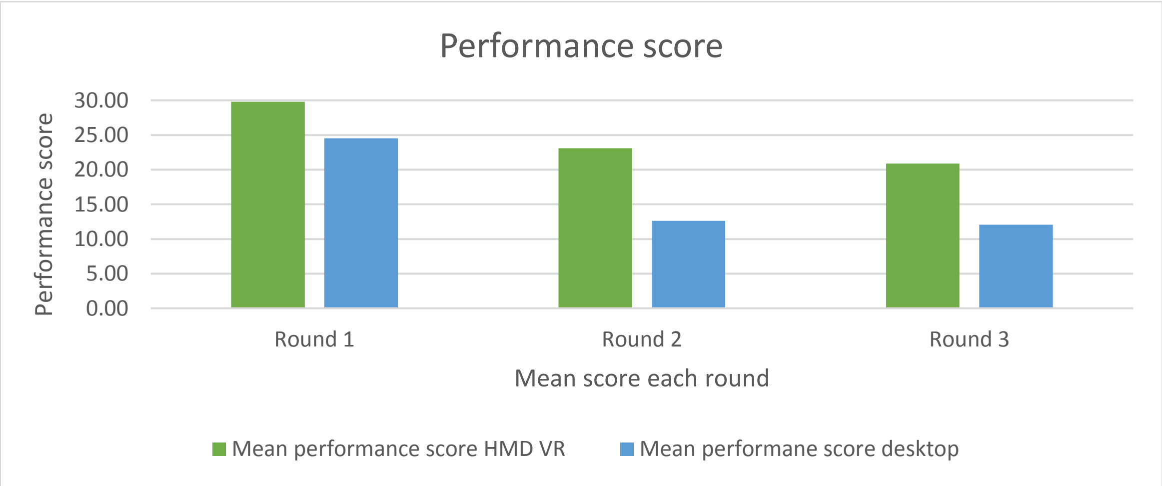

Figure 14: Mean performance score per experimental round, divided by experimental group (higher score indicates lower task performance)

Figure 14 visualize how performance score for desktop have a higher level of improvement from round one to two than HMD VR. It also indicates a higher level of stagnation in improvement from round two to round three in the desktop group than in HMD VR.

4.1.1. Between-group comparison

The Shapiro-Wilk tests were utilized to investigate the assumption of normal distribution of performance score for both groups in each round ($p>.05$). The test revealed that only scores for the HMD VR group for round one could be assumed normally distributed ($p=.625$). Normality is one of the assumptions most likely to affect the performance of the t-test, as it might result in false rejection of the null hypothesis 5% of the time (given a significance level of .05) (Barlett, 2013).

The lack of normal distribution in the datasets, even after outliers were removed, and the sample size being too small to depend on the central limit theorem would result in invalid results from the t-test (Barlett, 2013; Investopedia, 2018). Therefore, the non-parametric test Mann-Whitney U was initiated to investigate differences between performance score between the two groups. This test is stated to be the non-parametric equivalent to the independent samples t-test, investigating differences in group medians rather than means (Conover WJ, 1999; Lærd Statistics, 2018; MJ & TDV, 2009).

Round 1

The Mann-Whitney U test of performance score in the first experimental round indicated a significant difference between performance score for the HMD VR setup ($Mdn= 31.7$) and desktop setup ($Mdn=22.7$), $U= 198.5$, $p=.032$ (1-tailed), $r=.27$. Levene's test satisfied the assumption of homogeneity of variance ($F = .113$, $p = .738$) (adjusted df). As high performance scores reflect low task performance, the result point to a significantly greater task performance in the desktop setup for round one of the experiment.

Round 2

Running the test for the two groups in the second experimental round, indicated a significant difference between performance score for the HMD VR setup ($Mdn= 22.3$) and desktop setup ($Mdn=8.6$), $U= 111$, $p<.001$, $r=.53$. Levene's test satisfied the assumption of homogeneity of variance ($F = 2.77$, $p = .105$) (adjusted df). Thus, results indicate a significant difference between the two groups, with higher level task performance in the desktop setup in round two of the experiment.

Round 3

The Mann-Whitney U test for the two groups in the third experimental round, also indicated a significant difference between performance score for the HMD VR setup (Mdn= 16.4) and desktop setup (Mdn=8.8), U= 141.5, p=,002 (2-tailed), r=.44. Levene's test satisfied the assumption of homogeneity of variance (F = 1.59, p = .214) (adjusted df). Thus, similarly to previous rounds, results for this final round indicate a higher level of task performance in the desktop setup.

Figure 15 demonstrates the trends of task performance in the two groups, indicated in the above sections. It shows how performance score for desktop has a higher level of improvement from round one to two than HMD VR. It also indicates a higher level of stagnation in improvement from round two to round three in the desktop group than in HMD VR.

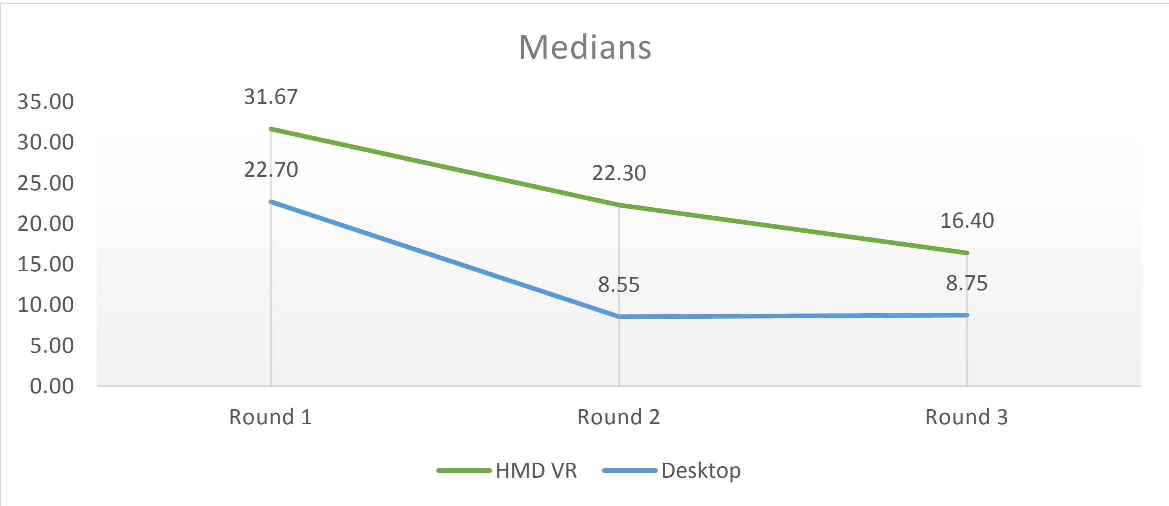

Figure 15: Median trends of performance score in both experimental groups (higher numbers indicate lower performance)

4.2. Previous Gaming Experience and Performance

Hypothesis 2 stated that prior gaming experience would affect task performance. To investigate this assumption, a comparison of performance means of each experience level was conducted. Table 9 presents performance means and standard deviations of each experience level.

Table 9: Central tendencies of performance score, divided by prior gaming experience level

	Performance score round 1			Performance score round 2			Performance score round 3		
	Low	Med.	High	Low	Med.	High	Low	Med.	High
N	15	13	20	15	13	20	15	13	20
Mean	35.20	23.94	23.17	23.42	18.32	13.48	24.35	12.09	13.39
Std. deviation	11.92	8.62	9.32	12.11	11.66	10.20	12.88	8.39	11.58

The means of performance score of all three experience levels are reduced from round one to two, indicating improved performance. From round two to the last round performance score in low experience level participants are increasing again, indicating worsened performance in the last round. The median performance score of the medium experience participants are greatly reduced round three, demonstrating steady improvement throughout the experimental rounds. The high-level experience performance score in the last round indicate a stagnation in performance, with a slight reduction in task performance. These trends in mean performance are visualized in the graph in Figure 16.

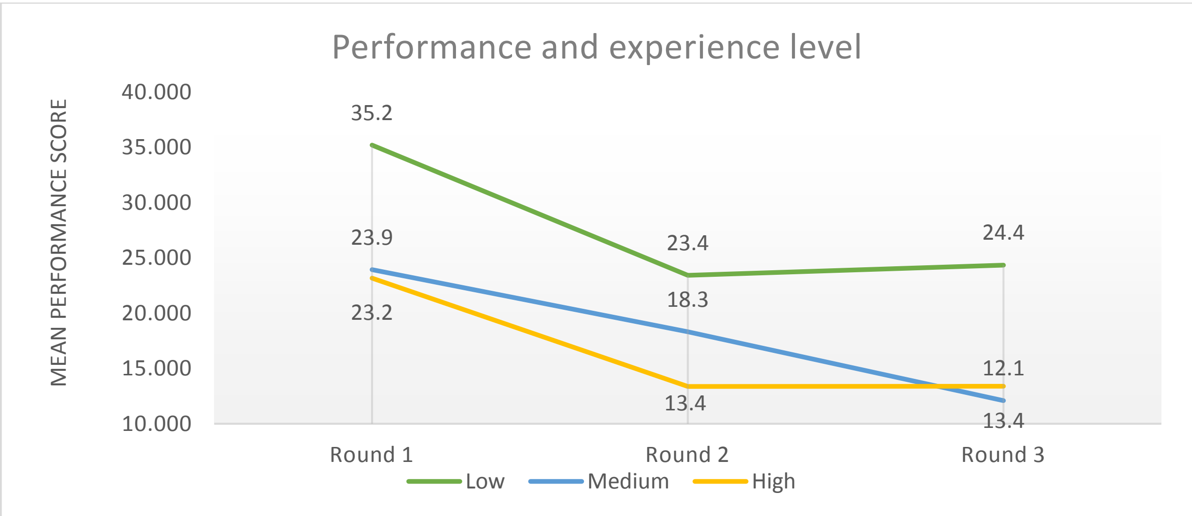

Figure 16: Performance trends of each level of prior gaming experience (higher score indicates lower task performance)

A Spearman's rho test was conducted to investigate correlations between gaming experience and performance. The test revealed a moderate negative correlation between previous gaming experience and total performance score round 1, $r_s(48) = -.393$, p=.006, performance score round 2, $r_s(48) = -.432$, p=.002, and performance score round 3, $r_s(48) = -.405$, p=.004. This indicated increased experience level to result in decreasing performance score. As low score relates to high level performance, these results show that participants with high level of previous gaming experience performed better than those with little or no experience.

