UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

ALINE MENIN

A Proposal of Design Guidelines for Immersive Serious Games

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Computer Science

Prof. Dra. Luciana Nedel
Advisor

Porto Alegre, November 2016

CIP – CATALOGING-IN-PUBLICATION

Menin, Aline

A Proposal of Design Guidelines for Immersive Serious Games / Aline Menin. – Porto Alegre: PPGC da UFRGS, 2016.

100 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2016. Advisor: Luciana Nedel.

1. Immersive serious games. 2. Perceptual learning. 3. Virtual reality. 4. Physical immersion. I. Nedel, Luciana. II. Título.

"Once we accept our limits,
we go beyond them."
— Sir Albert Einstein

ACKNOWLEDGMENTS

Firstly, a special thank you to my advisor, Luciana Nedel, and to my co-advisor, Rafael Torchelsen, for the opportunity of working with them and for all the support. Their hard work and commitment were impressive. I learned a lot with them and it was an amazing experience of work.

I would also like to acknowledge the members of my dissertation committee: Professors Fatima Nunes, Carla Freitas, and Marcelo Pimenta. They contributed with valuable suggestions to this work. And the Institute of Informatics (INF) that, with its infrastructure, made the execution of this research possible. Additionally, I also acknowledge CNPq-Brazil for the financial support.

To my family for the unconditional love and for always believe in my potential and support me in my life choices. Especially my parents for being my model of honest and kind people. I would not have gotten it without their special support. Finally, I acknowledge my colleagues of the Computer, Visualization and Interaction Lab that always helped me when I needed. I learned a lot with them and they became great friends in this last two years. Thank you all.

A Proposal of Design Guidelines for Immersive Serious Games

ABSTRACT

Nowadays, several application areas are adopting serious games as an ethical, safe and low-cost alternative to performing dangerous tasks, such as training of firefighters, and for medical surgeries. The employment of Virtual Reality in serious games helps to provide virtual environments much more realistic and better immersive experiences to the user. However, some VR systems can still cause sickness symptoms, such as nausea and headaches, and some of them can be expensive, such as CAVEs.

This thesis investigates the effects of physical immersion on serious games with perceptual learning purposes. Its main goal is to identify a set of design guidelines to help developers on choosing the appropriate level of immersion to be used in serious games in order to assure it is effective and comfortable for users. The main hypothesis is that higher levels of immersion improve serious games outcomes, but moderate immersion may be also adequate with the advantage of causing less simulation sickness on users.

We organized a survey of the literature to better understand how physical immersion is being currently used on serious games, how its effectiveness has been assessed, and how immersion impacts on the game outcomes and usability. Then, we conducted two empirical user studies looking for investigating the effects of the display, interaction and locomotion fidelity on users' perception and knowledge retention. For the experiments, we adapted a serious game previously developed for risk assessment and developed a new one to educate workers in electricity-line maintenance on safety procedures that need to be followed during electric installations.

Results showed that, in general, display fidelity has an effect on risk perception when searching for non-obvious risks. Higher display fidelity has better performance on complex risks identification. Interaction and locomotion fidelity did not show a significant difference in perceptual learning. Naturalness also presented higher workload, but in the same conditions the correctness of tasks was high and subjects recall the procedures in both post- and retention-test. Therefore, knowledge retention is not impacted by the workload imposed by the interactive technique. From these experiments, we elaborated and discussed a set of design guidelines that can be considered for the choice of the appropriate physical immersion to be used on the development of new serious games.

Keywords: immersive serious games, perceptual learning, virtual reality, display fidelity, interaction fidelity.

Uma Proposta de Orientações de Design para Jogos Sérios Imersivos

RESUMO

Esta dissertação investiga os efeitos da imersão física em jogos sérios com propósitos de aprendizado perceptivo. O principal objetivo é identificar um conjunto de orientações de *design* para ajudar desenvolvedores a escolher o nível apropriado de imersão a ser usado em jogos sérios a fim de assegurar que o jogo seja efetivo e confortável para os usuários. A hipótese principal é que altos níveis de imersão melhoram os resultados dos jogos sérios, mas níveis de imersão moderados também podem ser adequados com a vantagem de que causa menos mal-estar nos usuários.

Nós organizamos uma revisão da literatura para melhor entender como a imersão física está sendo usada atualmente em jogos sérios, como a efetividade tem sido avaliada, e como a imersão afeta os resultados do jogo e a usabilidade. Então, nós conduzimos dois estudos de usuário empíricos para investigar os efeitos da fidelidade de *display*, interação e locomoção na percepção dos usuários e retenção de conhecimento. Para os experimentos, nós adaptamos um jogo sério previamente desenvolvido para análise de percepção de riscos, e desenvolvemos um novo jogo para educar trabalhadores sobre procedimentos de segurança que precisam ser seguidos durante instalações elétricas na linha de manutenção elétrica.

Os resultados mostraram que, em geral, a fidelidade de *display* tem impacto na percepção de riscos quando procurando por riscos que não são óbvios. Alta fidelidade de *display* tem melhor desempenho na identificação de riscos complexos. A fidelidade de locomoção e interação não apresentou diferenças significantes no aprendizado perceptivo. Alta fidelidade também apresentou maior *workload*, mas nas mesmas condições a corretividade das tarefas foi maior e os participantes lembraram dos procedimentos tanto no pós-teste quanto no teste de retenção. Portanto, a retenção de conhecimento não foi afetada pelo *workload* introduzido pela técnica interativa. A partir destes experimentos, nós elaboramos e discutimos um conjunto de orientações de *design* que podem ser considerados para a escolha do nível apropriado de imersão física para ser usado no desenvolvimento de novos jogos sérios.

Palavras-chave: jogos sérios imersivos, aprendizado perceptivo, realidade virtual, fidelidade de exibição, fidelidade de interação.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

VR	Virtual Reality
SG	Serious Games
ISG	Immersive Serious Games
VE	Virtual Environment
IVE	Immersive Virtual Environment
FOV	Field of View
WIP	Walking-in-Place
CAVE	Cave Automatic Virtual Environment
HMD	Head Mounted Display
DOF	Degrees of Freedom
3D	Three dimensional
SSQ	Simulator Sickness Questionnaire
SUS	Slater, Usoh, and Steed
SCL	Skin Conductance Level
EDA	Electrodermal Activity
BVPA	Blood Volume Pulse Amplitude
PPG	Photoplethysmogram
EMG	Electromyography
FOV	Field of View
p	p-value from statistical analysis
HIHL	High Interaction and High Locomotion
HILL	High Interaction and Low Locomotion
LIHL	Low Interaction and High Locomotion
LILL	Low Interaction and Low Locomotion
NASA TLX	NASA Task Load Index
GEQ	Game Engagement Questionnaire

M	Mean
SD	Standard Deviation
mocap	Motion Capture

CONTENTS

1 INTRODUCTION

1.1 Motivation

Serious games are video games developed for training, learning, simulation or publicity purposes. It has the power to engage and charm the users to a specific proposal, such as the development of new skills (CORTI, 2006). It is more than just stories, arts, and software, but is the psychology involved that transform a video game into a serious game (ZYDA, 2005).

Using serious games is a safe, ethical and cost-effective alternative for training dangerous tasks. Wattanasoontorn et al. (WATTANASOONTORN et al., 2013), Rego et al. (REGO; MOREIRA; REIS, 2010), Marin et al. (MARIN; NAVARRO; LAWRENCE, 2011), for instance, report serious games for health-care for purposes of education, rehabilitation, and medical and psychological treatment. In the surgical field, they can be used to train both technical and non-technical skills, as showed by Graafland et al. (GRAAFLAND; SCHRAAGEN; SCHIJVEN, 2012). The fire service uses simulations for the training of firefighting skills and incident command coordination (WILLIAMS-BELL et al., 2015). Most of these games are played on standard desktops or mobile devices. However, the use of virtual reality in video games showed positive effects on users' performance (KULSHRESHTH; LAVIOLA JR, 2015), suggesting that the immersion and realism provided by the VR systems may improve the effectiveness of the serious game as well.

According to Chalmers et al. (CHALMERS; DEBATTISTA, 2009), serious games needs a high level of realism to ensure that training and education in virtual environments (VE) are equivalent to the real world. Furthermore, they should be able to simulate all five human body senses simultaneously. High levels of immersion are employed to simulate real world situations, in particular for purposes of training and learning. Chittaro et al. (CHITTARO; BUTTUSSI, 2015) presents a serious game for educating passengers about aviation safety that allows players to experience a serious aircraft emergency through an HMD with the goal of surviving it. The training of the firefighter skills is also dangerous and could be addressed by a serious game, such as the one presented by Backlund et al. (BACKLUND et al., 2007), which uses a CAVE based system and motion tracking to provide a free interaction with the game. Through these and other studies, it is possible to observe that higher levels of immersion improve the game effectiveness and user performance. However, it may be expensive and cause discomfort, such as *cybersickness*, which presents symptoms such as nausea, vomit, headache, and disorientation. Regarding these issues, it is relevant to investigate the effects of moderate levels of immersion in the effectiveness of immersive serious games to understand whether they can be an efficient alternative to introducing immersion and realism to serious games.

The research presented in this thesis examines these issues, considering physical immersion features, to define the appropriate VR system to provide immersion in serious games for learning, while being comfortable and efficient. The research includes two controlled experiments and a literature review that investigates how the levels of immersion affect the effectiveness, performance, and satisfaction of users of serious games. This investigation includes consideration for the features of display systems that could have an impact on the effectiveness of perceptual learning and the degree of interactive control for exploring the virtual world and acquire knowledge. By considering factors of characteristics of display systems, our main goal is to provide guidelines to support design decisions for physical immersion required for creating serious games for learning.

1.2 Background

All games, from video games to serious games, are defined by five components (WATTANASOONTORN et al., 2013). The *gameplay*, which sets the rules that connect the player and the game. The *challenge* which determines the obstacles that the player have to pass through to get the reward. The *interaction* represents the communication between the player and game, it can be physical, visual and audio. The *objective* is something that the player will accomplish after crossing through the challenges in the game. There are two types of objectives: the *explicit*, which addresses the entertainment, nature of every game, and the *implicit*, which includes increasing skills and abilities, gaining knowledge or acquiring experience. The serious games have implicit objectives, which differentiate them from games developed for entertainment.

Virtual reality has been widely used to make the serious games more realistic. It consists of a virtual world, immersion, sensory feedback, and interactivity (SHERMAN; CRAIG, 2003). A *virtual world* is a content of a given medium, it can exist without being displayed in a VR system (e.g., CAVEs and HMDs). The *immersion* can be mental and physical. The mental immersion refers a mental state, a feeling of being involved in the experience. The physical immersion refers to the properties of the VR system to replace or augment the stimulus of the participant's senses. These stimuli are the *sensory feedback*, an essential ingredient to virtual reality, providing senses such as sight, hearing, touch and smell. The feedback is predominantly provided through visual and auditory channels, and the addition of haptic feedback improves the sense of presence, enabling the user "feel" and "touch" virtual objects. The *interactivity* is in charge of communicating the VE with the user, which intends to be the most natural possible, such as through motion tracking.

The technology employed to develop VEs with interaction as natural as possible and with feedback for most of the human senses is still new and sometimes uncomfortable. This factor may add an extra cost to the user to achieve a particular level of performance in the proposed task, which is known as workload (HART; STAVELAND, 1988). Therefore is relevant that the VE be able to provide naturalness without increase the workload to reach the goals of the serious game.

From a literature review on immersive serious games, presented in Section 2, it was observed that most of the ISGs developed so far have purposes of learning and training. Learning is defined here as the process of educating users about some content through a game, while training is the process of practicing the knowledge acquired to improve users' skills. Both methods aim to provide knowledge for the user, which can be achieved by many ways.

Perceptual learning is the early form of learning that a person has contact. Perceptual development is a process of learning about affordances, becoming better able to detect appropriate supports and resources and discovering new affordances as action capabilities change (GIBSON, 1992). The acquisition of new skills – looking, reaching, walking, weightlifting, swimming, driving, sewing, handwriting – produces new affordances to be learned throughout life (ADOLPH; KRETCH, 2015). The acquisition of information can happen through different modalities (e.g., looking, hearing, smelling, and so on). Nedel et al. (NEDEL et al., 2016), for instance, present a serious game for risk perception analysis, in which it is highly relevant that the user can perceive dangerous situations in the virtual world. Another example of the importance of perception is showed in Ragan et al. (RAGAN et al., 2015), which present a serious game for military personnel training their ability of visually searching for signs of dangerous activity and threatening individuals while driving through urban streets. In both situations, the perception can be a way to evaluate the users' abilities and to training it, acquiring knowledge.

1.3 Research Questions

Though previous studies showed that higher immersion has positive effects on serious games for learning, some VR systems still are expensive and may be harmful to the users. Besides that, most of these works focus on the benefits of display fidelity. This thesis also focuses on the effects of interaction and locomotion fidelity. Our goals are to improve the understanding of how physical immersion impacts on the learning provided by serious games and identifies a set of design guidelines for low-cost, comfortable and efficient serious games. More specifically, this thesis is guided by the following research questions:

1. **What are the effects of display fidelity on user perception?**

The benefits of higher display fidelity on serious games had been already demonstrated ((NAZIR et al., 2012), (NAZIR; KLUGE; MANCA, 2014), (CHITTARO; BUTTUSSI, 2015), (HUPONT et al., 2015)). It also extends to memorization, in which it is was proved that higher visual immersion improves spatial knowledge and allows users to use a spatial memorization strategy for procedure memorization (HIROSE et al., 2009). It is also beneficial for target detection, as showed by Ragan et al. (RAGAN et al., 2015). As our study aims to investigate the effects of immersion on perceptual learning, it is relevant to understand how the display fidelity affects the user perception of the virtual environment. This research question also aims to answer the following issues:

(a) **Can moderate display fidelity provide as much effectiveness as higher display fidelity?**

Though higher display fidelity has benefits in serious games, some VR systems are expensive and cause symptoms of *cybersickness*. We believe that high display fidelity will present better user performance for perception, once that the field of view allow the user to have a larger consciousness of the environment.

Therefore, we aim to investigate whether the moderate display fidelity is effective for user perception, which allows using low-cost displays while maintaining the effectiveness.

(b) **How the levels of immersion affect the simulator sickness symptoms and how this affects user performance?**

One of our concerns is related to the comfort offered by the physical immersion. Therefore, we aim to investigate the impacts of display fidelity on the severity of simulator sickness symptoms to identify which level of immersion is more appropriate to provide comfortable serious games.

The user performance is defined here as trial completion time, the number of risks selected in the virtual environment, and path traveled. It is relevant to investigate how the severity of simulator sickness impacts these issues, to verify whether the sickness caused by the display fidelity is prejudicial for user performance.

2. **What are the effects of interaction and locomotion fidelity on knowledge retention?**

Though most of the previous works investigate the effects of visual immersion on serious games effectiveness, the interaction is also an important component for providing the engagement between player and game. Though it is not possible yet to provide, simultaneously, feedback for the five human senses, the use of highly natural interaction and locomotion may increase the immersion in the virtual environment. The naturalness, then, should improve the user performance in serious games. Previous works have shown that having a physical reference helps the user to be more precise on memory recall (MINE; BROOKS JR; SEQUIN, 1997) and that naturalness is not always a necessary component of an effective technique (BOWMAN; HODGES, 1997). Therefore, we aim to investigate the impacts of low and moderate interaction and locomotion fidelity on memorization for knowledge retention. The following questions support this research question:

(a) **How naturalness impact the simulator sickness symptoms and what are the effects of it on knowledge retention?**

As for display fidelity, it is important to guarantee that the interaction with the game and the navigation inside the virtual environment would not cause the player sick. It is a relevant issue for that the user can play the game and acquire knowledge. We aim to investigate how the naturalness of these techniques impact the sickness symptoms.

Knowledge retention is the recall of the acquired learning, in this case, the learning obtained through the serious game. Therefore, it is relevant to investigate the effects of simulator sickness symptoms to verify whether it may be prejudicial for it.

(b) **How the workload introduced by the interface impacts the user performance and perceptual learning?**

Interfaces with low naturalness may add workload to the game, which may be prejudicial for gaining of knowledge, considering that the player has to learn how to use a technique that is not natural in its everyday life. Although Mania et al. (MANIA et al., 2006) suggested that high workload may be beneficial for memorization, they did not prove it. Therefore, our study intends to investigate how the workload impacts on perceptual learning and user performance.

1.4 Contributions

Besides the validation of the hypotheses, the major contributions of this thesis are the following:

- A Literature Review that presents a comprehensive survey of investigations regarding the employment of immersive serious games, the physical immersion currently used, and the materials and methods applied to access aspects of effectiveness, usability and engagement of serious games (see Chapter 2).
- An investigation on the effects of physical immersion (e.g., display, interaction and locomotion) on user perception, *cybersickness* and knowledge retention (see Chapters 4 and 5.
- A set of design guidelines that help to chose the appropriate display, interaction, and locomotion fidelity to develop effective, comfortable and low-cost serious games for perceptual learning (see Chapter 6).

2 LITERATURE REVIEW

For the purpose of this thesis, a state-of-the-art review was conducted aiming to structure and organize the research concerning immersive serious games (ISG). Based on the concepts proposed by Keele (KEELE, 2007), we defined a review protocol, compose of four stages: background, research questions, search process and exclusion criteria. The description of each step of in the following subsections.

2.1 Background

In the context of a literature review, the background refers to the rationale for the survey. In summary, the content presented in this section introduces a thorough survey of the employment of immersive virtual reality in serious games and classify them towards a taxonomy of ISGs, which considers aspects of games' purpose, player and immersion level, and user studies features. Moreover, this review has provided the basis to build up this thesis.

2.2 Methodology

To define the scope of the study, we structured our analysis through a set of research questions (described below) and we follow some criteria to include papers in this review. It should present at least one immersive serious game, with or without validation through user studies. The game should give an enhanced virtual reality system according to description presented by Muhanna (MUHANNA, 2015). The VR system should provide body tracking to ensure physical immersion. The authors should appropriately describe the game/user study to allow the classification and analysis of the immersive serious game. The paper should not be a review of other papers.

The initial search for papers was based on keywords such as “serious games", “immersive serious games", “virtual environments", and “virtual reality". The databases searched were ACM (Association for Computing Machinery), IEEE (Institute of Electrical and Electronics Engineers), Springer, MIT Press, Europe, and Elsevier. Several application areas employ SGs, so we can find existing publications regarding this subject in Conferences, Workshops, and Symposium Proceedings, as well as in Journals that may be relevant for other areas besides virtual reality and games.

We analyzed each article according to its abstract and dropped or accepted it for a further analysis. The accepted ones were read to identify whether their content match to the scope of the study, and then to be cataloged. In a second search, we used the reference list presented in each relevant article. The last publications search date was December 2016.

2.2.1 Research Questions

1. How immersion influences serious game usability and effectiveness?

Hirose et al. (HIROSE et al., 2009) provide empirical evidences that higher levels of immersion can produce a measurable improvement in the performance of an abstract mental activity. However, because the VR displays still have some limitations, such as connection wires and low resolution, it may cause the user to feel sick or uncomfortable. Thus, we identified the current technological setting applied to provide immersion and their impacts on serious game usability and effectiveness.

2. How to conduct user studies to assess effectiveness, usability and immersion of SGs?

Previous works reported several benefits of using immersion to increase engagement and effectiveness. In order to validate such results and assure that immersion would not produce negative effects researchers frequently conduct user studies. The methods employed to run user studies are not standard, depending on the experiment/game objective, which might not be immersion related. We identified the experiment design and methods employed in the game's assessment, and how to evaluate immersion effects on effectiveness and usability.

3. Who is the public taking part on the user experiments? How the subject's profile affects the outcomes?

SGs have specific purposes, so their player should have a particular profile. A SG for treatment, for instance, requires patients with a given disease. However, recruiting people that match with the player profile is not an easy task, which cause the researchers to use university students on their experiments, particularly because they conduct the research in University research Labs and undergraduate students are easily available. Concerning the impact this profile difference may have on SG's validation, this paper correlates participants' profile, expected player's profile and experiment' results.

4. How immersive VR benefits and limits the SGs development?

The previous section pointed to some benefits of immersion, such as increasing engagement and sense of presence, and memorization improvement, as well as some advantages of using SGs for substituting or as an additional resource for training, learning, and treatment. Nevertheless, because immersive VR is a relatively new technology, it may present some constraints. Thus, we explore the current ISGs in order to identify the benefits and limitations of using immersive VR on SGs.

2.3 ISG Taxonomy

The Figure 2.1 presents the taxonomy built to classify the ISGs, which consists of three categories: ISG, Player, and User Study. The ISG category considers objective, level of immersion, and sensory feedback provided by the VR system:

- **Level of Immersion:** We defined three immersion levels based on the body parts tracked, the use of 3D stereo and the taxonomy for VR systems presented by Muhanna (MUHANNA, 2015) (see Table 2.1). To visualize the virtual world, the games use large-screens, three-walls, CAVE and HMD displays.

Body tracking grants interaction through physical immersion, which could be the head, the hands and fingers, the eyes, the torso, the feet and other body parts, or yet indirect tracking through props and platforms. A few games even provide interaction through touch or voice.

- **Feedback:** We identified three sensory feedback, i.e. haptic, visual, and aural, and performance feedback, which gives to the user its progress along the game. Sensory feedback is an essential element of virtual reality and depends on participants' physical position. Though it is mainly provided by audiovisual channels, haptic feedback has been applied in VR systems enabling the user to “touch" and “feel" virtual objects.

- **Objective:** The surveyed SGs have five purposes: to learn, to practice, to test, to treat, and for self-care. The outcomes of games to learn are knowledge acquisition or content understanding. Games for training aim to improve the user skills and to provide content understanding. Testing means to verify the user's reaction when facing a specific situation or to assess his/her knowledge about a given content. There are games for treatment of diseases, particularly focusing on rehabilitation and therapy. Finally, games for self-care frequently aims to encourage people to be more physically active.

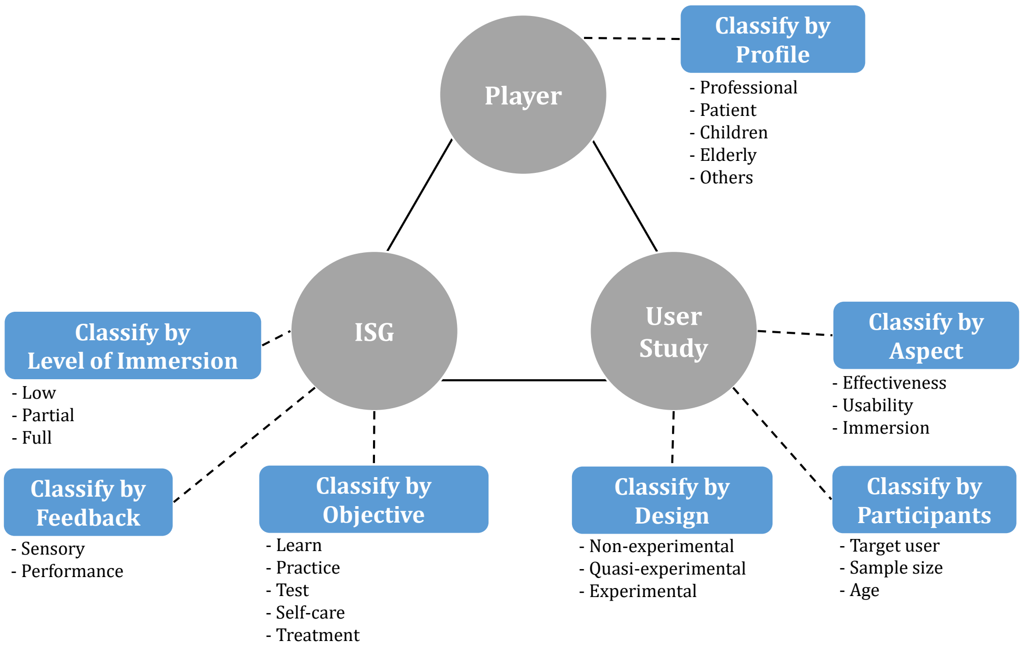


Figure 2.1 Classification of the ISGs.

According to the game objective we defined a set of profiles to classify the player: the professional category includes people such as machines and industrial operators, pilots, and surgeons; the patient refers to the people that have the specific disease treated by the game; children; elderly; and others, which includes people with a non-specific profile, such as the ones that play games for self-care.

We explore the user study under game's aspect assessed, experiment design and participants' profile:

Table 2.1 Levels of immersion classification based on the immersive VR features.

Level of Immersion	**Display**	**Direct body tracking**	**Head Tracking**	**3D Stereo**
Low	large-screen	no	no	no
	three-wall	no	no	no
Partial	large-screen	yes	yes/no	yes/no
	three-wall	yes	yes/no	yes/no
	CAVE	no	no	yes/no
Full	CAVE	yes/no	yes/no	yes/no
	HMD	yes/no	yes	yes

- **Aspect:** The features assessed on user studies depends on game's objective or study's hypotheses. We identified three aspects usually addressed: effectiveness, usability and immersion. Effectiveness refers to the users' ability to complete the tasks using the system, and the outcomes of such tasks. Usability concerns to the efficiency, i.e. the level of resources consumed to perform tasks, and satisfaction, i.e. the users' subjective reactions while using the system. Immersion refers to the engagement, i.e. a quality of user experience, a multidimensional construction characterized by aesthetic appeal, novelty, perceived challenge, feedback and control, attention, motivation, and affection (O'BRIEN et al., 2008).

- **Design:** There are typically three experiment designs to conduct user studies: the experimental design which allows comparison between groups random assigned; the quasi-experimental design which allows comparison between non-equivalent groups, i.e., there is no random assignment; and non-experiment design which has no comparison between groups and is usually used to study the effects of some intervention (TROCHIM; DONNELLY; ARORA, 2015). Among the studies ran under experimental and quasi-experimental designs, they usually compare features related to technology, participants' profile, and method to reach the game's objective (e.g., a traditional method of training vs. ISG).

- **Participants:** We crossed the participants' and the players' profile aiming to identify whether they were the game' target users. Particularly for games that do not require a specific player, i.e. the ones classified as "others", the user study may recruit students to take part in the experiment because of its feasibility and availability. In this case, we consider that the participants are *part of* the target public. Additionally, we explored their age and the sample size.

