Effects of fidelity in virtual reality simulators on learning technical surgical skills

Aylen Ricca

Thèse de doctorat

NNT: 2020UPASG031

université PARIS-SACLAY

Effets de la fidélité dans les simulateurs de réalité virtuelle sur l'apprentissage des compétences techniques en chirurgie

Effects of fidelity in virtual reality simulators on learning technical surgical skills

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 580, Sciences et Technologies de l'Information et de la Communication (STIC)

Spécialité de doctorat : Informatique

Unité de recherche : Université Paris-Saclay, Univ Evry, IBISC, 91020, Evry-Courcouronnes, France

Référent : Université d'Evry-Val-d'Essonne

Thèse présentée et soutenue à Evry, le 9 décembre 2020, par

Aylen Ricca

Composition du jury:

Dominique Bechmann Professeur, Université de Strasbourg	Présidente
Alexis Paljic Maître de Conférence HDR, MINES ParisTech	Rapporteur et Examinateur
Indira Thouvenin Maître de Conférence HDR, Université de Technologie de Compiègne	Rapporteur et Examinatrice
Caroline Cao Professeur, Wright State University	Examinatrice
Nicolas Férey Maître de Conférence, Université Paris-Saclay	Examinateur
Guillaume Moreau Professeur, IMT Atlantique	Examinateur
Samir Otmane Professeur, Université d'Evry, Université Paris-Saclay	Directeur
Amine Chellali Maître de Conférence, Université d'Evry, Université Paris-Saclay	Co-encadrant

"It is only through failure –and learning the cause of failure– that the true pathway to success lies"

–Richard Satava

Amanecer, Cerro Arequita, Lavalleja, Uruguay

To my beloved parents . . .

Acknowledgements

***

At the end of these years devoted to this thesis work, I would like to thank all the people who contributed to its progress in the best conditions and that supported me in this adventure.

First of all, I would like to thank Indira Thouvenin and Alexis Paljic for accepting to review my thesis. I also thanks Caroline Cao, Dominique Bechmann, Guillaume Moreau and Nicolas Férey for agreeing to be part of my jury. I feel honored.

I would like to hugely thank my thesis supervisor Amine Chellali, without whom this thesis would have never seen light. I feel very fortunate to have been supervised by such an exceptional researcher, whose human values I share, who I have been able to trust, and who has also been able to trust me. He gave me the space and the autonomy I needed to succeed in my research, and the advice and encouragement that helped stimulate my work all these years. I would also like to thank him for his time, patience and dedication, the insightful discussions about each experiment carried out during this thesis, and the time spent reading my manuscript.

I also thank the University of Evry Val d'Essonne and IBISC laboratory for their hospitality. Many thanks to all the IRA2 members for welcoming me to the research team, in particular to Guillaume Bouyer and Jean-Yves Didier. A special thanks to Ludovic David, my PhD colleague and friend. Thanks for his enthusiasm and kindness that made my time in the lab more enjoyable and challenging, and for motivating me in the most stressful moments. Thanks also to former PhD students and friends, Faouzi, Alia, Assia, with whom I have shared valuable time.

I would also like to thank all the people who contributed to the smooth running of all the experiments. My most sincere recognition to the technical team, Christophe, Olivier, Serge and Khaled, for their helpful work and everyday sympathy. Thanks to Ludovic David for his time and patience testing and debugging the prototypes, Fabien Bonardi and David Roussel for their help and advice for the image processing algorithms, Frédéric Davesne for his time and help with the evr@ platform, and Lorenzo for his time reviewing the output images. My gratitude to all the participants that volunteer to participate in the

experiments, a special thought for those that engaged in the two-week training experiment, the results of this work would not have been possible without their participation.

My warm gratitude to the teaching team colleagues at the University for the comfortable reception, and to the coffee-break people with whom I shared remarkable discussions. An immense appreciation to the skypéro team, Claire, Elisabeth, Fabien, Olivier, Christine, Frédéric, David and Jean-Yves, that added to this last confined year a warm space at the end of each week to share a convivial moment. A special thank you to Elisabeth and Olivier for their constant support and enjoyable time spent with me.

Finally, I would like to express my eternal appreciation and gratitude towards my parents who have always encourage me to pursue my goals. And to my brother and sisters, Lorenzo, Gisselle and Joseline, who have always been there for me no matter where I am. Thank you also to everyone outside the lab who supported me during this journey. A warm appreciation to Carolina, Rodrigo, Miguel, Yamil, Belén and all my friends and colleagues from Uruguay for their constant support and motivation through these years abroad and my life in general.

Once again, a huge thank you to everyone.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre: Effets de la fidélité dans les simulateurs de réalité virtuelle sur l'apprentissage des compétences techniques en chirurgie

Mots clés: Interaction Homme Machine, Réalité Virtuelle, Conception Centrée-Utilisateurs, Simulation Chirurgicale, Fidélité des Simulateurs, Formation des Compétences Motrices

Résumé: La réforme des études médicales en France réserve une place importante à la simulation dans la formation des professionnels de santé. Dans ce contexte, les simulateurs en réalité virtuelle peuvent jouer un rôle important. Cependant, la fidélité de ces systèmes reste une question ouverte avec un manque de lignes directrices pour déterminer les niveaux appropriés de leur fidélité pour une formation efficace. Dans cette thèse, nous nous intéressons à la fidélité des simulateurs immersifs pour l'apprentissage des compétences techniques en chirurgie. Plus particulièrement, nous cherchons à comprendre comment la fidélité des interactions et artefacts pourrait affecter l'efficacité de certaines tâches effectuées dans le simulateur. Sur le plan théorique, nous proposons une articulation des notions de fidélité des simulateurs à travers un modèle basé sur les différents composants de la fidélité identifiés dans notre revue de littérature, et des différentes interfaces et interactions utilisées dans les simulateurs virtuels pour la formation des compétences techniques. Deux études expérimentales ont été réalisées pour explorer les facteurs de la fidélité du simulateur associés à une tâche de navigation et à une tâche de manipulation d'un outil et leurs impacts sur l'apprentissage de deux gestes techniques en réalité virtuelle. Les résultats montrent que des techniques d'interaction et artefacts du simulateur avec une fidélité modérée peuvent supporter la réalisation de tâches secondaires pour une formation efficace. Les concepteurs des simulateurs virtuels peuvent s'appuyer sur les recommandations issues de nos travaux afin d'éviter d'inclure certains dispositifs et composants de la fidélité qui peuvent être encombrants et coûteux sans avoir un réel impact sur l'efficacité de ces simulateurs.

Title: Effects of fidelity in virtual reality simulators on learning technical surgical skills

Keywords: Human Computer Interaction, Virtual Reality, User Centered-Design, Surgical Simulation, Simulator Fidelity, Motor Skills Training

Abstract: The reform of medical studies in France attributes an important role to simulation in healthcare professionals' training. In this context, virtual reality simulators can be very useful. However, the fidelity of these systems remains an open question with a lack of guidelines for determining their appropriate levels of fidelity to support effective training. In this thesis, we are interested in the fidelity of immersive simulators for the training of technical skills in surgery. More particularly, we investigate how the fidelity of interactions and artifacts to perform certain tasks in the simulator could affect the efficiency of these systems. From a theoretical perspective, we propose an articulation of the concepts of simulator fidelity through a model based on the various fidelity components identified in our literature review, and the various interfaces and interactions used in virtual simulators for the training of technical skills. Two experimental studies were conducted to explore the factors of simulator fidelity associated with a navigation task and a tool handling task and their impacts on the learning of two technical skills in virtual reality. The results show that the simulator's interaction techniques and artifacts with a moderate fidelity can support the completion of secondary tasks for effective training. The designers of virtual simulators can rely on the recommendations resulting from our work in order to avoid including certain devices and fidelity components, which can be cumbersome and expensive without a real impact on the efficiency of these simulators.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

Table of contents

***

List of figures

***

List of tables

***

Introduction

***

Context

This thesis is in the field of Human-Computer Interaction (HCI) and Virtual Reality (VR). It aims to improve virtual surgical simulators' design from the simulator fidelity perspective by proposing recommendations based on theoretical considerations and supported by experimental studies.

This work was carried out at the "Informatique, BioInformatique & Systèmes Complexes" (IBISC) Laboratory, within the "Interaction, Réalité virtuelle & Augmentée, Robotique Ambiante" (IRA2) team.

The thesis was funded by the French "Ministère de l'Enseignement Supérieur et de la Recherche" (doctoral scholarship). The equipment used in this project was partially funded by Genopole and the "Paris Ile-de-France" Region in the context of the VR Skills Lab research project. This project aims to improve the design of virtual reality surgical simulators to guarantee effective training. More particularly, **my thesis work consisted of investigating how the concept of fidelity can influence the design of such simulators**. Therefore, starting from concrete situations, the idea was to determine which interactions and artifacts need improvement. Based on that, design decisions were taken, and experimental studies were used to determine the impact of these interactions and artifacts on performance and learning outcomes. Based on the results, design recommendations were built. This work will ultimately shed light on the fundamental needs that must be met to support the use of virtual reality technology in surgical simulators.

Motivation

A global estimation on surgery procedures shows that about 234 million major surgical operations are performed annually worldwide, from which almost 7 millions result in

complications and 1 million in deaths (Weiser et al., 2008). In Europe, the mortality rate of patients who undergo surgery was estimated to 4% for non-cardiac surgery (Pearse et al., 2012). One of the reasons for patient mortality is attributed to adverse events, which are unexpected consequences of the provided treatment rather than the disease itself (Kohn et al., 2000). Between 39.6% and 54.2% of adverse events happen in the operating room, which may derive from surgical errors, such as tissue damage or wrong instrument handling by the surgeon (Fecso et al., 2017). This information raises the importance of training technical skills in surgery (Kneebone Roger, 2003).

A technical skill refers to "any psychomotor action or related mental faculty acquired through practice and learning pertaining to a particular craft or profession" (Agha et al., 2015). The achievement of automaticity characterizes expertise in technical skills, meaning that, as expertise is gained, the type of processing performed by the operator moves from controlled processing (slower and demanding in terms of cognitive effort) to automatic processing (fast and performed with little conscious attention) (Dargar et al., 2015).