4.3. Motivation

Hypothesis 3 predicted a higher level of intrinsic motivation the HMD VR group than in the desktop group. To investigate whether the data results indicated the same, output from the intrinsic motivation questionnaire was analysed through comparison of mean measurements. Table 10 presents measures of central tendencies of the four subscales.

Table 10: Descriptive statistics per intrinsic motivation subscale, divided by experimental setup (VR=HMD VR)

	Interest/ Enjoyment		**Perceived competence**		**Effort/ Importance**		**Value/ Usefulness**	
	VR	Desktop	VR	Desktop	VR	Desktop	VR	Desktop
N	24	24	24	23	24	24	24	24
M	5.46	4.48	4.98	5.74	5.73	5.15	5.85	4.42
SD	1.05	1.76	1.02	.81	1.03	1.24	.95	1.69

The mean scores of the subcategories are generally higher in the HMD VR group than in desktop, except for in the second subcategory Perceived competence where scores are higher in the desktop setup. Figure 17 provides visual demonstrations of the subcategory means, of the two experimental groups.

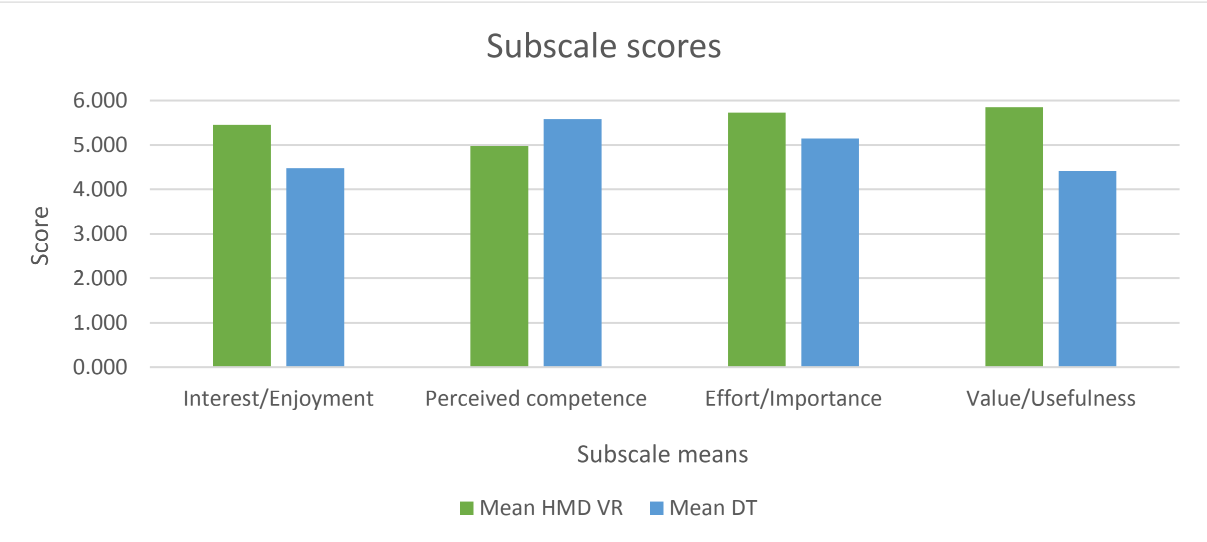

Figure 17: Mean score of each subcategory, presented divided by experimental group

4.3.1. Between-group comparison

Cronbach alpha revealed an above acceptable level reliability of the questionnaire, with a total score of (a=.709). The same test was completed for the subscales, indicating a high level of internal consistency of the subscales Interest/Enjoyment (a = .809), Perceived competence (a=.705), Effort/ Importance (a=.734), and Value/Usefulness (a = .736). The tests revealed alpha scores below the generally accepted limit of .7 for the two subscales Pressure/Tension (a = .403) and Perceived choice (a = .342) (IDRE, 2017). Thus, these subscales were excluded from further analysis performed in this study, resulting in a total of four subscales (see Appendix 8.4.1. for a total overview of questionnaire item results). Investigation of the four remaining subcategories reliability revealed a total reliability score of (a=.778), which is above the recommended level.

One outlier was detected in the desktop group in the subcategory Perceived competence, through the same investigations as used in outlier analysis of performance score data. To enable normal distribution of both datasets in this subcategory, the outlier was removed. After the outlier removal, three out of four subcategories met the assumption of normality for both experimental groups and were further analysed utilizing the independent samples t-test.

Before administering the t-tests, Shapiro-Wilk tests were utilized to investigate the assumption of normal distribution of both groups in each subcategory ($p>.05$). Effect sizes was analysed based on Cohen (1988) guidelines for interpreting power of correlation: small ($r=.10$), medium ($r=.30$) and large ($r = .50$) effect size (Gignac & T Szodorai, 2016).

Subcategory one: Interest/Enjoyment

Shapiro-Wilk revealed a non-significant difference between the normal distribution and the subscale Interest/Enjoyment for the experimental group ($p=.101$) and the control group ($p=.395$). Thus, the data in the two groups for this subcategory is considered normally distributed, as the null hypothesis of no difference cannot be rejected.

An independent t-test revealed a significant difference between the experimental group ($M= 5.46$) and the control group ($M= 4.48$) in the Interest/Enjoyment category $t(38) = 2.34$, $p=.025$ (equal variances not assumed), $d=.68$. Levene’s test of equality of variances showed a significant difference in variances ($F=7.15$, $p=0.1$). The large variance may have a negative effect on the validity of the t-value, indicating overlapping groups. Thus, the t-test indicated significantly higher Interest/Enjoyment score in the HMD VR group than in the desktop group, although the validity of the results may be questionable due to differences in variance.

Subcategory two: Perceived competence

For the second subcategory, Perceived competence, the Shapiro-Wilk test indicated normal distribution for the experimental group ($p=.56$), and for the control group ($p=.30$).

Thus, the assumption of normality was satisfied, and the t-test could be used with "high" validity.

An independent t-test revealed a significant difference between the experimental group (M= 4.98) and the control group (M=5.74) in the Perceived competence category $t(45) = 2.84$, p=.007, *d*=.83. Levene's test of equality of variances demonstrating homogeneity of variances (F=1.05, p=.31). Thus, the t-test indicated significantly higher score for the desktop group than the HMD VR group in the subcategory Perceived competence.

Subcategory three: Effort/Importance

The Shapiro-Wilk test results for the subcategory Effort/Importance, stated normal distribution in the experimental group (p=.055) and control group (p=.264), meeting the assumption of normality for parametric tests.

The independent t-test revealed a non-significant difference between the experimental group (M= 5,73) and the control group (M=5,15) in the Effort/Importance category $t(46)$ = 1,77, p=.083, *d*=.51. Levene's test of equality of variances revealed a non-significant difference in variances (F=.23, p=.64), demonstrating homogeneity of variances. The resulting output from the t-test did not demonstrate significantly higher performance in HMD VR than in desktop, even though the mean scores would suggest otherwise.

Subcategory four: Value/Usefulness

For the last subcategory, Value/Usefulness, the Shapiro-Wilk test revealed normal distribution for the control group (p=.092), but lacking normality for the experimental group (p=.027). Due to non-normality, the validity of an independent t-test would be low. Therefore, the Levene's test of homogeneity of variances was initiated to investigate whether the Mann-Whitey U test would provide valid results. Results were however poor with a significant difference in variances (F = 8.08, p = .007), indicating low-level validity for the Mann-Whitney U test as well. As no other test equivalent to the independent samples t-test were identified, the researcher decided to analyse and compare the results from both the parametric and non-parametric test, regardless of low validity.

The results of the Mann-Whitney U test indicated a significant difference between the two groups U=141, p=.002 (2-tailed), r=.45. However, these results have low validity because of the unequal variances. In comparison, the independent samples t-test were performed, disregarding the lack of normal distribution. The test indicated a significant difference between the experimental group (M=5.85) and the control group (M=4.42) in the Value/Usefulness subcategory $t(36)$, p=.001 (equal variances not assumed), d=1.043. Thus, both tests showed a significant difference in the two experimental groups, indicating higher results in the HMD VR group. However, the validity of these results is highly questionable because of the violations of critical assumptions.

4.3.2. Interview

An exit interview was initiated after experiment completion. A content analysis of audio and notes from the interviews was completed, revealing several patterns and themes in participants experience of the training setup. These main patterns are presented in Table 11.

Table 11: Results of content analysis of the post-experimental exit interviews conducted. Positive and negative remarks on the experience in the two experimental setups are listed.

	HMD VR	Desktop
Positive remarks	- Exiting view of the virtual world. - New and novel training method. - High level of immersion. - Great potential in training and learning practical tasks.	- Engaging task. - Familiarity of the system. - High interest of video gaming/ previous experience.
Negative remarks	- Confusing manoeuvring - Unfamiliar controllers - Some level of dizziness	- Task difficulty level (too easy) - Low interest of video gaming - Some level of discomfort

5. Discussion

The purpose of this quantitative study was to contribute information on the effectiveness of HMD VR simulators, by investigating elements of task performance and motivation for learning in HMD VR and desktop simulation. The discussion chapter is divided into four main sections, three related to the hypotheses of the study and the last focusing on overall implications for maritime training.

5.1. Task performance in skill acquisition

Task performance in skill acquisition was measured through the KPIs time to completion and number of figurines collected, used in similar studies to investigate performance and effectiveness of simulator training (Grubel, Thrash, Holscher, & Schinazi, 2017; Mitchell, Hopper, Daniels, George-Falvy, & James, 1994). Measurement of these KPIs indicated skill level of the participants and performance improvement throughout the experiment.

The results indicated significant differences in task performance between the two experimental groups. H1 predicted task performance to be better in HMD VR than in desktop setup, but results showed the opposite. The Mann-Whitney U tests indicated significantly higher performance in the desktop group and a trend in task performance in the tree experimental rounds (see Figure 15 for visualization of trend).

The overall lower level task performance in HMD VR may be due to familiarity with the system, as individual differences like former experience have been proven to influence task performance (Fabroyir & Teng, 2018; Mead & Fisk, 1998). Additionally, the low task performance may be related to the novelty of the technology. Many of the participants in this study had never utilized HMD VR simulators before. The new features of the virtual environment and experience may have distracted the participants from the main task, reducing performance. As the experimental procedure provided limited practice, the exposure time may have been insufficient to get used to the system. The experimental duration may thus have been too short to achieve the skilled performance which is the desired outcome of Langan-Fox et. al. (2002) three-stage process.