2.4 Classification Towards the ISG Taxonomy

2.4.1 Immersive Serious Games

This section presents the surveyed serious games with respect to the proposed classifications. Table 2.2 categorize the games and user studies according to the features addressed in the taxonomy presented before. From left to right, column (1) contains the reference of the article. Column (2) presents the level of immersion provided by the VR system. Column (3) presents the display device and column (4) informs whether it is 3D stereo. The following four columns are relative to the kind of feedback given to the user, i.e. haptic (H), visual (V), aural (A) and performance (P). The next two columns describe the tracked body parts. In the column (11) it is specified the game's objective. Column (12) classify the player into the categories presented earlier. Regarding the user study, column (13) presents the study design and column (14) presents the game's aspects evaluated. The final three columns introduce the participants' characteristics: sample size, target user and mean/range age, respectively. The participant could be considered as being (yes) or not (no) the target user, or it could be "part of" the target public. We based our analysis on 46 immersive serious games, collected from 51 papers. The majority of serious games (80%) are fully or partially immersive, from which 63% provides stereoscopic vision. However, almost half of them provide body tracking indirectly through specific tools, such as wands, treadmills, and walking platforms. The most tracked body parts are hands and torso, which helps to identify players' position and orientation.

In Figure 2.2d we can observe a special interest in developing serious games for practice and treatment, employing usually high levels of immersion. Themost used displays are HMDs and large-screens, being more common for treatment than practice. The games with other objectives are displayed mostly through CAVEs (Figure 2.2b). Regarding physical immersion, we observed that serious games for treatment use mainly direct body tracking, probably because its application on rehabilitation, which requires higher degrees of freedom so the user can move the damaged body part to recover it.

All games provide sensory feedback, at least visually. As presented in Figure 2.2e, most of them provide aural and visual, or only visual feedback and a few games support all three. Performance feedback is widely employed in order to help players improve their performance along the game, being included in 33 (71%) games. There are certain games designed for treatment that provide performance feedback only to the doctor, not to the user. Through the years, serious games have been adapted in different application areas aiming to improve or replace the traditional procedures. In health, they are applied to treat diseases such as epilepsy (GREWE et al., 2013), sclerosis (PERUZZI et al., 2013), (PERUZZI et al., 2016), autism (CAI et al., 2013), rehabilitation of stroke patients (SEO et al., 2014), (KIM et al., 2016), anxiety (CRESCENTINI et al., 2016), and child obesity (JOHNSEN et al., 2014), and to train surgeons (FERRACANI; PEZZATINI; DEL BIMBO, 2014). They are also employed to teach safety procedures to children (SMITH; ERICSON, 2009) and plane's passengers (CHITTARO; BUTTUSSI, 2015). In industry, to train industrial operators (NAZIR et al., 2012), (NAZIR; KLUGE; MANCA, 2014), (NAZIR; MANCA, 2015) and firefighters (BACKLUND et al., 2007). In sports, so that soccer (STINSON; BOWMAN, 2014) and rowing players (VARLET et al., 2013) can improve their skills.

Table 2.2 Classification of the immersive serious games toward the taxonomy presented in Section 2.3.

Reference	Immersive Serious Game											User Study				
				Feedback				Interaction						Participants		
	Level of Immersion	Device	3D	H	V	A	P	Head Tracking	Body Tracking	Objective	Player	Design	Aspect	Sample Size	Target User	Age
(CHITTARO; BUTTUSSI, 2015)	Full	HMD	Y	N	Y	Y	Y	yes	indirect	learn and practice	others	Experimental	Effectiveness and Immersion	48	Part of	24.19
(BACKLUND et al., 2007)	Full	CAVE	Y	N	Y	Y	Y	no	indirect, body movements and orientation	learn and practice	professional	Experimental	Effectiveness, Immersion and Usability	31	Yes	24
(BACKLUND et al., 2008)	Partial	CAVE	N	Y	Y	Y	N	no	indirect	learn	others	Experimental	Effectiveness	70	Yes	-
(SMITH; ERICSON, 2009)	Partial	CAVE	Y	N	Y	Y	Y	no	indirect	learn and practice	children	Experimental	Effectiveness and Usability	21	Yes	07 - 11
(ROUSSOS et al., 1999)	Partial	CAVE	Y	N	Y	Y	Y	no	indirect	learn	children	Non-experimental	Effectiveness, Immersion and Usability	52	Yes	-
(NAZIR et al., 2012) (NAZIR; KLUGE; MANCA, 2014) (NAZIR; MANCA, 2015)	Partial	CAVE	Y	N	Y	Y	Y	no	indirect	practice	professional	Experimental Experimental	Effectiveness Effectiveness	24 24	No No	19 - 22 22 - 24
(FERRACANI; PEZZATINI; DEL BIMBO, 2014)	Partial	large-screen	N	N	Y	N	Y	no	body position, hand gestures	practice	professional					
(NEDEL et al., 2016)	Low	three-wall	Y	N	Y	Y	N	no	indirect	practice and test	professional	Non-experimental	Effectiveness and Immersion	15	Yes	29
(BEDOYA CASTAÑO et al., 2014)	Partial	large-screen	N	N	Y	Y	N	no	arm's movements	treatment	patient	Experimental	Effectiveness	6	Yes	43,6
(MA; BECHKOUM, 2008)	Full	HMD	Y	Y	Y	Y	Y	yes	finger flex, hand postures, hand, arms and upper body movements	treatment	patient	Experimental	Effectiveness	8	Yes	56.4
(PEDRAZA-HUESO et al., 2015)	Partial	large-screen	N	N	Y	Y	Y	no	body movements	treatment	elderly					
(ORDAZ et al., 2015)	Partial	large-screen	N	Y	Y	Y	Y	no	hand and body position and movements, and indirect	practice	professional	Non-experimental	Effectiveness and Usability	10	No	23.2
(STACH; GRAHAM, 2011)	Low	large-screen	N	Y	Y	N	Y	no	indirect	self-care	others	Experimental	Effectiveness, Immersion and Usability	24	Part of	23
(PARASKEVOPOULOS et al., 2014)	Low	large-screen	N	N	Y	N	Y	no	indirect	treatment	elderly/patient	Non-experimental	Usability	5	Yes	68.25
(HERPERS et al., 2009)	Low	three-wall	N	Y	Y	Y	Y	no	indirect	learn	children					
(SCHÖNAUER; PINTARIC; KAUFMANN, 2011)	Partial	large-screen	N	N	Y	Y	Y	yes	body movements and indirect	treatment	patient	Non-experimental	Usability	10	Part of	22 - 60
(SCHÖNAUER et al., 2011)												Non-experimental	Effectiveness, Immersion and Usability	6	Yes	58
(GREUNKE; SADAGIC, 2016)	Full	HMD	Y	N	Y	Y	Y	yes	hands, fingers	practice	professional					
(ROUSSOU; SLATER, 2005) (ROUSSOU; OLIVER; SLATER, 2006)	Full	CAVE	Y	N	Y	Y	Y	yes	body orientation	learn	children	Experimental Experimental	Effectiveness and Usability Effectiveness and Usability	45 50	Yes Yes	8 - 12 8 - 12
(ADAMO-VILLANI, 2007)	Full	CAVE	Y	N	Y	N	Y	yes	steps, body position, hands motion, eyes position	learn	children	Non-experimental	Usability	12	Part of	5 - 11
(BAROT et al., 2013)	Partial	large-screen	Y	N	Y	?	Y	no	body motion	practice	professional	Experimental	Usability	15	No	-
(SEO et al., 2014)	Partial	large-screen	N	N	Y	Y	Y	no	arms movements and position	treatment	patient	Experimental	Effectiveness and Usability	17	Yes	49.6
(SCHWEBEL; GAINES; SEVERSON, 2008)	Low	three-wall	N	N	Y	Y	Y	no	indirect	learn	children	Quasi-experimental	Effectiveness and Immersion	102 72	Yes No	8.59 21.76
(KAMARAJ et al., 2016)	Low	three-wall	N	?	Y	?	Y	no	indirect	test	professional	Experimental	Usability	21	Yes	18 - 80
(DIERS et al., 2015)	Full	HMD	Y	N	Y	Y	?	yes	thumb and index finger	treatment	patient	Experimental	Effectiveness and Usability	20	No	31.26
(LORENZO et al., 2016)	Partial	CAVE	N	N	Y	Y	?	no	pose	practice	children/patient	Experimental	Effectiveness	40	Yes	7 - 12
(KINATEDER et al., 2015)	Partial	CAVE	Y	N	Y	Y	N	no	indirect	test	others	Experimental	Effectiveness and Immersion	40	Part of	25.4
(CRESCENTINI et al., 2016)	Full	HMD	Y	N	Y	Y	N	yes	indirect	treatment	patient	Experimental	Immersion and Usability	18	Yes	42.88
(GRABOWSKI; JANKOWSKI, 2015)	Full	HMD	Y	N	Y	N	?	yes	hands	practice	professional	Experimental	Effectiveness, Immersion and Usability	21	Yes	45.7
(KIM et al., 2016)	Partial	large-screen	N	N	Y	N	Y	no	arms	treatment	elderly/patient	Non-experimental	Effectiveness and Usability	15	Yes	52 - 77
(BALDOMINOS; SAEZ; POZO, 2015)	Full	HMD	Y	N	Y	?	Y	yes	upper trunk joints, i.e., the shoulder, the elbow and the hand	treatment	patient					
(PERUZZI et al., 2013) (PERUZZI et al., 2016)	Full	HMD	Y	N	Y	Y	Y	yes	feet	treatment	patient	Non-experimental Experimental	Usability Effectiveness and Usability	9 8	Yes Yes	44.3 44.3
(CAI et al., 2013)	Full	CAVE	Y	N	Y	Y	N	no	hand gestures, body position and orientation	treatment	children/patient	Non-experimental	Effectiveness and Usability	15	Yes	6 - 17
(GREWE et al., 2013)	Full	CAVE	N	N	Y	Y	N	no	body orientation	treatment	patient	Quasi-experimental	Effectiveness, Immersion and Usability	19 5	No Yes	23 35.04
(LEE; CHUN, 2014)	Partial	large-screen	N	N	Y	N	Y	no	arms movements and position	treatment	patient	Experimental	Effectiveness	59	Yes	60
(ALAHMARI et al., 2014)	Partial	CAVE	N	N	Y	Y	N	no	indirect	treatment	patient	Experimental	Effectiveness and Immersion	38	Yes	-
(CHAN et al., 2011)	Partial	large-screen	Y	N	Y	Y	Y	yes	body movement and orientation	practice	others	Experimental	Effectiveness and Immersion	8	Part of	21 - 30
(FINKELSTEIN et al., 2011)	Full	CAVE	Y	N	Y	Y	Y	yes	forehead, wrists, and waist	others		Non-experimental	Immersion and Usability	30	Yes	6 - 50
(JOHNSEN et al., 2014)	Partial	large-screen	N	N	Y	Y	Y	no	speech and gesture	self-care	children	Experimental	Effectiveness	61	Yes	-
(KILLANE et al., 2015)	Low	large-screen	N	N	Y	N	Y	no	indirect	treatment	patient	Quasi-experimental	Effectiveness	13 7	Yes No	64.2 64
(LUECKE, 2012)	Low	large-screen	N	N	Y	Y	Y	no	indirect	practice	professional					
(RAGAN et al., 2010)	Full	HMD	Y	N	Y	N	Y	yes	indirect	practice	professional	Experimental	Effectiveness	45	No	21
(RIZZO et al., 2014)	Full	HMD	Y	N	Y	Y	N	yes	indirect	treatment	patient					
(STANSFIELD et al., 2000)	Full	HMD	Y	N	Y	Y	N	yes	hands, back	practice	professional	Experimental	Effectiveness	23	Yes	18 - 45
(STINSON; BOWMAN, 2014)	Full	CAVE	Y	N	Y	Y	Y	yes	body movement and orientation	practice	professional	Experimental	Effectiveness, Immersion and Usability	25	No	22 - 32
(VARLET et al., 2013)	Low	large-screen	N	N	Y	Y	Y	no	indirect	practice	professional	Experimental	Effectiveness	17	No	21.4
(YAVRUCUK; KUBALI; TARIMCI, 2011)	Full	HMD	Y	Y	Y	Y	N	yes	hands and indirect	practice	professional					

2.4.2 User Studies

People use serious games in real life to train, heal and teach people, which makes its validation highly relevant in the implementation process. Indeed, the majority of papers (91%) consider the importance of this phase and conduct user studies to validate the game to reach its objectives. As we discussed before, there are three possible designs under which researchers could conduct the experiments: non-experiment, quasi-experiment, and experiment. Further, their format might be as pretest-posttest or posttest only. The pretest phase applies socio-demographics questionnaires and obtains baselines for physiological measures. The posttest stage consists of obtaining the user's impressions over the game/experiment and the game outcomes. Before making the intervention with the serious game, it might require sessions of hardware calibration and training. Only a few papers report the adoption of consent forms, in which participants agree to the use of their data in the study.

Methods applied in the pretest and posttest phases, as well as during the game intervention, to measure the dependent variables can be subjective and objective. The subjective ones are normally questionnaires and interviews that collect user impressions about the experiment and/or the game. The objective measures are the ones user independent, such as time to complete tasks, physiological data and other game's output variables.

Among the papers that conducted user studies, 26% (11) were non-experiments, from which ten experiments were pretest-posttest design and only two were posttest only. Three user studies were pretest-posttest nonequivalent quasi-experiment. From the real experiments, about 75% were pretest-posttest design.

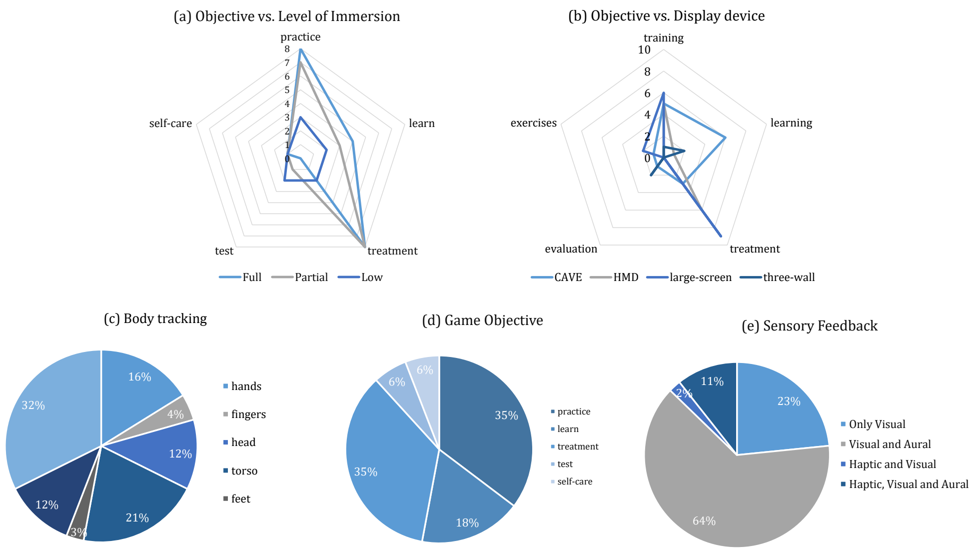


Figure 2.2 Classification of surveyed serious games with respect to visualization and interaction techniques and its relation with game objective.

To address game effectiveness, most studies (59%) perform comparisons between the method used to reach the game objective. For instance, Chitarro et al. (CHITTARO; BUTTUSSI, 2015) compared the ISG developed to educate passengers about safety procedures to the safety card traditionally used in airplanes. Immersion had been assessed comparing different levels of immersion such as in Stach et al. (STACH;

GRAHAM, 2011), in which they assess the benefits of using haptic feedback on three exergames (i.e. games developed to encourage people to be more physically active) by using versions of the games with and without feedback. The few studies, which implemented a quasi-experimental design, assessed the games' effects on participants with different profiles. This can be observed through the study administered by Crescentini, et al. (CRESCENTINI et al., 2016), in which they compared how mindfulness-oriented meditation (MOM) participants and people not involved in any meditation group respond to stressful situations. In the following, we present an overview of methods applied to measure effectiveness, usability and immersion.

Effectiveness is most assessed through user performance (90%), which depends on the game objective. In the evaluation of games to learn, the dependent variable might be knowledge acquisition, while for treatment games the dependent variable might be the health improvement of the patient, which vary according to the disease it treats. The usability is assessed by 76% of the studies through factors such as user experience, satisfaction, workload, and feasibility.

The usability of any system can be validated by the widely used System Usability Scale (SUS) questionnaire (BROOKE et al., 1996), which cover aspects such as need for support, training, complexity and satisfaction in a 5-point Likert scale. Satisfaction could still be more detailed if estimated by the Questionnaire for User Interaction Satisfaction (QUIS) (HARPER; NORMAN, 1993), which focuses on evaluating specific aspects of human/computer interface such as screen factors, terminology, system feedback, learning factors, and system capabilities. This questionnaire contains 80 items on a 9-point Likert scale, and the researcher has the option of selecting a subset of questions that are more interesting to the experiment.

Hart and Staveland's NASA Task Load Index (TLX) (HART; STAVELAND, 1988) is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales: Mental Demand, Physical Demand, Temporal Demand, Own Performance, Effort, and Frustration. The NASA TLX is a two-part evaluation procedure consisting of both weights and ratings. The first part contains fifteen pair comparisons of the six scales, and the rating form has six questions regarding each scale on a 10-point scale. Workload could still be evaluated through the Borg Rating Scale of Perceived Exertion Scale (BORG, 1982), which estimate and monitor the level of physical effort through a 10-point scale to rate the subjective load of the exercise or physical work. To apply it together with heart rate measures gives more accurate results.

The application of VR on serious games is still young, dating from about ten years ago, which cause to be highly relevant to measure the immersion impacts on game effectiveness and usability. From the papers surveyed, half of them had concerns about this feature while assessing engagement, physiologic arousal, sense of presence, simulator sickness, anxiety, and simulation realism.

The Simulator Sickness Questionnaire (SSQ) (KENNEDY et al., 1993) derives from the Pensacola Motion Sickness Questionnaire, which reports a single, composite score. The SSQ is composed of four subscales called nausea, disorientation, oculomotor distress, and total score, in which each symptom is rated in a 4-point scale (None, Slight, Moderate, Severe). The nausea scale contains symptoms such as increased salivation, nausea, stomach awareness, and burping. The oculomotor scale includes headache, eyestrain, and blurred vision. The disorientation scale considers symptoms such as nausea, dizziness, and vertigo. The subjects' tolerance in the virtual environments could also be measures by the Subjective Units of Discomfort (SUD) questionnaire (KAPLAN; SMITH; COONS, 1995) together with vital signs, such as blood pressure and pulse rate.

The questionnaire presented by Jennet et al. (TIJSD; WALTOND, 2008) consists of 32 questions on a 5-point scale addressing flow, cognitive absorption (CA), presence, and an extra question asking how immersed the participant felt in a 10-point scale. Flow is defined by a state provoking an intense and focused concentration on what one is doing, the sense time has passed faster than normal, and that the end goal is only what matters, justifying the entire process (NAKAMURA; CSIKSZENTMIHALYI, 2014). Cognitive absorption is a state of deep involvement with software, which appears through the temporal dissociation, attention focus, heightened enjoyment, control, and curiosity. Presence refers to the psychological sense of being in the VE (SLATER, 2003).

Furthermore, presence can be measured by the SUS Questionnaire (SLATER; MCCARTHY; MARINGELLI, 1998) of Slater, Usoh, and Steed, which consists of 6 questions on a 7-point Likert scale, addressing two factors on the subjective presence in VEs, the extent of body movement and the complexity of a task undertaken in the environment. The Presence Questionnaire (PQ) (WITMER; SINGER, 1998) consists of 32 questions on a 7-point scale to measure the degree to which individuals experience presence in a VE and the influence of possible contributing factors on the intensity of the experience such as scene realism, degree of movement perception, interface awareness, and meaningfulness of experience. The Immersive Tendencies Questionnaire (ITQ) (WITMER; SINGER, 1998) evaluates the tendency of subjects to be involved (or immersed) with the virtual environment using a set of 29 questions on a 7-point scale. The assessment of interaction between players serves to validate the multi-player approach in serious games supporting collaborative environments. In this context, the Social Presence in Gaming Questionnaire (SPGQ) (DE KORT; IJSSELSTEIJN; POELS, 2007) consists of 25 questions on a 5-point scale addressing aspects of co-presence, psychological involvement, behavioral interdependence, enjoyment of social context, connectivity, empathy, sympathy, admiration, jealousy, and revenge. This should prove users' awareness and involvement with their co-players.

O'Brien et al. (O'BRIEN et al., 2008) present a questionnaire focusing on the six factors of engagement: focused attention, usability, aesthetics, endurability, novelty, and involvement. Focused attention describes the perception of time passing and the users' degree of awareness about what was taking place outside of their interaction with the virtual environment. Usability consists of the emotions experienced by the respondents when completing their task, i.e. “annoyed", “frustrated", “stimulated", and “discouraged". Aesthetics concerns to specific features of the interface, such as the screen layout and graphics/images, and to respondents' overall aesthetic impressions of the environment attractiveness and sensory appeal. Endurability evaluates the participants' perceptions of whether the experience met their expectations of being “successful", “rewarding", “worthwhile", and working out as planned. Novelty assesses the curiosity evoked by the task or participants' interest in the interaction, while involvement pertained to respondents' perceptions of feeling drawn into and involved inside of the virtual environment.

Moreover, the Differential Emotional Scale (DES) ((IZARD, 1991)) addresses features of fear (scared, fearful, afraid), surprise (surprised, amazed, astonished), and interest (attentive, concentrating, alert) in a 5-point Likert scale. The following physiological data measure motion sickness and engagement:

- Skin conductance level (SCL) extracted from the electrodermal activity (EDA) signal: it is the most stable of the two components of the EDA signal and is typically used to measure the level of EDA during a given period of time (CHITTARO; BUTTUSSI, 2015). Authors applied it to measure fear and anxiety.

- The blood volume pulse amplitude (BVPA) can be used as an index of sympathetic arousal: a decrease in the BVPA indicates increased arousal (CHITTARO; BUTTUSSI, 2015). A photoplethysmograph (PPG) sensor provides this signal.

- Electromyography (EMG) measures the electrical activity produced by skeletal muscles.

- Cardiovascular measures, such as heart rate, has been used to indicate the user sense of presence in the VE.

In addition to the methods to get the dependent variables, the participants are a crucial element of the assessment. Serious games normally require a specific player' profile depending on its objective, then it would be appropriate to recruit its target public to take part in the experiment. However, conducting studies with real users, particularly if it involves an uncontrolled environment, is challenging because it is difficult to ensure the control of conditions, and the data gathering is not always available, as well as real users. Frequently participants are university students, because it might be difficult to recruit volunteers that match with player's profile or the player does not have a specific profile. In this last case, they belong to the target public, but they have often a different profile of people that are not in the University, which might add bias to the results.

In the assessment of immersive environments, even the controlled ones, there are yet other issues such as the safety and ethical issues. Steed et al. (STEED et al., 2016) conducted an "in the wild" experiment, i.e. an uncontrolled experiment with real users, to test, beyond other goals, the feasibility, and utility of running a study on virtual reality in the wild. Their results suggest that it is possible to use real users to assess simple immersive environments. Mostly of the studies recruited people that fit on players' profile (62%), 13% had part of the public and 25% had volunteers that were not the target public.

2.5 Summary

This paper presented a state-of-the-art review on immersive serious games and classified it towards a taxonomy. It also investigates under which methods and designs that structures the user studies. The classification towards the taxonomy shows us that a great portion of the games provides a virtual environment fully or partially immersive and their purposes are mainly treatment or practice. User study findings confirmed many benefits in using virtual reality in serious games to learn, practice, treatment, test, and self-care. Experiments comparing the serious games either against traditional methods or against a different game version/technology reported better or similar knowledge gain. Chitarro et al. (CHITTARO; BUTTUSSI, 2015) even showed that immersion can improve knowledge retention over time. Moreover, the user experience proved to be superior in the immersive environment, causing greater engagement through factors such as fear and sense of presence. Thus, immersion can be beneficial in two ways: providing greater learning gain and retention, and being more engaging.

These games are also motivating and encouraging for rehabilitation and therapy patients, presenting positive effects on clinical outcomes. In self-care field, serious games have proved helpful to develop and improve the deficits in emotional skills on children (LORENZO et al., 2016), having better results when applied to an immersive environment. They are also effective in increasing mindfulness and decreasing trait and anxiety levels (CRESCENTINI et al., 2016), providing psychological well-being. Studies demonstrate the potential advantages of recent technologies for clinical application, and sensory feedback increased presence and engagement. For instance, haptic feedback showed to be a useful tool for balancing exergames for people of disparate physical abilities (STACH; GRAHAM, 2011), visual feedback motivates participants and induces cortical reorganization (KIM et al., 2016), and performance feedback stimulates users to achieve better results and motion capture turn the activity more interesting (CHAN et al., 2011).

Technological innovation has grown quickly in the last decade, which made it possible to advance research on virtual reality, spreading it over several areas, including serious games. Though immersion VR introduces great benefits on serious games, the development of ISGs still face challenges, such as:

- *To provide effectively immersive environments comfortable and accessible to the public.*
  To support partial or full immersion in serious games, the publications reviewed employ immersive VR hardware being mostly HMDs and CAVEs. It restricts use to players who have access to such (not yet widespread) hardware, as well as it might still cause side effects, such as motion sickness. A few user studies related that participants felt nausea or other kind of discomfort after playing the game during times varying from 15 to 90 minutes (there was only one game that allows people to play the game up to 3 hours), and there were participants that did not even finish the experiment (BACKLUND et al., 2008; SMITH; ERICSON, 2009; ROUSSOS et al., 1999; NEDEL et al., 2016; ADAMO-VILLANI, 2007; KAMARAJ et al., 2016; CAI et al., 2013; GREWE et al., 2013; ALAHMARI et al., 2014; RAGAN et al., 2010). Though some of these user studies also claim the motion sickness did not significantly affects results and that after a number of interventions people got used to the system and the symptoms were reduced, this is still a great limitation in the field. As we discussed in this review, using immersion in serious games can be highly beneficial to increase learning gain, to improve rehabilitation outcomes, while being much more engaging, holding the user motivated to complete the game, which is fundamental to accomplishing their goals. While a desktop version of these games might be inferior to the immersive ones, it would have the advantage of being accessible to a much larger user population, while being less uncomfortable. However, serious games should be able to benefit from the advantages introduced by technological innovations, thus the challenge of offering effective serious games that benefit from immersion without causing

discomfort and being accessible to a larger public still remains.