Traditional methods for training surgical technical skills followed Halsted's apprenticeship model (Halsted, 1904): *"see one, do one, teach one"*, where trainees learn by observation of experts and training on real-patients and animals (*in-vivo*), or cadavers (*ex-vivo*). However, this approach raises ethical concerns (Balcombe, 2004; Coles et al., 2011) and patient-safety issues (Akhtar et al., 2014; Coles et al., 2011). To limit these issues, a new model based on practicing on simulators (Rodriguez-Paz et al., 2009; Vozenilek et al., 2004): *"see one, **practice safely**, do one, teach one"* emerged over the past years as a credible alternative approach. In this context, the US has been the first country to develop medical simulation centers (McGaghie et al., 2010) to train healthcare professionals before being exposed to patients. Other English-speaking countries widely followed them. In 2012, the French HAS ("Haute Autorité à la Santé") published a report whose conclusion was ***"never the first time on a patient"*** with substantial consequences in terms of training methods. As a result, in 2017, a new reform of the medical curricula emerged (JORF, 2017), which imposes an evolution of traditional education models and the strengthening of the role of simulation for training healthcare professionals. Therefore, **developing new and innovative simulation-based training techniques and programs become an actual need and an emergency**.

Simulation can be defined as "the technique of imitating the behavior of some situation or process by means of a suitably analogous situation or apparatus, especially for the purpose of study or personnel training" (Bradley, 2006). Physical part-task trainers, designed to replicate only a part of the environment (Forsslund et al., 2011), and integrated

simulators that combine a manikin with sophisticated computer controls (Maran and Glavin, 2003) are commonly used to train technical, procedural or psychomotor skills. Some inconveniences related to these types of simulators reside in the costs associated with consumables, which increase the annual training cost (Orzech et al., 2012), and the need for a teacher/evaluator to be present to ensure the correct execution of the task and to provide the necessary feedback to the student (Arikatla et al., 2013).

More recently, virtual reality technologies started gaining momentum in simulation-based training by allowing clinicians to practice their skills before performing real procedures. These systems provide a cost-effective (Balcombe, 2004; Chellali et al., 2016), controlled environment (Buckley et al., 2012), with objective measurement tools (Satava, 2001). They allow different scenarios configuration (Ullrich and Kuhlen, 2012), where students can train and repeat the exercises until reaching the level of proficiency required to perform the procedure on a real patient (Buckley et al., 2012).

To summarize, technical skills training on simulators is important in order to reduce surgical errors. The use of virtual reality simulators stands as a suitable and advantageous tool to train practitioners. However, to allow the acquisition of technical skills on these systems, effective interfaces should be designed, which remains a challenge.

Research problematic

The knowledge learned during training must be transferred to the real world setup, and this is, indeed, the primary key to determine the effectiveness of a virtual reality simulator (Liu et al., 2008; McElhinney et al., 2012). It has been shown that the knowledge transfer is enhanced when the training environment and the real world are closely matched (Hamblin, 2005). This introduces the concept of the fidelity of a simulator, which can be defined as "the similarity between the knowledge taught in a simulator to the one used in the real-world environment" (Stoffregen et al., 2003). In this line, it would be reasonable to think that if both, virtual and real systems, cannot be distinguished, then the transfer of knowledge is guaranteed. However, true fidelity is far into the future due to technology not yet being ready to accomplish full realism (Stoffregen et al., 2003). Moreover, there is a high cost associated with existing high-fidelity simulators. On the other hand, some low-fidelity systems have been shown to be sufficient for efficient training (Chellali et al., 2016; Kim et al., 2003). Nevertheless, there is still a lack of guidelines on how to achieve the appropriate levels of fidelity for surgical simulators.

The research problem of this Ph.D. thesis is **how to design and implement effective user interfaces and interactions for virtual reality simulators to train technical skills**

in the surgical domain. In order to tackle this issue, the objective of this work is to study virtual reality surgical simulators from both the simulator fidelity and interaction design points of view in order to draw design guidelines for such systems.

Approach

By following an iterative design approach (Maguire, 2001), we first studied existing surgical simulators to analyze the tasks to be trained as well as the interactions and artifacts necessary to accomplish effective training and ensure the transfer of skills to the real world.

In addition, we identified different components of simulator fidelity to elaborate a generic model that allowed us to classify the components of surgical simulators based on this fidelity concept.

Particular use-cases about two simulator components were then identified to design the appropriate interaction techniques and interfaces to perform them in virtual reality, evaluate them through user studies, and extract design guidelines that can then be generalized to other surgical simulators.

Manuscript organization

This thesis manuscript is divided into two main parts:

- The first part presents the study of existing surgical simulators and prototypes with a focus on the user interactions developed for them, as well as bibliographic research on the simulator fidelity concept.
- The second part presents the two main case-studies that we have conducted, the user's head movement for navigating in a needle insertion VR simulator and the user's hand movement visualization in an immersive tool-based motor skills VR simulator.

In *chapter 1*, we introduce technical skills training in surgery. We present a state of the art of existing VR surgical simulators and the associated design challenges. We also present how interactions are provided for these systems and where the focus is made. Indeed, most surgical simulators focus only on the main task to train, whereas secondary interactions might have an impact on the users' performance of the primary task and also on skills transfer.

In *chapter 2*, we present the concept of simulator fidelity and discuss the related existing frameworks. Indeed, while being studied for a long time, simulator fidelity has recently gained importance in the design of VR simulators. Moreover, depending on the application domain, different aspects of fidelity may have more or less relevance and/or impact on the design of effective virtual simulators. Although many researchers are making contributions to study the effects of some simulator fidelity components, it is not easy to generalize the results since the contexts, domains and tasks explored are numerous and sometimes completely different. Based on that, we present our contribution to extend and adapt an existing fidelity framework.

In *chapter 3*, we present our first case-study, a needle insertion prototype in VR. The first aspect we focus on in this work is how the interaction fidelity of secondary tasks (navigation) might impact the user's performance on the primary (needle insertion) task. For this purpose, we present two navigation techniques whose design is based on the analysis of actual needle insertion tasks. We then present a classification of their interaction fidelity levels through the state of the art framework (FIFA). Finally, we present two user studies carried out to compare them. The results of these studies indicate that a high-fidelity navigation technique is not required to perform the needle insertion task. A moderate navigation technique, which design rationale is based on observing real procedures, permits users to reach similar accuracy performance as the higher fidelity technique but faster, with fewer errors and higher safety. Finally, the system obtained a good usability score, as evaluated subjectively by users, supporting the design choices made for this system.

In *chapter 4*, we present the second case-study, a tool-based motor skill simulator in immersive VR. In the second part of the thesis, we explore how another fidelity aspect of the simulation, i.e., the user's hand control and representation, might influence the users' training of the primary motor task and the transfer of skills to the real-world setup. We present the design and implementation of two prototypes, an immersive VR simulator and a physical prototype, of a tool-based pick, transfer and place task. We also detail the two user studies, i.e., a validation and a longitudinal study, to determine the impact of user's hand movement visualization (kinematic fidelity) on performance, training outcomes, and skill transfer. The first study results indicate that the users reached an equivalent task performance regardless of the hand representation condition. This suggests that a high fidelity of the user's hand representation does not impact the user performance in the tool-based pick, transfer and place task. In addition, the longitudinal study shows that users improve their performance after a two-week training period in the VR simulator, regardless of the training condition (visualization/absence of the virtual hand), when compared to a control group. On the one hand, this confirms that training in a VR simulator allows

users to improve their performance as compared to no training. On the other hand, the results indicate that hand representation fidelity does not impact the learning curve of tool-based technical skills in the VR system. Finally, although the system was rated with a good usability score by the users, only a small transfer of skills to the real-world task is observed, regardless of the hand representation condition. This reveals a need to continue improving the VR simulators' user interface to ensure skills transfer to the real world.

Finally, the *conclusion* recapitulates the contributions of this thesis work, assesses the potential impact on the design of surgical virtual reality simulators, and speculates on the future directions of our research work in this area.

Chapter 1

Surgical simulators

***

Medical training has evolved in the past decades addressing the safety and ethical issues associated with previous paradigms. Nowadays, there is a growing interest in training through simulators.

This chapter introduces and defines the main concepts related to our study object: virtual reality (VR) simulators for learning technical surgical skills, commonly referred to as VR medical trainers. We present and also discuss the current state of the art of surgical simulators, with a strong focus on virtual reality trainers. After identifying some of the current issues related to the design of those systems, we present our research methodology based on the interaction design and simulator fidelity approaches.

Finally, we discuss the general interaction tasks involved in designing and evaluating these virtual reality simulators, and analyze surgical VR simulators by focusing on these interaction tasks to identify the research gaps in this domain.

1.1 The characteristics of surgical skills

It is important to distinguish between two different types of knowledge, the "declarative" and the "procedural" knowledge. The declarative knowledge "concerns knowing facts, theories, events and objects" (Jiamu, 2001). On the other hand, the procedural knowledge is "knowing how to do something, which includes motor skills, cognitive skills, and cognitive strategies" (Jiamu, 2001). In other words, when the learner needs to know something static, such as the anatomy of the human body, it refers to declarative knowledge. In contrast, when the learner is required to know how to do something, such as the steps to intubate a patient or perform the actual intubation, it refers to procedural knowledge. As mentioned before, procedural knowledge encompasses motor skills as well as cognitive skills and strategies. Therefore, it is important to distinguish between learning how to perform a surgical procedure (i.e., the steps to follow) and the actual execution of the procedure. Indeed, practitioners first need to learn the different steps to perform, get to know the tools they are going to employ and how to wear/manipulate them, and then learn to perform the different parts of the procedure (specific gestures, movements and actions), to finally combine all into a whole procedure.

In this thesis, the focus is made on surgical skills. For those skills, two main procedural skills to train can be identified, technical and non-technical skills. In this work, the focus is more particularly made on technical skills. These skills are defined as "any psychomotor action or related mental faculty acquired through practice and learning pertaining to a particular craft or profession" (Agha et al., 2015). More specifically, we concentrate on the motor skill component of this type of knowledge and not on the procedure needed to perform or the cognitive aspects required during the task. For this type of skill, proceduralization is particularly desirable. Indeed, a complex technical skill can be subdivided into shorter and simpler tasks to train.

Technical surgical skills are complex by nature. Even the most basic tasks, such as suturing, will require several skills such as precision, bimanual dexterity, hand-eye coordination, and force control. During this task, clinicians need to safely manipulate the tools while controlling the applied forces and looking at the tissue and structures. This makes learning these skills complex and highlights the need to design adapted tools to support this learning process.