5.1.1. Experimental rounds

Task performance was lowest for both groups in round one, but significantly higher in the desktop group. Langan-Fox et. al. (2002) stated the performance in the first stage of skill acquisition requires information processing and is highly cognitive demanding. Thus, participants may have struggled with task understanding and concentration as it was unfamiliar and cognitively demanding. Additionally, the effect size of the test results was rather low, indicating weakness of results. These weak results may be related to the fact that even though those familiar with the system may adapt more easily to the scenario, it was still unfamiliar. All participants had to learn the objectives of the task. This may have led to diverse but low-level task performance in the first round, regardless of simulator setup.

Both groups improved the second round, but the desktop group improved considerably more. This may again be due to the desktop being a more known tool, resulting in easier adaption to task and scenario. There was a definite increase in effect size from round one to round two, indicating the effectiveness of training and continued exposure to the virtual environment. Thus, the desktop group seems to have progressed to the second stage of skill acquisition in round two, which is less cognitive demanding and more automated enabling participants to focus on the task itself (Langan-Fox et al., 2002).

In the last round, performance continued to improve in HMD VR. Relating to the theory of Langan-Fox et- al. (2002) performance in HMD VR now reached a level indicating progression to the second stage of more automatic and skilled performance. In the desktop group however, the performance stagnated and task performance level was slightly reduced. A possible explanation of this stagnation is that participants had already reached “top-level” performance in round two as the task was the same each round. Ludolph, Giese, and Ilg (2017) suggest that the most effective learning occurs with increasing task difficulty, matching individual skill level. Thus, the result may be due to the limited challenge of the experimental tasks.

Another aspect, observed during the experiments, which may explain the low-level performance, was that some participants got overly confident after the second round. As their performance increased to the second round, they got overly confident due to their previous performance, and ended up with lower scores as a result of personal expectations. They attempted to exceed their previous performance, but instead their eagerness to succeed led to mistakes reducing results.

5.2. Previous gaming experience and performance

Previous gaming experience was measured through elements of the introduction questionnaire. After organizing the data into three levels of gaming experience, high – medium – low, investigations of task performance in each experience level was conducted.

H2 predicted that level of previous gaming experience would affect task performance. Investigating performance score levels of participants in the three experience levels, revealed results indicating a significant effect on performance. Additionally, investigation of the trend in mean performance score for each experience level indicated clear differences in learning curves in the three conditions (See Figure 16 for the graph of trend).

5.2.1. Effect of previous gaming experience

Münzer & Zadeh (2016) stated that it would be easier for an individual with prior gaming experience to adapt to simulated scenarios, as simulators present properties similar to video games regarding manoeuvring and system understanding. This might explain why the performance of those with low-level experience was so poor the first round. Participants with low-level experience may have struggled to understand the system, resulting in low-level results.

Spearman's correlation revealed a significant relationship between level of previous gaming experience and performance. These results indicated that when level of experience increased, level of task performance increased correspondingly. This was also prominent in investigating central tendencies in the data, which revealed that participants with high-level experience performed best in the first round. This may be due to an exceeding number of hours video gaming changing the individual's perceptual and cognitive abilities, enabling quick understanding and easy adaption to similar scenarios (Münzer & Zadeh, 2016; Richardson et al., 2011).

Investigation of the trends in the data revealed differences in learning curves of the three experience levels. Performance increased from round one to two in all three experience levels, suggesting a general improvement in task understanding and performance. In round three overall task performance of the participants with medium experience continued to increase. Experimental observations indicated this to be a result of the continued learning curve of these participants.

In both the low- and the high-level experience participants however, performance was reduced. This may be related to several factors, some mentioned in the previous section on task performance like over-confidence, or it may be a result of the participant losing interest. Observation and responses from the exit interview indicated that those with little or no experience, had generally lower interest in the outcome of the task. Some got bored by the end of round two and did not persist for good results in the last round. In addition, some of these participants expressed low competitiveness and interest in results, as they generally were uninterested in video gaming, which may have affected the outcome.

5.3. Motivation for learning

Motivation for learning was measured through an intrinsic motivation questionnaire investigating six subcategories of intrinsic motivation, two of which were removed, complemented with additional data from exit interviews. A comparison of responses from the two experimental groups were conducted, investigating the effect of simulator immersion.

Eagerness to expand knowledge is important for an effective learning process, as motivated learners are interested and engaged in the task at hand and eager to achieve good results (Ai-Lim Lee et al., 2010; Garris et al., 2016). Salas and Cannon-Bowers (2001) identified several studies indicating a strong correlation between trainees' motivation for learning and skill acquisition. If intrinsic motivation for learning could be lifted through implementation HMD VR simulators, it may positively influence training results. H3 assumed that this would be the case, that results would show intrinsic motivation to be higher in HMD VR than in desktop setup. The following sections present the experimental results and implications for simulator training.

5.3.1. Interest/Enjoyment

The analysis of the first subcategory Interest/Enjoyment showed generally high scores for both groups. Brazil et. al. (2018) stated that gamification elements in simulator training may positively affect the engagement of the user. As the tasks of the experimental procedure were highly game related, the participants may have experienced this positive effect. Additionally, participants expressed associations to other games, often reminiscing periods in their youth. Thus, many associated the virtual scenario to positive experiences with video gaming. This may have affected their interest and enjoyment, increasing their motivation.

The results showed significantly higher scores in the HMD VR group than in the desktop group, indicating that the participants in the HMD VR setup experienced a higher level of interest for the tasks and enjoyment in completing them. Jensen and Konradsen (2017) found level of simulator immersion to affect the engagement of the user. Thus, the increased engagement resulting from the higher-level immersion of HMD VR may explain the results. If the increased immersion positively affects user interest and enjoyment, HMD VR may positively influence learning outcome and effectiveness of simulator training.

5.3.2. Perceived competence

The Perceived competence was the only one of the four subcategories rated lower in the HMD VR group than desktop. This low rating may be due to the novelty of the system. As computer technology is familiar to most, the desktop group may have been less dependent on cognitive resources to understand the system and perform the task (Münzer & Zadeh, 2016). Thus, they may perceive their competence as good, as they easily understood the task and were able to complete it. Those conducting the experiment in HMD VR however, had little or no knowledge of this technology and may have experienced difficulties familiarizing with it. Because of the short experimental period, they may not have had enough time to master the manoeuvring technique perceiving their competence as low.

The expectations and goals set by participants may also have affected their rating of this subcategory. Kernan and Lord (1989) found evidence that goal setting may affect individual performance evaluation. They identified how those without specific goals and expectations tended to judge their performance as better than those with specific, as they exceeded own expectations simply by completing the task. Several participants expressed low expectations for performance and were therefore positively surprised even if they did not locate all ten figurines. Thus, a balance in goals and expectations may be important for the effectiveness in HMD VR simulators in training, as the user will require a certain level of expectations to reach goals promoting progress and success (Langan-Fox et al., 2002).

5.3.3. Effort/Importance

The results of this subcategory showed no significant difference in ratings in the two groups. Experimental observations and interview responses revealed that effort initiated in task performance was highly individual, and often related to individual competitiveness rather than training condition. This might explain why there was no significant difference in the ratings, as individual differences within both groups affected results.

Lai (2011) presented research indicating task difficulty to affect student engagement and perceived competence. Thus, an explanation of the indifference in perceived importance may be due to the difficulty level of the task and individual perception of the task. Several participants expressed that the task was somewhat limited in terms of excitement, but that they would have liked to continue for further rounds if the task had evolved. In addition, some of the participants disclaimed the task to be easy, reducing their interest after a while. Too easy tasks may reduce motivation because of diminished experience of competence and accomplishment, and exceedingly difficult tasks may induce challenge avoidance or resigning (Lai, 2011). Therefore, a balanced task difficulty, compatible with user abilities, will be important in designing training scenarios to obtain positive learning outcome through HMD VR simulators.

5.3.4. Value/Usefulness

Investigation of the results of this subcategory indicated significantly higher scores in the HMD VR group, although the validity of these results is questionable. Considering these results as valid, they provide a demonstration of how HMD VR may be perceived as a more valuable and useful tool in training increasing users interest in obtaining knowledge and experience.

The increased rating of this subcategory may be explained by several aspects, but the most prominent aspects identified through observations and interviews were system novelty and immersion. The interviews revealed excitement related to HMD VR, as this was a technology few had utilized before. Many also expressed a vision of opportunities for utilizing such systems in several training areas. Thus, the high rating may be a result of eagerness and individual anticipation of the technology. The interviews also revealed some expecting simulator sickness to negatively affect their experience. These expectations related to results presented by Howard (2017), indicating that medium level of immersion induces user discomfort. The immersion of the scenarios utilized in this study however, seemed sufficiently realistic not to induce simulator sickness, resulting in participants perceiving the system as usable. These results point to the importance of providing simulator training with high-level immersion, and how this may increase user comfort and motivation for learning.

5.4. Implications for training and education

Results showed an overall lower task performance in HMD VR than in desktop. Nevertheless, they indicated significant differences in task performance in the two groups, which in itself was an interesting outcome indicating how the outcome of training seems to depend on proficiency in the virtual system. As a substantial part of the practical maritime educational training is conducted in simulators, the knowledge of these systems will greatly influence the learning outcome. Thus, an extensive period of system familiarisation should be included in the training procedure, to obtain optimal results utilizing HMD VR in maritime training.

The difference in task performance between the two experimental groups may be due to several aspects, but analysis indicated previous experience and duration of exposure to be essential to the outcome. This effect was most prominent in the initial period of exposure to the virtual scenario. As participants became familiar with the scenario and simulator system, results were indicated to be less dependent on previous experience and more on engagement. Thus, the effect of individual differences like technological novelty and familiarity is related to exposure duration and may be mitigated through continued HMD VR utilization.

The virtual scenario utilized in the experimental task was based on a video game, including gaming elements like time limit and score. Studies have found that the implementation of such elements may increase individual engagement and interest, because of the prospect of direct information of training outcome (high/low score) (Buttussi & Chittaro, 2018). This inclusion of gaming elements may have enhanced the task performance of participants, by increasing learning and task understanding (Garris et al., 2016). This indicates affordances of incorporating gaming elements into simulated training procedures, which may enhance outcome when implementing HMD VR simulators in maritime training.