- *To properly conduct user studies that validate the ISGs.*
  The greatest part of the publications surveyed conduct user studies to validate the serious game presented. Nevertheless, there is a portion of them that introduce a new SG and do not validate its effectiveness or usability (FERRACANI; PEZZATINI; DEL BIMBO, 2014; PEDRAZA-HUESO et al., 2015; HERPERS et al., 2009; LUECKE, 2012; GREUNKE; SADAGIC, 2016; BALDOMINOS; SAEZ; POZO, 2015; RIZZO et al., 2014; YAVRUCUK; KUBALI; TARIMCI, 2011). Some of these papers introduce serious games applied to health, such as rehabilitation and self-care, which are sensitive issues and make it crucial to validate the game before patients can play it, to assure that it will not be prejudicial to them. While the assessment is conducted in a similar way as for non-immersive serious games (BELLOTTI et al., 2013), not all user studies follow the same protocol, they employ different methods to measure the same dependent variables, and sometimes without concern for ethical issues (e.g., do not apply consent forms). We presented in the previous section questionnaires and others methods used to measure the impacts of immersion through motion sickness, sense of presence, engagement and so on. Although there are subjective and objective ways to measure such variables, most of the studies apply only the subjective methods, which by itself may introduce bias to the results. Thus, would be relevant to investigate how the studies can benefit from the combination of such methods to better validate ISGs outcomes. User studies should consider issues such as the number and profile of participants because sometimes the sample size might not be significantly enough and the participants do not have the profile of the game player (see the discussion in the next item).

- *To recruit volunteers with the player's profile to take part in the experiments.*
  Most studies recruit target users, especially when it includes very specific players, such as patients with a determined disease. A few studies compare the impacts of the game on participants, recruiting people with the player's profile and people without it. Schwebel et al. (SCHWEBEL; GAINES; SEVERSON, 2008) present a game that aims to understand and prevent children's pedestrian injuries. They compare children against adults, from which they observe typical differences, such as the adults had safer behavior nearly all measures of interest, the 7-year-old children had significantly more hits than adults did, and they waited longer to cross the street. Grewe et al. (GREWE et al., 2013) present and evaluate a serious game for cognitive neurorehabilitation, and they compared a sample of healthy people against a small sample of patients with focal epilepsy. The healthy group executed the tasks better and faster than the target public. Killane et al. (KILLANE et al., 2015) present a game for treatment of Freezing of Gait (FOG), and they compared a group of people with FOG and another group of non-FOG people. Results showed that the improvements were better for the group with FOG than for the other one. The findings of these studies reveal relevant differences in performance between groups, which implies that assessing the game only with the data collected by the non-target public would give a mistaken conclusion about its effectiveness. Moreover, the sample size for these experiments may be considered

small (M=28.7 participants, SD=23.72) taking into account that this assessment validates serious games to train surgeons, treat diseases, learn content to children, and so on.

- *To explore how the serious games can benefit from different levels of immersion.*
  Immersion can provide great gain to serious games on several aspects, but it also has limitations, such as small access to VR hardware by population and discomfort. While most user studies have experimental or quasi-experimental designs, the approach focuses on comparing the serious game against the methods traditionally employed for the same goal, which allows validation for its purpose. Assessing only one specific level of immersion, the limitations of the technology might influence the experiment's results considering that traditional methods are well known and then present better outcomes even the game being effective. Therefore, varying the immersion aspects (e.g., VR system, FOV, FOR, sensory feedback) allows them to eliminate the technological issues and explore different immersive setups beneficial to the game that might reduce motion sickness and discomfort. A few papers addressed this issue varying aspects such as sensory feedback, FOV, FOR, and level of realism (STACH; GRAHAM, 2011; ROUSSOS et al., 1999; ROUSSOU; SLATER, 2005; KAMARAJ et al., 2016; GRABOWSKI; JANKOWSKI, 2015; RAGAN et al., 2015). Their findings show that haptic feedback is beneficial to exergames, but for some participants the haptic clues were distracting, and haptics could be difficult to implement in games where there is no physical contact with hardware. Considering visual fidelity, results showed that the level of FOV did not have significant effect on assessment target detection or strategy usage and that training systems for visual scanning and similar tasks should when possible, use a level of visual complexity that is as close to the real environment as possible in order to ensure good transfer.

3 DESIGN DECISIONS

Our approach was based on the analysis of previous works on the field, as well as on the conduction of two user studies. Firstly, we conducted an exhaustive state of the art review (see Chapter 2) aiming to identify the current levels of immersion used on serious games and which displays are used to reach that immersion. We also verified how the effectiveness of immersion and serious games is evaluated. We surveyed 28 immersive serious games, and through them, we defined our research questions and the procedures to conduct our empirical studies. We chose to investigate two points: the effects of display fidelity, defined here as the display device and the level of immersion used to allow the user to be isolated from the real world; and the interaction and locomotion fidelity, defined by the naturalness of the interface.

Higher levels of immersion provide better performance in serious games (NAZIR et al., 2012; NAZIR; KLUGE; MANCA, 2014; CHITTARO; BUTTUSSI, 2015; HUPONT et al., 2015). The experimental design of these studies, mostly, consider only two levels of immersion (e.g., low and high), or compare the traditional methods of reaching the serious game goals (e.g., through slides presentation or cards, in the cases of training and learning) against the game itself, showing that the serious game is more effective. Higher display fidelity also improves spatial knowledge and allows users to use a spatial memorization strategy to do procedure memorization (HIROSE et al., 2009). Although the game reaches its goals, some VR systems, such as HMDs, still may cause symptoms of *cybersickness*, which could be avoided by using intermediate display fidelity.

Ragan et al. (RAGAN et al., 2015) have explored the idea of using serious games for scanning tasks, considering the effects of visual immersion and levels of realism on the user performance. Their findings showed that higher FOVs leads to better training trial detection. Visual realism affects strategy transfer and training task performance, and that training with higher FOV does not improve the real-world task performance more than training with a lower FOV. They also showed that there is no correlation between the user performance on the VR system and on the real-world.

We observe that previous works still do not use VR systems that provide feedback for all five human body senses, simultaneously. There are, though, techniques such as motion tracking, that allows us to interact and navigate in the VE using natural movements. The use of real walking provides naturalness to navigation. However, it requires a virtual environment that has the same size of the real environment tracked. Redirection techniques (HODGSON; BACHMANN; WALLER, 2011) can solve this issue, and they are significantly better at navigating than walking-in-place (WIP) or joystick navigation (PECK; FUCHS; WHITTON, 2011). Still, it requires a widely tracked area to enable redirection and is expensive to implement. Even though WIP has been shown inferior performance to real walking techniques, it is better than a joystick for exploration human-scale spaces and spatial orientation, provides a heightened sense of presence (USOH et al., 1999), and is inexpensive to implement.

Motion tracking can also provide a natural interaction using gestures, such as in the serious game presented by Greunke et al. (GREUNKE; SADAGIC, 2016). Previous works, though, shown that having a physical reference helps the user to be more precisely on memory recall (MINE; BROOKS JR; SEQUIN, 1997) and that naturalness is not always a necessary component of an effective technique (BOWMAN; HODGES, 1997). Mania et al. (MANIA et al., 2006) also suggests that low interaction/display fidelity may enhance the memorial experiences by adding attentional demand to the cognitive system. However, results have not achieved statistical significance to prove it. Motion tracking may provide a natural form of interaction, but can also insert noise in the data, which may worsen the accuracy of the interaction.

We assumed that the purpose of a serious game based on real world simulation is to prepare the user for the real situation that has high visual complexity. Therefore, we developed two serious games for training and learning by adapting virtual environments with high realism (JORGE et al., 2013; NEDEL et al., 2016). For the implementation of the games, we add a set of audio instructions to guide the player through the game, and the interaction to move the objects through the environment in the game for safety procedures learning. These games were used in two empirical studies to investigate the effects of display fidelity in user perception, and the effects of locomotion and interaction fidelity on knowledge retention. From the results of the first user study, we selected the level of display fidelity to implement the second experiment.

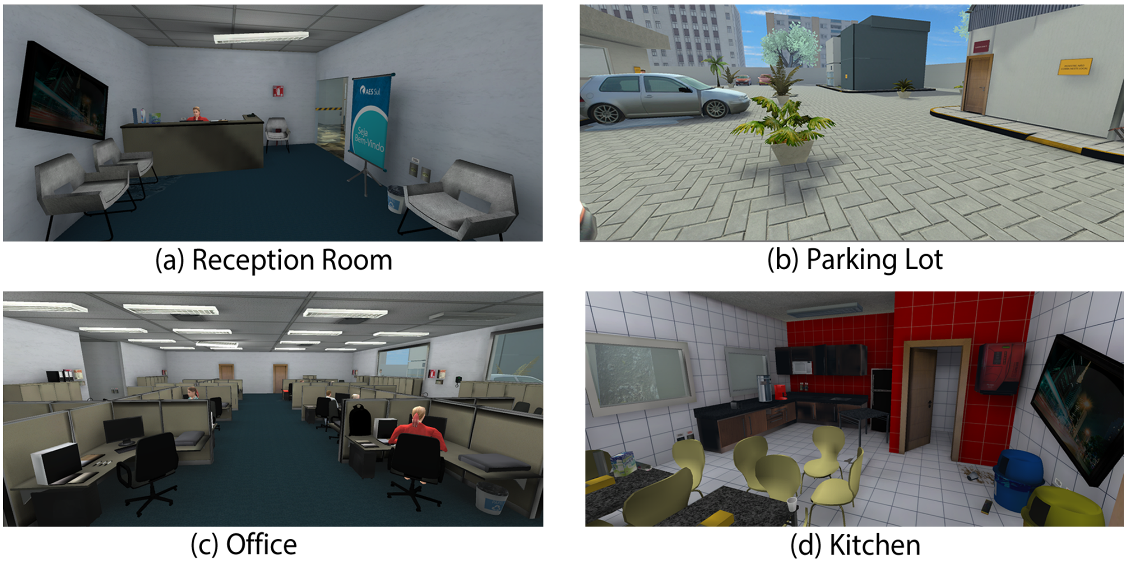
For both studies, we followed a well-known procedure consisting of: the agreement with a consent form in which participants allow us to anonymously use their data; the pre-test phase, in which we applied a demographic questionnaire and get the baseline of sickness and knowledge about the content of the serious game; the training phase, in which the instructor explains the interface and gives a time to the user get familiar with them; the test phase in which the user plays the serious game; and the post-test phase, where we applied questionnaires to get the subjective effectiveness through sickness, workload, presence, usability, engagement, and knowledge acquired. The measure of the user performance consists of the serious game outcomes, completion time and path traveled in the VE.

Though we have mentioned the importance of use target users in the experiments, most of our participants are university students, especially of computer science. Our experiments were designed to investigate issues about immersion and interaction, and each one has a considerable number of experimental conditions. Thus, to investigate the impacts of using target users in the experiments would require a larger set of experiments and subjects, involving different profiles of participants, and a longer time to execute the studies.

From the empirical studies we identified a set of design guidelines for the development of serious games for learning (see Chapter 6). In the following we present the SGs and the materials and methods applied to measure the dependent variables in the user experiments.

3.1 Risk Perception Training

Figure 3.1 Places of the virtual building from player point of view.



Source: Compiled by author.

This game is dedicated to training users' ability on detecting elements of risk in their work environment. The virtual environment consists of a building with a reception room (Figure 3.1(a)), a parking lot (Figure 3.1(b)), an office (Figure 3.1(c)) and a kitchen (Figure 3.1(d)). It contains 53 objects pre-determined as potential risks and distributed as showed in Figure 3.2. There are two categories of risks: simple and composite (see Table 3.1). Simple risks are objects that represent immediate risks, such as a wet floor and a fire extinguisher blocked. Composite risks are objects that may introduce an eventual danger in the future, such as a flashing lamp and the alarm off.

There is a narrator that guide the user through the game with verbal instructions, which are activated automatically and manually. In between of executing the given tasks, the player should search for possible risk objects, as a secondary task. The instructions are the following:

1. Hello! I'll give you instructions on how to proceed in the scene. Go inside of the building, to the center of the reception room.

2. Stop! Stay in this location and see around. Point to the possible risk elements that you can see, for how long you want. When you have finished, activate the next instruction.

3. Go to your car in the parking lot through the door next to the reception desk. Your car is the one flashing in red.

4. Inside of this building has a kitchen. Go there and get a coffee.

5. Now, go to your desk. It is the one flashing in red.

Figure 3.2 Map of the virtual environment with the position of the potential risks.

Source: Compiled by author.

6. Stay in this location and see around. Point to the possible elements of risk that you can see. When you have finished, activate the next instruction.

7. Meet with your coworkers outside of the office, next to the parking lot.

8. Well done! You finished the simulation!

The messages contain directives that guide the users to navigate passing through the largest possible area of the virtual environment. Each instruction leads the user towards some location in one of the places of the virtual building, where there is a target object flashing in red to help find it quickly. When the player reaches the target, the next message is automatically played, except for the instructions #2 and #6. In these instructions the user should stay in its current location and search for potential risks around him/her, for how long s/he wanted and, then, play the next instruction. These instructions were included to reduce the difference between the display devices, forcing the users to do the same movements in these moments of the game. Also, it serves to investigate how much focused in the instructions they were instead searching for potential risks as instructed by the researcher at the beginning of the session.

Table 3.1 Potential risks categories

Simple	Composite
Cabinet with open doors	Fallen frame
Rope	Power outlet not signaled
Broken tile	Flashing lamp
Rolled up carpet	Lamp off
Object blocking the passage	Door not signaled
Wet floor	Door without doorknob
Cleaning products	Wet stairs
Carpet stain	Alarm off
Fire extinguisher blocked	Open drawer
Hole	Fallen lamp
Cup on the table	
Outlet without cover	
Puddle	
Trash on the floor	
Irregular connection	
Equipments in the wrong place	
Bag	

Source: Compiled by author.

3.2 Learning Basic Safety Procedures

This game was designed for educating workers on basic safety procedures for electrical installations on public utility poles. The virtual environment represents a town street with a replacement of lightning rod been performed. The player should follow audio instructions given by a narrator, basically positioning the objects in specific locations to make the environment safe to work. The game consists of three stages, one for learning and two for evaluation, which differ in the objects position and the set of instructions. In the learning stage, each instruction is presented highlighting the objects and showing the target locations with arrows (see Figure 3.3). The instructions were the following:

1. Hello! In this game, you will learn basic safety procedures, which has to be adopted when performing a replacement of a lightning rod. Your task is to organize the materials to left the environment safe and correctly signaled.

2. First, it is important to ensure that the truck is parked correctly and stuck with shims at the wheels. Place one shims in each wheel, which is indicated by the red arrows. When you have finished, activate the next instruction.

3. The workers need to be equipped with scrap gloves, safety belt, and helmet. Besides that, they have to carry the service order. Select the objects representing this equipment. They are flashing in red. When you have finished, activate the next instruction.

4. In some cases, it should be used canvas or stand for sticks to accommodate the materials on the ground, which has to stay inside of the signaled area. Place

the materials that are flashing in red on the canvas that is indicated by the red arrow. When you have finished, activate the next instruction.

5. Before start the installation, the workspace has to be signaled through signposts. Place the signposts in visible locations, as indicated by the red arrows, signaling the pedestrians. When you have finished, activate the next instruction.

6. Isolate the work area, including the vehicle, using the cones and ropes. First, place the cones around the area where the installation will happen, which is between the spaces indicated by the red arrows. When you have finished, activate the next instruction.

7. Now that you place the cones put the ropes between them. To do this, select one rope and two cones, one at the time. Follow doing this procedure until the area is properly isolated. Remember of leaving an opening for that the workers can transit. When you have finished, activate the next instruction.

8. Very well! You finished the learning stage. Now you will practice the acquired knowledge.

In both evaluation levels, the narrator gave the following instructions to the player:

1. Now that you have learned the necessary procedures for signaling and safety, remember of the received instructions and signals the environment, leaving it safe for the workers and pedestrians. When you have finished, activate the next instruction.

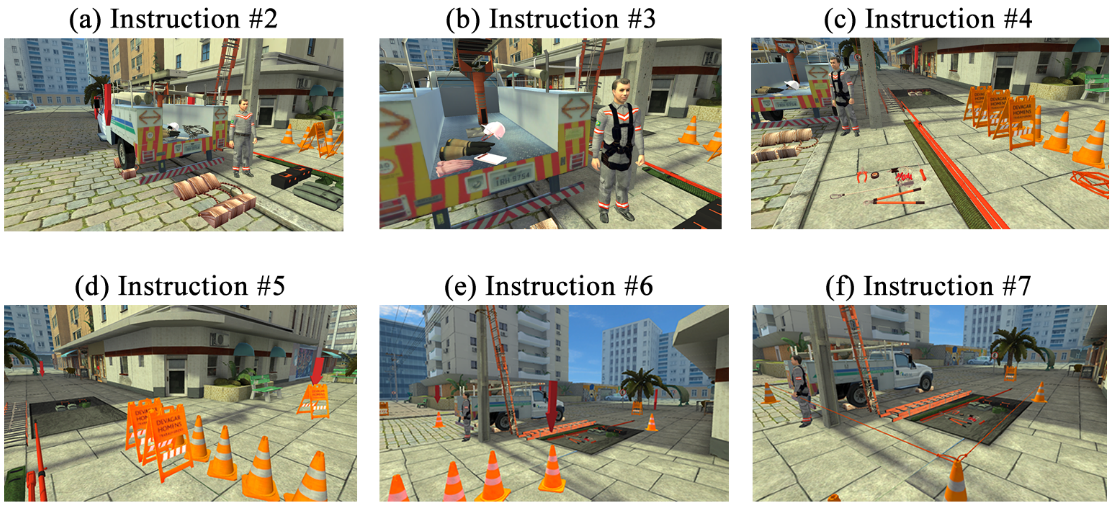
2. Very well! You finished the game!

The player should execute the tasks according to its interpretation. For the instruction #6 they had eight cones to be arranged in the middle of four arrows, so they could place only one cone in each arrow or distribute the eight cones in the provided space. In the next instruction (#7) they had seven ropes available to be tied between the cones. Once again, they could choose the desired number of ropes to isolate the place. After tying the rope between two cones, the player could no to pass through the cord. We used this as a strategy to remember the user that s/he should leave an opening between the cones for that the workers could walk in and out the isolated area.

In the learning stage, the objects were allowed to be selectable only after play the relative instruction. Then, it remains selectable for the rest of the game. In both evaluation stages, the objects were selectable since the begin, so the player could choose the best strategy to organize the workplace. In this phases, there was not either highlight objects or indicated locations. The correctness of the tasks are assessed partially or entirely, according to the following criteria:

- *Instruction #2*: Each shim correctly placed (+0.25).
- *Instruction #3*: All equipment should be selected.
- *Instruction #4*: Each set of tools correctly placed on the canvas (+0.33).

Figure 3.3 Virtual environment from player point of view for each instruction. The arrows indicate the target location of the objects (e.g., in the figure is showed the visual tips for the first instruction). In the evaluation levels, the arrows are not showed.



Source: Compiled by author.

- *Instruction #5*: Each signpost placed in a visible location (+0.33).
- *Instruction #6*: Users should place the cones around a minimum area (e.g., including the canvas and the truck).
- *Instruction #7*: First, the ropes should be tight between cones isolating the area. Second, they should leave an opening for that the workers could transit.

3.3 Measures

3.3.1 Simulator Sickness

To investigate the severity of the simulator sickness symptoms experienced by the participants in each condition we applied the Simulator Sickness Questionnaire (SSQ) (KENNEDY et al., 1993) (see Appendix F). It consists of sixteen symptoms rated in a 4-item scale (e.g., "none", "slight", "moderate", "severe"). It gives scores for nausea, oculomotor, disorientation and a total simulator sickness score.

To obtain the SSQ scores we assigned values to the severity of the symptoms. Thus, rating "none" for some symptom would be equal of 0 for that symptom, "slight" would be equal to 1, "moderate" would be equal to 2, and "severe" would be equal to 3. The values for the symptoms included in the particular scale are summed and then applied a conversion formula presented by Kennedy et al. (KENNEDY et al., 1993). Every scale contains seven symptoms with weight equal 1. The nausea scale includes symptoms of general discomfort, increased salivation, sweating, nausea, difficulty concentrating, stomach awareness and burping. The oculomotor scale contains symptoms of general discomfort, fatigue, eyestrain, headache, difficulty focusing, difficulty concentrating, and blurred vision. The disorientation scale has weight 1 for the symptoms of difficulty focusing, nausea, the fullness of head, blurred vision, dizzy (eyes open and closed), and vertigo.

To verify how much harmful was the scores obtained, it was calculated the maximum values for each scale, considering the maximum severity of all symptoms. Therefore, nausea could present a maximum score of 267, the oculomotor scale could be top 212, the maximum disorientation score could be 389, and the total score, considering the maximum score of all scales, could be 2437.88.

3.3.2 Presence

We measured the subjective sense of presence using the SUS (Slater-Usoh-Steed) presence questionnaire (USOH et al., 2000) (see Appendix G). It consists of six questions addressing the presence in the places of the virtual environment. Participants should rate their experience on a 7-point Likert scale. Further, we also asked them on a 10-point scale "how much immersed did you feel?", to measure their subjective sense of immersion.

3.3.3 Usability

In the Experiment I we asked users to answer an opinion questionnaire, in which they should agree or disagree using a 5-point Likert scale with the following affirmations about the serious game for risk perception training:

- The serious game was developed with the goal of analyzing the abilities of the users on finding possible risk objects in the workplace. The game is attending this objective.
- It was easy to play the game.
- I had fun playing the game.

3.3.4 Workload

To measure the workload of each technique we applied the Hart and Staveland's NASA Task Load Index (TLX) (HART; STAVELAND, 1988), which is a multi-dimensional rating procedure that provides an overall workload score based on a weighted average of ratings on six subscales: Mental Demand, Physical Demand, Temporal Demand, Own Performance, Effort, and Frustration. The NASA TLX is a two-part evaluation procedure consisting of both weights and ratings. We applied the first part with fourteen pair-wise comparisons of the six scales, and the rating form consists of six questions regarding each scale in a 10-point scale.

To obtain the workload scores for the scales, we count how many times users chose each one in the weights part, and multiply this value by the related rating given in the second part of the questionnaire. To get the total workload score, we add the scales' scores and split by fourteen.

3.3.5 Engagement

We applied a modified version of the Game Engagement Questionnaire presented by McMahan et al. (MCMAHAN et al., 2012) (see Appendix E) to measure the user engagement with the game in each experimental condition. It addresses aspects of immersion, presence, flow and psychological absorption. Immersion describes the experience of becoming engaged in the game-playing experience while retaining some awareness of one's surroundings. Presence has the definition of being in a normal state of consciousness and having the experience of being inside a virtual environment. Flow describe the feelings of enjoyment that occurs when it has achieved a balance between skill and challenge in the process of performing an intrinsically rewarding activity. Psychological absorption represents the total engagement in the present experience.

The questionnaire consists of 15 statements that the participants should rate their experience of the sensation while playing the game on a scale from 1 (did not experience) to 5 (definitely experienced). To obtain the engagement score, we summed the ratings for the following questions and divided by fifteen.

3.3.6 Knowledge

To measure the participant's knowledge we prepared a test with seven questions addressing the safety procedures given in the serious game. The questionnaire was composed of essay questions, presented to the user one at a time, to avoid suggesting possible answers (e.g., as a multiple-choice questionnaire would do). The questions were the following:

1. In general, what are the procedures that should be adopted to make the workplace safe for workers and pedestrians?
2. Once the vehicle has been parked, what is the next thing to do?
3. Before making the place safe, it is important that worker is wearing appropriate equipment. What is this equipment?
4. Besides the safety equipment, what is the other item that the worker need to carry?
5. Where and how the work tools should be placed?
6. What is the first step to signal the workplace in order to make it safe for the pedestrians?
7. How do you isolate the workplace?

Considering the range of answers for each question, we split some of them into several parts in a way that answers are found partially correct, and the maximum rating is 1 for each question. We followed a code-book that indicates the possible correct answers and the value for each one (Figure 3.2).

3.4 Statistical Analysis

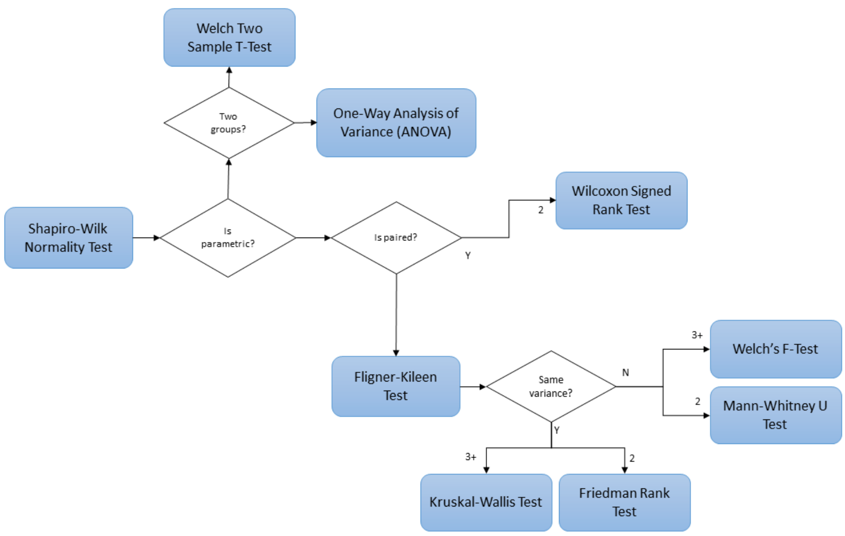
To perform the analysis of the data, the significance level was set as p-value = 0.05. The statistical significance is reported in the charts through the following symbols: 'ns' for p-value > 0.05, '*' for p-value < 0.05, '**' for p-value < 0.01, and '***' for p-value < 0.001. In the Figure 3.4 is presented the process to choose the appropriate test to perform the statistical analysis in the data.

Table 3.2 Codebook for possible correct answers in the knowledge test.