Conclusions on surgical skills characteristics

Human knowledge is divided into two different components. In this thesis, we will mainly focus on the technical skills component necessary to perform surgical procedures. These

skills are complex and require practice before being mastered. In the following section, we will expose the existing models for learning these skills.

1.2 Learning models for surgical skills

The two types of knowledge presented above imply different learning processes. On the one hand, basic declarative knowledge units are learned relatively easily and can be modified (or reorganized) quickly (Jiamu, 2001). On the other hand, procedural knowledge can only be acquired slowly, but once automated, it is tough to modify (Jiamu, 2001). The mastery of this type of skill "depends on the ability to perform it unconsciously with speed and accuracy while consciously carrying out other brain functions" (Bloom, 1986). Indeed, motor skills can be performed unconsciously with great speed and accuracy only when reaching the highest level of automaticity, allowing the practitioners to focus their attention on the other factors of the procedure.

Traditional methods for training surgical technical skills followed Halsted's apprenticeship model (Halsted, 1904): "see one, do one, teach one", where trainees learn by observation of experts and by training on real-patients and animals (in-vivo), or cadavers (ex-vivo). However, this approach raises ethical concerns (Balcombe, 2004; Coles et al., 2011) and patient-safety issues (Arikatla et al., 2013; Coles et al., 2011).

To limit these issues, a new model based on practicing on simulators (Rodriguez-Paz et al., 2009; Vozenilek et al., 2004): "see one, *practice safely*, do one, teach one" emerged over the past years as a credible alternative approach. In this context, the US has been the first country to develop medical simulation centers (McGaghie et al., 2010) to train healthcare professionals before being exposed to real patients. Other English-speaking countries widely followed them. In 2012, the French HAS ("Haute Autorité à la Santé") published a report whose conclusion was "never the first time on a patient" with substantial consequences in terms of training methods. As a result, in 2017, a new reform of the medical curricula emerged (JORF, 2017), which imposes an evolution of traditional education models and the strengthening of the role of simulation for training healthcare professionals.

Simulation has been defined as: "The technique of imitating the behavior of some situation or process by means of a suitably analogous situation or apparatus, especially for the purpose of study or personnel training" (Bradley, 2006). One of the first significant events in medical simulation history is the *Resusci-Anne* manikin simulator developed in 1960 (Cooper and Taqueti, 2004; Grenvik and Schaefer, 2004). It was designed for the practice of mouth-to-mouth breathing and later evolved to incorporate a spring

in the chest for the practice of cardiopulmonary resuscitation. Since then, simulation has improved, and new technologies are being integrated to support simulation-based medical skills learning.

Conclusions on surgical skills learning models

To overcome the issues related to the traditional methods for learning surgical skills, simulation evolves as a reliable and promising model. The ultimate goal of surgical simulation is to improve patients' outcomes by allowing residents to train and test their surgical performance in a safe and standardized environment, ensuring surgical competence before performing live procedures on patients.

In the next section, we present a review of surgical simulators, which are being used to master technical skills in this domain.

1.3 Classification of surgical simulators

Simulation provides a safe, educational environment. It allows novices and experts to practice and develop their skills, knowing that mistakes carry no penalties and that no harm can be made to patients. It also allows learners to set their own pace, facilitating on-demand learning, allowing repetitions, generating different tasks and scenarios, and providing automatic feedback. Furthermore, it enhances standards against which learners' performance can be compared, helping to better identify the educational needs (Bradley, 2006).

Simulators are designed to reproduce some aspects of the working environment, which can vary from replicating a single aspect of a task to the recreation of the entire working environment, such as the operating theater.

Maran and Glavin (2003) propose a classification system for medical simulators:

- Computer-based systems
- Simulated patients
- Simulated environments
- Integrated simulators
- Part-task trainers
- Virtual reality and haptic systems

1.3.1 Computer-based systems

Computer systems are used to model aspects of human physiology, pharmacology, or simulated tasks and environments. The interaction is acquired through a computer interface, and the focus of learning is on using the information to make treatment decisions and observe their results. An example of this type of medical simulator is the *Gas-Man* system (Med Man Simulations, Inc., Boston, MA) (Simulations, 2018), which is a pharmacological modeling program that allows the user to train for anesthesia dosage (see Figure 1.1a).

These simulators are useful for acquiring basic skills concerning treatment decisions by observing the results in action. They are relatively inexpensive and allow for independent learning. However, they are limited to only presenting decisions and actions through a computer interface, not allowing the experience of realistic situations and complex procedures.

1.3.2 Simulated patients and simulated environments

The simulated patient is a professional actor trained to present a clinical case and sometimes to mimic physical signs, or can be a patient previously trained to present the medical history (i.e., facts and symptoms) in a reliable manner. They are commonly used to train communication skills (Bradley, 2006).

Simulated environments refer to the recreation of an environment in which the clinical activity will occur (e.g., an operating theater). They expect to increase the learner's engagement and are commonly used for team training, allowing learners to examine their roles within a team (Maran and Glavin, 2003).

These simulation types are more related to non-technical skills (e.g., communication, decision-making, team management, and leadership) and are time-consuming. They usually involve debriefing sessions with instructors and require the coordination of a whole team and a dedicated operating room.

1.3.3 Integrated simulators

Integrated simulators combine a manikin with sophisticated computer controls in order to provide some physiological parameter outputs. These outputs can be physical (e.g., pulse rate and respiratory movements) or electrical (e.g., monitor readouts).

The earliest of this type of simulators is the *Sim One* simulator, developed by Abrahamson and Denson in the late 1960s (Abrahamson et al., 2004). This manikin's main features are breathing, heartbeat, temporal and carotid pulse, and blood pressure. It responds to

intravenously administrated drugs and gas administration, and the physiologic responses are rendered in real-time through a computer program. More recently, the *SimMan* simulator (Laerdal Medical, Stavanger, Norway) (Medical, 2018) has become available at a much lower cost (see Figure 1.1b).

On the one hand, these simulators allow the learners to train in pre-configured scenarios with realistic models and outcomes. Both technical and non-technical skills are put into practice under stressful simulations. On the other hand, most of these simulators are relatively expensive in terms of purchasing the simulator and the associated running costs. Because of this, they are located in specialized centers. Some less expensive simulators exist, with limited computer programs. Finally, a debriefing session is required to provide feedback on the learner's performance.

1.3.4 Part-task trainers

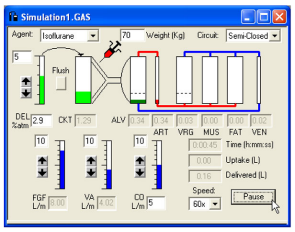
Part-task trainers are designed to replicate only some aspects of the environment. They are commonly used to train technical and psychomotor skills, such as bimanual dexterity, suturing, tissue dissection, eye-hand coordination, catheterization, cannulation, or venepuncture (Forsslund et al., 2011).

These simulators can mainly be divided into physical trainers and virtual reality trainers (virtual reality examples will be discussed later in section 1.5).

In general, physical trainers incorporate synthetic models of tissue and organs and use real instruments for interacting with them. The objective is to master specific individual motor skills or subtasks of a whole surgical procedure. Once mastered, those skills can then be combined in order to perform the full procedure.

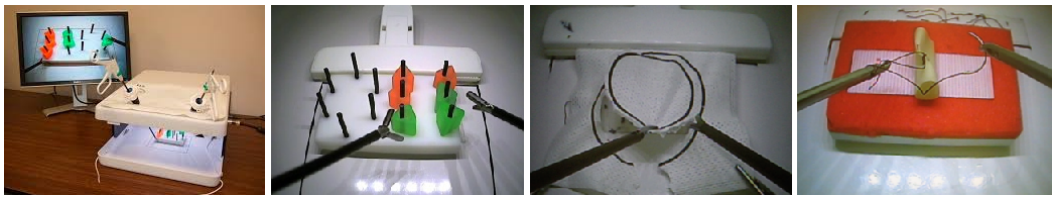
An example of a physical part-task trainer is the *Fundamentals of Laparoscopic Skills* (FLS) box-trainer, which consists of five standard tasks: the peg transfer, the pattern cutting, the ligating loop, and intracorporeal and extracorporeal suturing (Fried et al., 2004) (see Figure 1.1c). This system has been validated and adopted in North America as the standard for training and assessing basic psychomotor skills needed in laparoscopic surgery (Palter, 2011; Pitzul et al., 2012; Skinner et al., 2013; Valdivieso and Zorn, 2014).

To summarize, part-task trainers allow training several basic skills separately to include them in a complex full procedure. Besides, physical part-task trainers purchasing is relatively inexpensive, allowing centers to acquire several simulators. However, the consumables used in these simulators elevate the associated training cost. Finally, independent learning is possible, but no performance feedback is automatically provided, requiring an instructor to be present for feedback communication.



(a) Computer-based system: "Gas-Man".

(b) Integrated simulator: "SimMan".



(c) Physical part-task trainers: Fundamentals of Laparoscopic Skills (FLS) box-trainer.

Fig. 1.1 Surgical simulators classification

1.3.5 Virtual reality and haptic systems

Virtual reality simulators use computer modeling and complex software to simulate a part-task or a full procedure. These systems have physical interfaces, which may integrate sensing instruments to obtain the user's movements. The software is in charge of computing the changes happening within the model and rendering the correct response, which can be visual, auditory, haptic, or multi-modal. These systems provide three dimensional (3D) imaging to make the environment more realistic and intuitive. They can be provided with real-time interactivity in the full 3D space. This allows the user to explore the model's representation from different viewpoints and to query specific or individual structures to understand the model better while performing the task.

Some VR simulators integrate haptic (from the Greek word, *"haptesthai"*, which means "related to the sense of touch" (Salisbury Jr, J. Kenneth, 1999)) feedback, producing in the user the illusion of being in contact with the physical model. Based on the underlying neural inputs, the sense of touch can be divided into cutaneous (employs receptors integrated in the skin), kinesthetic (employs receptors which are located in muscles, tendons, and joints), and haptic systems (employs both cutaneous and kinesthetic receptors) (Wang et al., 2014). Through haptic devices, the user can "feel" the patient representation in addition to seeing it.

Virtual reality has proven to reduce risks to both medical students and patients by allowing them to train in a controlled environment before the real practice (Buckley et al., 2012). It also provides the objective measurement tool through which the evaluation process can be quantified and validated (Satava, 2001), which is not the case with physical

simulators, where a teacher/evaluator needs to be present to ensure that the task is performed correctly, give feedback to the trainees and globally evaluate them (subjective interpretation of the results).