Immersion has been stated to increase engagement of the user, and affect the user's perception of the value of the training exercise (Jensen & Konradsen, 2017). As the level of immersion of simulators presently used in maritime training is varied, the engagement of the user may be limited. The results of this study indicated that the HMD VR group might have experienced a higher level of immersion. Thus, the application of such high-level immersion systems in maritime training may increase effectiveness, as a result of level of immersion leading to increased intrinsic motivation.

5.5. Summary

Even though former research has indicated higher immersion to positively affect performance, the results from this study do not indicate higher effectiveness of training in HMD VR than in desktop. However, as these results are related to individual differences expected to be reduced in time, the applicability of HMD VR simulators may still be high in maritime training. One important thing to consider when implementing these simulators in maritime training will be to familiarize the user to the system, providing a scenario with increasing difficulty level appropriate to the user's abilities, and creating a scenario maintaining the engagement and interest of the user.

5.6. Limitations

Most HMD VR aspects addressed in this study are relatively new to the field of research and little testing of this type of training equipment has been performed. An experimental design including a commercial HMD system was applied in this study. Some limitations in the experimental design will exist regardless of the technology, sampling methods and measures.

A between-subjects design was chosen for this study. Within-subjects design could also have been used. Utilizing the within-group, the sample would not have been divided, but all participants would have received both treatments (Psychology World, 1998b). This could have provided insight into individual's comparative perception of the two experimental setups. However, as the study intended to investigate several experimental rounds, the within-subjects design would extend experimental duration.

The sampling method used in this study was convenience sampling, aided by snowball sampling. This method resulted in a sample mostly consisting of students and faculty at USN and Vestfold Campus, including outsiders recruited through the researchers' connections. In addition, the majority of the participants were Norwegian. Thus, the sample may be limited with regards to the population, as a sample with a wider spread may have better represented it. However, the sample size was relatively large and consisted of students and faculty of different domains, ages and gender, and therefore the sample may be considered a strong representation of the population (Frankfort-Nachmias et al., 2015).

Only subjective measurements of intrinsic motivation were included in the experiment, which may lower the validity of the results as subjective measurements are prone to personal bias (Bollen & Paxton, 1998). Additionally, one could argue that it would have been valuable to administer questionnaires between the experimental rounds, as it would provide detailed information on each round. This was not executed in this study, because of the already extensive duration of each experimental session.

One of the most influential limitations of the study was connected to the inability to control for previous gaming experience and the effect it may have on performance. It was evident that this aspect affected test results and should therefore have been controlled for. Previous experience may also have affected the level of engagement and excitement, as the difficulty level of the tasks was too low to reflect all participants needs (Garris et al., 2016). Thus, varied level of experience resulted in a scenario that was perceived as somewhat dull for some (Merchant et al., 2014).

The limitation of the duration of the experimental study may have affected the outcome of learning and skill acquisition, as the participant had little time for performance improvement. Ideally, another round of experiments would have been conducted after about a month, to investigate skill retention and long-term learning (Buttussi & Chittaro, 2018). It would also have been interesting to continue further rounds, with increasing difficulty level, to investigate trends in performance over time.

5.7. Recommendations for further research

Although this study included several aspects of HMD VR and its effect on training and education, additional aspects were identified as highly interesting for further research. Main areas identified which should be investigated further are presented below:

- Investigating how HMD VR affects additional aspects of skill acquisition, like memory retention or transferability of simulator-trained skills into real-life activities;
- Testing the long-term effects of HMD VR on the process of skill acquisition;
- Further analysis of the effect of previous gaming experience on training in HMD VR simulator, and methods of controlling for this effect;
- Investigating time perception in HMD VR training compared to other simulators and analyse the effect time perception may have on training effectiveness;
- Identifying methods of objective measurements of motivation.

6. Conclusion

Hypothesis one predicted performance to be higher in HMD VR than in desktop. The results of this study did however not support this hypothesis. Task performance was higher in the group performing the experiment on desktop setup, as these participants generally collected more figurines in less time. The prediction of hypothesis two was supported by the results, as level of previous gaming experiences was shown to affect task performance. Finally, the results of the analysis of intrinsic motivation of the two experimental groups indicated higher level intrinsic motivation in the HMD VR group than in the desktop group. Although one subcategory favoured desktop, the overall results were in accordance with the prediction of hypothesis three.

Even though this study does not support the effectiveness of HMD VR in task performance in skill acquisition, positive aspects to implementing HMD VR in maritime training, like increased intrinsic motivation, have been identified. These initial results from the comparison of HMD VR and desktop are considered a promising foundation for further research on the effectiveness of such systems in maritime training.

7. References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*(3), 288-318.

Ai-Lim Lee, E., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education, 55*(4), 1424-1442. doi:10.1016/j.compedu.2010.06.006

Alhalabi, W. S. (2016). Virtual reality systems enhance students' achievements in engineering education. *Behaviour & Information Technology, 35*(11), 919-925. doi:https://doi.org/10.1080/0144929x.2016.1212931

Alshaer, A., Regenbrecht, H., & O'Hare, D. (2017). Immersion factors affecting perception and behaviour in a virtual reality power wheelchair simulator. *Appllied Ergonomics, 58*, 1-12. doi:10.1016/j.apergo.2016.05.003

Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2000, 2000). *Measuring the effects of interruptions on task performance in the user interface.* Paper presented at the Systems, Man, and Cybernetics, 2000 IEEE International Conference on.

Barlett, J. (2013). *The t-test and robustness to non-normality.* Retrieved from http://thestatsgeek.com/2013/09/28/the-t-test-and-robustness-to-non-normality/

Barnes, J., Burns, J., Nesbitt, C., Hawkins, H., & Horgan, A. (2015). Home virtual reality simulation training: The effect on trainee ability and confidence with laparoscopic surgery. *Journal of Surgical Simulation*, *2*, 53-59. doi:10.1102/2051-7726.2015.0012

Barsan, E. (2009). Sea Service Equivalency for Full Mission Simulators Training. *Maritime Transport & Navigation Journal, 1*(1), 14-23.

Basak, T., Unver, V., Moss, J., Watts, P., & Gaioso, V. (2016). Beginning and advanced students' perceptions of the use of low- and high-fidelity mannequins in nursing simulation. *Nurse Education Today, 36*, 37-43. doi:10.1016/j.nedt.2015.07.020.

Berg, L. P., & Vance, J. M. (2017). Industry use of virtual reality in product design and manufacturing: a survey. *Virtual Reality*, *21*(1), 1-17. doi:10.1007/s10055-016-0293-9

Bertrand, J., Bhargava, A., Madathil, K. C., Gramopadhye, A., & Babu, S. V. (2017). *The Effects of Presentation Method and Simulation Fidelity on Psychomotor Education in a Bimanual Metrology Training Simulation*. Paper presented at the 2017 IEEE Symposium on 3D User Interfaces (3DUI), Los Angeles, USA.

Bethesda Game Studios. (2018). *Fallout 4.* Retrieved from https://fallout.bethesda.net/

Bollen, K. A., & Paxton, P. (1998). Detection and determinants of bias in subjective measures. *American Sociological Review, 63*(3), 465-478.

Bowman, D. A., Johnson, D. B., & Hodges, L. F. (2001). Testbed evaluation of virtual environment interaction techniques *Presence: Teleoperators and Virtual Environments, 10*, 75-95. doi:http://dx.doi.org/10.1162/105474601750182333

Brazil, A. L., Conci, A., Clua, E., Bittencourt, L. K., Baruque, L. B., & Silva Conci, N. d. (2018). Haptic forces and gamification on epidural anesthesia skill gain. *Entertainment Computing, 25*, 1-13. doi:10.1016/j.entcom.2017.10.002

Bui, Y. N. (2014). *How to Write a Master's Thesis* (2nd ed.). California, USA: SAGE Publications, Inc.

Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology* (Vol. 1). Hobroken, New Jersey: John Wiley & Sons.

Buttussi, F., & Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Trans Vis Comput Graph, 24*(2), 1063-1076. doi:10.1109/TVCG.2017.2653117

Castells, M. (2009). An Introduction to the Information Age. In S. Thornham, C. Bassett, & P. Morris (Eds.), *Media Studies: A Reader* (pp. 152-165). Washington Square, New York: New York Universtiy Press.

Castells, M. I., Ordás, S., Barahona, C., Moncunill, J., Muyskens, C., Hofman, W., . . . Skorokhodov, S. (2015). Model course to revalidate deck officers' competences using simulators. *WMU Journal of Maritime Affairs, 15*(1), 163-185. doi:10.1007/s13437-015-0092-2

Chessa, M., Maiello, G., Borsari, A., & Bex, P. J. (2016). The Perceptual Quality of the Oculus Rift for Immersive Virtual Reality. *Human–Computer Interaction, 00*, 1-32. doi:10.1080/07370024.2016.1243478

Cibulka, J., Komulainen, T. M., Mirtaheri, P., Nazir, S., & Manca, D. (2016). Virtual Reality Training Simulators in the Process Industry: A Review of Existing Systems and the Way Towards ETS. *IEEE Computer Society*, 458-464. doi:DOI 10.1109/EUROSIM.2016.26

Cohen, N., & Arieli, T. (2011). Field research in conflict environments: Methodological challenges and snowball sampling. *Journal of Peace Research, Vol 48*(4), 423 - 435. doi:10.1177/0022343311405698

Conover WJ. (1999). *Practical nonparametric statistics* (3rd ed.). New York: John Wiley & Sons.

Cummings, M. L., & Quimby, P. W. (2018). The power of collective intelligence in a signal detection task. *Theoretical Issues in Ergonomics Science, 19*(3), 375-388. doi:10.1080/1463922X.2017.1381776

Deb, S., Carruth, D. W., Sween, R., Strawderman, L., & Garrison, T. M. (2017). Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics, 65*, 449-460. doi:10.1016/j.apergo.2017.03.007

Deci, E., Eghrari, H., Patrick, B., & Leone, D. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*(119-142). doi:10.1111/j.1467-6494.1994.tb00797.x

Deci, E., & Ryan, R. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior* (Vol. 3). New York: Plenum Press.