Question	Answer	Score
1	Place the shims in the truck wheels	0.2
	Put on the personal safety equipment	0.2
	Signaled the workplace with signposts	0.2
	Place the work tools inside the marked area over a canvas	0.2
	Isolate the area with cones and ropes or ribbons	0.2
2	Isolate or signaled the workplace	0.5
	Ensure that the truck doesn't move putting shims in the wheels	1
3	Safety belt	0.33
	Helmet	0.33
	Gloves	0.33
4	Order of service	1
5	Inside the signaled area over a canvas	1
6	Place the signposts in strategic locations without disturbing the pedestrians traffic	1
7	Using cones	0.5
	Using ropes or ribbons	0.5

Source: Compiled by author.

Figure 3.4 The flowchart depicts the process for choosing the suitable method for the statistical analysis. The numbers represents the number of samples for the analysis.



Source: Compiled by author.

Firstly, was performed a Shapiro-Wilk Normality test of the null hypothesis that the data came from a normal distribution. The data was collected between-subjects to compare the performance among different setups and within-subjects to verify the acquired knowledge and the effect of each condition on the simulator sickness symptoms. For the parametric data, it was applied a One-Way Analysis of Variance

(ANOVA) for three or more non-paired groups, and a Welch Two Sample t-test for two independent groups.

For the non-parametric data, we performed a Fligner-Killeen test of the null that the variances in each of the groups are the same. Then, for three or more groups and homogeneous data, it was applied a Friedman Rank Test for paired groups, and a Kruskal-Wallis Test for non-paired groups. In the cases that the data was heterogeneous, a Welch's F-Test was performed to test the equality of three or more non-paired groups. In comparisons involving two groups, we applied a Mann-Whitney U Test for independent groups and a Wilcoxon Signed Rank Test for paired groups.

The data was submitted to a mixed design, in which the experimental conditions was the between-subjects variable. Because both games aim to educate the player about some content, we chose to apply this conditions between-subjects to avoid that the participant is influenced by the previous condition. In the second experiment, the mean time to complete the test (e.g., about 45 minutes) was also a criterion to chose the between-subjects approach. Though it avoided the influence in the game outcomes, the statistical analysis required different tests and the samples not always had the same variance, which may introduce noises in the results. The within-subjects variable was the time of measurement (before and after the trial) for the SSQ and knowledge data and for the knowledge data, the levels of the serious game, the places of the VE and the instructions.

The analysis of the questionnaires and logs, along with the charts were built through R statistics. The scripts and data used for the analysis are available on https://goo.gl/twqfGj.

4 EXPERIMENT I - DISPLAY FIDELITY AND USER PERCEPTION

In this experiment, we investigated the effects of visual immersion on a serious game that uses perception to learn to users. The organization of this chapter is as follows. Section 4.1 presents the hypotheses that conducted this experiment. Section 4.2 shows the materials and methods of the experiment, which includes the participants involved, and the apparatus used for interaction and display the serious game. Section 4.3 present the results obtained from the experiment, which are discussed on Section 4.4.

4.1 Hypotheses

This experiment investigates the effects of visual immersion on users' perception of potential risks in a working place, and how the immersion impacts the symptoms of simulator sickness. The overarching hypothesis was that a serious game that uses perception to provide learning could be effective, comfortable and offer a good VR experience even in a moderate display fidelity. On a more particular level, this experiment tests the following hypotheses:

- **H1. Higher and moderate levels of immersion will not show significant difference on user performance.**
  As mentioned before, previous works suggest that higher levels of immersion improve user performance on serious games. Intermediate display fidelity may also provide a good sense of immersion for users, so we believe that the user performance will be similar in both moderate and high display fidelity.

- **H2. Higher levels of immersion will cause higher simulator sickness.**
  The *cybersickness* presents symptoms of motion sickness, such as nausea, headache, and dizziness, and these occur because the body is standing while the world is moving around us. The navigation in regular video games may also cause these symptoms because the virtual world is moving and the player is standing. The setup used for higher display fidelity (e.g., HMD) has also the issue that the player is not able to see the real world while playing the game. Therefore, we believe that this configuration will increase the severity of the simulator sickness symptoms.

- **H3. The simulator sickness will not significantly affect the user performance.**

Although we believe that higher display fidelity will increase the simulator sickness, we also believe that it will not influence the result of the serious game negatively. Therefore, regardless the severity of the sickness symptoms, the user performance will not be impaired.

4.2 Materials and Methods

4.2.1 Experimental Conditions and Apparatus

As previous works showed that the stereo cues not reliably influence either the accuracy or phenomenology of memory for the VE simulation and that the consistency of the objects with the scenario context improves the user performance (BENNETT; COXON; MANIA, 2010), we focused on providing a consistent virtual environment. For this experiment we varied the display fidelity considering three experimental conditions:

- **Low-display**: This condition offers the lowest immersion regarding the display. We used the desktop monitor that do not isolate the user from the real world.
- **Medium-display**: This condition intend to provide a medium level of immersion, using a large-screen that allows the user to focus more in the game than in the real world around him/her.
- **High-display**: This condition has the highest level of immersion, once the user is isolated from the real world and is able to focus only in the game with a higher sense of being in the environment.

The three display devices used were: a 23" LCD display with a refresh rate of 120Hz and a resolution of 1,920 x 1,080 pixels (Figure 4.1(a)); a set of twelve 22" LCD displays, with an aspect ratio of 16:10 each and resulting in a display wall with a total dimension of 244 x 108 centimeters, and a resolution of 3,200 x 1,800 pixels (Figure 4.1(b)); and an Oculus Rift DK2 device, with a resolution of 1,920 x 1,080 pixels per eye, and a refresh rate of 75Hz (Figure 4.1(c)). Participants were also equipped with a headphone to listen the instructions.

Figure 4.1 The levels of physical immersion used in the experiment.

(a) Low Fidelity (b) Moderate Fidelity (c) High Fidelity

Source: Compiled by author.

The interaction with the game was through a game controller and head tracking. The controller input used was the following: the 5th axis for locomotion forward and backward; button 0 for select objects; button 1 for activating the next instruction, and button 2 for repeat the instruction.

It was used a crosshair to targeting objects in the scene and moving the virtual camera. In low immersion, we used the movement of the 4th axis of the game controller. For moderate and high immersion, we did the movement of the crosshair and the virtual camera through head tracking. We used the head tracking provided by the Oculus Rift sensors, and in the moderate immersion, we equipped the participants with a helmet that had a Sixense Razer Hydra controller fixed on top of it, in a way that the orientation of the controller serves as the head orientation.

4.2.2 Procedure

We split the participants into three groups, one for each condition. We told them that the experiment goal was to verify the impact of the display device on their ability of find dangerous objects in a regular office environment. They signed a consent term of participation. Then, it was applied a characterization form (age, gender, education level, vision problems, experience with virtual reality, 3D video games, and safety work).

Figure 4.2 Top view of the virtual environment used in the serious game.

Source: Compiled by author.

After the user had answered the pre-test questionnaires, the researcher showed them a top view picture of the virtual environment (Figure 4.2) for that they could be familiar with the virtual world. Following a specific order of conditions, the user would play the game using one of the displays. In the case that the visualization would be through the Oculus Rift, the instructor used a demo scene to calibrate it, making the user comfortable with the equipment. Then, the researcher explained the controllers, making sure that the user understood how to interact with the game.

Finally, the instructor put the headphone on the user's head.

At the end of the game, the participant was instructed to answer a post-test questionnaire. It contains the SSQ, the presence questionnaire and some questions about its experience.

4.2.3 Subjects

The experiment involved a sample of 61 volunteers (15 female). They were students and university personnel, with an age range between 19 and 63 years (M=28.67, SD=10.12). They were separated randomly in three groups, one for each experimental condition. All participants gave a written informed consent before the study (See Appendix A). Most of the participants were taking or had finished the graduate studies.

Subjects were naïve concerning the purpose of the experiment (i.e., work safety and virtual reality). The experience in video games 3D was varied, about 42% of the participants considered themselves experienced, while about 37% considered themselves non-experienced, and approximately 20% were neutral in their answer. More than half of the subjects (56%) reported some vision problem, from which most of them had myopia. For the experiment, they could use their glasses along with the displays, even with the Oculus Rift, to correct their vision. The possible discomfort caused by using the glasses did not affect the results.

4.3 Results

4.3.1 User Performance vs. Display Fidelity

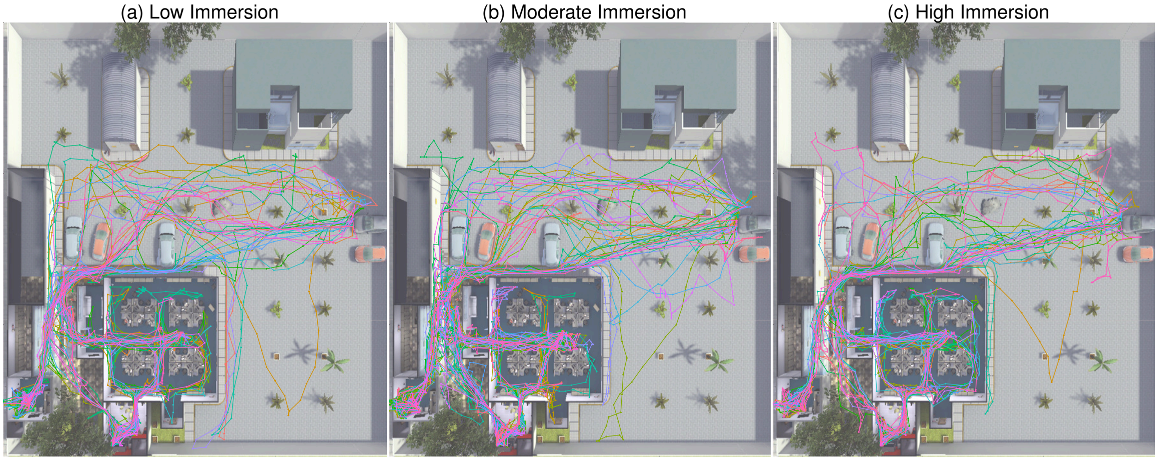
To understand the behavior of the participants, a map of the path traveled was logged for each one. In Figure 4.3 it is possible to observe that the users present a similar behavior, following a logical path to reach the target objects mentioned in the instructions (e.g., the car at the parking lot, the coffee maker at the kitchen, the desk at the office, and the co-workers outside of the office). Although we did not observe extraordinary differences among the setups, it seems that the users with high display fidelity followed a path slightly less varied than the ones using low and moderate immersion. It may be a consequence of the fact that participants felt more present and immersed in the high display fidelity condition, allowing them to walk more strictly through a logic path.

The completion time showed that subjects of moderate display fidelity took longer than other conditions to finish the trial, with $p < 0.001$ and mean time of 744.15 seconds. The lower mean time registered was 361 seconds for high display fidelity. This result may be caused by that the participants found harder to play the game in the moderate display condition.

We recorded information about the selected objects, such as name, time, and position to calculate the risk perception rate for each user. Mean scores suggest that subjects of low display fidelity (M=24.71, SD=2.89) found more risks than moderate and high fidelity (M=18.3, SD=2.76), respectively. However, it was not found statistical significance to support these results ($p = 0.28$).

Based on the elements of risk directly visible in each instruction from the pattern path followed by the users (see Figure 4.3), it was possible to infer that the participants selected higher number of potential risk objects in the instruction #2, in

Figure 4.3 Path traveled by each user, represented by the colorful lines, in the virtual environment. Low display fidelity (a). Moderate display fidelity (b). High display fidelity (c).



Source: Compiled by author.

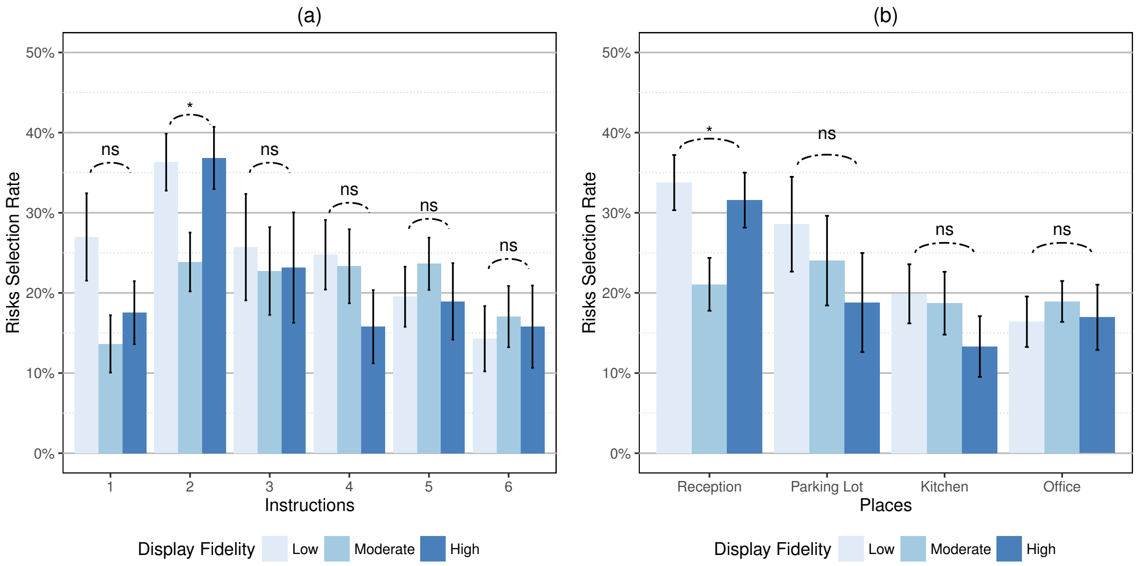
which the narrator explicitly ask them to search for risks. In this case, the number of risks selected was similar for low and high immersion, and lower for moderate display fidelity ($p < 0.025$) (see Figure 4.4(a)). In low and high display fidelity, subjects found more elements of risks in this instructions than in the others, with $p < 0.015$ and $p < 0.009$, respectively. In the moderate display fidelity, we did not observe significance differences among the instructions.

Considering the risks selected in each place of the virtual building we noted that in the reception room, subjects of low display fidelity found more risks than high and moderate display fidelity, respectively, with $p < 0.025$. In the reception room was also the place where the subjects found more risks against the other places of the virtual environment in the low display fidelity ($p < 0.004$) and high fidelity with $p < 0.007$ (see Figure 4.4(b)). For this case, we observe that the sense of presence had no influence, as the subjects felt more present in the parking lot than at the reception room.

As most of the participants consider themselves non-experienced in safety work, we analyzed the number of risks found in each category, to verify which kind of risks they were able to identify. Figure 4.5 presents the percentage of objects selected by category in each experimental condition. It is possible to observe that participants found significantly more simple than composite risks in all experimental conditions (see Figure 4.5), with $p < 0.001$ for low and moderate display fidelity, and $p < 0.01$ for high fidelity. However, subjects of high display fidelity found a significantly higher number of composite risks than low and moderate fidelity conditions, with $p < 0.0018$.

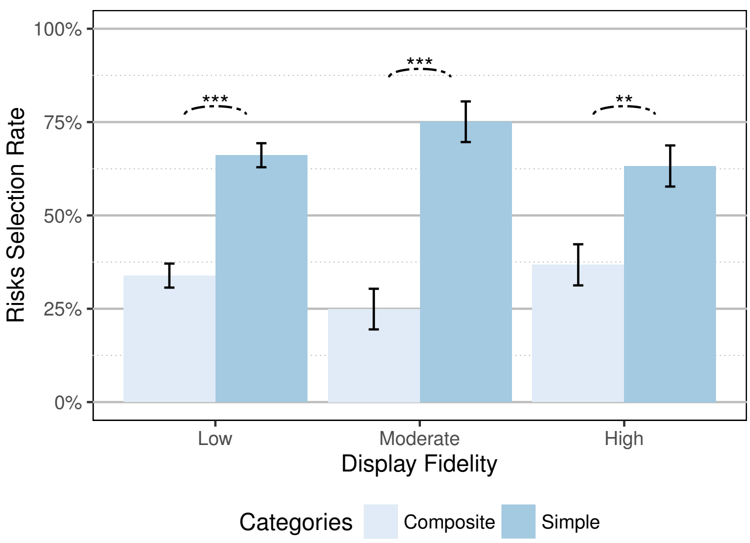
Though the completion time was considerably higher for moderate display fidelity than for high condition, the risk perception was not affected by the level of immersion. However, we found that for composite elements of risk, the subjects of high display fidelity had better performance than the users of the moderate condition. We cannot consider H1 correct, because even that we did not observe

Figure 4.4 Percentage of risks found in each experimental condition. (a) In each instruction. (b) In each place. Statistical significance: '*' for $p < 0.05$, and 'ns' for $p > 0.05$.



Source: Compiled by author.

Figure 4.5 Percentage of risks found by category in each experimental condition. Statistical significance: '***' for $p < 0.001$ and '**' for $p < 0.01$.



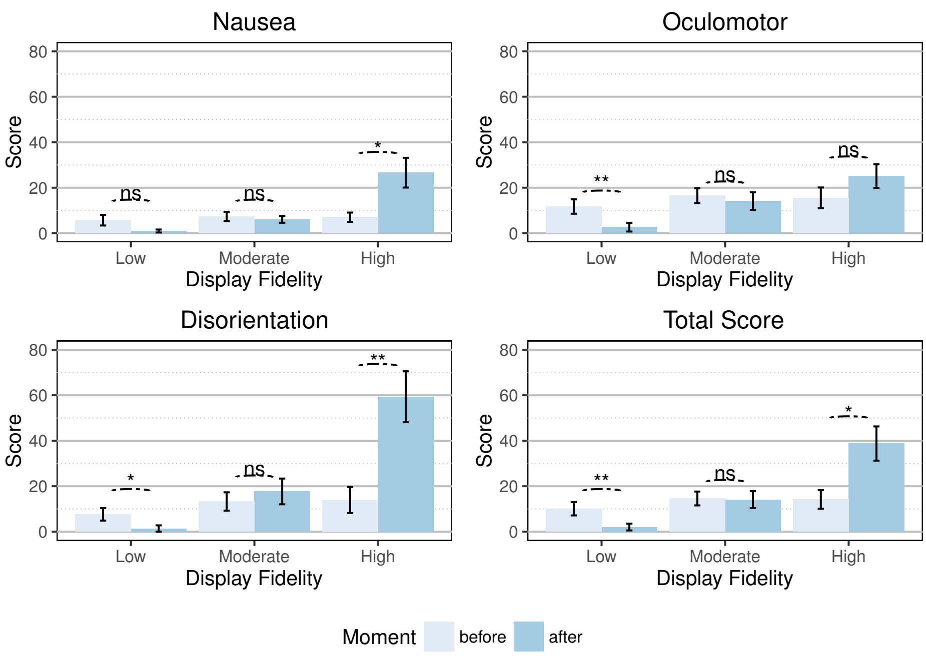
Source: Compiled by author.

significance differences between the overall number of risks found, the completion time was significantly higher in the moderate display fidelity. Therefore, it seems that display fidelity has negative impacts on some aspects of user performance.

4.3.2 Simulator Sickness vs. Display Fidelity

We administered the SSQ twice: before trying the experimental condition (pre-test), and immediately after trying it (post-test). Mean pre-test score showed that the participants were not feeling entirely well, presenting mostly symptoms of fatigue, eyestrain, difficulty focusing, and difficulty concentrating. The obtained SSQ scores were as follows: Nausea (M=6.72, SD=9.43), Oculomotor (M=14.66, SD=16.36), Disorientation (M=11.63, SD=19.21), and Total Score (M=12.99, SD=15.04). An independent Kruskal-Wallis rank test showed no significance differences in the initial scores between the groups (p = 0.47 for Nausea, p = 0.47 for Oculomotor, p = 0.7 for Disorientation, and p = 0.39 for Total Score).

Figure 4.6 Simulator sickness effects measured before and after the trial for each scale (e.g., Nausea, Oculomotor, and Disorientation) and the total score in each experimental condition. Statistical significance: '**' for $p < 0.01$, '*' for $p < 0.05$, and 'ns' for $p > 0.05$.

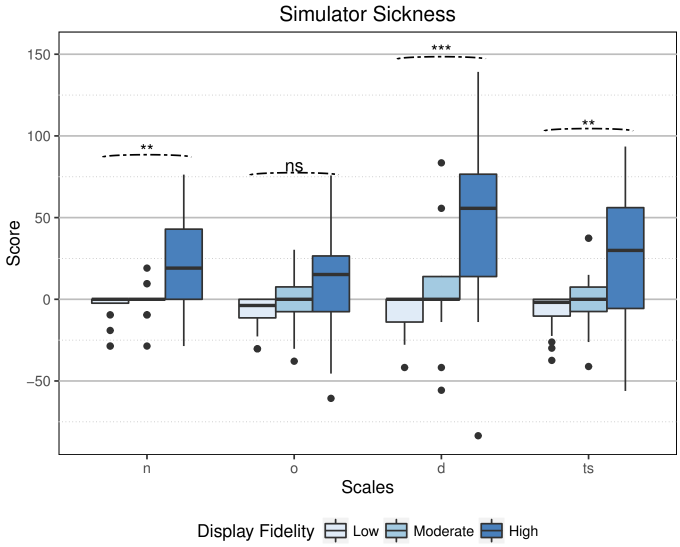


Source: Compiled by author.

The Simulator Sickness Questionnaire showed sickness effects for all experimental conditions, as observed in Figure 4.6. In high display fidelity, the severity of symptoms increased significantly after the trial for nausea, disorientation, and total score scales. While in low display fidelity, we observed that the sickness effects decreased after the test. In moderate display fidelity, we did not observe any significance differences between before and after the trial.

The most harmful condition was high display fidelity, as we can see in Figure 4.7, in which we found statistical significance for nausea ($p < 0.008$), disorientation ($p < 0.001$), and for total score ($p < 0.004$). Based on this result, we consider the H2 true. The highest SSQ score registered was for disorientation (M=45.42, SD=53.68) in high fidelity. Note that this score is very low, as the highest possible total SSQ score is 2437.88. Therefore, though high display fidelity has the higher SSQ scores,

Figure 4.7 Comparison of the simulator sickness effects among the experimental conditions. Statistical significance: '***' for $p < 0.001$, '**' for $p < 0.01$, and 'ns' for $p > 0.05$



Source: Compiled by author.

it does not block the user of using the serious game.

4.3.3 Simulator Sickness vs. User Performance

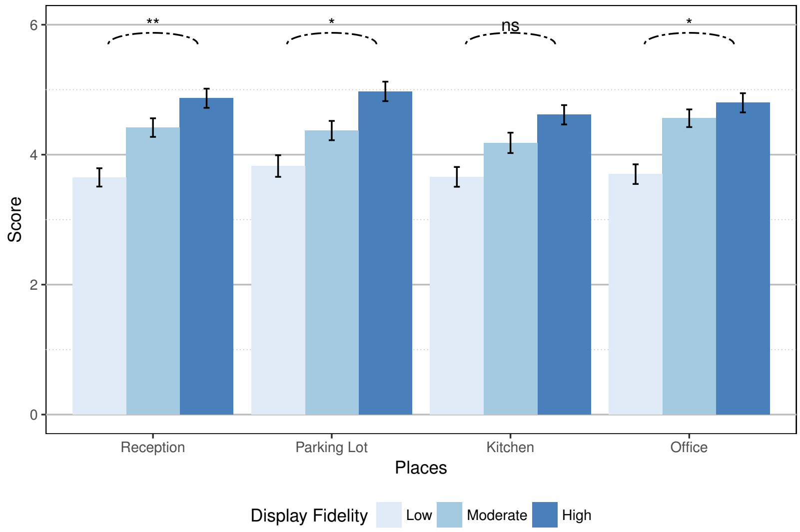
As showed before, the number of risks found was not significantly different among the experimental conditions. Although, for composite risks, subjects of high display fidelity had better performance than the users of other conditions. The high display fidelity also presented higher sickness effects against other conditions. It means that simulator sickness has no adverse effects on risk perception. For the completion time, we observed that subjects of moderate display fidelity took longer to finish the test. However, this condition presented lower sickness effects than high display fidelity, which had the lowest completion times. We can observe that simulator sickness is not affected by the execution times and vice verse. Based on these results, we consider H3 confirmed.

4.3.4 Other Results

Based on the subjective sense of presence, it was possible to verify that users of high display fidelity condition felt more present in all areas than the subjects of other conditions, as observed in Figure 4.8(a), with $p < 0.002$ among the conditions for the reception room, $p < 0.014$ for the parking lot, $p < 0.039$ for the kitchen, and $p < 0.013$ for the office. We also observed that, in high display fidelity, the highest sense of presence was reported at the parking lot, with $p < 0.009$.

They also classified their sense of immersion, in which we notice that the subjects of high display fidelity condition felt more immersed (M=4.8, SD=1.58) than moderate and low display fidelity, respectively, with $p < 0.001$. This result is consistent with the sense of presence indicated by the participants. Though they felt more

Figure 4.8 Comparison of the subjective sense of presence of participants in each place of the virtual environment among the experimental conditions. Statistical significance: '**' for $p < 0.01$, '*' for $p < 0.05$, and 'ns' for $p > 0.05$



Source: Compiled by author.

present and immersed in the high display condition, these values are considerably low, as the greatest possible value for immersion is 10, and for presence is 7.

4.4 Discussion

This experiment investigates the effects of different levels of display fidelity in the users' perception in a serious game for risk perception analysis. We observed the performance of the user through the number of risks found and the time s/he took to play the game, the sickness effects caused by the immersion, and the participants' sense of presence.

We hypothesized that higher and moderate levels of immersion would not differ significantly on user performance. In general, it was not found any difference in the number of risks selected between the experimental conditions. However, comparing the total of selected risks by category (e.g., simple and composite), we found that subjects of high display fidelity found a greater number of risks than the ones in low and moderate conditions, respectively. Though participants took longer to complete the trial with the moderate display than with other conditions, the total number of objects selected was similar between the experimental conditions. Subjects considered themselves non-experienced with work safety, which is coherent with the low amount of possible risks found, as well as the higher percentage for simple than composite risks.

The subjective sense of presence was greater for the higher immersion condition, which may suggest that participants selected more composite risks in this condition because they felt like they were in the virtual environment and, therefore, they could stay more focused on the further risks offered by some objects. Another factor that may contribute to this result is that most of the users (94%) found easier to use the simulator under the high immersion condition, which suggests that they could be more focused on the task instead of navigating and interact with the serious game.