1.4 Evaluation of surgical simulators

Satava et al. (2003) have proposed a taxonomy of metrics for evaluating surgical abilities and skills. This taxonomy includes two concepts: *validity* and *reliability*.

Validity refers to accepting a test if it complies with five validation steps. Reliability refers to the consistency of the results of a test if this test is performed multiple times. In the following, we will focus more particularly on the concept of validity, which is usually used to assess different aspects of surgical simulators. The five steps required to validate a surgical simulator (Satava et al., 2003) are: face, content, construct, concurrent and predictive validity.

Face validity (validity of appearance) is an aesthetic validation based on the appearance of the simulated task interface and its difference with the real device. Content validity assesses the accuracy and relevance of the content proposed by the simulator and is performed by expert surgeons based on the detailed examination of the system content. Construct validity aims to determine the system's ability to differentiate a novice's performance from that of an expert. Concurrent validity determines whether the simulator is similar to its competitors in the field. Prediction validity verifies that the performance obtained on the simulator is similar to that of the real situation, i.e., whether the evaluated trainee will have the same performance level in a real task.

Each validation step checks an essential aspect of the simulator, and by fulfilling all of them, its complete validation is obtained.

Conclusions on surgical simulators classification and evaluation

To summarize, the leading simulators used for training technical surgical skills are part-task trainers and VR-based simulators. Compared to physical simulators, VR-based simulators are cost-effective. Indeed, most physical trainers are resource-intensive in terms of consumables, which increases the annual training cost. Moreover, VR systems allow exposing the user to many different scenarios and easily configuring different properties of human bodies (Chellali et al., 2016). Finally, the users can train on a VR-based simulator until they reach the target level of proficiency required to perform the procedure on a real patient. As (Satava, 2001) summarizes: "The greatest power of virtual reality is the ability

to try and fail without consequence to animal or patient. It is only through failure –and learning the cause of failure– that the true pathway to success lies".

However, although several studies have been performed on existing VR simulators and prototypes (Ayodeji et al., 2007; Beyer et al., 2011; Duffy et al., 2005; Eriksen and Grantcharov, 2005; Gallagher et al., 2001; Schijven and Jakimowicz, 2003; Sherman et al., 2005; Van Dongen et al., 2007; Våpenstad et al., 2013; Zhang et al., 2008), a lack of formal validation is observed, which limits the adoption of these systems in the training curricula. Indeed, most of the works presenting construct validity of VR-based simulators are questionable in terms of power analysis (i.e., number of subjects and profiling), vague in terms of system description (limiting the study replication), not unified in terms of performance evaluation measures, and contradictory in their results findings (Våpenstad et al., 2013). In addition, predictive validity studies are missing due to the ethical and patient safety issues associated with this type of validation, i.e., it requires testing the transfer of skills on real patients. Therefore, further evaluation of these systems is required to assess their effectiveness as surgical trainers and to include them in the surgical education curricula. Particular attention is required for assessing the transfer of skills to real-world tasks. The concept of transfer will be further discussed in chapter 2.

The following section presents examples of existing research prototypes and commercially available VR simulators for the surgical training to analyze later VR surgical technical simulators from an interaction design perspective.

1.5 Examples of VR surgical simulators

Virtual reality for surgical training was first conceptualized by Reznick (1993), and a VR surgical simulator was first described by Satava (1993) the same year. This simulator, strongly inspired by the pilots' training program, consisted of 3D images of the abdomen organs and surgical tools. The virtual environment (VE) was displayed in a head-mounted display (HMD), and simple interactions were provided through a data glove. No force or tactile feedback was present (see Figure 1.2).

Since then, many VR surgical simulators have appeared, both as research prototypes and as commercial products to be acquired and incorporated into the surgeon's curricula. Besides, with the appearance of VR technologies as a new training modality, the number of works and strategies that focus on teaching and assessing surgical skills increased. In the following sections, we present examples of different surgical and medical procedures and the associated VR simulators. These examples were chosen because they are directly related to the systems designed during this thesis work.

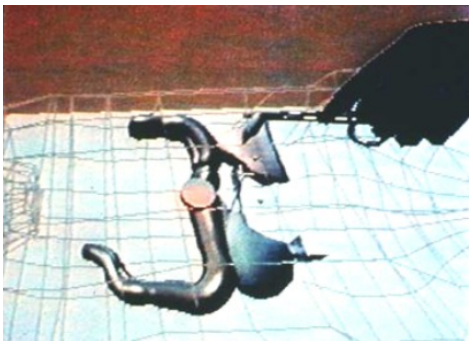


Fig. 1.2 First surgical virtual reality simulator (Satava, 1993)

1.5.1 Minimally invasive surgery simulators

Minimally invasive surgery (MIS) requires surgeons to operate with special tools inserted into the patient through small incisions. They manipulate the handles of these tools from outside the patient without entering in contact with the diseased tissues or organs. In most cases, without having a direct vision of the inside body situation (Cao et al., 1996). This surgical procedure has many advantages for the patient, such as the avoidance of large wounds, less postoperative pain, and faster recovery (Kim et al., 2003). However, it is more demanding for clinicians, who do not have a direct vision while performing the procedure, and who must learn to manipulate a variety of tools for each particular procedure. In this context, a clinician is required to master several technical skills in order to perform MIS, e.g., hand-eye coordination, tool handling/manipulation, grasping an object, tying a knot, needle insertion, tissue dissection, cutting, catheterization, cannulation, venipuncture, suturing, and ambidexterity.

Simulators to train MIS vary from systems replicating a single aspect of a task to those recreating the entire surgical environment (Maran and Glavin, 2003). Both commercial and research prototypes of VR surgical simulators for such procedures exist. Some examples are discussed hereafter.

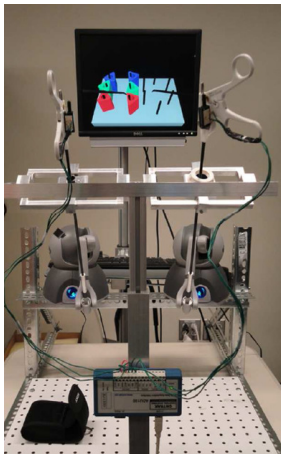
The *LapSim* simulator (Surgical Science, Göteborg, Sweden) (Science, 2018) contains modules for training laparoscopic surgery skills and tools manipulation. It consists of a software application displayed on a monitor and two laparoscopic handles with haptic feedback (see Figure 1.3a). The different basic skills tasks are camera navigation, instrument navigation, coordination, grasping, lifting and grasping, cutting, and clipping and cutting (Van Dongen et al., 2007). This simulator also provides whole procedure modules such as cholecystectomy and appendectomy.

The *VBLaST* is a VR research prototype for laparoscopic skills training (Arikatla et al., 2013; Chellali et al., 2015), consisting of computational software to simulate the FLS indi-

vidual procedures (the FLS box-trainer was described in subsection 1.3.4). It is provided with a physical user interface to connect two laparoscopic graspers handles to two haptic devices (3 degrees of freedom (DOF) for force feedback) (see Figure 1.3b). It has been evaluated through experimental studies and compared to the FLS box-trainer (prediction validity). The results of these studies show that the system has a surgical performance not significantly different from that of the FLS box-trainer in assessing laparoscopic skills (Chellali et al., 2016; Fu et al., 2019; Linsk et al., 2018; Zhang et al., 2013).

Similar commercial simulators are also available, presenting the same setup than previous simulators, e.g., *LAP-Mentor* (Simbionix USA, Cleveland, OH) (Simbionix, 2018), *LapVR* (CAE Healthcare Inc., Sarasota, FL) (Healthcare, 2018), *Lap-X* (Medical-X, Rotterdam, the Netherlands) (Medical-X, 2018), etc. The reader can refer to Goodman et al. (2019) and Våpenstad and Buzink (2013) for a more extensive review of existing endoscopic and laparoscopic simulators.

(a) Lap-Sim simulator (Science, 2018).



(b) VBLaST simulator (Chellali et al., 2016).

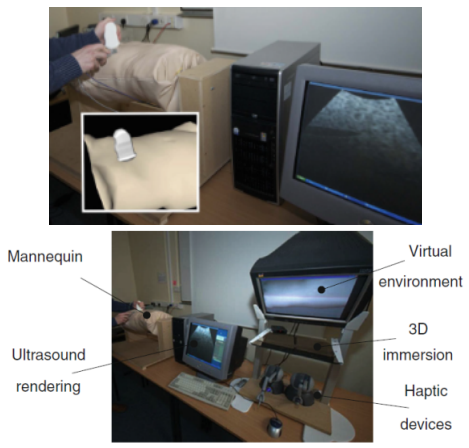
Fig. 1.3 Surgical simulators for Minimally Invasive Surgery

1.5.2 Needle insertion simulators

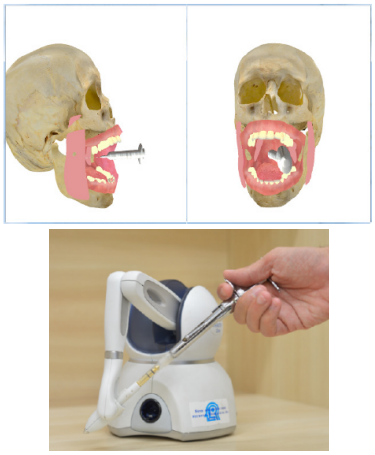
Needle insertion procedures consist of introducing a needle into the patient's body with the help of a medical imaging system (e.g., CT-scan, scanner X, ultrasound) to reach a target area (e.g., tumor or organ) and perform a tissue sampling or a treatment. This kind of procedure requires the surgeon to master haptic perception skills and three-dimensional spatial abilities. They are sometimes considered as MIS procedures. Several researchers have focused on developing VR trainers for needle insertion procedures.

Villard et al. (2014) proposed a simulator for ultrasound-guided needle puncture skills through a liver biopsy task. This procedure requires hand-eye coordination of ultrasound imaging and needle manipulation to reach a liver mass for biopsy diagnosis. It is designed through two workstations. The first one uses a manikin to determine the incision and insertion point. The second one is a VR simulator with a first haptic device for force feedback and needle manipulation, and another one to use an ultrasound probe for needle guidance (see Figure 1.4a).