DNV GL. (2011). STANDARD FOR CERTIFICATION No. 2.14. In *MARITIME SIMULATOR SYSTEMS*.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53* 109–132. doi:10.1146/annurev.psych.53.100901.135153

Edwinson, A. E. (2018). *Investigating User Experience of VR Head-Mounted Display in Maritime Simulator Training.* (Master thesis), University of South-Eastern Norway,

Elliott, D., Hayes, S. J., & Bennett, S. J. (2012). 125 Years of Perceptual–Motor Skill Research. *The American Journal of Psychology, 125*(1), 9-23. doi:10.5406/amerjpsyc.125.1.0009

Everitt, B. S. (2002). *The Cambridge Dictionary of Statistics*. Cambridge, UNITED KINGDOM: Cambridge University Press.

Fabroyir, H., & Teng, W.-C. (2018). Navigation in virtual environments using head-mounted displays: Allocentric vs. egocentric behaviors. *Computers in Human Behavior, 80*, 331-343. doi:https://doi.org/10.1016/j.chb.2017.11.033

Fay, M. P., & Proschan, M. A. (2010). Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. . *Statistics Surveys, 4*, 1-39. doi:http://doi.org/10.1214/09-SS051

Filak, V. F., & Sheldon, K. M. (2010). Student Psychological Need Satisfaction and College Teacher-Course Evaluations. *Educational Psychology, 23*(3), 235-247. doi:10.1080/0144341032000060084

Filipovsk, J., Ducimetiere, P., & Safar, M. E. (1992). Prognostic significance of exercise blood pressure and heart rate in middle-aged men. *Hypertension, 20*(3), 333-339.

Fonseca, A. M., & Brito, A. P. (2001). Psychometric properties of the Portuguese version of Intrinsic Motivation Inventory (IMIp) in contexts of sports and physical activity. *Análise Psicológica, 19*(59-76). doi:10.14417/ap.344

Forester, T. (1990). *High-Tech Society, The Story of the Information Technology Revolution*. Great Britain: Basil Blackwell Ltd.

Frankfort-Nachmias, C., Nachmias, D., & DeWaard, J. (2015). *Research Methods in the Social Sciences*. New York: Worth Publishers.

Garris, R., Ahlers, R., & Driskell, J. E. (2016). Games, Motivation, and Learning: A Research and Practice Model. *Simulation & Gaming*, *33*(4), 441-467. doi:10.1177/1046878102238607

Gartner, K. (2016, 13.10.3016). Consumerism, Mass Extinction and our Throw-Away Society. *The Art Of*. Retrieved from https://www.theartof.com/articles/consumerism-mass-extinction-and-our-throw-away-society

Gignac, G., & T Szodorai, E. (2016). *Effect size guidelines for individual differences researchers* (Vol. 102).

Goran. (2016). Virtual Reality: A Brief History. *Use of Technology*. Retrieved from https://www.useoftechnology.com/virtual-reality-history/

Grolnick, W., & Ryan, R. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology, 52*, 890-898. doi:10.1037/0022-3514.52.5.890

Grubel, J., Thrash, T., Holscher, C., & Schinazi, V. R. (2017). Evaluation of a conceptual framework for predicting navigation performance in virtual reality. *PLOS one, 12*(9), 1-22. doi:10.1371/journal.pone.0184682

Guay, F., Chanal, J., Ratelle, C. F., Marsh, H. W., Larose, S., & Boivin, M. . (2010). Intrinsic, identified, and controlled types of motivation for school subjects in young elementary school children. *British Journal of Educational Psychology, 80*(4), 711–735. doi:https://doi.org/10.1348/000709910X499084

Gutiérrez, M., Ruiz, L., & López, E. (2010). Perceptions of motivational climate and teachers' strategies to sustain discipline as predictors of intrinsic motivation in physical education. *The Spanish Journal of Psychology, 13*, 597-608. doi:10.1017/S1138741600002274

Hanafi, H. F., Said, C. S., Wahab, M. H., & Samsuddin, K. (2017). Improving Students' Motivation in Learning ICT Course With the Use of A Mobile Augmented Reality

Learning Environment. *IOP Conference Series: Materials Science and Engineering, 226*. doi:10.1088/1757-899x/226/1/012114

Heiling, M. (1960). USA Patent No.

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the Academically Unmotivated: A Critical Issue for the 21st Century. *Review of Educational Research, 70*(2), 151-179. doi:10.3102/00346543070002151

History. (2010). *The Invention of the Internet.* Retrieved from https://www.history.com/topics/inventions/invention-of-the-internet

Hoaglin, D. C., & Iglewicz, B. (1987). Fine-Tuning Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association, 82*(400), 1147-1149.

Hoaglin, D. C., Iglewicz, B., & Tukey, J. W. (1986). Performance of Some Resistant Rules for Outlier Labeling. *Journal of the American Statistical Association, 81*(396), 991-999. doi:10.1080/01621459.1986.10478363

Holst, P. A. (1982). George A. Philbrick and Polyphemus-The First Electronic Training Simulator. *Annals of the History of Comptuing, 4*(2), 143-156. doi:10.1109/MAHC.1982.10021

Honda, Y., Watanabe, T., Otaki, Y., Tamura, H., Nishiyama, S., Takahashi, H., . . . Kubota, I. (2018). Gender differences in the impact of anemia on subclinical myocardial damage and cardiovascular mortality in the general population: The Yamagata (Takahata) study. *International Journal of Cardiology, 252*, 207-212. doi:https://doi.org/10.1016/j.ijcard.2017.11.019

Hontvedt, M. (2015). Professional vision in simulated environments — Examining professional maritime pilots' performance of work tasks in a full-mission ship simulator. *Learning, Culture and Social Interaction, 7*, 71-84. doi:10.1016/j.lcsi.2015.07.003

Howard, M. C. (2017). Investigating the simulation elements of environment and control: Extending the Uncanny Valley Theory to simulations. *Computers & Education, 109*, 216-232. doi:10.1016/j.compedu.2017.03.005

HTC Corporation. (2018). *Vive VR System.* Retrieved from https://www.vive.com/us/product/vive-virtual-reality-system/

IDRE. (2017). *What Does Cronbach's Alpha Mean? | SPSS FAQ.* Retrieved from https://stats.idre.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/

IMO. (2011). International Convention on Standards of Training, Certification and Watchkeeping for Seafarers (STCW) : including 2010 Manila amendments. In *STCW Convention and STCW Code (3rd consolidated ed. ed.)*. London: IMO.

Investopedia. (2018). *Central Limit Theorem - CLT*. Retrieved from https://www.investopedia.com/terms/c/central_limit_theorem.asp

ISO. (1991a). ISO 9241-11. In *Ergonomic requirements for office work with visual display terminals, ISO 9241*. https://.www.iso.org.

ISO. (1991b). ISO DIS 9126. In *Software product evaluation - Quality characteristics and guidelines for their use, ISO DIS 9126*.

Janßen, D., Tummel, C., Richert, A., & Isenhardt, I. (2016). Towards Measuring User Experience, Activation and Task Performance in Immersive Virtual Learning Environments for Students. In *Immersive Learning Research Network* (pp. 45-58).

Jenkins, S. (2017, 02.02.2017). We're over the digital revolution. This is the age of experience. *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2017/feb/02/digital-revolution-age-of-experience-books-vinyl

Jensen, L., & Konradsen, F. (2017). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*. doi:10.1007/s10639-017-9676-0

Jørgenrud, M. (2018, 19.02.2018). Elever forserer ett års matematikk på få uker ved hjelp av norsk VR-spill. *Digi*. Retrieved from https://www.digi.no/artikler/elever-forserer-ett-ars-matematikk-pa-fa-uker-ved-hjelp-av-norsk-vr-spill/430638

Kanfer, R. (1991). Motivational theory and industrial and organizational psychology. In M. Dunnette, L. Hough, & M. Leaetta (Eds.), *Handbook of Industrial and Organizational Psychology*, (Vol. 2, pp. 75-170). Palo Alto: CA: Consult.

Kent State University Libraries. (2018). *SPSS Tutorials: Independent Samples t Test.* Retrieved from https://libguides.library.kent.edu/SPSS/IndependentTTest

Kernan, M. G., & Lord, R. G. (1989). Effects of explicit goals and specific feedback on escalation processes. *Journal of Applied Social Psychology, 19*, 1125-1143. doi:https://doi.org/10.1111/j.1559-1816.1989.tb01243.x

Kim, K., & Ahn, S. J. (2017). The Role of Gamification in Enhancing Intrinsic Motivation to Use a Loyalty Program. *Journal of Interactive Marketing, 40*, 41-51. doi:10.1016/j.intmar.2017.07.001

Kolo, K. (2017, 13.12.17). Virtual Reality: The Next Generation Of Education, Learning and Training. *Forbes*. Retrieved from https://www.forbes.com/sites/forbesagencycouncil/2017/12/13/virtual-reality-the-next-generation-of-education-learning-and-training/#7b9ae56d733f

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of Cognitive, Skill-Based, and Affective Theories of Learning Outcomes to New Methods of Training Evaluations. *Journal of Applied Pshycology, 78*(2), 311-328. doi:10.1037//0021-9010.78.2.311

Lai, E. R. (2011). *Motivation: A Literature Review*. Retrieved from Pearson: https://images.pearsonassessments.com/images/tmrs/Motivation_Review_final.pdf

Langan-Fox, J., Armstrong, K., Salvin, N., & Anglim, J. (2002). Process in Skill Acquisition: Motivation, Interruptions, Memory, Affective States, and Metacognition. *Australian Psychologist, 37*(2), 104-117. doi:10.1080/00050060210001706746

Lee, G. I., & Lee, M. R. (2018). Can a virtual reality surgical simulation training provide a self-driven and mentor-free skills learning? Investigation of the practical influence of the performance metrics from the virtual reality robotic surgery simulator on the skill learning and associated cognitive workloads. *Surg Endosc, 32*(1), 62-72. doi:10.1007/s00464-017-5634-6

Leng, E. Y., Ali, W. Z. b. W., Baki, R., & Mahmud, R. (2010). Stability of the Intrinsic Motivation Inventory (IMI) For the Use of Malaysian Form One Students in ICT Literacy Class. *Eurasia Journal of Mathematics, Science & Technology Education, 6*(3), 215-226. doi: https://doi.org/10.12973/ejmste/75241

Levac, D., Driscoll, K., Galvez, J., Mercado, K., & O'Neil, L. (2017). OPTIMAL practice conditions enhance the benefits of gradually increasing error opportunities on retention of a stepping sequence task. *Human Movement Science, 56*(Pt B), 129-138. doi:10.1016/j.humov.2017.10.017

Li, C., Liang, W., Quigley, C., Zhao, Y., & Yu, L. F. (2017). Earthquake Safety Training through Virtual Drills. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, 23*(4), 1388-1397. doi:10.1109/TVCG.2017.2656958

Loup, G., Serna, A., Iksal, S., & George, S. (2016, 2016//). *Immersion and Persistence: Improving Learners' Engagement in Authentic Learning Situations.* Paper presented at the Adaptive and Adaptable Learning, Cham.