The novelty of the immersive setup may cause distractions, making the participants forget to search for risks, only following the storyteller instructions, which can also explain the low number of risks found. It was observed by the researcher while applying the test that some participants completely forgot about doing the secondary task or simply did knew what is considered a potential risk. In the comparison of the number of risks found in between the instructions, it is possible to observe that the users found more risks when the storyteller said to do so, which suggests that they focused on following the instructions. We found statistical significance for this issue in the lower and higher immersion.

For the simulator sickness symptoms, results showed that the high display fidelity presented higher sickness scores. The data also shows that with low immersion, the users felt better after the exposure to the serious game. In this case, we consider that the setup had no effect on the simulator sickness symptoms. Though the higher immersion was more harmful than other conditions, it had no effect on the general risk perception rate. Besides that, in high fidelity condition, subjects were able to select more composite risks than in other conditions, which showed that sickness did not introduce adverse effects on this issue.

The game duration time ranges from 6 to 15 minutes, been higher in the moderate display fidelity condition. This difference of exposure time did not show any effects on perception rate or simulator sickness severity. This result, along to the impact on perception, allows choosing the visual immersion offering more or less immersion without being prejudicial to the user's performance.

The results of this experiment demonstrate that the visual immersion had no significant effects on the users' perception in a serious game for risk perception analysis. Although the results may be a consequence of the low experience of the users, it was a constant, so when observing the simple risks found (e.g., the ones that were common for all the participants), most of them had a similar performance. It is also relevant to observe that the player had two tasks to execute in the game, and in the cases that the primary and secondary assignments were the same (instructions #2 and #6) they found more potential risks than in the remaining instructions. It suggests that the participants are more likely to following orders and may not be able to multitasking.

5 EXPERIMENT II - INTERACTION FIDELITY AND KNOWLEDGE RETENTION

This experiment focuses on investigating the effects of interaction and locomotion on perceptual learning in a serious game dedicated to educating workers on basic safety procedures for electrical installations on public utility poles, more specifically, for the replacement of lightning rod.

The organization of this chapter is as follows. Section 5.1 presents the hypotheses that guided this experiment. Section 5.2 show the materials and methods employed. It includes the subjects that took part of the test, the apparatus used to implement the interaction and locomotion techniques, and the procedure that leads to the application of the experiment. Section 5.3 presents the results obtained. Finally, in Section 5.4 the results are discussed.

5.1 Hypotheses

This experiment investigates the effects of locomotion and interaction on perceptual learning. The overarching hypothesis is that learning may be affected when the interaction and locomotion techniques add too much workload to the game or cause much discomfort. On a more particular level, this experiment tests the following hypotheses:

- **H1. The repeated exposure to the serious game will improve the user's performance.**
  Every new exposure of the user to the serious game, s/he will be training its abilities about the content of the game as well as about the technique of interaction and locomotion. Therefore, we believe that, even for the conditions more difficult, they will learn how to use the technique and will execute the tasks in the game more correctly and faster.

- **H2. Higher interaction and locomotion fidelity will improve the user performance.**
  Higher interaction and locomotion fidelity are supposed to be natural to the user because it uses movements similar to s/he would do in real world. For the WIP, the user will walk as in real life, but without leaving the place, and for the high interaction, it will use the arms movements for moving the virtual hands. Considering that these movements are natural for the user, s/he will be able to concentrate on the task, which will improve the correctness and completion time.

- **H3. User performance will be affected by the workload.**
The techniques with high workload will require that the user stays more focused on learning how to use the technique than on performing the tasks in the game. Therefore, we believe that workload will affect the correctness of the tasks and completion time negatively.

- **H4. The knowledge retention will be affected by the workload.**
As for user performance, we believe that higher workload will worsen the knowledge retention. The user will be much more focused on use the interaction and locomotion techniques than in the safety procedures given by the narrator.

- **H5. Learning will not be significantly affected by the simulator sickness symptoms.**
Though we believe that higher locomotion and interaction fidelity will increase the severity of simulator sickness symptoms, it will not affect the gaining of knowledge. Considering that in Experiment I, the sickness symptoms had no effect on user perception and that audio instructions give the safety procedures, we suppose that this issue will not be prejudicial.

5.2 Materials and Methods

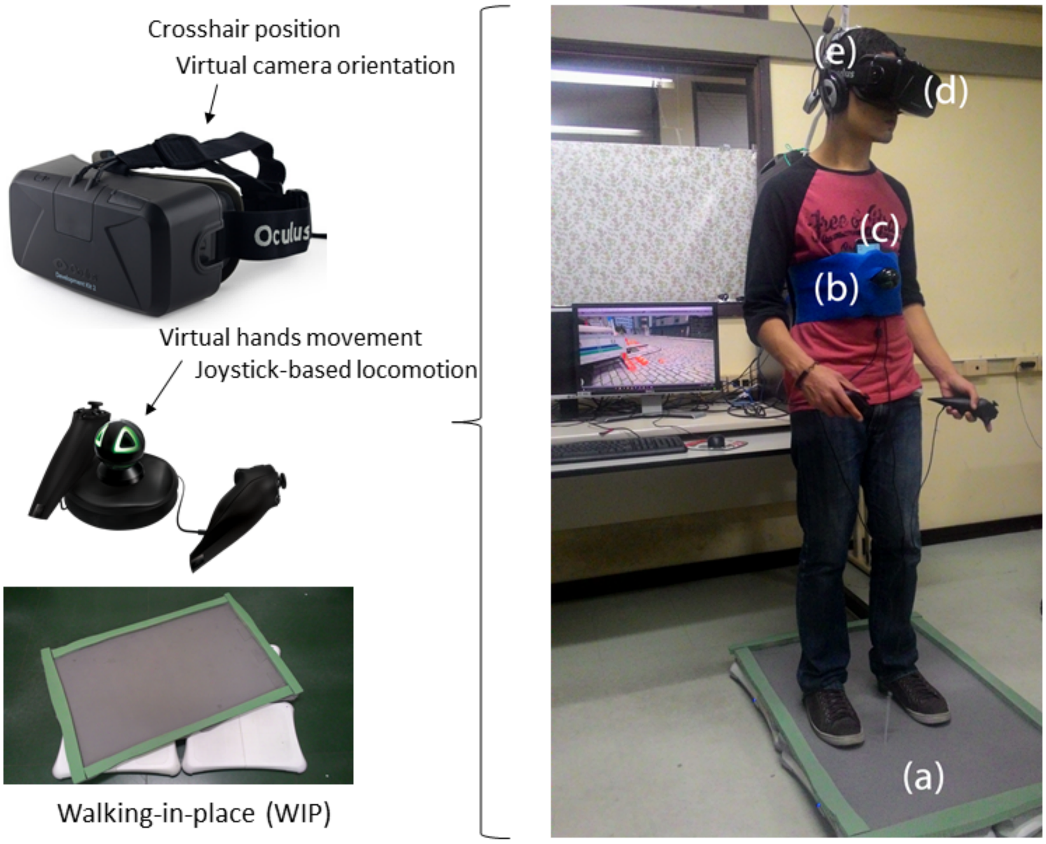
5.2.1 Experimental Conditions and Apparatus

We equipped the participants with an Oculus Rift DK2 device, with a resolution of 1920 x 1080 pixels per eye and a refresh rate of 75Hz, a headphone, a Sixense Razer Hydra, and they stay on top of a WIP platform, built with four Wii Fit Balance Boards (Figure 5.1). The Razer Hydra device, used for the non-natural locomotion and both interaction techniques, has a base station with a low-power magnetic field that gets the controllers' position and orientation. Although it does not require a line of sight to the controllers, the range area is quite small, which cause damage to the tracking when the user turns back to the base station. It was used a band to hold the base station along to the user's body (see Figure 5.1(b)). Between the base station and the user's body, we fixed a smartphone from which we get the sensors information to obtain its orientation and, then, infer the user's body orientation. Using this configuration, we were able to update the virtual hands according to the base station position and orientation, following the player's movements.

For this experiment we varied the interaction and locomotion fidelity considering four experimental conditions. Regardless the condition, users were standing on the WIP platform.

High-interation, high-locomotion (HIHL): This condition is the combination that represents the most natural communication with the virtual environment. The locomotion is through the WIP technique, the speed is controlled by the user according to the pressure on the balances and is limited by a pre-determined threshold. Participants only could walk towards their point of view, and they could not walk back. The interaction used virtual hands that mimic the movement of Hydra's controllers (i.e., the movement of the user's hands). To select the objects a button in the controllers should be pressed to make the index finger point to the target object. While they were pressing the button, the object would move along with the player. To release the object, they should release the button.

Figure 5.1 Physical setup. (a) Walking plataform. (b) Band to hold the Razer Hydra base station along to the user's body. (c) A smartphone to get the user's orientation. (d) Oculus Rift. (e) Headphones.



Source: Compiled by author.

High-interaction, low-locomotion (HILL): As in the HIHL, the interaction was through virtual hands. The locomotion, however, was done using the controllers' analogic joystick. To have equality between the techniques, the player only could move forward, always towards its point of view. The speed was constant during the game and the same for all participants.

Low-interaction, high locomotion (LIHL): This condition uses the WIP locomotion mode, as in HIHL, and a low interaction fidelity. The interaction consists of targets the objects with a crosshair that follows the head orientation.

Low-interaction, low-locomotion (LILL): This condition is the less natural combination of interaction and locomotion levels of fidelity. It uses the crosshair to target the objects, as in the LIHL, and the joystick to locomotion, as in the HILL.

5.2.2 Procedure

We told the participants that the experiment goal was to explore the effects of locomotion and interaction on learning obtained through serious games that educate the user about necessary safety procedures to the replacement of lightning rod, which also serves for any other electrical installations on public utility poles. They agreed to a consent term for participating in the experiment. Then, it was applied a characterization form (age, gender, education level, vision problems, experience with virtual reality, 3D video games, and safety work). It was also applied a pre-test questionnaire about self-efficacy memory (ZELINSKI; GILEWSKI, 2004), the SSQ, and the knowledge test.

After the user had answered the pre-test questionnaires s/he was invited to stay on top of the walking platform. The experimental condition that the user would use to play the game was defined randomly. The researcher then fixed the band with the Hydra's base station on the user's chest (a little down for female) and explained how to interact with the game. The Oculus Rift was put on, and the instructor opened a demo scene to calibrate it, making the user comfortable with the equipment. Finally, the researcher put the headphone on the user.

The trial contains four stages: training, learning, evaluation level 1, and evaluation level 2. In the training phase, the user could train the interaction and locomotion. For users of LIHL and HIHL conditions, it is also defined a comfortable threshold for the WIP technique, according to the user's preferences. We used a simple scenario containing two objects and two arrows (see Figure 5.2), the user should put each object in the position indicated by the arrow. In the learning stage, s/he would learn the safety procedures. And, in both evaluation stages, s/he would practice the knowledge acquired.

Figure 5.2 Virtual environment from user point of view used for the Training stage.

Source: Compiled by author.

At the end of the game, the participant was instructed to answer a post-test questionnaire. It contains the NASA TLX workload form, the SSQ, the GEQ, and the knowledge test. Then, s/he received a candy as gratitude for the participation.

5.2.3 Subjects

The experiment involved a sample of 60 participants, from which 13 were not able to complete the experiment because of simulation sickness. Thus, we had a sample of 46 valid participants (4 female), which an age range between 18 and 33 years (M=23.21, SD=3.39), mostly computing students. Almost 20% of them also took part in the previous experiment. They were split into four groups, one for each experimental condition. Generally, they consider themselves experienced on video games 3D and having little experience with virtual reality, having no significant difference between the groups.

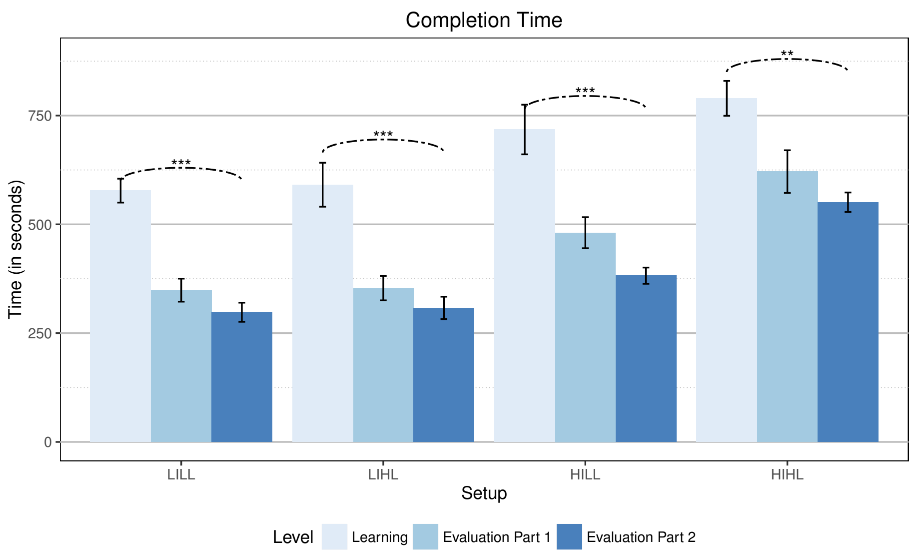
As on the previous experiment, we asked the participants whether they have any vision problems. There were 43% of the users that reported myopia, 30% have astigmatism, and one person has hypermetropia. Also, four people needed to use their glasses along with the Oculus Rift. None reported extra discomfort while playing the game or after.

5.3 Results

5.3.1 User Performance vs. Repeated Exposure

We assessed user performance through the completion time, correctness of execution of tasks in the evaluation stages, and the number of tries necessary to place the object in the desired position. For all experimental conditions, we observe that the completion time was higher for learning stage, and reduced significantly as well as subjects played the next level, with $p < 0.001$ (see Figure 5.3). The highest times were registered for HIHL condition, with 789.86 seconds for the learning stage, and 550.82 seconds for the second level of evaluation.

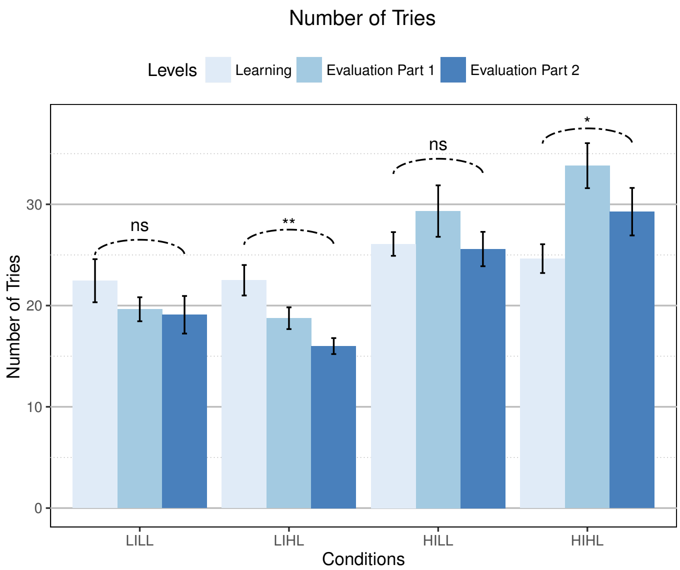
Figure 5.3 Completion time for the levels of the game in each experimental condition. Statistical significance: '***' for $p < 0.001$, '**' for $p < 0.01$.



Source: Compiled by author.

In general, participants performed the tasks correctly in almost every trial. The instruction #2 had 97% of correctness, and the instruction #3 was conducted successfully in all trials. In the instruction #4 the users put the tools correctly on the canvas in about 93% of the trials. The instruction #5 had 95% of hits, and we observed that in the evaluation stages, about 88% of the users put the signposts in the places that were indicated by the arrows in the learning stage. The instructions #6 and #7 were evaluated observing whether the users isolated the minimum area and whether they left an opening for that the workers could transit. The participants put the cones correctly around the workplace in 78% of the trials, but they left an opening in only 36% of the cases. This rate of correctness did not change significantly in the last level of evaluation.

Figure 5.4 Comparison of the number of tries to place the objects between levels at each experimental condition. Statistical significance: '**' for $p < 0.01$, '*' for $p < 0.05$, and 'ns' for $p > 0.05$



Source: Compiled by author.

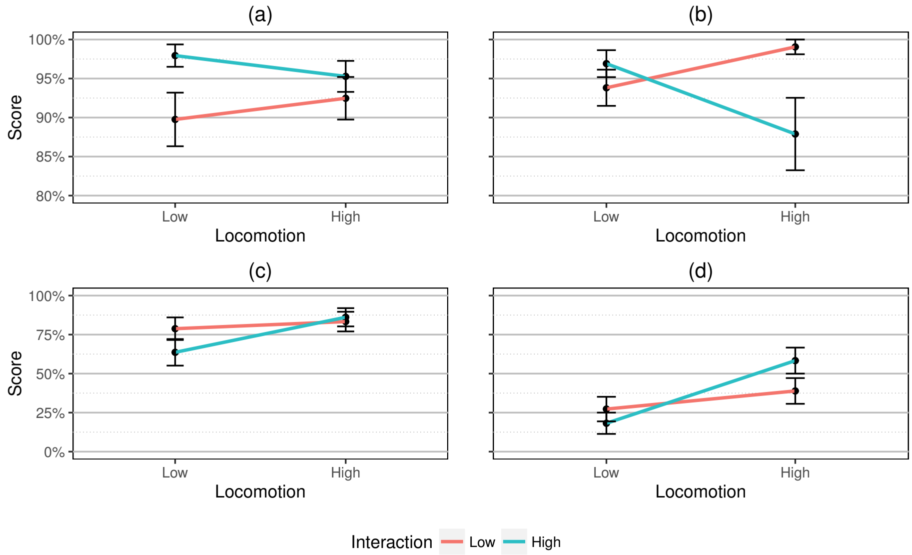
The number of tries necessary to place the objects in the desired position and reach this level of accuracy is presented in Figure 5.4. We found that the number of tries have reduced over the levels for two experimental conditions: LIHL with $p < 0.031$ and HIHL with $p < 0.011$. We also can observe that for the conditions that uses high interaction fidelity the subjects made more tries in the first stage of evaluation than in the learning stage, but this number reduce for the last evaluation phase. This result may be consequence of the fact that in the first stage of evaluation they should remember which tasks they should do and, at the same time, be used to the technology.

It is possible to observe that for all conditions, the number of tries reduced in the evaluations stages. Thus, it allows us to conclude that subjects improved their performance when they played the next level, reducing the completion time and the number of tries, and maintaining the level of correctness of the tasks. Therefore, we consider the H1 confirmed.

5.3.2 User Performance vs. Interaction and Locomotion Fidelity

Though users performed the tasks correctly practically every trial, we found some significance differences in the user performance between the experimental conditions. As observed in Figure 5.5(a), subjects of HILL performed the task of putting the tools over the canvas more correctly than the users of LILL group, with $p < 0.038$. The instruction #5 was accomplished more rightly by subjects of LIHL than users of HIHL condition ($p < 0.023$). It was also executed better using the HILL than the LILL condition, with $p < 0.036$ (see Figure 5.5(b)).

Figure 5.5 Percentage of correctness of tasks in each experimental condition. (a) Place tools on the canvas. (b) Place the signposts in visible places. (c) Isolate the workplace. (d) Opening.



Source: Compiled by author.

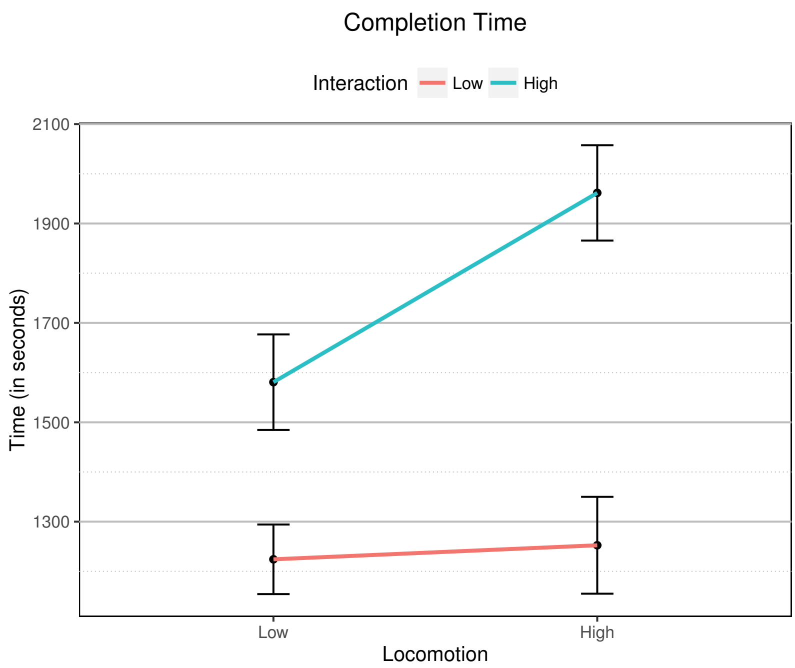
We observed that the task of isolate the workplace was executed more correctly by subjects of HIHL than the ones of HILL condition ($p < 0.032$). Also, users left an opening more times in the HIHL than in the HILL condition ($p < 0.001$). The number of tries necessary to place the objects correctly was higher for HIHL than HILL condition, with $p < 0.001$, and participants also needed more attempts in the LIHL than LILL condition, with $p < 0.001$.

Figure 5.6 present the total completion time for each experimental condition. We observed that participants of HIHL condition took longer to complete the trial than subjects of HILL condition, with $p < 0.012$. The time was also higher for high interaction fidelity than low interaction for both conditions with low locomotion fidelity ($p < 0.001$) and high locomotion ($p < 0.008$).

The completion time of each instruction showed differences between the experimental conditions (see Figure 5.7). In general, the task #2 (place the shims at the truck wheels) and #7 (tie ropes between cones) took longer to be completed, with $p < 0.001$, being the highest time observed for the instruction #7 in the HIHL condition (M=224.37, SD=39.08). From this result, we may suppose that the task of place the shims in the truck wheels took longer because it was the first instruction, and subjects were not used to the environment and the interfaces of interaction and locomotion. The task of isolate the workplace also present high completion time, probably because it requires that users to select only one rope at a time and carry it to the cones, repeating this procedure until the area be properly isolated. The data also showed that the instruction #6 was the fastest one to be performed in all experimental conditions, with the lowest time found for HILL condition (M=47.12, SD=13.57).

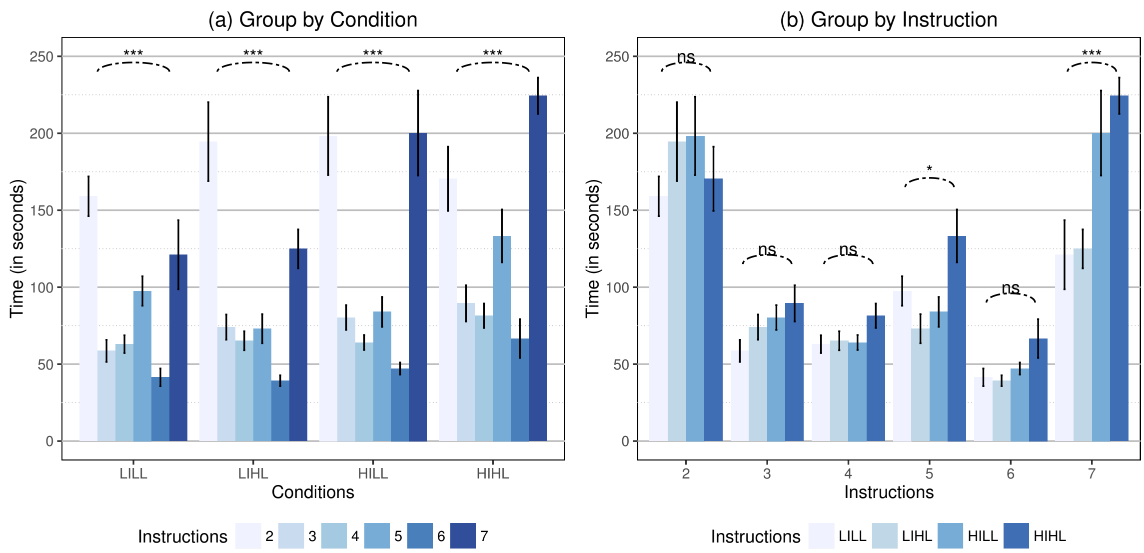
Considering each instruction individually, we observed that subjects of HIHL took longer to complete the instruction #5 than other conditions, with $p < 0.05$,

Figure 5.6 Completion time in the experimental conditions.



Source: Compiled by author.

Figure 5.7 Comparison of the completion time between instructions in the learning level for each experimental condition. Statistical significance: '***' for $p < 0.001$, '*' for $p < 0.05$, and 'ns' for $p > 0.05$



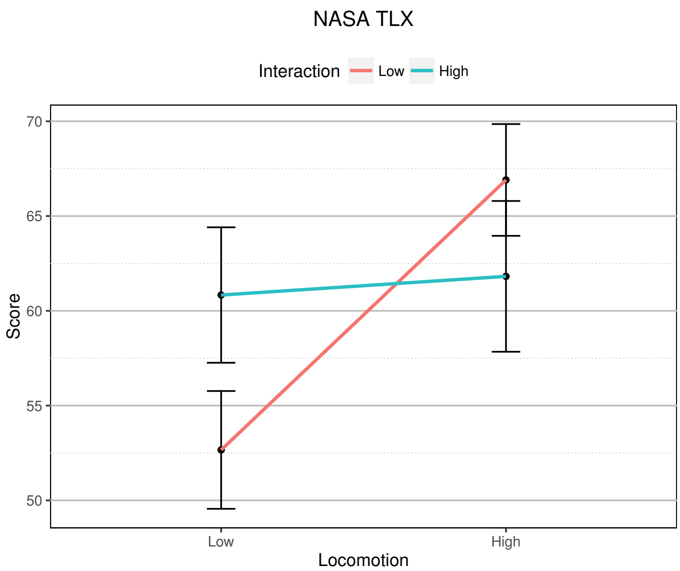
Source: Compiled by author.

while the lower completion time for this task was for LIHL condition. The instruction #7 also took longer to be completed in the HIHL condition, while it was faster using the LILL condition ($p < 0.001$).

These results show that naturalness increases the completion time and the number of tries required to place the objects correctly. Locomotion fidelity has positive impacts in the correctness of the task in which users were asked to isolate the workplace, leaving an opening to workers transit. The interaction fidelity improved the task of place the tools on the canvas. However, it had negative impacts in the task of place the signposts in visible locations. We observe that naturalness has positive effects in some cases, especially for tasks that require higher attention from the user, but it worsened the general user performance. Thus, we cannot consider H2 correct.