Corrêa et al. (2014) present the development and preliminary evaluation of a VR-based simulator for dental anesthesia training. The simulated task is the manipulation of the syringe, evaluated through several measurements (velocity of manipulation, regions reached by the needle, translation and rotation movements, and completion time). The system uses a haptic device to render the forces, and it was implemented in a VR framework allowing 3D synthetic models, collisions, and deformation (see Figure 1.4b).



(a) Liver biopsy trainer (Villard et al., 2014).



(b) Dental training (Corrêa et al., 2014).

Fig. 1.4 Surgical simulators for needle insertion

Henshall et al. (2015) presented their work in progress for a VR environment for training a kidney biopsy procedure. The system consists of a high-quality stereoscopic display with head-tracking and haptic feedback. They combined two hardware components to achieve the biopsy needle behavior. The G-Coder Simball 4D joystick enables the needle insertion in the virtual body accurately. The Geomagic Touch haptic device (3 DOF force feedback) was connected to the previous device and used as a force feedback interface (see Figure 1.4c). In-vivo measurements generated the force model with a glove using a force sensor resistor from a commercially available prostate biopsy. Their software was developed using the Unity3D game engine. The anatomical models of the spine, liver, and kidneys were created by segmenting medical images from a patient's abdomen CT scan.

The *EYESI* system (VRmagic Holding AG, Mannheim, Germany) (VRmagic, 2018) uses computer-generated images, which can be seen through a realistic stereoscopic microscope. It combines haptic feedback and realistic instruments to allow trainees to practice intraocular surgery (see Figure 1.4d).

The reader can refer to Abolhassani et al. (2007) and Corrêa et al. (2018) for a more extensive review of needle insertion trainers.

(c) Kidney tumor biopsy simulator (Henshall et al., 2015).

(d) The EYESI system (VRmagic, 2018).

Fig. 1.4 Surgical simulators for needle insertion (*Cont.*)

1.5.3 Ear surgery simulators

Ear surgery requires the surgeon to observe the surgical area on a microscope while using different instruments (e.g., drills, tweezers, suction tubes). The main challenges of this type of simulation are the magnification of the surgical area through a surgical microscope, the use of specialized precision instruments, and master the different specific operating techniques.

The *VOXEL-MAN* Tempo Surgery Simulator is a temporal bone virtual reality simulator (Arora et al., 2012). Volumetric high-resolution computed tomography images of the temporal bone are used to produce a 3-dimensional representation. It uses stereoscopic rendering of virtual simulation through shutter glasses, and haptic feedback is provided through a desktop device, which is also used to manipulate the virtual drill. A pedal serves as a device for switching the tool. Besides, performance measures are recorded during the whole procedure (see Figure 1.5a).

The *Mediseus Surgical Simulator* (CSIRO/University of Melbourne Temporal Bone Simulator) consists of a simulated operating microscope (Wijewickrema et al., 2018, 2015). The user interacts with a 3D volumetric virtual representation of a cadaver temporal bone

using two haptic devices (see Figure 1.5b). Color is applied to the bone model parts to represent their hardness, and both visual and tactile cues are rendered to simulate the real surgical procedure. In addition, physiological functions, such as bleeding, are also included through auditory and visual feedback.

(a) The VOXEL-MAN simulator for temporal bone (Arora et al., 2012).

(b) The Mediseus surgical simulator (Wijewickrema et al., 2018).

Fig. 1.5 Surgical simulators for ear surgery

Conclusions on VR surgical simulators

To summarize, several surgical VR simulators, commercial and research prototypes, exist for training technical skills in this domain.

However, their development is complex and potentially expensive. They require the inclusion and integration of several interfaces and devices to provide multi-modal feedback, increasing the entire simulator cost. Indeed, to allow the acquisition of technical skills on these systems, effective interfaces should be designed, which remains a challenge (Chellali et al., 2016). Indeed, complex and inappropriate user interfaces may make an interactive system underutilized or misused (Maguire, 2001). Therefore, it is necessary to focus their design and use on the actual training requirements of the healthcare professionals by determining which are the interaction tasks to simulate and how to choose the best interaction techniques and technologies to guarantee effective training. In the following section, we expose the approaches traditionally used for designing VR surgical simulators and present the approach followed during this thesis work.

1.6 Surgical simulators design approaches

As mentioned above, our object of study in this thesis is *VR surgical simulators*. These systems can be viewed at the same time as a simulation system and as an interactive system.

There is a strong belief that the development of effective VR simulators is solely an engineering challenge. In fact, much of the engineering approach is technology-driven and has been focusing on high visual realism (Drews and Bakdash, 2013). On the other hand, the multidisciplinary iterative design approaches permit to design interactive systems that better fit the end-users' needs. There is a growing interest in using this approach to design virtual reality surgical trainers (Forsslund et al., 2011). According to those approaches, building an efficient interactive training surgical system requires (Chellali et al., 2016; Forsslund et al., 2009):

1. Defining the learning objectives of the system. This point is specific to each skill and procedure to be trained using the system. According to the user-centered design approach, this requirement is defined during the task analysis phase, performed during the early stages of the design process.

2. Defining the most important part of the surgical task to simulate, and to what extent it can be simulated successfully. This point is related to the simulator fidelity concept and will be discussed more in-depth in chapter 2.

3. Designing effective user interfaces and interactions to encourage trainees to use the system. Previous research has shown that complex and inappropriate user interfaces and interactions make an interactive training system likely to be underused or misused, with frustrated trainees not acquiring the targeted skills when using these systems or preferring to maintain their current training methods (Maguire, 2001). Therefore, it is crucial to investigate how interactions are supported and designed on those systems. This is discussed more in-depth in the following section.

The approach we have followed during this thesis work is mainly based on the HCI iterative design methodology coupled with a technology-based approach tackled from the simulator fidelity concept perspective. In fact, there is a considerable amount of research on simulator fidelity that needs to be considered when designing such interactive systems. The objective was to take advantage of both approaches in order to better choose appropriate design components that might help ensure system usability, user performance, and skills transfer of these systems. The concept of simulator fidelity, as well as its link with interactions' design, will be discussed more in-depth in chapter 2.

In the following section, VR simulators will be regarded as a 3D interactive system where the user (trainee) interacts with a system (virtual environment) through a set of user interfaces and interaction techniques. This will permit us to analyze these systems from an HCI perspective in order to investigate their interactions' design.

1.7 3D interactions in VR systems

A common way of describing human-computer interfaces is in terms of the communication process between a user and the system (Bowman et al., 2004) (see Figure 1.6).

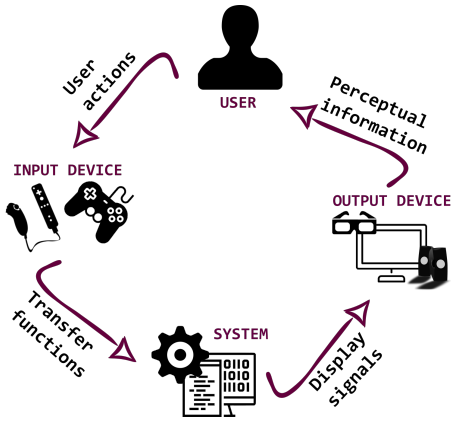


Fig. 1.6 Human-system communication design

Firstly, there is an action performed by the user as a result of translating his goals and intentions. This action corresponds to a direct manipulation of an input device (physical interface). The input device data is then computed to generate an instruction or raise an event in the simulator core. The models, rules, and algorithms implemented are used to evoke an answer to this event or instruction. This answer is then translated into meaningful signals to be displayed by some output devices as a way to respond to the user's action. Finally, the stimuli produced by these devices are perceived by the users (e.g., sound, light). These stimuli allow them to determine whether the interaction was fulfilled and translated into a meaningful semantic representation.

In the context of VR simulators, we are more particularly interested in 3D interactions performed through input and output devices and interaction techniques.

Interaction techniques are defined as the methods used to accomplish a given task via the interface (Bowman et al., 2004) and include both hardware and software components. 3D interaction techniques can be divided into four basic tasks that the user can perform when using a virtual reality simulator: selection, manipulation, navigation, and system control.

- **Selection:** This task consists of specifying one (or more) object from a set of objects.
- **Manipulation:** consists of modifying object properties, such as its position or orientation, scale, color, shape, behavior.

- **Navigation:** is the task of moving in and around the environment and includes travel and wayfinding. Travel (or viewpoint control) refers to the user's low-level actions to control the position and orientation of the viewpoint. Wayfinding is the cognitive aspect of the navigation, which involves identifying the current position and desired goal location, planning the trajectory to reach it, and making a mental representation of the environment map.
- **System control:** encompasses the changing of the system mode or state, usually achieved through menus or commands.

This classification is continually revisited, proposing, for instance, to consider selection and manipulation as one task, since they are usually analyzed together; to split navigation task into its two components: travel or viewpoint control, and wayfinding, since many individual solutions exist for each of them; and to add the user's avatar representation and control as a new basic task. In the next section, we perform a more in-depth analysis of each of these tasks and their use within VR surgical simulations.

1.8 3D interaction techniques in VR surgical simulators

This section aims to analyze the interfaces and interaction tasks included in surgical virtual reality simulators. We discuss the use of each of the basic tasks in general systems, and more specifically, in existing surgical VR simulators.

For the purpose of the design of surgical trainers, we will not focus on analyzing system control since it is inherent to VR simulation. Hence, not playing, from our point of view, a role in the training of technical surgical skills. Manipulation and selection are discussed together. A greater focus is made on manipulation, since during surgical simulation, the trainees are holding tools during the whole task, making the selection task less relevant. In addition, when discussing navigation, the focus is made on travel (i.e., how to change the point of view of the virtual camera). Indeed, during surgical procedures, the practitioners are either seated or standing facing the working site. In both cases, with a limited requirement to change their position. Hence the travel component is limited to changing the perspective (i.e., the point of view) to visualize the tools, hands, and the patient's body. Finally, we discuss the user's avatar control as we hypothesize that it can impact motor skills learning. This particular hypothesis is investigated more in-depth in chapter 4. It is to be noted that the borders between the different tasks are sometimes skinny and that some interaction techniques can support several tasks at a time.

1.8.1 Manipulation

Interactions in surgical simulators for training technical skills generally involve manipulating tools and instruments to interact with the different objects of the virtual environment. They are the primary operative task in these systems.