Ludolph, N., Giese, M. A., & Ilg, W. (2017). Interacting Learning Processes during Skill Acquisition: Learning to control with gradually changing system dynamics. *Sci Rep, 7*(1), 13191. doi:10.1038/s41598-017-13510-0

Lærd Statistics. (2018). *Mann-Whitney U Test using SPSS Statistics.* Retrieved from https://statistics.laerd.com/spss-tutorials/mann-whitney-u-test-using-spss-statistics.php

Mallam, S., Nazir, S., Renganayagalu, S. K., Ernstsen, J., Veie, S., & Edwinson, A. E. (2018). *A Pilot Study Investigating User Motivation and User Experience in Virtual Reality*. Paper presented at the International Ergonomics Association - 20th Congress, Florence, Italy.

Mallaro, S., Rahimian, P., O'Neal, E. E., Plumert, J. M., & Kearney, J. K. (2017). *A comparison of head-mounted displays vs. large-screen displays for an interactive pedestrian simulator*. Paper presented at the 23rd ACM Conference on Virtual Reality Software and Technology, VRST, Gothenburg, Sweden. www.scopus.com

Mania, K., Troscianko, T., Hawker, R., & Chalmers, A. (2003). Fidelity Metrics for Virtual Environment Simulations Based on Spatial Memory Awareness States. *Precence, 12*(3), 296-310. doi:10.1089/109493101300117938

Marks, S., Estevez, J. E., & Connor, A. M. (2014). *Towards the Holodeck: Fully Immersive Virtual Reality Visualisation of Scientific and Engineering Data.* Paper presented at the 29th International Conference on Image and Vision Computing New Zealand.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confi rmatory factor analysis. *Research Quarterly for Exercise and Sport, 60*, 48-58. doi:10.1080/02701367.1989.10607413

Mead, S., & Fisk, A. D. (1998). Measuring skill acquisition and retention with and ATM simulator: The need for Age-Specific Training. *Human Factors;, 40*(3), 516-123. doi:10.1518/001872098779591313

Meng-Lin, W., & Voicu, P. (2017). Efficient VR and AR Navigation through Multiperspective Occlusion Management. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, PP*(99), 1-12. doi:DOI 10.1109/TVCG.2017.2778249

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education, 70*, 29-40. doi:10.1016/j.compedu.2013.07.033

Miles, H. C., Pop, S. R., Watt, S. J., Lawrence, G. P., & John, N. W. (2012). A review of virtual environments for training in ball sports. *Computers & Graphics, 36*(6), 714-726. doi:https://doi.org/10.1016/j.cag.2012.04.007

Mineev, G. (2017). *The Impact of Immersive Virtual Reality on Effectiveness of Educational Games.* (Master), Utrecht University, Retrieved from https://pdfs.semanticscholar.org/d310/1ddb13be01836341c291f1542af8d0a91d1f.pdf

Mitchell, T. R., Hopper, H., Daniels, D., George-Falvy, J., & James, L. R. (1994). Predicting Self-Efficacy and Performance During Skill Acquisition. *Journal of Applied Pshycology, 79*(4), 506-517. doi:10.1037/0021-9010.79.4.506

MJ, C., & TDV, S. (2009). *Statistics at Square One* (11th ed.). Wiley-Blackwell: BMJ Books 2009.

Monteiro, V., Mata, L., & Peixoto, F. (2015). Intrinsic Motivation Inventory: Psychometric Properties in the Context of First Language and Mathematics Learning. *Psicologia: Reflexão e Crítica, 28*(3), 434-443. doi:10.1590/1678-7153.201528302

Moore, P. (1995). Learning and teaching in virtual worlds: Implications of virtual reality for education. *Australian Journal of Educational Technology, 11*(2), 91-102. doi:https://doi.org/10.14742/ajet.2078

Murias, K., Kwok, K., Castillejo, A. G., Liu, I., & Iaria, G. (2016). The effects of video game use on performance in a virtual navigation task. *Computers in Human Behavior, 58*, 398-406. doi:https://doi.org/10.1016/j.chb.2016.01.020

Münzer, S., & Zadeh, M. V. (2016). Acquisition of spatial knowledge through self-directed interaction with a virtual model of a multi-level building: Effects of training and individual differences. *Computers in Human Behavior, 64*, 191-205. doi:http://dx.doi.org/10.1016/j.chb.2016.06.047

Nazir, S., Sorensen, L. J., Øvergård, K. I., & Manca, D. (2015). Impact of training methods on Distributed Situation Awareness of industrial operators. *Safety Science, 73*, 136-145. doi:https://doi.org/10.1016/j.ssci.2014.11.015

Newman, D. (2016, 08.03.2016). In The Age Of Experience: The Best Brands Tell Better Stories And Make You A Part Of Them. *CMO Network*. Retrieved from https://www.forbes.com/sites/danielnewman/2016/03/08/in-the-age-of-experience-the-best-brands-tell-better-stories-and-make-you-a-part-of-them/#39ab8ebf7799

Niehorster, D. C., Li, L., & Lappe, M. (2017). The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research. *i-Perception*, 1-23. doi:10.1177/2041669517708205

Ocampo, J. R., Esparragoza, I. E., Rodríguez, J., Ivashyn, U., Sacchelli, C., & Vigano, R. (2017). *The effect of gender on the motivation of engineering students participating on multinational design projects*. Paper presented at the IEEE Global Engineering

Education Conference (EDUCON), Athens, Greece. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7942860

Payatagool, C. (2008). Theory and Research in HCI: Morton Heilig, Pioneer in Virtual Reality Research. *Telepresence Options*. Retrieved from http://www.telepresenceoptions.com/2008/09/theory_and_research_in_hci_mor/

Psychology World. (1998a). *2x2 Between Subjects Factorial Design.* Retrieved from https://web.mst.edu/~psyworld/between_subjects.htm#

Psychology World. (1998b). *Within-Subjects Design.* Retrieved from https://web.mst.edu/~psyworld/within_subjects.htm

Ragan, E. D., Bowman, D. A., Kopper, R., Stinson, C., Scerbo, S., & McMahan, R. P. (2015). Effects of Field of View and Visual Complexity on Virtual Reality Training Effectiveness for a Visual Scanning Task. *IEEE Trans Vis Comput Graph, 21*(7), 794-807. doi:10.1109/TVCG.2015.2403312

Rahm, S., Wieser, K., Wicki, I., Holenstein, L., Fucentese, S. F., & Gerber, C. (2016). Performance of medical students on a virtual reality simulator for knee arthroscopy: an analysis of learning curves and predictors of performance. *BMC Surgery, 16*(14), 1-8. doi:10.1186/s12893-016-0129-2

Rasheed, F., Onkar, P., & Narula, M. (2015). *Immersive virtual reality to enhance the spatial awareness of students*.

Ray, A. B., & Deb, S. (2016, 2-4 Dec. 2016). *Smartphone Based Virtual Reality Systems in Classroom Teaching — A Study on the Effects of Learning Outcome.* Paper presented at the 2016 IEEE Eighth International Conference on Technology for Education (T4E).

Reiners, T., Wood, L. C., & Gregory, S. (2014). *Experimental study on consumer-technology supported authentic immersion in virtual environments for education and vocational training*. Paper presented at the The 31st Annual Conference of the Australian Society for Computers in Tertiary Education, ASCILITE 2014., Australia.

Richardson, A. E., Powers, M. E., & Bousquet, L. G. (2011). Video game experience predicts virtual, but not real navigation performance. *Computers in Human Behavior, 27*(1), 552-560. doi:10.1016/j.chb.2010.10.003

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist, 55*(1), 68-78. doi:10.1037110003-066X.55.1.68

Salas, E., & Cannon-Bowers, J. A. (2001). THE SCIENCE OF TRAINING: A Decade of Progress. *Annual Review of Psychology, 52*(1), 471.

Sansone, C., Harackiewicz, J. M., & Sansone, D. C. (2000). *Intrinsic and Extrinsic Motivation : The Search for Optimal Motivation and Performance*. San Diego, UNITED STATES: Elsevier Science & Technology.

Savjani, K., Haseeb, F., & Reay, M. (2018). Measuring quality and outcomes in intensive care. *Surgery (Oxford), 36*(4), 196-200. doi:https://doi.org/10.1016/j.mpsur.2018.01.008

Sellberg, C. (2017). Simulators in bridge operations training and assessment: A systematic review and qualitative synthesis. *WMU J Marit Affairs, 16*, 247-293. doi:10.1007/s13437-016-0114-8

Shahriari-Rad, A., Cox, M., & Woolford, M. (2017). Clinical Skills Acquisition: Rethinking Assessment Using a Virtual Haptic Simulator. *Technology, Knowledge and Learning, 22*(2), 185-197. doi:10.1007/s10758-017-9308-1

Small, M. L. (2011). How to Conduct a Mixed Methods Study: Recent Trends in a Rapidly Growing Literature. *Annual Review of Sociology, 37*(1), 57-86. doi:10.1146/annurev.soc.012809.102657

Soares, E. d. S., Corazza, S. T., Piovesan, A. C., Azevedo, R. P. d., & Vasconcellos, S. J. L. (2016). Creation, validation, and reliability of a shooting simulator instrument for reaction time evaluation. *Motriz: Revista de Educação Física, 22*(4), 277-282. doi:10.1590/s1980-6574201600040010

Society, V. R. (2017). *History Of Virtual Reality*. Retrieved from https://www.useoftechnology.com/virtual-reality-history/

Tannenbaum, S., & Yukl, G. (1992). *Training and Development in Work Organizations* (Vol. 43).

Turner, J. (1995). The Influence of Classroom Contexts on Young Children's Motivation for Literacy. *Reading Research Quarterly, 30*(3), 410-441. doi:10.2307/747624

Tvedt, S., Oltedal, H., Batalden, B. M., & Oliveira, M. (2018). Way-finding on-board training for maritime vessels. *Entertainment Computing, 26*, 30-40. doi:10.1016/j.entcom.2018.01.002

VanLehn, K. (1996). Cognitive Skill Aquisition. *Annual Review Psychology, 47*, 513-539. doi:https://doi.org/10.1146/annurev.psych.47.1.513

Venkatesh, V., Thong, J., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly, 36* (1), 157-178.