5.3.3 User Performance vs. Workload

Figure 5.8 Workload scores for each experimental condition.



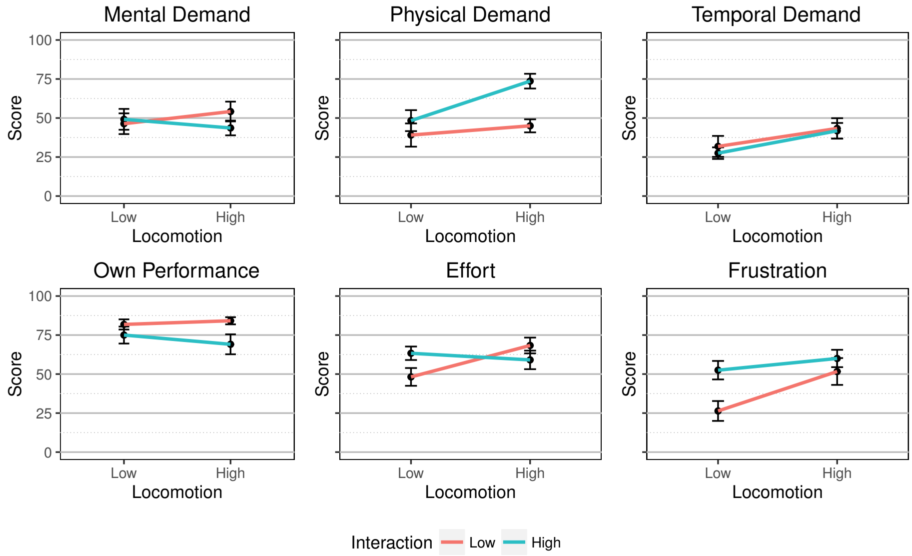
Source: Compiled by author.

Figure 5.8 presents the mean scores for all experimental conditions. Mean scores show that LIHL condition had a higher workload LILL condition, with $p < 0.0032$. We did not observe statistical significance between the remaining of the techniques.

We evaluated the effects of the experimental conditions in each subscale of the NASA TLX questionnaire (Figure 5.9). In the groups with high locomotion fidelity, the high interaction fidelity was more physically demanding than low interaction ($p < 0.001$). For the conditions with high interaction fidelity, the high locomotion presented higher physical demand than low locomotion ($p < 0.006$). Temporal demand was higher for high locomotion in the conditions with high interaction fidelity ($p < 0.033$). Participants found more efforting to navigate with high locomotion in the conditions with low interaction ($p < 0.012$). The effort was also higher for high interaction in the conditions with low locomotion ($p < 0.036$). High locomotion fidelity was more frustrating than low locomotion in the condition with low interaction ($p < 0.048$), and the frustration was also higher for high interaction in the conditions with low locomotion ($p < 0.006$). We did not observe statistical significance for the remaining of techniques and scales.

From the results of user performance presented in the previous subsections, we noted that the experimental conditions that require higher number of tries to place

Figure 5.9 Workload scores for the experimental conditions in each subscale of the NASA TLX.



Source: Compiled by author.

the objects correctly are also the ones that introduced higher scores of workload and was considered more frustrating than other conditions. In general, we found that locomotion and interaction fidelity increase workload. For some tasks, such as place the signposts in visible positions, the workload had no negative impacts. However, considering completion time, the number of tries and the result of the remaining tasks, workload has adverse effects on user performance. Based on this, we can consider H3 confirmed.

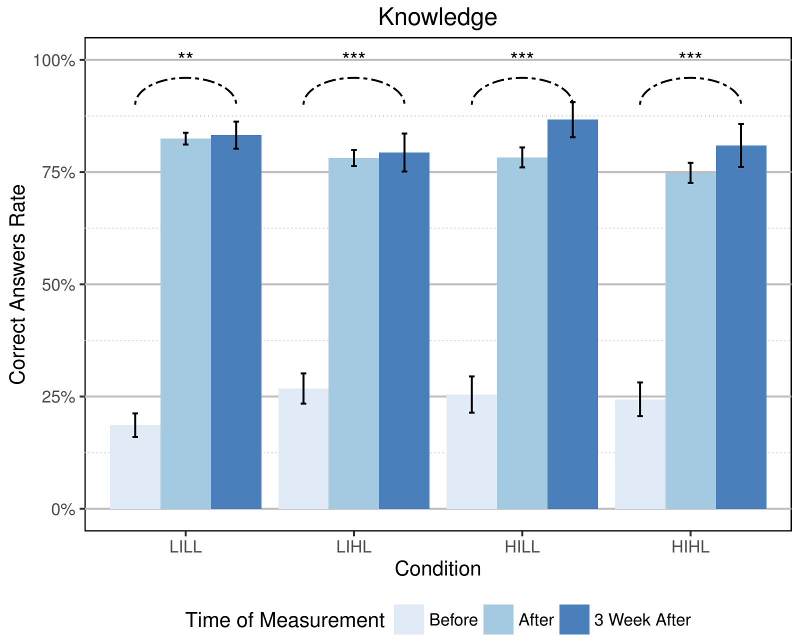
5.3.4 Workload vs. Knowledge Retention

We administered the knowledge test three times: before trying the serious game (pre-test), immediately after trying it (post-test), and three weeks after the trial (retention-test). Mean pre-test score showed that, before trying the experimental conditions, participants were able to provide about 24% of correct answers (M=0.24, SD=0.11). An independent One-Way ANOVA showed no significance differences in knowledge before the trial between the four groups (p-value = 0.7091).

Three weeks after the trial, we sent the knowledge test to the participants by e-mail, from which we had 41 answers. Thus, the knowledge data was analyzed over 89% of users. Mean post-test scores showed that the participants were able to provide 78% of correct answers. In the retention test, subjects were able to provide 82% of correct answers. We observed that users acquired knowledge through the serious game, with $p < 0.01$ for the LILL condition, and $p < 0.001$ for the remaining of experimental conditions.

Figure 5.10 shows that subjects acquired knowledge through the serious game and were able to recall it in the third application of the knowledge test. We observe that in the HILL condition participants were able to remember significantly more information about the safety procedures in the retention test than in post-test ($p < 0.027$). The remaining of the conditions also presented higher scores in retention test

Figure 5.10 Comparison of the knowledge scores in each experimental condition between times of measurement. Statistical significance: '***' for $p < 0.001$, and '**' for $p < 0.01$.



Source: Compiled by author.

than in the post-test. However, we did not observe statistical significance between them. Mean scores of post- and retention test did not present statistical significance between experimental conditions.

We observed that the LILL condition had lower workload than other conditions, and allow users to recall the information in the retention test even better that they did in the post-test. In this case, low workload showed positive effects on knowledge retention. However, even in the experimental conditions that presented higher workload, participants were able to remember the knowledge acquired in the serious game. Therefore, the high workload has no negative impacts on knowledge retention. From this results, we can't confirm, because participants were able to retain the knowledge acquired in all experimental conditions, despite the workload introduced by the interface. Our results suggest that there is no relation between the workload and knowledge retention.

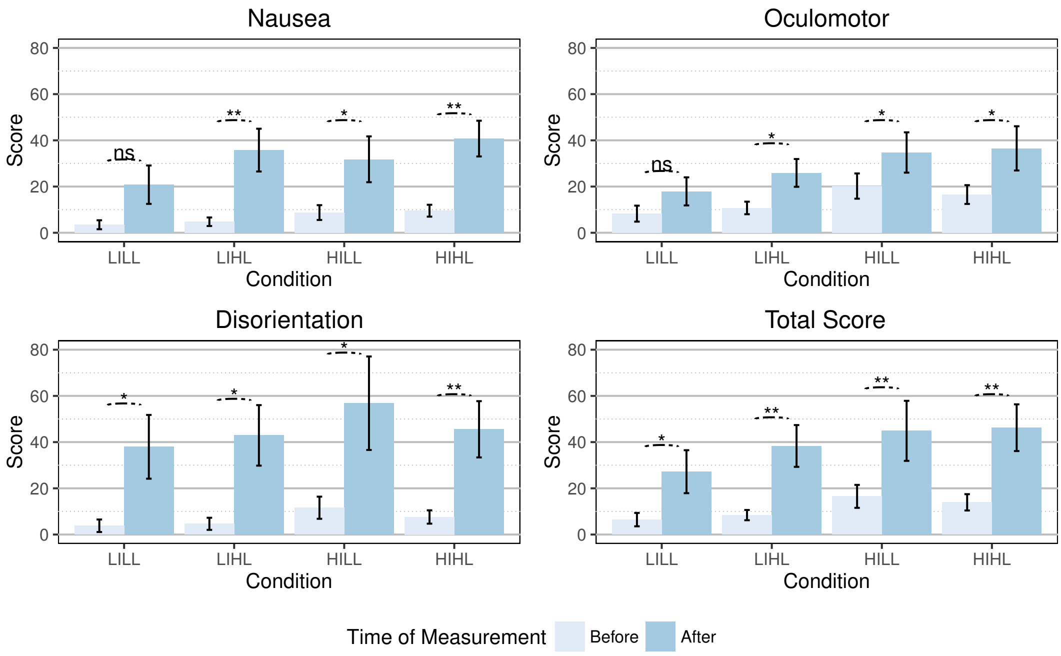
5.3.5 Simulator Sickness vs. Knowledge Retention

We administered the SSQ in two times: before trying the experimental condition (pre-test), and immediately after trying it (post-test). Mean pre-test score showed that the participants were not feeling completely well, presenting mostly fatigue, headache, and eyestrain. The SSQ scores: Nausea (M=6.53, SD=8.5), Oculomotor (M=14, SD=14.2), Disorientation (M=6.96, SD=11.64), and Total Score (M=11.38, SD=12.44). An independent Kruskal-Wallis rank test showed no significance differences in the initial scores between the groups (p-value = 0.19 for Nausea, p-value = 0.17 for Oculomotor, p-value = 0.45 for Disorientation, and p-value = 0.14 for Total

Score).

Simulator Sickness scores showed that all experimental conditions were harmful to the participant (see Figure 5.11). Table 5.1 presents the range of scores between before and after the trial and the p value found in the statistical analysis for each scale. Subjects of HIHL showed higher symptoms of nausea and oculomotor scales. Symptoms of disorientation scale were greater in the HILL condition. The LILL condition presented the lowest scores for symptoms in all three scales. We observe the same behavior for the total score of sickness. Note that these scores are very low, as the highest possible total SSQ score is 2437.88. Thus, we did not observe statistical significance between the experimental conditions.

Figure 5.11 Simulator sickness effects measured before and after the trial for the scales of Nausea, Oculomotor and Disorientation, and the Total Score of sickness. Statistical significance: '**' for $p < 0.01$, '*' for $p < 0.05$, and 'ns' for $p > 0.05$.



Source: Compiled by author.

Table 5.1 Simulator Sickness scores. Mean scores, standard deviation and p-value found in the analysis between before and after the trial for each scale and the total score.

Condition	Nausea			Oculomotor			Disorientation			Total Score		
	M	SD	p	M	SD	p	M	SD	p	M	SD	p
LILL	20.81	27.58	0.099	17.91	20.13	0.2	37.96	45.78	0.04	27.2	30.84	0.041
LIHL	35.77	32.06	0.003	25.89	20.81	0.023	42.92	45.38	0.013	38.33	31.33	0.002
HILL	31.8	34.35	0.016	34.74	30.22	0.02	56.84	70.09	0.014	44.88	45.02	0.005
HIHL	40.76	25.63	0.005	36.52	31.76	0.02	45.55	40.38	0.008	46.24	33.54	0.005

Source: Compiled by Author

As showed before, the retention test had even better results than the post-test, particularly for the HILL condition. The sickness effects were not significantly different between the experimental conditions, as well as retention test results. Considering that simulator sickness scores were very low and retention test scores were high for all conditions, we conclude that simulator sickness has no adverse effects on knowledge retention. Thus, we consider H5 confirmed.

5.4 Discussion

In this experiment, we investigated the effects of interaction and locomotion on perceptual learning. We used a serious game developed with the goal of educating workers about work safety, more specifically, which procedures should be adopted to execute installations of lightning rods. Two techniques of locomotion and interaction were compared using a between-subjects design. The game has three levels: the first one instructs the player about the safety procedures, and the other two are used to evaluate the learning obtained.

Results showed that the completion time reduced when the user played the next level. We observe the same behavior for the number of tries to place the objects in the final positions. In this case, the amount was reduced in the conditions that use high locomotion fidelity. Subjects complete the level faster, and the number of hits remains similar, which means that they were able to repeat the task more quickly and efficiently.

High levels of naturalness presented higher workload than non-natural techniques. The workload introduced by the interaction and locomotion interfaces had negative impacts on user performance, increasing the completion time and the number of attempts to place the objects correctly. The task to put the signposts in visible locations also was affected by the workload, in a way that in conditions with high locomotion, subjects performed the task more correctly with low interaction. However, the workload was not an issue to that users isolate the workplace and place the tools on the canvas correctly. This results may suggest that even naturalness presented higher workload and higher completion time and the number of attempts, it allows users to execute most of the tasks more correctly.

The number of attempts was higher for high interaction in the conditions that use low locomotion fidelity. This information is consistent with the result in which participants felt more frustrating using the HILL condition than LILL. It may suggest that the frustration was also caused by the repeated tries to execute the tasks.

The cybersickness was a feature addressed by this thesis because it is still a problem in VR systems. Indeed, some participants were not able to finish the experiment because they felt sick. Assessing the simulator sickness scores for the users that completed the experiment, we observe that all experimental conditions increased the sickness symptoms. However, the scores were very low and we did not found any significance difference between the interfaces.

The mean completion times ranged between 18 and 35 minutes, and even that time was significantly higher in some experimental conditions, as in HIHL condition, it did not cause the sickness symptoms to be more severe.

High locomotion fidelity presented higher completion times than low fidelity in the conditions with high interaction. For the interfaces with low interaction fidelity, high locomotion fidelity did not increase the completion time. Therefore, it is possible to observe that more natural conditions, although improve the user performance in more thinking tasks, also increase the time to complete it. For instance, the isolation of the area was executed more correctly in the HIHL condition and also took longer in this condition than in the others. Otherwise, for tasks that do not demand much attention, such as placing signposts in visible locations, the user performance was better using low interaction.

The obtained knowledge was similar for all experimental conditions. The retention test presented even better results than the post-test, particularly for HILL experimental condition. Though simulator sickness had no adverse effects on knowledge retention, results suggest that in the post-test user had lower scores than retention test because they were under the sickness effects.

In general, the knowledge was not significantly affected by the interaction and locomotion techniques. Results indicate that subjects were able to recall the instructions received. However, they were not always able to execute the tasks correctly in the serious game, either because they did not understand the instruction or the workload introduced by the interface was prejudicial. In this case, we observe adverse effects only for the task to place the signposts in visible locations.

6 DESIGN GUIDELINES

This thesis aims to identify a set of design guidelines to develop serious games for learning that take advantage of the benefits of immersion while being comfortable and efficient. To reach this goal, we conducted two empirical studies that investigate the effects of the display, interaction and locomotion fidelity in the serious game effectiveness and user performance, addressing issues such as simulator sickness, workload, user performance and knowledge retention.

The serious games used in the user experiments belong to a very specific knowledge area. In the first game, participants should execute a scanning task (e.g., look over the place looking for objects that fit the given description) searching for objects that represent some dangerous, but they had little or none experience on work safety. The second game could be considered more generic because the subjects learned the content through the game, though the content is very specific.

The guidelines address the idea of having complex and simple objects and tasks in the serious games. The complex objects in a scanning task are the ones that do not fit directly with the description of the search, such as the complex risks presented in the first game. The simple objects are the ones that the description fits immediately. Similarly, the complex tasks are the ones that require a higher level of attention to be accomplished, and the simple ones can be executed quickly and do not require much attention.

Use high display fidelity for scanning tasks involving complex objects. Composite risks require additional attention to be identified because it is necessary to consider further events that may trigger an accident because of that possible element of risk. For instance, a flashing lamp may not present immediate risk, however, in the future it may cause an accident due to a short-circuit. Simple risks require low attention because the threat presents itself immediately. For instance, a hole in the ground can cause direct personal accidents. The results of the first experiment showed that, in general, the number of risks found was not very different between the conditions. However, the subjects were able to select significantly more composite than simple risks with the high display. Thus, if the scanning tasks consist of many complex objects is advisable to use high immersion.

Use high display fidelity for serious games that require a high level of realism. We showed in Experiment I that the users' sense of presence was greater using high display. We also verified that in the same condition they were able to select a higher number of composite risks. Thus, the fact that they felt embedded in the virtual environment may help them to analyze the objects with higher judgment identifying the further risks that it may offer, than in the conditions that they did not feel so present. Considering that high display fidelity presented lower completion time and a higher number of composite risks than other conditions, we can suggest that with high display combined with high realism will provide better user performance.

Use low or moderate display fidelity for scanning tasks involving simple objects. Considering the number of risks found by category, we found that the number of simple risks was significantly higher than composite risks in all experimental conditions. Though high display fidelity may improve the user performance in some situations, as showed before, it also increases the sickness effects. Thus, if the task did not involve complex objects is more convenient to use low or moderate display, which presents a good user performance and is not harmful to the user.

Use high display fidelity for serious games that the primary goal is a scanning task. As previously presented by Ragan, et al. (RAGAN et al., 2015), higher FOVs lead to better target detection in a serious game that the challenge was to scan a street searching for dangerous situations. Our experiment also showed that high display fidelity is better to scan composite risks. However, in the serious game that we used in the experiments, the scanning task should be performed between the execution of some instructions, being a secondary task in the game. This fact along with the low experience of the users with work safety may explain the low number of risks selected. Therefore, if the primary goal of the game is the scanning task, it is more suitable to use high display. It will provide a higher field of view and will help the user to me more concentrate in the task, once that it will be more embedded in the game.

Use high interaction and locomotion fidelity for perceptual learning based on interaction. In the second experiment we addressed the naturalness of interaction and locomotion to execute an interactive task in the serious game. The results showed that interaction and locomotion fidelity has no impact on knowledge retention, once that the users could recall the content learned in the serious game even three weeks after the trial. However, we observed that the naturalness of the techniques had effects on user performance (e.g., completion time and correctness of the tasks). Higher interaction and locomotion fidelity affected the completion time negatively and increased the number of tries to place the objects, but it also increased the correctness of most of the tasks. As the primary goal of the serious games is to provide knowledge, the correctness of the tasks is a relevant issue, regardless of the completion time. Thus, it is more appropriate to use interfaces as natural as possible.

Use low interaction and locomotion fidelity to accomplish simple tasks fast and efficiently. We showed before that most of the tasks were accomplished more correctly using interfaces closer to the natural way of interaction and locomotion. However, we observed that for some simple tasks, as place the signposts in visible locations and put the tools on the canvas, the low fidelity of interaction or locomotion presented better results. It took less time to complete the task correctly and less attempts to place the objects in the final position than using high fidelity. Therefore, in serious games that contain tasks that do not require high attention it is more appropriate to use low interaction or locomotion. Low locomotion will also reduce the workload introduced by the interface.

Use repeatedly the serious game to improve the user performance. From the results, we observed that when the user played the next level its completion time reduced significantly. The number of tries to place the objects in the final position also reduced, especially for the evaluation stages. Though the time and attempts reduced, the correctness of the tasks remained similar between the game levels. It suggests that the user can execute the tasks faster and efficiently in each new try. Considering these results, it is possible to observe that the repeated exposure to the serious game has benefits in the user performance. Therefore, it is relevant that the game is played several times and that it has several stages of learning. This guideline can be applied to serious games with different purposes.

Considering the number of experimental conditions in each user experiment and that it was assessed between-subjects, we can see that the number of participants was low, especially in the second experiment. Besides that, a considerable number of subjects was not able to complete the trial because of sickness effects, which was prejudicial to find more volunteers, once that we have already used the time slot reserved for the experimentation. The profile of the participants was mostly students with low or none experience in work safety. This characteristic of the population was prejudicial especially for the first experiment.

7 CONCLUSION

The research presented in this thesis focused on investigating which is the best apparatus for physical immersion to provide perceptual learning through serious games. The goal has been to improve the design of physical immersion by studying the impacts of the display, interaction, and locomotion fidelity on user perception and knowledge retention. We also verified how the workload and *cybersickness* of each experimental condition affect these issues. Based on the data collected from two empirical studies, we identified a set of design guidelines for the use of physical immersion to develop serious games for perceptual learning. The guidelines (presented in Chapter 6) include considerations for navigation, interaction and display choices.

7.1 Summary of Findings

This section summarizes the findings of the research questions that were presented and explained in Section 1.3.

1. **What are the effects of display fidelity on user perception?**
We found that, in general, the display fidelity has no effect on user perception. For all experimental conditions, participants found a similar number of elements of risk. However, separating the risks in categories (e.g., simple and composite), it was observed that higher immersion has better results for composite risks, while moderate display fidelity was more efficient for finding simple risks. As users had a higher sense of presence with high display fidelity, it may suggest that subjects could focus themselves on the further risks that objects could offer, once that situation seemed almost real for them.

   (a) **Can moderate display fidelity provide as much effectiveness as higher display fidelity?**
   In general, the risk perception rate was similar for both experimental conditions. However, in high display fidelity, users selected higher number of composite risks than in moderate fidelity. For simple risks, we also did not observe any significance difference between the conditions. Therefore, it is possible to suppose that both medium and high display fidelity had similar performance for scanning task when it involves simple objects.

   (b) **How the levels of immersion affect the simulator sickness symptoms and how this affects user performance?**
   Higher display fidelity increases the severity of simulator sickness symptoms.

However, we did not find differences in risk perception caused by this issue. The completion time was higher for moderate display fidelity, but it did not aggravate the sickness effects, once that moderate condition showed lower sickness scores than high display fidelity. The completion times were low for all conditions (e.g., between 6 and 15 minutes), though. In general, we can suggest that display fidelity has no impact on user perception.

2. **What are the effects of interaction and locomotion fidelity on knowledge retention?**
Between before and after the trial, the participants had a significant gaining of knowledge. It was retained by them to the last knowledge test applied three weeks after the trial. The retention test had higher scores than post-test, and we found statistical significance for this difference in the LILL experimental condition. The scores of post- and retention-test did not present differences between the conditions.

(a) **How naturalness impact the simulator sickness symptoms and what are the effects of it?**
Results showed that all experimental conditions had sickness effects. However, we did not observe significance differences between conditions. Experiment I showed that higher levels of immersion present higher severity of simulator sickness symptoms. Therefore, we found that the sickness effects introduced by the experimental conditions may be a consequence of using HMDs, and the interaction and locomotion fidelity may not impact this issue. Simulator sickness did not affect user performance and knowledge retention.

(b) **How the workload affects the user performance and perceptual learning?**
Naturalness presented higher workload than the non-natural interfaces. The workload had effects on user performance increasing the completion time and the number of tries to place the objects. However, it did not affect the correctness of most tasks negatively. The task of place the signposts on visible locations presented better performance with low interaction in the conditions that uses high locomotion. The remaining tasks had better outcomes when users used high interaction. Thus, the workload impacts the user performance partially. The gaining and retention of knowledge were not affected by the workload.

Our research aid to improve the understanding of the effects of the VR systems in the effectiveness of the serious game, helping to choose the best approach to developing effective and comfortable serious games. We proposed a classification of the current immersive serious games considering the game characteristics and some relevant issues about the assessment of its effectiveness and usability. We observed that most of the studies to not consider different levels of immersion in the evaluation, assuming that a high level of immersion will be always the best choice. As we showed in this dissertation, it is not true for all situations, especially considering that higher levels of immersion may increase the severity of the sickness effects. Thus, our research fills this gap in the assessment comparing several levels of immersion and interaction to identify the appropriate level of immersion for serious games for learning, considering issues such as sickness, workload, and engagement.

7.2 Contributions

During the Master program, we participated in several projects along to other members of the research laboratory, besides the research presented in this dissertation. These projects were important so we could acquire experience to develop and execute the empirical studies conducted in this thesis. It was published three works, explained in the following.

We proposed an augmented mouse to aid people to stay more focused on their main work. The idea is that the user can do secondary tasks, such as media control, performing a short air gesture, without the need of switching between applications. We conducted two user experiments in order of developing an approach based on the prospect users' needs. From the results of the first experiment, we chose the gestures and tasks suitable for the proposal. The second experiment was conducted to validate our proposal. In 2016, we published a poster entitled '3D Gesture Mouse: Being multitask without losing the focus' on *IEEE Symposium on 3D User Interfaces* (FRANZ; MENIN; NEDEL, 2016a) and a full paper entitled 'Lossless multitasking: Using 3D gestures embedded in mouse devices' on *SBC Symposium on Virtual and Augmented Reality* (FRANZ; MENIN; NEDEL, 2016b).

Another project proposed the use of simulators for risk perception assessment and training. We conducted a user evaluation with target users to measure the effectiveness of the psychological risk VR simulator in helping recruiters assess the psychological profile of candidate workers. In 2016, we published a paper entitled 'Using Immersive Virtual Reality to Reduce Work Accidents In Developing Countries' on *IEEE Computer Graphics and Applications* (NEDEL et al., 2016).

In June of 2016, we presented the proposal of this Master thesis in a Workshop of Thesis and Dissertations on the *SBC Symposium on Virtual and Augmented Reality*, through a paper entitled 'Designing and Assessing Immersive Serious Games'. We obtained valuables considerations from the examination committee that helped in the execution of this dissertation.

7.3 Future Work

This research has contributed to a greater understanding of how physical immersion affects outcomes, user performance and simulator sickness in user perception and knowledge retention, but this is a small set of studies. Replication of these experiments is essential to validate (or refute) the observed findings and increase (or decrease) the strength of the claims made in this thesis. Further, our results do not provide a definitive set of outcomes for the design factors that we focused on (i.e., physical immersion). It is needed additional studies for a complete understanding of how these factors influence user perception and knowledge retention.

The empirical studies performed in this research varied several aspects of physical immersion. However, the serious game used to investigate the effects of this factors has a very particular application, which is work safety. Experiments addressing games with other applications are necessary to understand whether the design guidelines identified for this serious game is applicable for other purposes. These guidelines also have to be validated in the future works.

In the literature review, we addressed the importance of investigating the effects of using target users in the experiments to validate the serious games. In our studies, the majority of participants were computing students with minimal experience in work safety, which is the application area of the serious games used in this research. This profile of the subjects has impacts, especially for user perception. Therefore, additional studies using target users as subjects are relevant to verify the findings of this research with a public highly experienced in the application area of the serious game.