Manipulation in virtual environments

Manipulation involves changing the characteristics of a selected object. Usually, this change consists of applying a spatial transformation, but it can also be performed by applying a deformation or changing the object's visual properties (e.g., color) or behaviors. In VR systems, this process is usually performed by means of an interaction device.

Different spatial transformations can be performed on an object, but the main relevant ones are translation, rotation and scaling (similar to the real-world physical interactions) (Mendes et al., 2019). Each transformation can be applied in 3 dimensions, which will determine the number of degrees of freedom of the interaction. Indeed, these transformations DOF are generally the result of mapping the input device used for interacting. It is possible to combine more than one transformation (i.e., perform a translation and rotation simultaneously).

Three types of mapping can be identified for these transformations:

1. a direct mapping (1 : 1 control of the virtual object), which means that the input device DOF is directly mapped to the virtual object movement;
2. an indirect mapping, for instance, when we use different devices (i.e., mouse, keyboard) to carry out the transformations, or when we use 2D tracking devices to generate 3D manipulations;
3. and a combination of direct with indirect mappings.

Three-dimensional manipulations based on keyboard and mouse input are used for several tasks, such as architectural design and virtual modeling. An example of 3D manipulation using the mouse is Houde (1992) proposition of a bounded box with handles for performing translations and rotations. Other works have focused on using touch interactions (Hancock et al., 2007; Martinet et al., 2010), and finger identification with different gestures transformations (Goguey et al., 2016). In this matter, Wu et al. (2017) proposed two manipulation techniques using finger identification for the different axis and using the LeapMotion device to track them while interacting on a touch-pad display (Wu et al., 2016, 2017). This type of device (i.e., LeapMotion and Kinect) also allows performing mid-air gesture-based touchless interactions (Hurstel and Bechmann, 2019).

Moreover, other direct manipulation techniques have been explored, allowing to manipulate objects in 3D with a more natural input mapping, similar to what one can do in the physical world. Examples of such techniques include the tracking of the users' hands or the use of wearable devices. These techniques are particularly used in immersive VR simulators. Tracked devices being hold by users are current solutions proposed by the gaming domain due to the availability of commercial products, such as the HMD joysticks (e.g., HTC VIVE, Oculus Touch). These devices provide 6 DOF tracking capabilities, plus several buttons that increase the total DOF of the interaction technique. Haptic desktop devices are also used as direct object manipulation devices, where users manipulate the stylus to interact with the objects in the virtual scene. DOF may vary depending on the sensing capabilities of the device. Furthermore, other solutions were proposed for hand-tracking based manipulations. Mapes and Moshell (1995) proposed to use gloves to track both user hands, allowing them to manipulate the object located in the hands' middle position. In this proposition, translations were applied when both hands moved in one direction. Rotations were applied when moving the hands around the center of the object. Finally, by modifying the distance between hands, scaling operations were applied. Other propositions for hand-tracking are based on camera and inertial sensors tracking capabilities (De Araújo et al., 2013; Hachet et al., 2011)

The reader can refer to Bowman et al. (2004), Argelaguet and Andujar (2013), Jankowski and Hachet (2015), and Mendes et al. (2019) works for an exhaustive review on selection and manipulation techniques.

To summarize, different techniques can be used for manipulating objects inside VEs. The choice between them will depend on the application to simulate and the level of accuracy and effectiveness required by the task and the simulator itself. It is necessary to determine which techniques are more appropriate for manipulation in VR surgical simulators. As we mentioned before, surgical procedures require the user to manipulate virtual tools (e.g., tweezers, scalpel, needle, forceps) to interact with the virtual objects in the scene (e.g., tissue, organs, bones, gaze, thread). Since technical skills learning requires practicing the same gestures and actions until reaching the highest level of automaticity, the principles of direct manipulation of virtual tools appear to be the most suitable choice in this case. This hypothesis is further discussed hereafter.

Manipulation in surgical simulators

As mentioned before, surgeons' hands indirectly manipulate tissue and organs through instruments. However, it is not completely understood how these instruments interfere with the sensory perception of the surgeons (Westebring – van der Putten et al., 2008).

Usually, surgeons and radiologists need to feel forces and position information (kinesthetic perception) generated by the instruments which are in contact with the tissue to control them better, understand what they are doing, and avoid fatal mistakes (Basdogan et al., 2004; Zhou et al., 2012). The perception through instruments is reduced compared to bare hands, although they are able to distinguish the shape, size, and consistency of the tissue. Previous research has shown that haptic feedback is an important requirement for surgical VR trainers (Drews and Bakdash, 2013; Kim et al., 2003). The use of haptic feedback can be explained through Martin Heidegger's theory, presented in his book *"Being and Time"* (Heidegger, 1927). He explains how we encounter the world and act through it, and distinguishes between *"ready-to-hand"* and *"present-at-hand"*. By *"ready-to-hand"* he describes an entity, a tool, that has become invisible, which resides in the background of the work and we are no longer conscious of it, allowing an ordinary involvement and interaction with the task we are performing, as opposed to the tools we are using to perform this task. But when that entity or tool fails to achieve its purpose, since a part or feedback is missing, which is required for the entity to function as we expect, it comes to the foreground as an entity that is *"present-at-hand"*, making the focus of the attention and obstructing the concentration towards the task. Hence, haptic interaction and well-functioning tools in motor skills training may allow the users to be entirely concentrated on the task they are asked to perform. In other words, a simulator without force feedback may not train the students to cope with the interference of the tools and organs, hence not allowing them to master haptic skills. Therefore, haptic feedback appears as a requirement for designing virtual reality surgical trainers if we expect trainees to better transfer their skills to the real procedure, as supported by previous literature (Chellali et al., 2015; Dang et al., 2001; Panait et al., 2009; Reich et al., 2006; Seymour et al., 2002; Ström et al., 2006).

A detailed summary of haptic simulators in the medical domain can be found on Escobar-Castillejos et al. (2016). Most of the works presented in that review implement force feedback simulations (3 or 6 DOF for force feedback) through haptic interfaces (Falcon and Haptic Omni devices in most of the works). However, some works focus on providing feedback in the form of vibrations (e.g., through the use of joysticks like the Wiimote).

In addition, many researchers use haptic devices coupled with a shape-alike tool to achieve a correct representation and by reproducing the same manipulation interaction as expected in the real procedure (Chellali et al., 2016; Henshall et al., 2015; Ullrich and Kuhlen, 2012). Commercially available VR simulators for laparoscopic surgery, such as *LAP-Mentor* (Simbionix USA, Cleveland, OH) (Simbionix, 2018), *LapVR* (CAE Healthcare Inc., Sarasota, FL) (Healthcare, 2018), *Lap-X* (Medical-X, Rotterdam, the Netherlands)

(Medical-X, 2018), and *LapSim* (Surgical Science, Göteborg, Sweden) (Science, 2018) are also some of the examples that follow this principle.

Finally, other systems use classical VR joysticks to carry-out the manipulation tasks, sometimes by guiding the interaction or restricting the movement to fewer degrees of freedom than the real world task demands, e.g., *SimforHealth* VR simulator (SimforHealth, Bordeau, France) (SimforHealth, 2018) and *Vtopia* (Trois Prime Lab, Paris, France) (Lab, 2018). However, these systems were not evaluated, and the absence of haptic feedback may impact their effectiveness as training tools.

To summarize, we can observe that most of the existing surgical simulators focus their attention on simulating the manipulation task of instruments and tools. Indeed, this task is the primary operative task to train for technical skills learning. Most of the solutions propose to use one (or two for bi-manual tasks) haptic device to manipulate the virtual tool in the VE, allowing to implement a 6 DOF manipulation technique with a $1:1$ or $1:N$ (e.g., for magnification movement under a microscope) mapping between the input device and the tool's movements in the virtual scene. In addition, the haptic device also renders the collision forces between the virtual tool and the virtual objects. Indeed, haptic feedback was reported to be useful for motor skills training in surgery. However, since no studies have shown how many DOF are required for haptic feedback, nor if it is required during the whole training procedure or specific tasks, this remains an open research question. Answering this research question is considered as out of the scope of this thesis. In fact, since most previous work suggests that haptic interactions for manipulating surgical instruments and perceiving feedback from the environment are essential requirements for surgical simulators design, we have chosen to include this component in all the prototypes we have designed in this thesis work. Nevertheless, further investigations on this specific research question are required in the future.

1.8.2 Navigation

Beyond the manipulation tasks that are usually considered the primary tasks to be simulated in surgical simulators for training technical skills, the other interaction tasks can be seen as secondary. For instance, navigation tasks are usually overshadowed during the design of current VR surgical simulators. Nevertheless, as discussed earlier, the non-appropriate design of such tasks may impact the system usability and the users' performance when performing the primary tasks. Therefore, it is crucial to investigate how these tasks are implemented in VR simulators and whether there is an impact of the associated techniques on user performance. As mentioned before, we will only discuss

viewpoint control due to surgical procedures' constraints, rendering less relevant the trajectory planning aspects of the navigation.

Viewpoint control in virtual environments

Navigation involves the extraction of visual information from the environment in order to create a mental representation to use for route and trajectory planning or distance estimation. In VR systems, this process is usually performed by controlling the user's viewpoint (camera) of the VE.

Seven parameters are generally employed to model viewpoint control (Marchal et al., 2013): 3 dimensions for positional placement and 3 dimensions for angular placement of the camera, and the camera field of view (FOV). The high number of parameters to control appears as a difficulty in developing effective techniques for 3D navigation in VR (Jankowski and Hachet, 2015). Through these parameters, the camera can be manipulated in the scene to respond to the viewer's desire to explore the world. This can vary from large-scale exploration of a 3D environment to a closer inspection of a 3D object.

When a particular object in the world is being explored, the viewer's goal is to see it from different angles, which can be achieved by orbiting the camera around the object's position (i.e., a combination of rotations and translations around the object), and by modifying the camera's FOV to get a closer/distant look (Ortega et al., 2015). To design effective navigation techniques in VR surgical simulators, it is essential first to identify how navigation is performed in real surgical procedures and, subsequently, which camera parameters need to be manipulated using the interaction technique.

All surgical procedures need a certain degree of viewpoint control to perform the operative task. During technical skills execution, clinicians must efficiently visualize the operative field to better identify the tool movements and interactions with the organs and tissues in order to avoid errors (Jarc and Curet, 2017). In this case, the virtual scene is generally composed of the medical tool(s) and patient/organs representation(s). The trainee's objective in terms of navigation is to visualize the 3D objects from different angles to extract information about the tools' position with respect to them. This suggests that the principles of orbiting around the 3D object (point of interest) would be suitable in this case.