Vincenzi, D. A., Mouloua, M., Wise, J. A., & Mouloua, M. (2008). *Human Factors in Simulation and Training* Baton Rouge, UNITED STATES: Chapman and Hall/CRC.

Vissgren, J. (2017, 14.03.2017). Virtual Reality skal gjøre norske barn flinkere i matematikk. *InnoMag*. Retrieved from http://www.innomag.no/virtual-reality-skal-gjore-norske-barn-flinkere-i-matematikk/

Wadhera, M. (2016, 09.05.2016). The Information Age is over; welcome to the Experience Age. *TC Crunch Network*. Retrieved from https://techcrunch.com/2016/05/09/the-information-age-is-over-welcome-to-the-experience-age/

Wang, T. (2017). *The Effect of Virtual Reality on Learning Motivation and Academic Performance.* (Master of Science), Emporia State University, Retrieved from http://hdl.handle.net/123456789/3565

Webster, R. (2015). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments, 24*(6), 1319–1333. doi:10.1080/10494820.2014.994533

Whitler, K. A. (2014, 17.07.2014). Why Word Of Mouth Marketing Is The Most Important Social Media. *CMO Network, Forbes*. Retrieved from https://www.forbes.com/sites/kimberlywhitler/2014/07/17/why-word-of-mouth-marketing-is-the-most-important-social-media/#33d3679054a8

Wingrove, M. (2018, 02.01.2018). Ten Technologies to Shake Up Maritime in 2018. *Marine Electronics and Communications*. Retrieved from http://www.marinemec.com/news/view,ten-technologies-to-shake-up-maritime-in-2018_50317.htm

Wyner, G. (2014, March 2014). The Power of Word of Mouth. *Marketing News*. Retrieved from https://www.ama.org/publications/MarketingNews/Pages/The-Power-of-Word-of-Mouth.aspx

8. Appendix

8.1. Poster/Flyer

8.2. Experimental Script(s)

8.2.1. Script for experimental group (HMD VR setup)

Task	Script	Time
Entry/ welcome	Hi and welcome to the TARG lab. Thank you for participating. This is where the VR experience will take place. *introduction* In this VR exercise you will be asked to fill out some information forms before and after the experience. You will participate in a VR exercise lasting approximately 20 min. The entire process will take approx. 90 min. Before entering the VR exercise, you will go through a familiarization period to become familiar with the system and the controls. We will be measuring your heart rate by a belt worn around your torso and skin conductance by a wrist band. After you have completed the 20 min in VR, you will fill out a few short questionnaires and a short exit interview will be conducted to hear more about your experience. We will explain more details as we go along. Please set your phone in airplane mode and put it in your jacket, to avoid distractions.	5 min
Consent form	Before we begin, we must ask you to fill out this consent form. This is to confirm that you are aware of what you are participating in, and that we will be collecting data from this experiment to use in our papers. Please notice the part about not speaking about the experiment with anyone else, as they can be potential participants.	10 min
Questionnaire	Next is a short entry questionnaire to get some basic information about you, your previous gaming experience and health. Please fill this out as correctly as possible.	5 min
Fitting	To enable measurement of your bodily response to the exercises, we will need you to wear this wristband/ watch (right hand) and the belt.	5 min

	I will help you put on the wrist band (attaching the wristband). And I will show you how to put on the belt, so you can do it yourself. This is a standard fitness device. Just make sure that the sensor is positioned on the centre of your torso, right below your chest. The rubber part of the belt will be moist, to achieve the best contact with your skin. If you want some assistance positioning the belt, I can help you.	
Baseline	To be able to analyse and compare the measurements we get from the equipment you are now wearing, we need to measure your base line. This means that we need to measure how your body is when you are relaxed and not doing anything. To get accurate measurements, we will need you to sit as still as possible for 5 minutes. Feet flat on the floor, and arms resting on the chair or in your lap. Please do not use your phone or any other distracting device. You can study the picture of the controllers. ***Press wrist band button***	5 min
Familiarisation	***Press wrist band button*** Short explanation of controllers: Left: The round button works much like a mouse pad, you can swipe and also click with it. The trigger button is used for moving around by teleportation. When you hold the button, an icon will appear on the ground in either green or blue to indicate where you end up. When you release the button, you move. The button on the side is the “back-button” for exiting menu’s aso. Right: Here you only need to worry about the round touch-pad. By pointing the controller toward objects, you can pick things up or open drawers by clicking the middle of the pad once. Now you will go through a trial period in VR, to familiarise with the system and controls. Do not do anything, until we tell you to. Afterwards you will be able to do what you want for a few minutes.	10 min

	Please put the headset on. (Assisting with putting on headset, and correct the lenses/straps) Before we begin: You are now standing in the middle of the room. If you raise your hands out to the side, you will see the grid of the room. This is the limits of your space in the real world. Please move over to the wall/grid, just to see where the wall is… Position yourself in the middle of the room again, before we begin. Introduction to location: You are, as you see, in a deserted town. In this task there are no other people or creatures. *Explain the scenario: * Your first task is to navigate across the street. By pressing the trigger button on the left controller and point in front of you. Please navigate to the house across the street and enter it. In the middle of the room there is a table and a shelf with cans. Position yourself behind the table facing the shelf with the cans. As this is a virtual environment, you can see that you cannot physically interact with the table, your hand will go right through it. Use the gun in your right hand to shoot down the cans with the trigger button on the right controller. Now, walk into the hallway and into the room on your right. In this room there is a suitcase and a cabinet. With your right hand, point at the suitcase, you can see there is something inside. If you look at the circle initiated below the inventory area, the position of your thumb on the pad is indicated. You can use this to ensure that your thumb is in the centre of the button. Press the middle of the round touch-pad on the right controller to pick up the item and get it in your inventory. As you can see, the suitcase is now empty. Behind you, on top of the cabinet, there is a bobble head figure to pick this up, aim at the figure with the right hand and press once in the middle of the touchpad. If you press up on the	

	touch-pad, you open a menu. This might happen when you are trying to push the middle to pick things up. This menu is hard to use, thus the best thing to do when this happens is to exit the menu by pressing one of the side buttons. Raise your left arm. Press up on the touch pad and swipe up to get to misc. Here you can see the bobble heads in your inventory. Look at the cabinet and aim with your right hand. If you miss the middle of the touchpad and press up, you will enter a menu. Press up on the touchpad to see this. If you at any time have pushed a button opening a new screen or something similar, you can press the button on the side of your right controller (with your middle finger). This will take you back. Now you can leave the house. Teleport over to the big tree. From here you see a blue house. Please enter the house. Here you can play around, search and find things, collect or just pick them up. If you click one short click, it collects things into your inventory. If you hold down the touch-pad button, you lift things up without collecting them. So be sure to only press once and not hold. You can find things inside boxes or cabinets here as well. ***Press wrist band button*** If you are hot, please remove a layer of clothing before we start. You will not be able to do it during the experiment. **How long does it feel like you have been in the game now?**	
Description	***Press wrist band button*** Explain the scenario and tasks: You will start with a navigation task in a small labyrinth. In this labyrinth there are ten figures hidden. Your task is to walk around, find and collect as many of these bobble heads as possible, within five minutes. Lift your left arm and navigate to misc in inventory, so that you can at any time lift your left arm to see how many figures you have collected. You will complete three rounds. The score and time from each round will be collected and compared.	5 min

	You can end the experiment at any time, if needed. There will be a two-minute break between each round. Now, given the “training” you have had so far. How do you think you will perform in this task on a level from 1 to 5? (1= very bad, 5= very good) ***Note down answer*** ***Start recording PC***	
Experiment	We/I will not be assisting you at all through the experiment, verbally or otherwise. So, you will have to navigate on your own, without our help. If you are struggling, try to go slow, and remember what you have learned in the familiarization. You have five minutes to find the figures, from now. ***Press wrist band button*** Finished ***Press wrist band button*** There will now be a two minute break Please stand and look straight ahead so I can load the scenario. The time starts now. ***Press wrist band button*** Finished There will now be a two minute break ***Press wrist band button*** Please stand and look straight ahead so I can load the maze scenario again. The time starts now. ***Press wrist band button*** Finished.	20 min
Collection	***Press wrist band button*** The experiment is now completed. Please remove the equipment and sensors.	3 min

Questionnaire	Now as the experiment is completed, we want you to fill out some questionnaires about your experience and how you felt about the exercise. 1. Questionnaire subconscious learning 2. Quality of experience questionnaire. 3. Skill acquisition and motivation questionnaire. 4. Simulator sickness questionnaire.	10 min
Exit interview	The very last thing we wish you to participate in is a short exit interview. This is to get additional information and feedback on your experience. This will be audio recorded.	10 min
“Warning”	Please do not talk about the procedure and tasks in the experiment with any other potential participant, as this may affect the outcome of this study.	
Complete		90 min

8.2.2. Script for control group (Desktop setup)

Task	Script	Time
Entry/ welcome	Hi and welcome to the TARG lab. Thank you for participating. This is where the experiment will take place. *introduction* In this exercise you will be asked to fill out some information forms before and after the experience. You will participate in an exercise lasting approximately 20 min. The entire process will take approx. 90 min. The first part will be on desk top, and the second part in VR. Before entering the exercise, you will go through a familiarization period to become familiar with the system and the controls. We will be measuring your heart rate by a belt worn around your torso and skin conductance by a wrist band. After you have completed the 20 min exercise, you will fill out a few short questionnaires and a short exit interview will be conducted to hear more about your experience. We will explain more details as we go along. Please set your phone in flight-mode and place it in you jacket, to avoid distractions.	5 min
Consent form	Before we begin, we must ask you to fill out this consent form. This is to confirm that you are aware of what you are participating in, and that we will be collecting data from this experiment to use in our papers. Please notice the part about not speaking about the experiment with anyone else, as they can be potential participants.	10 min
Questionnaire	Next is a short entry questionnaire to get some basic information about you, your previous gaming experience and health. Please fill this out as correctly as possible.	5 min
Fitting	To enable measurement of your bodily response to the exercises, we will need you to wear this wristband/ watch (right hand) and the belt.	5 min