REFERENCES

ADAMO-VILLANI, N. A virtual learning environment for deaf children: design and evaluation. **International Journal of Human and Social Sciences**, [S.l.: s.n.], v.2, n.2, 2007, p.123–128.

ADOLPH, K. E.; KRETCH, K. S. Gibson's theory of perceptual learning. **i: Keller H.(Developmental Section red.) International Encyclopedia of the Social and Behavioral Sciences,**, [S.l.: s.n.], v.10, 2015, p.127–134.

ALAHMARI, K. A. et al. Comparison of virtual reality based therapy with customized vestibular physical therapy for the treatment of vestibular disorders. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, [S.l.: s.n.], v.22, n.2, 2014, p.389–399.

BACKLUND, P. et al. Sidh-a game based firefighter training simulation. In: INFORMATION VISUALIZATION, 2007. IV'07. 11TH INTERNATIONAL CONFERENCE. **Proceedings. . .** [S.l.: s.n.], 2007. p.899–907.

BACKLUND, P. et al. Games for traffic education: an experimental study of a game-based driving simulator. **Simulation & Gaming**, [S.l.: s.n.], 2008.

BALDOMINOS, A.; SAEZ, Y.; POZO, C. G. del. An approach to physical rehabilitation using state-of-the-art virtual reality and motion tracking technologies. **Procedia Computer Science**, [S.l.: s.n.], v.64, 2015, p.10–16.

BAROT, C. et al. V3S: a virtual environment for risk-management training based on human-activity models. **PRESENCE: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.22, n.1, 2013, p.1–19.

BEDOYA CASTAÑO, J. et al. Shoulder flexion rehabilitation in patients with monoparesia using an exergame. In: SERIOUS GAMES AND APPLICATIONS FOR HEALTH (SEGAH), 2014 IEEE 3RD INTERNATIONAL CONFERENCE ON. **Proceedings. . .** [S.l.: s.n.], 2014. p.1–5.

BELLOTTI, F. et al. Assessment in and of serious games: an overview. **Advances in Human-Computer Interaction**, [S.l.: s.n.], v.2013, 2013, p.1.

BENNETT, A.; COXON, M.; MANIA, K. The Effect of Stereo and Context on Memory and Awareness States in Immersive Virtual Environments. In: SYMPOSIUM ON APPLIED PERCEPTION IN GRAPHICS AND VISUALIZATION, 7., New York, NY, USA. **Proceedings. . .** ACM, 2010. p.135–140. (APGV '10).

BORG, G. A. Psychophysical bases of perceived exertion. **Med sci sports exerc**, [S.l.: s.n.], v.14, n.5, 1982, p.377–381.

BOWMAN, D. A.; HODGES, L. F. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In: INTERACTIVE 3D GRAPHICS, 1997. **Proceedings. . .** [S.l.: s.n.], 1997. p.35–ff.

BROOKE, J. et al. SUS-A quick and dirty usability scale. **Usability evaluation in industry**, [S.l.: s.n.], v.189, n.194, 1996, p.4–7.

CAI, Y. et al. Design and development of a virtual dolphinarium for children with autism. **IEEE transactions on neural systems and rehabilitation engineering**, [S.l.: s.n.], v.21, n.2, 2013, p.208–217.

CHA, M. et al. A virtual reality based fire training simulator integrated with fire dynamics data. **Fire Safety Journal**, [S.l.: s.n.], v.50, 2012, p.12–24.

CHALMERS, A.; DEBATTISTA, K. Level of realism for serious games. In: GAMES AND VIRTUAL WORLDS FOR SERIOUS APPLICATIONS, 2009. VS-GAMES'09. CONFERENCE IN. **Proceedings. . .** [S.l.: s.n.], 2009. p.225–232.

CHAN, J. C. et al. A virtual reality dance training system using motion capture technology. **IEEE Transactions on Learning Technologies**, [S.l.: s.n.], v.4, n.2, 2011, p.187–195.

CHITTARO, L.; BUTTUSSI, F. Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. **Visualization and Computer Graphics, IEEE Transactions on**, [S.l.: s.n.], v.21, n.4, 2015, p.529–538.

CORTI, K. Games-based Learning: a serious business application. **Inform of PixelLearning**, [S.l.: s.n.], v.34, n.6, 2006, p.1–20.

CRESCENTINI, C. et al. Psychological and physiological responses to stressful situations in immersive virtual reality: differences between users who practice mindfulness meditation and controls. **Computers in Human Behavior**, [S.l.: s.n.], v.59, 2016, p.304–316.

DE KORT, Y. A.; IJSSELSTEIJN, W. A.; POELS, K. Digital games as social presence technology: development of the social presence in gaming questionnaire (spgq). **Proceedings of PRESENCE**, [S.l.: s.n.], v.195203, 2007.

DIERS, M. et al. Illusion-related brain activations: a new virtual reality mirror box system for use during functional magnetic resonance imaging. **Brain research**, [S.l.: s.n.], v.1594, 2015, p.173–182.

FERRACANI, A.; PEZZATINI, D.; DEL BIMBO, A. A Natural and Immersive Virtual Interface for the Surgical Safety Checklist Training. In: ACM INTERNATIONAL WORKSHOP ON SERIOUS GAMES, 2014. **Proceedings. . .** [S.l.: s.n.], 2014. p.27–32.

FINKELSTEIN, S. et al. Astrojumper: motivating exercise with an immersive virtual reality exergame. **Presence: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.20, n.1, 2011, p.78–92.

FRANZ, J.; MENIN, A.; NEDEL, L. 3D gesture mouse: being multitask without losing the focus. In: IEEE SYMPOSIUM ON 3D USER INTERFACES (3DUI), 2016. **Proceedings...** [S.l.: s.n.], 2016. p.241–242.

FRANZ, J.; MENIN, A.; NEDEL, L. Lossless multitasking: using 3d gestures embedded in mouse devices. In: VIRTUAL AND AUGMENTED REALITY (SVR), 2016 XVIII SYMPOSIUM ON. **Proceedings...** [S.l.: s.n.], 2016. p.109–116.

GIBSON, E. J. How to think about perceptual learning: twenty-five years later. , [S.l.: s.n.], 1992.

GRAAFLAND, M.; SCHRAAGEN, J.; SCHIJVEN, M. P. Systematic review of serious games for medical education and surgical skills training. **British journal of surgery**, [S.l.: s.n.], v.99, n.10, 2012, p.1322–1330.

GRABOWSKI, A.; JANKOWSKI, J. Virtual Reality-based pilot training for underground coal miners. **Safety science**, [S.l.: s.n.], v.72, 2015, p.310–314.

GREUNKE, L.; SADAGIC, A. Taking Immersive VR Leap in Training of Landing Signal Officers. , [S.l.: s.n.], 2016.

GREWE, P. et al. Learning real-life cognitive abilities in a novel 360-virtual reality supermarket: a neuropsychological study of healthy participants and patients with epilepsy. **Journal of neuroengineering and rehabilitation**, [S.l.: s.n.], v.10, n.1, 2013, p.42.

HARPER, B. D.; NORMAN, K. L. Improving user satisfaction: the questionnaire for user interaction satisfaction version 5.5. In: ANNUAL MID-ATLANTIC HUMAN FACTORS CONFERENCE, 1. **Proceedings...** [S.l.: s.n.], 1993. p.224–228.

HART, S. G.; STAVELAND, L. E. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. **Advances in psychology**, [S.l.: s.n.], v.52, 1988, p.139–183.

HERPERS, R. et al. FIVIS–A Bicycle Simulation System. In: WORLD CONGRESS ON MEDICAL PHYSICS AND BIOMEDICAL ENGINEERING, SEPTEMBER 7-12, 2009, MUNICH, GERMANY. **Proceedings...** [S.l.: s.n.], 2009. p.2132–2135.

HIROSE, M. et al. Higher levels of immersion improve procedure memorization performance. , [S.l.: s.n.], 2009.

HODGSON, E.; BACHMANN, E.; WALLER, D. Redirected walking to explore virtual environments: assessing the potential for spatial interference. **ACM Transactions on Applied Perception (TAP)**, [S.l.: s.n.], v.8, n.4, 2011, p.22.

HUPONT, I. et al. How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift. In: QUALITY OF MULTIMEDIA EXPERIENCE (QOMEX), 2015 SEVENTH INTERNATIONAL WORKSHOP ON. **Proceedings...** [S.l.: s.n.], 2015. p.1–6.

IZARD, C. E. **The psychology of emotions**. [S.l.]: Springer Science & Business Media, 1991.

JOHNSEN, K. et al. Mixed reality virtual pets to reduce childhood obesity. **IEEE transactions on visualization and computer graphics**, [S.l.: s.n.], v.20, n.4, 2014, p.523–530.

JORGE, V. A. et al. Interacting with danger in an immersive environment: issues on cognitive load and risk perception. In: ACM SYMPOSIUM ON VIRTUAL REALITY SOFTWARE AND TECHNOLOGY, 19. **Proceedings...** [S.l.: s.n.], 2013. p.83–92.

KAMARAJ, D. C. et al. Interrater Reliability of the Power Mobility Road Test in the Virtual Reality–Based Simulator-2. **Archives of physical medicine and rehabilitation**, [S.l.: s.n.], v.97, n.7, 2016, p.1078–1084.

KAPLAN, D. M.; SMITH, T.; COONS, J. A validity study of the subjective unit of discomfort (SUD) score. **Measurement and Evaluation in Counseling and Development**, [S.l.: s.n.], 1995.

KEELE, S. Guidelines for performing systematic literature reviews in software engineering. In: **Technical report, Ver. 2.3 EBSE Technical Report. EBSE**. [S.l.: s.n.], 2007.

KENNEDY, R. S. et al. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. **The international journal of aviation psychology**, [S.l.: s.n.], v.3, n.3, 1993, p.203–220.

KILLANE, I. et al. Dual motor-cognitive virtual reality training impacts dual-task performance in freezing of gait. **IEEE journal of biomedical and health informatics**, [S.l.: s.n.], v.19, n.6, 2015, p.1855–1861.

KIM, J. et al. Feasibility of an individually tailored virtual reality program for improving upper motor functions and activities of daily living in chronic stroke survivors: a case series. **European Journal of Integrative Medicine**, [S.l.: s.n.], v.8, n.5, 2016, p.731–737.

KINATEDER, M. et al. The effect of dangerous goods transporters on hazard perception and evacuation behavior–A virtual reality experiment on tunnel emergencies. **Fire Safety Journal**, [S.l.: s.n.], v.78, 2015, p.24–30.

KULSHRESHTH, A.; LAVIOLA JR, J. J. Exploring 3D user interface technologies for improving the gaming experience. In: ANNUAL ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 33. **Proceedings...** [S.l.: s.n.], 2015. p.125–134.

LEE, S. J.; CHUN, M. H. Combination transcranial direct current stimulation and virtual reality therapy for upper extremity training in patients with subacute stroke. **Archives of physical medicine and rehabilitation**, [S.l.: s.n.], v.95, n.3, 2014, p.431–438.

LORENZO, G. et al. Design and application of an immersive virtual reality system to enhance emotional skills for children with autism spectrum disorders. **Computers & Education**, [S.l.: s.n.], v.98, 2016, p.192–205.

LUECKE, G. R. Greenspace: virtual reality interface for combine operator training. **Presence: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.21, n.3, 2012, p.245–253.

MA, M.; BECHKOUM, K. Serious games for movement therapy after stroke. In: SYSTEMS, MAN AND CYBERNETICS, 2008. SMC 2008. IEEE INTERNATIONAL CONFERENCE ON. **Proceedings...** [S.l.: s.n.], 2008. p.1872–1877.

MANIA, K. et al. The effect of visual and interaction fidelity on spatial cognition in immersive virtual environments. **IEEE Transactions on Visualization and Computer Graphics**, [S.l.: s.n.], v.12, n.3, 2006, p.396–404.

MARIN, J. G.; NAVARRO, K. F.; LAWRENCE, E. Serious games to improve the physical health of the elderly: a categorization scheme. In: INTERNATIONAL CONFERENCE ON ADVANCES IN HUMAN-ORIENTED AND PERSONALIZED MECHANISMS, TECHNOLOGIES, AND SERVICES. BARCELONA, SPAIN. **Proceedings...** [S.l.: s.n.], 2011.

MCMAHAN, R. P. et al. Evaluating display fidelity and interaction fidelity in a virtual reality game. **Visualization and Computer Graphics, IEEE Transactions on**, [S.l.: s.n.], v.18, n.4, 2012, p.626–633.

MINE, M. R.; BROOKS JR, F. P.; SEQUIN, C. H. Moving objects in space: exploiting proprioception in virtual-environment interaction. In: COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 24. **Proceedings...** [S.l.: s.n.], 1997. p.19–26.

MUHANNA, M. A. Virtual reality and the CAVE: taxonomy, interaction challenges and research directions. **Journal of King Saud University-Computer and Information Sciences**, [S.l.: s.n.], v.27, n.3, 2015, p.344–361.

NAKAMURA, J.; CSIKSZENTMIHALYI, M. The concept of flow. In: **Flow and the Foundations of Positive Psychology**. [S.l.]: Springer, 2014. p.239–263.

NAZIR, S. et al. Performance comparison of different training methods for industrial operators. In: HUMAN FACTORS: A VIEW FROM AN INTEGRATIVE PERSPECTIVE,. PROCEEDINGS HFES EUROPE CHAPTER CONFERENCE TOULOUSE, 2012. **Proceedings...** [S.l.: s.n.], 2012. p.235–242.

NAZIR, S.; KLUGE, A.; MANCA, D. Can immersive virtual environments make the difference in training industrial operators. **Proceedings of the Human Factors and Ergonomics Society Europe**, [S.l.: s.n.], 2014, p.251–265.

NAZIR, S.; MANCA, D. How a plant simulator can improve industrial safety. **Process Safety Progress**, [S.l.: s.n.], v.34, n.3, 2015, p.237–243.

NEDEL, L. et al. Using Immersive Virtual Reality to Reduce Work Accidents in Developing Countries. **IEEE computer graphics and applications**, [S.l.: s.n.], v.36, n.2, 2016, p.36–46.

O'BRIEN, H. L. et al. Developing and evaluating a reliable measure of user engagement. **Proceedings of the American Society for Information Science and Technology**, [S.l.: s.n.], v.45, n.1, 2008, p.1–10.

ORDAZ, N. et al. Serious Games and Virtual Simulator for Automotive Manufacturing Education & Training. **Procedia Computer Science**, [S.l.: s.n.], v.75, 2015, p.267–274.

PARASKEVOPOULOS, I. T. et al. Design guidelines for developing customised serious games for Parkinson's Disease rehabilitation using bespoke game sensors. **Entertainment Computing**, [S.l.: s.n.], v.5, n.4, 2014, p.413–424.

PECK, T. C.; FUCHS, H.; WHITTON, M. C. An evaluation of navigational ability comparing Redirected Free Exploration with Distractors to Walking-in-Place and joystick locomotio interfaces. In: IEEE VIRTUAL REALITY CONFERENCE, 2011. **Proceedings. . .** [S.l.: s.n.], 2011. p.55–62.

PEDRAZA-HUESO, M. et al. Rehabilitation Using Kinect-based Games and Virtual Reality. **Procedia Computer Science**, [S.l.: s.n.], v.75, 2015, p.161–168.

PERUZZI, A. et al. Feasibility and acceptance of a virtual reality system for gait training of individuals with multiple sclerosis. **European International Journal of Science and Technology**, [S.l.: s.n.], v.2, n.6, 2013, p.171–181.

PERUZZI, A. et al. Effects of a virtual reality and treadmill training on gait of subjects with multiple sclerosis: a pilot study. **Multiple sclerosis and related disorders**, [S.l.: s.n.], v.5, 2016, p.91–96.

RAGAN, E. D. et al. The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments. **Presence: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.19, n.6, 2010, p.527–543.

RAGAN, E. D. et al. Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. **IEEE transactions on visualization and computer graphics**, [S.l.: s.n.], v.21, n.7, 2015, p.794–807.

REGO, P.; MOREIRA, P. M.; REIS, L. P. Serious games for rehabilitation: a survey and a classification towards a taxonomy. In: INFORMATION SYSTEMS AND TECHNOLOGIES (CISTI), 2010 5TH IBERIAN CONFERENCE ON. **Proceedings. . .** [S.l.: s.n.], 2010. p.1–6.

RIZZO, A. et al. Virtual Reality Exposure Therapy for Combat-Related Posttraumatic Stress Disorder. , [S.l.: s.n.], n.July, 2014, p.31–37.

ROUSSOS, M. et al. Learning and building together in an immersive virtual world. **Presence: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.8, n.3, 1999, p.247–263.

ROUSSOU, M.; OLIVER, M.; SLATER, M. The virtual playground: an educational virtual reality environment for evaluating interactivity and conceptual learning. **Virtual reality**, [S.l.: s.n.], v.10, n.3-4, 2006, p.227–240.

ROUSSOU, M.; SLATER, M. A virtual playground for the study of the role of interactivity in virtual learning environments. **perspectives**, [S.l.: s.n.], v.8, 2005, p.9.

SCHÖNAUER, C. et al. Chronic pain rehabilitation with a serious game using multimodal input. In: VIRTUAL REHABILITATION (ICVR), 2011 INTERNATIONAL CONFERENCE ON. **Proceedings...** [S.l.: s.n.], 2011. p.1–8.

SCHÖNAUER, C.; PINTARIC, T.; KAUFMANN, H. Full body interaction for serious games in motor rehabilitation. In: AUGMENTED HUMAN INTERNATIONAL CONFERENCE, 2. **Proceedings...** [S.l.: s.n.], 2011. p.4.

SCHWEBEL, D. C.; GAINES, J.; SEVERSON, J. Validation of virtual reality as a tool to understand and prevent child pedestrian injury. **Accident Analysis & Prevention**, [S.l.: s.n.], v.40, n.4, 2008, p.1394–1400.

SEO, K. et al. RehabMaster: a pervasive rehabilitation platform for stroke patients and their caregivers. **Pervasive health—state-of-the-art and beyond. London: Springer**, [S.l.: s.n.], 2014.

SHERMAN, W. R.; CRAIG, A. B. Understanding Virtual Reality—Interface, Application, and Design. **Presence: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.12, n.4, 2003, p.441–442.

SLATER, M. A note on presence terminology. **Presence connect**, [S.l.: s.n.], v.3, n.3, 2003, p.1–5.

SLATER, M.; MCCARTHY, J.; MARINGELLI, F. The influence of body movement on subjective presence in virtual environments. **Human Factors: The Journal of the Human Factors and Ergonomics Society**, [S.l.: s.n.], v.40, n.3, 1998, p.469–477.

SMITH, S.; ERICSON, E. Using immersive game-based virtual reality to teach fire-safety skills to children. **Virtual reality**, [S.l.: s.n.], v.13, n.2, 2009, p.87–99.

STACH, T.; GRAHAM, T. N. Exploring haptic feedback in exergames. In: **Human-Computer Interaction–INTERACT 2011**. [S.l.]: Springer, 2011. p.18–35.

STANSFIELD, S. et al. Design and implementation of a virtual reality system and its application to training medical first responders. **Presence: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.9, n.6, 2000, p.524–556.

STEED, A. et al. An 'In the Wild'Experiment on Presence and Embodiment using Consumer Virtual Reality Equipment. **IEEE transactions on visualization and computer graphics**, [S.l.: s.n.], v.22, n.4, 2016, p.1406–1414.

STINSON, C.; BOWMAN, D. A. Feasibility of training athletes for high-pressure situations using virtual reality. **IEEE transactions on visualization and computer graphics**, [S.l.: s.n.], v.20, n.4, 2014, p.606–615.

TIJSD, T.; WALTOND, A. Measuring and defining the experience of immersion in games. **Int. J. Human-Computer Studies**, [S.l.: s.n.], v.66, 2008, p.641–661.

TROCHIM, W.; DONNELLY, J. P.; ARORA, K. **Research methods**: the essential knowledge base. [S.l.]: Nelson Education, 2015.

USOH, M. et al. Walking> walking-in-place> flying, in virtual environments. In: COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 26. **Proceedings. . .** [S.l.: s.n.], 1999. p.359–364.

USOH, M. et al. Using presence questionnaires in reality. **Presence: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.9, n.5, 2000, p.497–503.

VARLET, M. et al. Virtual reality as a tool to learn interpersonal coordination: example of team rowing. **PRESENCE: Teleoperators and Virtual Environments**, [S.l.: s.n.], v.22, n.3, 2013, p.202–215.

WATTANASOONTORN, V. et al. Serious games for health. **Entertainment Computing**, [S.l.: s.n.], v.4, n.4, 2013, p.231–247.

WILLIAMS-BELL, F. et al. Using serious games and virtual simulation for training in the fire service: a review. **Fire Technology**, [S.l.: s.n.], v.51, n.3, 2015, p.553–584.

WITMER, B. G.; SINGER, M. J. Measuring presence in virtual environments: a presence questionnaire. **Presence: Teleoperators and virtual environments**, [S.l.: s.n.], v.7, n.3, 1998, p.225–240.

YAVRUCUK, I.; KUBALI, E.; TARIMCI, O. A low cost flight simulator using virtual reality tools. **IEEE Aerospace and Electronic Systems Magazine**, [S.l.: s.n.], v.26, n.4, 2011, p.10–14.

ZELINSKI, E. M.; GILEWSKI, M. A 10-item Rasch modeled memory self-efficacy scale. **Aging & mental health**, [S.l.: s.n.], v.8, n.4, 2004, p.293–306.

ZYDA, M. From visual simulation to virtual reality to games. **Computer**, [S.l.: s.n.], v.38, n.9, 2005, p.25–32.

APPENDIX A INFORMED CONSENT FORM USED FOR EXPERIMENT I

You are been invited to participated of a research about the influence of immersion in a serious game for risk perception analysis. This document has the purpose of explain ti you the research goals, the procedures, the risks and how the experiment will be conducted. I request that you read this document carefully and take your doubts before consent your participation.

Objective: The objective of this research is to investigate the effects of display fidelity in a serious game that has the goal of trining the workers ability of detecting possible elements of risks in the work environment.

Procedures: You will play the game guided by a narrator that will give instructions to the player execute determined tasks. While executing the tasks, you has to search for possible elements of risk.

Before the experiment you has to answer a characterization form and the simulator sickness questionnaire. In the end, you will answer a questionnaire about the user experience. The data obtained in the experiment (images and data about the serious game outcomes) will be used in this study, only, completely anonymous. The completion time will be 15 minutes top.

You can leave the experiment any time, without any loss, if you want.

Risks and benefits: The experiment can cause some symptoms as dizziness and headache, because the VR system used (e.g., Oculus Rift). The benefits are the opportunity of experience technologies of virtual reality, and contributed a research about the effects of interactive techniques in the learning process through serious games.

- I accept to participate of the experiment. I was correctly informed by the researcher about the research, the procedures that it contains, and the risks and benefits that I am exposed to. I was guaranteed the secrecy of the in formations and that I can take off my consent any time.

APPENDIX B CHARACTERIZATION FORM USED FOR BOTH EXPERIMENTS

1. Name
2. E-mail account
3. Age
4. Gender (Female, Male, other)
5. Education level (No schooling, Incomplete elementary school, Complete elementary school, Incomplete high school, Complete high school, Incomplete undergraduate, Complete undergraduate, Incomplete post-graduation, Complete post-graduation)
6. Profession
7. Do you have any vision problem? (Myopia, Astigmatism, Hypermetropia, Daltonism, other)
8. How much experienced do you consider yourself in 3D video games? (1 - very inexperienced, 5 - very experienced)
9. How much experienced do you consider yourself in virtual reality? (1 - very inexperienced, 5 - very experienced)
10. Have you received any training about safety work? (Yes, No)
11. How much experienced do you consider yourself in safety work? (1 - very inexperienced, 5 - very experienced)

APPENDIX C INFORMED CONSENT FORM USED FOR EXPERIMENT II

You are been invited to participated of a research about the influence of immersion in a serious game for educate workers about safety procedures for electric installations. This document has the purpose of explain ti you the research goals, the procedures, the risks and how the experiment will be conducted. I request that you read this document carefully and take your doubts before consent your participation.

Objective: The objective of this research is to investigate the effects of locomotion and interaction fidelity in a serious game that has the goal of educate workers about safety procedures for installations of lighting rod.

Procedures: You will play the game guided by a narrator that will give instructions to the player execute determined tasks. Before the experiment you has to answer a characterization form, the self-efficacy of everyday memory questionnaire, and the simulator sickness questionnaire. In the end, you will answer a questionnaire about the user experience and a knowledge test. The data obtained in the experiment (images and data about the serious game outcomes) will be used in this study, only, completely anonymous.

The experiment consists of two stages, with two weeks of distance of each one. The first stage contains four sub-stages: training of the interaction techniques, play the serious game guided by audio instructions, and play the game without the instructions (two times). The completion time will be 45 minutes top. In the second stage you will answer a knowledge test once again.

You can leave the experiment any time, without any loss, if you want.

Risks and benefits: The experiment can cause some symptoms as dizziness and headache, because the VR system used (e.g., Oculus Rift). The benefits are the opportunity of experience technologies of virtual reality, and contributed a research about the effects of interactive techniques in the learning process through serious games.

- I accept to participate of the experiment. I was correctly informed by the researcher about the research, the procedures that it contains, and the risks and benefits that I am exposed to. I was guaranteed the secrecy of the in formations and that I can take off my consent any time.

APPENDIX D KNOWLEDGE TEST USED TO ASSESS THE ACQUIRED KNOWLEDGE IN THE EXPERIMENT I

D.1 In Portuguese

1. De um modo geral, quais são os procedimentos que devem ser adotados para deixar o ambiente de trabalho seguro para os trabalhadores e para os pedestres?
2. Uma vez que o veículo foi estacionado, qual o próximo passo?
3. Antes de deixar o local seguro, é importante que você esteja usando equipamentos adequados. Que equipamentos são estes?
4. Além dos equipamentos de segurança, que outro item o trabalhador deve levar consigo?
5. Onde e como as ferramentas de trabalho devem ser colocadas?
6. Qual o primeiro passo executado para sinalizar o local deixando-o seguro para os pedestres?
7. Como você isola o local de trabalho?