Orbiting around a 3D object can be achieved, for instance, by tracking the user's head position and updating the camera position accordingly, allowing a first-person perspective on VEs. Head-tracking based techniques are commonly used in VR systems and were first reported in the late 1960s by Sutherland (1968). Since then, different techniques and algorithms have been proposed to improve the user's experience (Murphy-Chutorian

and Trivedi, 2009). These techniques can be viewed as naturalistic techniques since they mimic human viewpoint control in the real world.

Besides the naturalistic techniques, metaphoric-based techniques are also used for changing the point of view. For instance, touch-based interactions were used for close-object exploration with three main approaches: direct touch manipulation in the screen space, the use of widgets, and a combination of these two approaches with metaphorical gestures (Kulik et al., 2018). Edelmann et al. (2009) have worked on the DabR system, a computer-vision based system for multi-touch interaction. The implemented techniques consist of a pan/tilt of the camera around its position (one-finger movement), camera zooming (two fingers moving apart), and camera movement parallel to the image plane (two-finger parallel movement). Marchal et al. (2013) have proposed the Move&Look technique. This technique maps single-touch movements to camera displacement along a path (up-down gesture) and egocentric one-dimension rotation (left-right gesture); and multi-touch gestures to rotate around a pointed pivot (first contact point determines the pivot, second contact point movement modifies the rotation angle), scrutinize (pinch gesture), and egocentric two-dimension rotation (left-right and up-down gesture with all contact fingers).

In addition to these techniques, other means for close-object inspection were proposed (Khan et al., 2005; Ortega, 2013). The reader can refer to Hand (1997), Bowman et al. (2004), Christie and Olivier (2009), and Jankowski and Hachet (2015) works for an exhaustive review on navigation techniques.

To summarize, different techniques can be used for controlling the point of view of the virtual camera in VEs. They vary from large-environment exploration techniques to object inspection techniques. The analysis of surgical navigation tasks suggests that VR surgical simulators should rather use object inspection techniques. Object inspection techniques can be divided into two types: naturalistic and metaphoric techniques. In the following section, we will analyze examples of viewpoint control techniques included in existing surgical simulators to investigate whether a consensus exists on choosing one type of technique over the others.

Viewpoint control in surgical simulators

Means for changing the point of view have already been included in surgical simulators. For instance, Henshall et al. (2015) VR prototype for kidney biopsy allows visualizing anatomical models of the spine, liver and kidneys to perform a needle insertion. For this system, the user's head position was tracked to provide a first-person perspective of the operational field. Corrêa et al. (2014) have presented a VR simulator to train anesthesia

procedures in the dentistry area. Users were able to visualize different models through different points of view using the keyboard.

Fortmeier et al. (2016) have presented a VR trainer for percutaneous intervention, for which they use the haptic interface to manage both the manipulation of the needle and navigation of the ultrasound probe. However, the use of the haptic device for both tasks is questionable and may negatively impact the system's usability.

Other works focus on implementing navigation techniques to train surgical skills that rely mainly on the visualization of structures, such as organs and tissues. For example, Tang et al. (2007) have built a VR trainer for neuro-endoscopy surgery. It provided different visualization modes of magnetic resonance imaging (MRI) data to better plan the surgical procedure. They have used the mouse position to allow the user to navigate inside the 3D model. However, this system was not designed for training technical skills.

To summarize, although several VR surgical simulators include viewpoint control techniques to provide a usable system, none of the previous works have discussed the design of the included techniques or investigated their impact on the user's performance of the primary tasks trained on these systems. This suggests that there are currently no guidelines to help designers choose an appropriate viewpoint control technique for these systems. Therefore, this remains an open research direction to investigate that will be discussed in chapter 3.

1.8.3 Avatar representation and control

Another important factor to consider during the design of interaction tasks for VR systems is the self-avatar representation control. The self-avatar is the visual representation of the user's body inside a 3D environment (Lok et al., 2003). Although many researchers study users' virtual representation and control in VR, little attention has been paid to motor skills training implications. In motor skills training, such as assembling mechanical parts, instrument handling, or many surgical procedures, the avatar representation concerns the user's hands primarily. In fact, the hands are used for manipulating the surgical tools and obtaining spatial cues. In this line, previous researchers identified that the user's hand avatar might play a role during the training of motor skills (Sankaranarayanan et al., 2016; Van Nguyen et al., 2015). Therefore, one important aspect to consider when designing surgical simulators for technical skills is how surgeons visualize and control their fingers/hands. In the following sections, we present a summary of research on hand representation and control in virtual environments and the potential impact on motor skills tasks.

Hand representation in virtual environments

The users' body representation has been studied for some time. The rubber hand illusion, initially presented by Botvinick and Cohen (1998), is a simple experiment performed to understand how our brain resolves visual and perceptual stimuli, which lead to the appropriation of a rubber limb. Recent works have revisited this experiment (Blanke et al., 2015; Ehrsson, 2009) to explain how the visual and tactile cues influence the users' sense of embodiment. Kilteni et al. (2012b) define the sense of embodiment towards a body as the sense that emerges when that body's properties are processed as if they were properties of one's own body. It can be decomposed into three different dimensions: the sense of body *ownership* (feeling that the artificial body is one's own body and the source of sensations), the sense of *agency* (sense of having motor control), and *self-location* (perceived location space of one's body) (Kilteni et al., 2012a). It has also been shown that there is a high correlation between presence ("sense of being in another environment" (Biocca et al., 1997)) and embodiment (Usoh et al., 1999).

The rubber hand experiment has had a significant impact on understanding the virtual body, and many recent experiments are based on it. Mohler et al. (2010) conclude that if we experience a conflict between the visual and proprioceptive position of our hand, we will strongly accept it as being placed where it is seen. Argelaguet et al. (2016) have shown that in VR, the sense of agency improves with the increase of the virtual hand control capabilities, and the sense of ownership is related to the visual appearance of the virtual hand. González-Franco et al. (2014) investigated brain activity in response to pain observation when the user's virtual hand was attacked with a knife. The results allowed them to conclude that when users are in an immersive VE, they have body ownership towards their body representation, up to the point of experiencing observed events to take place in reality.

It has also been shown that the virtual hand's structural and appearance differences might affect the sense of ownership. Having an extra finger can be accepted without limiting its controllability and reports high levels of body ownership (Hoyet et al., 2016). However, having fewer fingers in a virtual hand reduces the feeling of presence for realistic hands, but not for abstract ones (Schwind et al., 2017a). Besides, human hands evoke higher body ownership than abstract representations (Lin and Jörg, 2016; Ogawa et al., 2019). However, inverting gender models can decrease the users' level of acceptance and presence (Schwind et al., 2017b). In particular, a personalized virtual hand (real hand projection in the virtual environment) improves object size estimation accuracy. It also increases the sense of ownership compared to a generic hand model (Jung et al., 2018).

Moreover, users' object-size perception in VR is affected by anthropomorphic hands' size, but not for non-anthropomorphic ones (Ogawa et al., 2019).

To summarize, several works have focused on studying the user's hand representation and its influence on the sense of presence and embodiment. Findings support the inclusion of virtual hands to enhance the user experience in a VR simulator. In addition, the virtual hands' appearance has been shown to have an impact on performance in immersive VR applications. However, it is necessary to understand the impact of including these hand representations and how to control them on motor skills tasks.

Impact of virtual hand representation and control on motor skill tasks

The user's avatar visualization can also influence motor skills. For instance, Ossmy and Mukamel (2017) studied the influence of hand's size on short-term motor skill learning of a finger movements sequence. They concluded that the hand's size affected the performance, observing that the performance increases with a larger virtual hand and with a hand with an almost 1 : 1 ratio size. However, these results were observed only when the user controlled the virtual hand and not when the user was playing a spectator role during training. This result highlights the importance of agency on motor skills performance, which also proved to be true for distance-estimation tasks performance (Mohler et al., 2010). Another example was found for a needle insertion VR trainer, where authors investigated the virtual hand's influence on the users' feeling of accuracy and sense of realism (Van Nguyen et al., 2015). This study suggests that having a static virtual hand increases the self-perception of accuracy but limited the environment's overall realism, pointing out that users' real hand posture and movement were missing. Moreover, it has been shown that visualization of limb movements while learning can improve the motor tasks (Castro-Alonso et al., 2014). This suggests that if users visualize their own hands' movements while performing the task, they could accelerate the motor skills learning process.

On the other hand, the self-avatar fidelity was analyzed for a block arrangement task during interaction with real objects (Lok et al., 2003). The results suggest that to increase the sense of presence inside the VE, the avatar's kinematic fidelity (hand movement) was more important than the avatar's visual fidelity. However, the avatar fidelity did not influence the users' performance for this task.

Moreover, performance on a pick and place task was evaluated for different hand representations in virtual grasping. Tracked hands representations (hands that can pass through the objects which the user interacts with) improved task completion time. Nonetheless, users preferred to have more realistic interactions where the virtual hand physics with the

virtual objects reflects a real interaction (i.e., the hand visualization does not penetrate the manipulated object) (Canales et al., 2019).

One crucial aspect to consider for representing the user's hands in the VE is the technology employed to control it. In general, we can differentiate two main approaches to capture the user's hand movements: through optical trackers, such as the Microsoft Kinect or the LeapMotion, or by using inertial trackers or data gloves (many commercial solutions available) (Dargar et al., 2015)

The advantages of vision-based systems, such as the LeapMotion, is that they are cheap and easy to use (plug and play). They are usually mounted on HMDs to track the user's hands during interactions in immersive VR systems. However, they may not be usable in all the applications. For instance, they may not be suited for detecting the hand holding a tool due to occlusions generated by the tool itself. On the other hand, data gloves are more expensive, require proper calibration and data filtering to obtain acceptable tracking performance. Depending on the model, and mostly on the price, some provide only one sensor for each finger blending detection, requiring a new solution for positional tracking of the hand. However, they offer the advantage of being robust to occlusion issues when the user manipulates a tool. This suggests that data gloves might be more appropriate for tracking the user's hands to support hand control in immersive VR environments when the task requires manipulating a tool or when occlusion issues may occur during interactions.