	I will help you put on the wrist band (attaching the wristband). And I will show you how to put on the belt, so you can do it yourself. This is a standard fitness device. Just make sure that the sensor is positioned on the centre of your torso, right below your chest. The rubber part of the belt will be moist, to achieve the best contact with your skin. If you want some assistance positioning the belt, I can help you.	
Baseline	To be able to analyse and compare the measurements we get from the equipment you are now wearing, we need to measure your base line. This means that we need to measure how your body is when you are relaxed and not doing anything. To get accurate measurements, we will need you to sit as still as possible for 5 minutes. Feet flat on the floor, and arms resting on the chair or in your lap. Please do not use your phone or any other distracting device. You can study the picture of the controllers. ***Press wrist band button***	5 min
Familiarisation	***Press wrist band button*** Short explanation of keyboard: W is moving forward. A and D is moving sideways. S is moving backwards. To pick up things you press E. TAB is inventory. Esc is the back-button. You can change your point of view/ look around by moving the mouse. Now you will go through a trial period, to familiarise with the system and controls. Do not do anything, until we tell you to. Afterwards you will be able to do what you want for a few minutes. Please sit down. Introduction to location:	10 min

	You are, as you see, in a deserted town. In this task there are no other people or creatures. *Explain the scenario: * Your first task is to navigate across the street. This is done by using the keys to move, and the mouse to change view point. Please navigate to the house across the street and enter it. In the middle of the room there is a table and a shelf with cans. Position yourself behind the table facing the shelf with the cans. Use the gun in your right hand to shoot down the cans, by pointing the mouse towards it and pressing the left mouse button. Now, walk into the hallway and into the room on your right. In this room there is a suitcase and a cabinet. If you use the mouse to “point” toward the suitcase, you can see there is something inside. Press the E button to pick it up and put it in your inventory. As you can see, the suitcase is now empty. Behind you, on top of the cabinet, there is a bobble head figure to pick this up, aim at the figure with the mouse and press E. To see into your inventory, press TAB. In MISC you can see the bobble heads in your inventory. If you at any time have pushed a button opening a new screen or something similar, you can press the ESC button to go out of the menu. Now you can leave the house. Walk over to the big tree. From here you see a blue house. Please enter the house. Here you can play around, search and find things, collect or just pick them up. If you click one short click, it collects things into your inventory. If you hold down the E button, you lift things up without collecting them. So be sure to only press once and not hold, if you wish to save them in your inventory. You can find things inside boxes or cabinets here as well. ***Press wrist band button*** If you are hot, please remove a layer of clothing before we start. You will not be able to do it during the experiment.	

	How long does it feel like you have been in the game now?	
Description	***Press wrist band button*** Explain the scenario and tasks: You will start with a navigation task in a small labyrinth. In this labyrinth there are ten figures hidden. All figures are hidden inside the walls of the maze, no figures are placed on the outside. Your task is to walk around, find and collect as many of these bobble heads as possible, within five minutes. Press TAB and navigate to misc in inventory, so that you can at any time press TAB to see how many figures you have collected. You will complete three rounds. The score and time from each round will be collected and compared. There will be a two-minute break between each round. You can end the experiment at any time, if needed. Now, given the “training” you have had so far; How do you think you will perform in this task on a level from 1 to 5? (1= very bad, 5= very good) ***Note down answer*** ***Start recording PC***	5 min
Experiment	We/I will not be assisting you at all through the experiment, verbally or otherwise. So, you will have to navigate on your own, without our help. If you are struggling, try to go slow, and remember what you have learned in the familiarization. You have five minutes to find the figures, from now. ***Press wrist band button*** Finished ***Press wrist band button*** There will now be a two-minute break Please stand and look straight ahead so I can load the scenario. The time starts now. ***Press wrist band button***	20 min

	Finished There will now be a two-minute break ***Press wrist band button*** Please stand and look straight ahead so I can load the maze scenario again. The time starts now. ***Press wrist band button*** Finished.	
Collection	***Press wrist band button***	3 min
Questionnaire	Now we want you to fill out some questionnaires about your experience and how you felt about the exercise. 1. Questionnaire subconscious learning 2. Quality of experience questionnaire. 3. Skill acquisition and motivation questionnaire. 4. Simulator sickness questionnaire.	10 min
Testing VR	Now you can go through a round in the VR system before we conduct the exit interview. If we have more time, you can also try another game. The experiment is now completed. Please remove the equipment and sensors.	5 min
Exit interview	The very last thing we wish you to participate in is a short exit interview. This is to get additional information and feedback on your experience. This will be audio recorded.	10 min
“Warning”	Please do not talk about the procedure and tasks in the experiment with any other potential participant, as this may affect the outcome of this study.	
Complete		90 min

8.3. Introduction Questionnaire

Introduction Questionnaire

PERSONAL DEMOGRAPHICS

1. Name
2. Gender
3. Age
4. Nationality
5. E-mail address
6. Current work status (i.e. student, employed, self-employed, etc.)

GAMING EXPERIENCE

1. Do you have previous experience with video games? (Computer-based, Console-based or other)

No experience	Very little	Some experience	Play regularly (once a month)	Play often	Play several times a week	A lot of experience (play every day)
0	1	2	3	4	5	6

2. Did you play more before than you do now?
   a. No
   b. Yes

   c. If yes, how many years ago was the period you played most?

3. Do you enjoy puzzle games (strategy, problem solving, maze)?
   a. Yes
   b. No

4. How often have you played video games during the last 12 months?

Never	Once every other month	About once a month	About once a week	Several times per week	Every day for less than two hours	Every day for more than two hours
0	1	2	3	4	5	6

a. If other than 0, approximately how many individual playing sessions per week?

b. If other than 0, approximately how many total hours per session do you play?

5. What gaming systems have you used personally? (Please circle)
   a. Nintendo 64/ X-box/ Atari
   b. VR systems
   c. PC
   d. Wii/ Play Station Move
   e. Mobile phone games
   f. Other (please specify):____

PREVIOUS SIMULATOR EXPERIENCE

6. Have you used any type of simulator before?
   a. No
   b. Yes
      i. What kind/for what purpose?

      ii. What was the exercise/game?

7. Have you previously used a head mounted virtual reality system (VR headset/goggles)?
   a. No
   b. Yes, I have used:
      i. HTC Vive
      ii. Occulus Rift
      iii. Google Cardboard

iv. Samsung Gear VR
v. Playstation VR
vi. Not sure witch
vii. Other:____________________

HEALTH BACKGROUND

8. Have you ever suffered from epileptic seizures?
   a. Yes, sometimes
   b. No
      i. If yes, when did this happen?

9. Have you suffered from motion sickness before?
      i. Yes
      ii. No
   b. If yes, please explain:

10. Do you easily get car-sick?
   a. Yes
   b. No

EXPECTATIONS FOR STUDY

11. How do you think this experience will be?

Boring	Not very interesting	Of little impact to me	Ok	Somewhat interesting	Highly interesting	Life altering
0	1	2	3	4	5	6

12. What do you expect from this study?
   a. Please explain with a few words:__________________________________

   ____

13. What motivated you to participate in this study?

8.4. Intrinsic motivation Questionnaire

Questionnaire

1. I was pretty skilled at this activity. (Please circle)

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

2. While I was doing this activity, I was thinking of how much I enjoyed it.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

3. It was important for me to do well in this task.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

4. I found the task easy to understand.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

5. I felt I was part of the virtual world.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

6. I did not feel nervous at all while doing this.

Not at all true			Somewhat true			Very true

1	**2**	**3**	**4**	**5**	**6**	**7**

7. I am satisfied with my performance in this task.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

8. I enjoyed doing this activity very much.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

9. I put a lot of effort into this.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

10. I did this activity because I wanted to.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

11. I felt pressured while doing these.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

12. I was disappointed when the time was up, as I wish I could have continued.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

Comment: ______________________________________

13. I didn’t really have a choice about doing this task.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

14. I would be willing to do this again because it has some value to me.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

Comment: __________________________________

15. The experience lived up to my expectations.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

Comment: __________________________________

16. I think that doing this kind of activity is useful for training/learning.

Not at all true			Somewhat true			Very true
1	**2**	**3**	**4**	**5**	**6**	**7**

17. What was the most interesting/exiting aspect of the experience?

Please describe: ____

18. Please define your overall experience in one word:___________________

8.4.1. Results intrinsic motivation measurements

Table 12: Measures of central tendencies: Intrinsic Motivation Questionnaire results (all 12 items included)

Subscale	Item	Treatment	Mean	Variance	SD	Range
Interest/ Enjoyment	2	HMD VR	4,875	2,375	1,541	1-7
		Desktop	4,000	3,391	1,842	1-7
	6	HMD VR	6,042	,824	,908	3-7
		Desktop	4,958	3,607	1,899	1-7
Perceived competence	1	HMD VR	4,708	1,346	1,160	3-7
		Desktop	5,333	1,797	1,341	2-7
	5	HMD VR	5,250	1,239	1,113	3-7
		Desktop	5,833	1,536	1,240	2-7
Effort/ Importance	3	HMD VR	5,750	1,239	1,113	4-7
		Desktop	5,167	1,971	1,404	2-7
	7	HMD VR	5,708	1,172	1,083	3-7
		Desktop	5,125	2,288	1,513	1-7
Pressure/ Tension	4	HMD VR	2,583	2,514	1,586	1-6
		Desktop	3,083	3,906	1,976	1-7
	9	HMD VR	2,417	5,123	2,264	1-7
		Desktop	2,583	3,819	1,954	1-7
Perceived choice	8	HMD VR	6,583	,862	,929	3-7
		Desktop	6,000	2,261	1,504	1-7
	10	HMD VR	6,167	3,275	1,810	1-7
		Desktop	6,292	1,346	1,160	4-7
Value/ Usefulness	11	HMD VR	5,750	1,587	1,260	2-7
		Desktop	3,958	4,303	2,074	1-7
	12	HMD VR	5,958	1,346	1,160	4-7
		Desktop	4,875	2,723	1,650	2-7

8.5. Exit Interview Guide

Exit interview Guide

1. How was the experience?
    a. What emotions did you feel? (scared/happy/stressed/etc.)
    b. Was it enjoyable? If so why?
    c. What aspects did you not enjoy?
    d. Did you learn anything?
    e. Did you have any strategy in the task?

2. What do you feel could be done to improve the experience?

3. How did the equipment feel?
    a. Any discomfort?

4. What effect do you think VR can have on education in the future, if any?

5. Do you see any other applications for a VR system?

6. Any other comments on the experiment in general?