APPENDIX E GAME ENGAGEMENT QUESTIONNAIRE USED IN THE EXPERIMENT II

For each of the following, please rate your experience of the sensation while playing the game, on the following scale from 1 (did not experience) to 5 (definitely experienced).

1. I lost track of time.
2. Things seemed to happen automatically.
3. I felt different.
4. I felt scared.
5. The game felt real.
6. I felt tense.
7. Time seemed to stand still or stop.
8. I felt unaware of my surroundings.
9. Playing seemed automatic.
10. My thoughts were fast.
11. I forgot I was in a virtual environment.
12. I played without thinking about how to play.
13. Playing made me feel calm.
14. I really got involved with the game.
15. I did not want to stop playing.

APPENDIX F SIMULATOR SICKNESS QUESTIONNAIRE

Indicate how much each symptom below is affecting you right now in the following scale: None, Slight, Moderate, Severe

1. General discomfort
2. Fatigue
3. Eye strain
4. Difficulty focusing
5. Salivation increasing
6. Sweating
7. Nausea
8. Difficulty concentrating
9. Fullness of head
10. Blurred vision
11. Dizziness with eyes open
12. Dizziness with eyes closed
13. *Vertigo
14. **Stomach awareness
15. Burping

* Vertigo is experienced as loss of orientation with respect to vertical upright.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

APPENDIX G SUS PRESENCE QUESTIONNAIRE USED FOR EXPERIMENT I

1. Please rate your sense of being in the spaces below, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place. (1 - Not at all. 7 - Very much.)
   (a) In the reception room
   (b) In the parking lot
   (c) In the kitchen
   (d) In the office
2. To what extent were there times during the experience when the places below was a reality for you? (1 - At no time. 7 - Almost all the time.)
   (a) In the reception room
   (b) In the parking lot
   (c) In the kitchen
   (d) In the office
3. When you think back about your experience, do you think the places below more like images that you saw, or more like somewhere that you visited? (1 - Images that I saw. 7 - Somewhere that I visited.)
   (a) In the reception room
   (b) In the parking lot
   (c) In the kitchen
   (d) In the office
4. During the time of the experience, which was strongest on the whole, your sense of being in the places below, or of being elsewhere? (1 - Being elsewhere. 7 - Being in the virtual space.)
   (a) In the reception room
   (b) In the parking lot
   (c) In the kitchen
   (d) In the office

5. Consider your memory of being in the places below. How similar in terms of the structure of the memory is this to the structure of the memory of other places that you have been today? By "structure of the memory" consider things like the extent to which you have a visual memory of the virtual space, whether the memory is in color, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such structural elements. (1 - Not at all. 7 - Very much so.)
   (a) In the reception room
   (b) In the parking lot
   (c) In the kitchen
   (d) In the office
6. During the time of the experience, did you often think to yourself that you were actually in the places below? (1 - Not very often. 7 - Very much so.)
   (a) In the reception room
   (b) In the parking lot
   (c) In the kitchen
   (d) In the office

APPENDIX H AUDIO INSTRUCTIONS IN PORTUGUESE USED IN THE SERIOUS GAME FOR EXPERIMENT I

1. Olá! Darei instruções sobre como você deverá prosseguir na cena. Entre no prédio e vá até o centro da recepção.
2. Pare! Fique onde está e olhe ao redor. Aponte para os possíveis elementos de riscos que você vê, pelo tempo que quiser. Ao terminar, ative a próxima instrução.
3. Saia pela porta ao lado do balcão da recepção e vá até o seu carro no estacionamento. Seu carro é o que está piscando em vermelho.
4. Neste prédio há uma cozinha. Vá até lá e pegue um café.
5. Agora vá até a sua mesa. É aquela que está piscando em vermelho.
6. Fique onde está e olhe ao redor. Aponte para os possíveis elementos de risco que você vê.
7. Encontre os seus colegas de trabalho do lado de fora do escritório, próximo ao estacionamento.
8. Muito bem! A simulação está concluída!

APPENDIX I AUDIO INSTRUCTIONS IN PORTUGUESE USED IN THE SERIOUS GAME FOR EXPERIMENT II

1. Olá! Neste jogo você irá aprender alguns procedimentos de segurança, que devem ser adotados ao realizar a instalação de para raios. Sua tarefa é organizar os materiais de forma a deixar o ambiente seguro e bem sinalizado.

2. Primeiro, é importante que o caminhão esteja bem estacionado e com calços nas rodas. Coloque um calço em cada roda indicada pelas setas vermelhas. Ao terminar, ative a próxima instrução.

3. Os trabalhadores devem estar equipados de luvas de raspa, cinto de segurança e capacete. Além disso, devem sempre levar consigo a ordem de serviço. Selecione os objetos que representam estes equipamentos. Eles estão piscando em vermelho. Ao terminar, ative a próxima instrução.

4. Quando necessário, deve-se utilizar lona ou cavalete para bastões, para acomodar materiais no solo, os quais devem ficar dentro da área sinalizada. Coloque os materiais, que estão piscando em vermelho, sobre a lona indicada pela seta. Ao terminar, ative a próxima instrução.

5. Antes de iniciar a instalação, a área de trabalho deve ser sinalizada através de placas de sinalização. Posicione as placas em locais visíveis, conforme indicado pelas setas, sinalizando os pedestres. Ao terminar, ative a próxima instrução.

6. Isole a área de trabalho, incluindo o veículo, utilizando os cones e as cordas. Primeiro posicione os cones ao redor da área, entre os espaços indicados pelas setas, onde será realizada a instalação do para raios. Ao terminar, ative a próxima instrução.

7. Agora que os cones foram posicionados, amarre as cordas entre eles. Para isso, selecione uma corda e dois cones, um de cada vez. Siga fazendo o mesmo procedimento até que a área esteja devidamente isolada. Lembre-se de deixar um espaço aberto para que os trabalhadores possam transitar. Ao terminar, ative a próxima instrução.

8. Muito bem! A fase de aprendizado está concluída! Agora você passará para a fase onde irá praticar os conhecimentos adquiridos.

APPENDIX J UMA PROPOSTA DE ORIENTAÇÕES DE DESIGN PARA JOGOS SÉRIOS IMERSIVOS

Resumo da Dissertação em Português

J.1 Introdução

Jogos sérios são jogos desenvolvidos para propósitos como treinamento, educação e publicidade. O uso desses jogos têm crescido nos últimos anos por se tratar de uma alternativa de baixo custo, segura e ética de realizar o treinamento de tarefas perigosas, como treinamento de bombeiros (WILLIAMS-BELL et al., 2015), e simulações de cirurgia ((WATTANASOONTORN et al., 2013) (REGO; MOREIRA; REIS, 2010) (MARIN; NAVARRO; LAWRENCE, 2011)). Embora, a maioria destes jogos são jogados em computadores ou dispositivos móveis, o uso de realidade virtual tem apresentado efeitos positivos no desempenho do usuário (KULSHRESHTH; LAVIOLA JR, 2015).

Segundo Chalmers et al. (CHALMERS; DEBATTISTA, 2009) é necessário um alto nível de realismo para garantir que o treinamento ou aprendizado no ambiente virtual seja equivalente ao mundo real. Portanto, deveria ser possível simular os cinco sentidos humanos simultaneamente. O uso de altos níveis de imersão para simular situações do mundo real tem apresentado efeitos positivos na efetividade e desempenho do usuário, principalmente em jogos sérios para aprendizado (p. ex. (BACKLUND et al., 2007, 2008; CHA et al., 2012), treinamento (p. ex. (NAZIR et al., 2012; NAZIR; KLUGE; MANCA, 2014; NAZIR; MANCA, 2015), e tratamento (p. ex. (MA; BECHKOUM, 2008; PEDRAZA-HUESO et al., 2015; PARASKEVOPOULOS et al., 2014). Chitarro et al. (CHITTARO; BUTTUSSI, 2015) apresenta um jogo sério que educa passageiros sobre segurança na aviação, permitindo que os usuários enfrentem uma emergência aérea com o objetivo de sobreviver. Bcklund et al. (BACKLUND et al., 2007) apresenta um jogo sério para treinamento das habilidades de bombeiros, que usa um sistema baseado em CAVE e interação através do rastreamento de movimentos. Apesar de apresentar resultados satisfatórios, o uso de sistemas de Realidade Virtual (RV) podem ainda ser caros e/ou causar sintomas de *cybersickness*, como náusea, vômito, dores de cabeça e desorientação. Considerando estas questões, se torna relevante investigar os efeitos de níveis moderados de imersão na efetividade de jogos sérios.

A pesquisa apresentada nesta dissertação se preocupa em investigar tais tópicos, no que se refere a aspectos de imersão física, buscando identificar o sistema de RV apropriado para criar jogos sérios imersivos, confortáveis e efetivos. Para isso, o estudo é guiado pelas seguintes questões de pesquisa:

1. Quais são os efeitos da fidelidade de *display* na percepção do usuário?

   (a) Níveis de fidelidade médios podem resultar em efetividade equivalente à alta fidelidade de *display*?

   (b) Como os níveis de fidelidade de *display* afetam os sintomas de *cybersickness* e como isso afeta o desempenho do usuário?

2. Quais são os efeitos da fidelidade de interação e locomoção na retenção de conhecimento?

   (a) Como a naturalidade da técnica afeta os sintomas de *cybersickness* e quais os efeitos disso no aprendizado perceptivo?

   (b) Como a carga de trabalho afeta o desempenho do usuário e o aprendizado perceptivo?

Para verificar tais questões, a pesquisa inclui dois estudos empíricos e uma revisão da literatura que investiga como os níveis de imersão afetam a efetividade, desempenho e satisfação dos usuários em jogos sérios. O estudo considera aspectos dos sistemas de RV que poderiam afetar a efetividade do aprendizado perceptivo, e o grau de controle interativo necessário para explorar o ambiente virtual e adquirir conhecimento. A partir destes estudos são identificadas orientações de *design* para imersão física em jogos sérios para aprendizado perceptivo.

J.2 Experimento I - Fidelidade de *Display* e Percepção do Usuário

Neste experimentos foram investigados os efeitos da fidelidade de *display* na efetividade e desempenho do usuário em um jogo sério para análise de percepção de riscos. A hipótese geral é que um jogo sério que utiliza percepção para prover aprendizado pode ser efetivo, confortável e oferecer uma experiência de RV boa também com moderada fidelidade de *display*. Em um nível mais específico, este experimento testa as seguintes hipóteses:

- H1. Níveis de imersão médios e altos não apresentarão diferença significativa na percepção de riscos.

- H2. Níveis de imersão mais altos irão apresentar maior severidade nos sintomas de *sickness*.

- H3. A severidade dos sintomas de *sickness* não irá afetar significativamente a percepção de riscos.

- H4. A maioria dos usuários estará mais focado em executar as instruções transmitidas pelo narrador do que em selecionar os riscos no ambiente.

A variável independente é a fidelidade de *display*, conforme apresentado na Figura J.1. Os *displays* utilizados foram um monitor regular, um *display-wall* formado por doze monitores regulares, e um dispositivo Oculus Rift DK2, para baixa, média e alta fidelidade, respectivamente. As variáveis dependentes foram *cybersickness*, medida subjetivamente com o SSQ (KENNEDY et al., 1993), o senso de presença, medido através do questionário de presença SUS (Slater-Usoh-Steed) (USOH et al., 2000), e desempenho do usuário, medido através dos resultados do jogo sério, como tempo de execução, número de riscos encontrados e caminho percorrido.

Figure J.1 Níveis de fidelidade de *display* variados no experimento.

(a) Fidelidade Baixa (b) Fidelidade Moderada (c) Fidelidade Alta

Fonte: Próprio autor.

O jogo sério utilizado neste experimento consiste de um ambiente virtual com uma recepção, uma cozinha, um estacionamento e um escritório. Há cinquenta e três elementos de risco espalhados pelo ambiente que devem ser selecionados pelo usuário enquanto executa tarefas repassadas por um narrador (veja Anexo H)).

J.2.1 Resultados e Discussão

O experimento envolveu 61 participantes (15 mulheres), com idade entre 19 e 63 anos (M=28.67, DP=10.12). Todos receberam um formulário de consentimento. A maioria dos participantes eram estudantes universitários e inexperientes em relação ao propósito do experimento.

Para realizar a análise de dados, o nível de significância foi estabelecido em $p = 0.05$. Os dados paramétricos foram submetidos a Análise de Variância One-Way (ANOVA One-Way) para três ou mais grupos, e ao teste T de Welch para duas amostras independentes. Os dados não paramétricos foram submetidos a um teste de Fligner-Killeen para verificar a homogeneidade dos dados. Para três ou mais grupos homogêneos, foi aplicado o teste de Friedman para amostras dependentes e um teste Kruskall-Wallis para amostras independentes. Nos casos em que os dados são heterogêneos, foi realizado um teste F de Welch para três ou mais amostras independentes. Para dois grupos independentes foi realizado um Teste U de Mann-Whitney, e um teste de sinais de Wilcoxon para amostras pareadas.

Os resultatos do SSQ mostraram que a alta fidelidade de *display* foi mais prejudicial aos participantes do que as demais (H2) ($p=0.008$ para náusea, $p < 0.001$ para desorientação e $p=0.004$ para pontuação total), exceto para escala de *oculomotor*.

O caminho percorrido pelos participantes no ambiente virtual não variou significantemente entre as condições, embora com nível de imersão mais alto, o caminho percorrido foi levemente mais preciso. Considerando, então, os riscos diretamente visíveis no caminho percorrido pelos usuários, observou-se que nas instruções em que foi explicitamente pedido para que procurassem por riscos, a quantidade selecionada foi maior do que nas demais instruções, indicando que os participantes ficaram mais focados em seguir as instruções do que em realizar a busca por riscos (H4). Em geral, no entanto, não foi observado significância estatística no número de riscos encontrados entre as condições experimentais (H1).

Considerando as categorias dos riscos, simples e compostos, observou-se que os participantes selecionaram maior número de riscos simples do que compostos em todas as condições experimentais. No entanto, o número de riscos complexos foi maior na condição com alto nível de imersão ($p < 0.05$). E a quantidade de riscos simples foi maior com imersão moderada ($p < 0.05$). Os participantes se consideram inexperientes em segurança do trabalho, o que ajuda a explicar o baixo número de riscos selecionados. Além disso, os riscos simples podem ser considerados mais fáceis de serem detectados, pois o perigo de acidentes pode ser verificado diretamente.

Os participantes se sentiram mais presentes no ambiente virtual na condição em que a fidelidade de *display* foi mais alta ($p < 0.01$ para recepção, e $p < 0.05$ para o estacionamento e o escritório), não foi observado significância estatística na pontuação para a cozinha. Considerando este resultado, podemos supor que os usuários encontraram mais riscos complexos com alto nível de imersão por sentirem como se estivessem no ambiente e, assim, puderam analisar com maior atenção os riscos que os objetos poderiam oferecer a longo prazo.

O tempo de execução foi significativamente maior para o nível de imersão moderado ($M=744.15$, $DP=277.55$, $p < 0.001$). Este resultado, no entanto, não influenciou a severidade dos sintomas de *sickness*, nem o número de riscos encontrados. Da mesma forma, a severidade dos sintomas de *sickness* não influenciaram o número de riscos selecionados (H3).

J.3 Experimento II - Fidelidade de Interação e Retenção de Conhecimento

Neste experimento são analisados os efeitos da fidelidade de interação e locomoção na retenção de conhecimento em um jogo sério que ensina os usuários procedimentos básicos de segurança para instalação de para raios. A hipótese geral é que o aprendizado pode ser afetado nos casos em que as técnicas de interação e locomoção possuem muita carga de trabalho ou causem altos níveis de *sickness*. Em um nível mais específico, este experimento testa as seguintes hipóteses:

- H1. A exposição repetida ao jogo sério irá melhorar o desempenho do usuário.
- H2. O desempenho do usuário será afetado pelo *workload*.
- H3. O aprendizado não será significantemente afetado pelos níveis de *sickness*.
- H4. Alta fidelidade de locomoção e interação irá melhorar a performance do usuário.
- H5. A retenção de conhecimento será afetada pelo *workload*.

A variável independente é a fidelidade de locomoção e interação, que é variada em quatro níveis. Para a baixa fidelidade de interação (BI) é usado uma mira que acompanha o movimento da cabeça do usuário. Alta fidelidade de interação (AI) é feita usando mãos virtuais que acompanham o movimento das mãos do usuário. Baixa fidelidade de locomoção (BI) é alcançada através da navegação com um controle de jogo. E alta fidelidade de locomoção (AL) é feita através da técnica *Walking-in-place*. O experimento foi executado entre grupos, havendo quatro condições experimentais: BIBL, BIAL, AIBL e AIAL.

As variáveis dependentes são *sickness*, medida através do SSQ (KENNEDY et al., 1993), carga de trabalho, medida através do questionário NASA TLX (HART; STAVELAND, 1988), envolvimento do usuário com o jogo, medido através do questionário apresentado por McMahan et al. (MCMAHAN et al., 2012), e desempenho do usuário, medido através de um teste de conhecimento (veja Anexo D), tempo de execução e correção das tarefas executadas no jogo.

O jogo sério utilizado neste experimento consiste de um ambiente virtual onde ocorre a instalação de um para raios, o objetivo dos jogadores é deixar o ambiente seguro para os trabalhadores e pedestres. Há um narrador que passa os procedimentos de segurança via mensagens de áudio (veja Anexo I). O jogo contém três fases: uma fase de aprendizado, guiada pelas instruções do narrador, e duas fases de avaliação, sem instruções para avaliar o conhecimento adquirido.

J.3.1 Resultados e Discussão

O experimento envolveu 46 participantes (4 mulheres) com idade entre 18 e 33 anos (M=23.21, DP=3.39), sendo a maioria estudantes de computação. Todos receberam um formulário de consentimento, e eram inexperientes no que se refere ao propósito do experimento.

Os dados foram submetidos aos mesmos testes usados no Experimento I (veja Seção J.2.1), conforme a origem dos dados e quantidade de grupos.

Os resultados do SSQ mostraram que, embora a exposição ao jogo sério em todas as condições experimentais aumentam a severidade dos sintomas de *sickness*, não há diferença significativa entre elas.

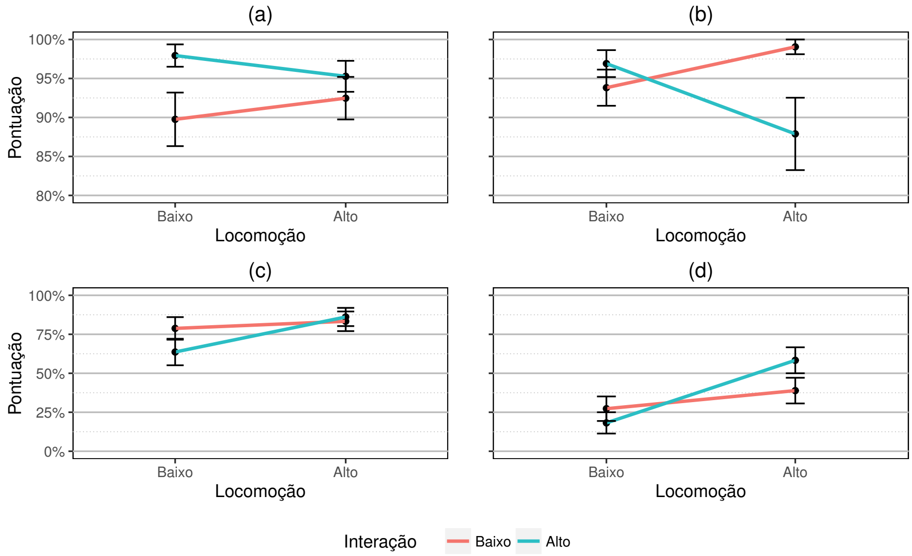
O teste de conhecimento foi aplicado duas vezes, antes e depois do jogo. Os resultados mostram que, antes do experimento os participantes acertaram aproximadamente 25% das respostas, enquanto que ao final o acerto foi de, em média, 77%. Não foi observada diferenças significativas entre as condições experimentais ($p = 0.28$).

O desempenho do usuário durante o jogo foi avaliado separadamente para cada tarefa. Na tarefa que pedia aos usuários para colocar as ferramentas sobre a lona dentro da área sinalizada (Figura J.2(a)), nas condições com alta fidelidade de interação e baixa fidelidade de locomoção, isso foi feito mais corretamente ($p = 0.038$), do que com alta fidelidade de locomoção.

Na tarefa de colocar as placas de sinalização em locais visíveis, os usuários completaram a tarefa mais corretamente com baixa fidelidade de interação nas condições com alta fidelidade de locomoção ($p = 0.023$). E nas condições com baixa fidelidade de locomoção, a tarefa foi executada mais corretamente usando alta fidelidade de interação ($p = 0.036$).

Na tarefa de isolar o local de trabalho (Figura J.2(c)), foi observado que, nas condições com alta fidelidade de interação, a tarefa foi executada mais corretamente usando alta fidelidade de locomoção ($p = 0.032$). Nesta mesma condição, os usuários também deixaram mais vezes a abertura corretamente entre os cones ($p < 0.001$) (ver Figura J.2(d)).

Figure J.2 Resultado das tarefas entre as condições experimentais. (a) Colocar as ferramentas na lona. (b) Colocar as placas de sinalização em locais visíveis. (c)Isolar o local de trabalho. (d) Deixar abertura para os trabalhadores transitarem.



Fonte: Próprio autor.

O *workload* geral foi significativamente maior para a alta fidelidade de locomoção do que para baixa locomoção, nas condições em que o usuário interagiu com baixa fidelidade ($p = 0.0032$). Analisando as escalas do *workload* separadamente, pode-se observar que nas condições com baixa fidelidade de locomoção, o uso de alta fidelidade de interação foi mais frustrante ($p = 0.048$). Nesta mesma condição, os usuários precisaram realizar maior número de tentativas para posicionar os objetos nos locais desejados ($p < 0.001$).

As condições com alta fidelidade de interação e locomoção apresentaram maior tempo de execução, no entanto isso não afetou a severidade dos sintomas de *sickness* ou a corretude das tarefas.

Posicionar objetos em locais determinados, como as placas de sinalização, pode ser considerada uma tarefa que demanda menos atenção. Enquanto, isolar o local de trabalho, requer que o jogador analise o tamanho mínimo da área a ser isolada e onde posicionar os cones de forma estratégica. Além disso, eles devem lembrar-se de deixar um espaço aberto para que os trabalhadores possam transitar. Dessa forma, os resultados sugerem que o *workload* tem efeitos negativos nas tarefas que demandam mais atenção, enquanto que não afetam as tarefas menos exigentes.

Finalmente, observa-se que as condições experimentais não tiveram efeitos sobre a obtenção de conhecimento, mas foram observados diversos impactos na aplicação do conhecimento, ou seja, no desempenho do usuário ao realizar as tarefas aprendidas.

J.4 Orientações de *design*

Considerando os resultados dos experimentos realizados nesta pesquisa, foram identificadas algumas orientações para imersão física no desenvolvimento de jogos sérios para aprendizado perceptivo observando aspectos de conforto, efetividade e usabilidade.

A fidelidade de *display* impacta a detecção de elementos de risco nos casos em que essa é a tarefa principal, de forma que altos níveis de imersão apresentam melhores resultados. No caso em que a procura por elementos de risco é feita enquanto o jogador possui outras tarefas para realizar, a fidelidade de *display* não apresenta diferenças significativas no desempenho do usuário, isso ocorre por que os participantes geralmente estão mais focados em realizar a tarefa principal que lhes foi designada, deixando a procura por riscos em segundo plano.

Embora altos níveis de imersão aumentam a severidade dos sintomas de *sickness*, os usuários se sentiram mais presentes no ambiente quando usando alta fidelidade de *display* e selecionaram maior número de riscos complexos do que nas demais condições, além de apresentar tempo de execução menor. Usar fidelidade moderada de *display* não afeta os resultados do jogo e reduz os sintomas de *sickness*, no entanto, requer mais tempo de jogo.

A fidelidade de interação e locomoção não possui efeitos na obtenção do conhecimento. Embora, o *workload* apresentado afete negativamente o desempenho do usuário nas tarefas que exigem maior concentração. De forma que usar locomoção e interação mais natural resulta em melhor desempenho em tais tarefas, enquanto que a baixa fidelidade de interação apresenta maior workload e melhor desempenho nas tarefas que não exgem tanta atenção.

Todos as condições experimentais aumentam os sintomas de *sickness*, mas não há diferença significativa entre elas. O tempo de execução do jogo é maior para as condições que usam interação ou locomoção mais natural, no entanto, isso não afeta os sintomas de *sickness* ou o desempenho do usuário durante o jogo.

Os resultados sugerem que para jogos sérios com objetivo de aprendizado não são impactados pela forma de interação e locomoção. No entanto, jogos sérios para treinamento podem ser afetados no desempenho do usuário ao aplicar o conhecimento adquirido.

J.5 Conclusões

A pesquisa apresentada nesta dissertação investiga os efeitos da fidelidade de *display*, interação e locomoção na efetividade e usabilidade de jogos sérios para aprendizado perceptivo. Em geral, a fidelidade de *display* não possui efeitos sobre a percepção do usuário, mas altos níveis de imersão podem ser melhores para tarefas de detecção de objetos mais complexos, e níveis moderados de imersão apresentaram melhor desempenho para procura de riscos simples. Embora altos níveis de imersão apresentem mais sintomas de *sickness*, os usuários se sentiram mais presentes no ambiente virtual quando usando altos níveis de imersão e realizaram a tarefa mais rapidamente.

A forma de interação e locomoção não possui efeitos sobre a obtenção de conhecimento. No entanto, os resultados mostram que o *workload* afeta o desempenho do usuário em tarefas que exigem maior concentração, de forma que a aplicação do conhecimento adquirido pode ser afetado pela forma de interação e locomoção. Não há diferença significativa nos sintomas de *sickness* entre as formas de interação e locomoção, embora todas as condições experimentais aumentarm o mal-estar dos usuários. Naturalidade pode ser relevante em situações que o usuário precisa estar mais focado na tarefa do que em aprender a interagir com o jogo.

Esta pesquisa se limita ao passo que os jogos sérios usados nos experimentos foram desenvolvidos com propósitos específicos, voltados a segurança do trabalho envolvendo energia elétrica, e foram realizados apenas dois estudos empíricos para investigar as questões de pesquisa. Outro fator é o perfil restrito da população participante dos experimentos, constituída em sua maioria de estudantes da computação.