Finally, other works have used hand visualization in training motor skills, without assessing its influence on performance and skills transfer to the real world (Batmaz et al., 2018; Carlson et al., 2015; Pulijala et al., 2018; Sportillo et al., 2015).

To summarize, motor skills training is influenced by visual feedback in general, and particularly by the self-avatar representation. The users' hand representation and control are important design choices to enhance users' participation and state of presence in a VR simulation. They may influence motor skills performance and training in such systems. However, its impact on tool-based motor skills training in general and on technical surgical skills in particular remains an open question to investigate. This particular aspect is investigated in chapter 4.

1.9 Conclusions

In this chapter, we have introduced and defined the main concepts related to our object of study, virtual reality simulators for learning technical surgical skills. We have also presented an overview of existing surgical simulators, both research prototype trainers and commercially available solutions. Furthermore, considering the VR surgical simulator as an interactive system, we have described the communication process between the user and the system to extract the 3D interaction techniques needed to support VR simulation. Examples of different techniques used to support 3D interactions in VR systems and surgical simulators were presented.

In summary, beyond the main tool manipulation tasks, there are other interactions and factors worth being analyzed for certain procedures. For instance, navigation tasks are not the main focus of current VR surgical simulators, sometimes leaving just a fixed point of view or allowing the use of mice or tactile displays to change the virtual camera position. Moreover, their impact on users' performance on the main tasks remains unknown.

Furthermore, the user's avatar representation and control in surgical simulators, which concerns mainly the user's hands, is an issue that has not been studied in immersive surgical simulators. Again, the existing systems either do not include any hand avatar or include a static hand. The impact of these representations on users' performance on the main tasks also remains unknown.

The common inherent question of these identified issues is the level of fidelity of the interactions and artifacts to include in the simulators. In fact, the used navigation techniques range from metaphoric to naturalistic techniques. Moreover, the used hands' avatar representations range from impersonal and static representations to realistic and fully controlled virtual hands. However, the current research does not indicate what levels of fidelity a designer can use to include these components in a VR surgical simulator.

In order to give recommendations regarding the appropriate techniques to design for VR surgical simulators, we must specify the tasks to perform and analyze the possible impact of the fidelity of interaction techniques to implement them in the simulator. To proceed with this analysis, we will study the existing state of the art models of simulator fidelity. The next chapter summarizes simulator fidelity models and then discusses and analyzes the different interactions and components needed to improve users' performance and their importance for training technical skills in virtual reality simulators.

Chapter 2

Simulator fidelity

***

In the first chapter, we have presented a description of existing applications of virtual reality surgical simulations for technical skills training, as well as the design approaches that can be used to build VR simulations. Our analysis suggests that questions regarding the interactions' design and the fidelity level of these interactions are still open in order to design effective VR surgical simulators.

This chapter aims to present a review of the literature on the simulator-fidelity concepts focusing on the fidelity of interactions to identify the potential directions to follow for a proper design and evaluation of interactions for VR surgical simulators.

A simulator fidelity model, based on display, interaction and scenario fidelity components, is retained for this thesis work. A discussion of the different components, their application, and limitations are presented. Finally, propositions are made to use this model to evaluate two distinct aspects of surgical VR trainers, with a particular focus on the interaction fidelity component.

2.1 Transfer of skills and simulator fidelity

The development of virtual reality simulators for training skills is based on the assumption that the skill(s) acquired in those systems should be transferred to the real world. Therefore, it is essential to define the concept of transfer. Training transfer is defined as the degree to which learned skills or knowledge can be applied to another situation (Ragan et al., 2015). It is important to notice here that training in a simulator has value only if it permits to train the skills that are required in the operational environment (real-world). As discussed in chapter 1, the predictive validity evaluates this requirement by looking at the training transfer, i.e., assessing the performance of the corresponding real-world task (issue of the simulation) after the training on the simulator. This means that for surgical training, for instance, to fully assess the system efficiency, one needs to evaluate the performance in a real surgical procedure (i.e., in patients) after training in the simulator. However, as discussed earlier, this is not always possible due to the associated ethical and patients safety issues.

There is the belief that to design and build simulators that allow training transfer, the training environment and the real world should be closely matched (Hamblin, 2005). This introduces the notion of simulator fidelity. Taking into account training objectives, we can define simulator fidelity as the similarity between the knowledge taught in a simulator and the one used in the real-world environment (Stoffregen et al., 2003). However, it is important to note that the published literature has not yet revealed a direct relationship between the level of simulation fidelity and training effectiveness (Beaubien and Baker, 2004). There are, however, some works that point out that novices experience the most rapid skill acquisition with lower levels of fidelity, while experts require high-fidelity simulators to improve their skills (Arbogast and Rosen, 2012) (see Figure 2.1).

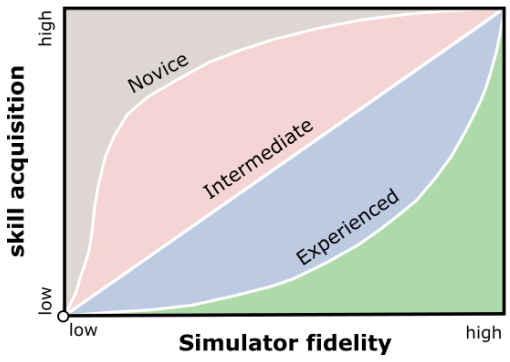


Fig. 2.1 The relationship between fidelity of simulator and acquisition of skills at three levels of expertise: novice, intermediate and experienced (Arbogast and Rosen, 2012)

Accepting that there is a strong relationship between the level of fidelity and the training effectiveness encourages the development of a virtual environment with the highest possible fidelity. Indeed, it would be reasonable to think that if both virtual and real systems cannot be distinguished, then the transfer of knowledge is guaranteed. However, *true fidelity*, conveying that reality is the highest possible fidelity, is unachievable. Indeed, the sensory stimulation in a simulator cannot be identical to that available in the system being simulated (Stoffregen et al., 2003). Moreover, high-fidelity VR simulators require technologies such as HMDs, high-precision 3D tracking systems, and advanced input and output devices that are quite expensive (Duncan, 2006). Furthermore, as previously mentioned, other researchers have shown that low-fidelity simulators may be enough for effective and efficient training, particularly for novices (Arbogast and Rosen, 2012; Chellali et al., 2016; Kim et al., 2003). Nonetheless, this raises some questions, such as how low can we go in terms of fidelity? What is important to include in a simulator?, and whether all aspects of the simulator need to be of equal fidelity. Finally, will less fidelity in some aspects affect the users' performance, training, and transfer of skills to the real situation?, and will less fidelity affect users' perception of the simulation?

Although many recent research contributions on the effects of simulator fidelity, it is not easy to generalize the results since the contexts, domains and tasks explored are numerous and very different. In addition, the use of the terminology may vary among different works, and some aspects defined may also overlap, which adds more variety and complexity to the topic. In the following section, we discuss the simulator fidelity literature, including definitions, models, and components. This will help us make some choices regarding the fidelity frameworks that can guide us in the design of VR surgical simulators.

2.2 Simulator fidelity definition and dimensions

Since the origin of the term "fidelity", many researchers have proposed different dimensions to evaluate their simulators. The variety of domains in which this term is employed and the different nature of the simulators proposed make it difficult to reach a classification which embraces all the dimensions found in the literature. In addition, depending on the application, different aspects of fidelity may have more or less relevance to the objectives and/or performance of the virtual reality simulator. Finally, no consensus has been reached to determine the exact role of simulator fidelity in medical training, and more importantly, which of these dimensions play an important role to ensure a positive

training and transfer of technical surgical skills. This section aims to examine the existing definitions proposed in the literature and summarize the contributions to this topic.

2.2.1 Simulator fidelity

We previously defined simulator fidelity in terms of training objectives as the similarity of the knowledge taught in a simulator to the one used in the real-world environment (Stoffregen et al., 2003). Most works suggest that there are two major aspects of simulator fidelity to consider. Notably, they distinguish between the *experiential fidelity* (fidelity of subjective experience) and the *action fidelity* (fidelity of performance) (Stoffregen et al., 2003). In the same line, Fuchs (2006) agrees with this classification, but uses the terms *perceptual fidelity* and *psychological fidelity* for experiential and action fidelity, respectively.

On the other hand, Gerathewohl (1969) defines fidelity in terms of the quality of the stimulus produced and not as a human perception of the simulation: "the degree to which a device accurately reproduces a specific effect" (Gerathewohl, 1969). Indeed, no consensus is reached to determine which type of "realism" is being addressed when discussing simulator fidelity. However, there is a general agreement on the need to describe, in a subtle and precise way, the relationship between the simulacrum generated by the virtual environment and aspects of the real world (Fuchs, 2006).

In the following section, we discuss the different models and dimensions of simulator fidelity presented in previous works.

2.2.2 Simulator fidelity dimensions

The first mention of a fidelity framework comes back to the early 1980s, with the works of Hays (1980) and Baum et al. (1982) for the aviation domain. In the following, we will present the most common simulator fidelity models found in the literature and discuss their dimensions and underlying concepts.

Experiential and action fidelity

The first model we will discuss was presented by Stoffregen et al. (2003) and distinguishes between two components of fidelity *experiential fidelity* and *action fidelity*.

Experiential/perceptual fidelity *Experiential fidelity* is defined as "the extent to which a simulation gives rise to a subjective experience of being there" (Stoffregen et al., 2003).

It is associated with the sense of presence and immersion.

Immersion can be objectively assessed, and it is defined as what technology exhibits from an objective point of view (Slater, 2003). So, the more immersive a system is, the most faithful the real-world sensory simulation becomes. On the other side, *presence* is a human reaction to a system with a certain level of immersion, the point at which simulated sensory data plus perceptual processing produces a coherent place in which one can be and act (Slater, 2003). A clear sign of presence is, for example, having the user behave in the virtual environment in the same or similar way to how he/she would behave in the analogous real situation.

Presence can be decomposed into three types:

- Environmental Presence: is determined by the extent to which the environment recognizes the user's existence and reacts to it.
- Social Presence: states that greater interaction of presence of others will lead to higher engagement of the individual with the simulation and the group.
- Personal Presence: describes the extent of, and reasons for, a person's feeling like he/she is "in" a virtual environment.

Slater (2003) stands that the presence in a real-life situation can be simulated by a virtual reality that delivers extremely poor sensory data in relation to physical reality. He remarks at the same time that the important issue is to use the knowledge of the perceptual system to determine what is really important in the representation of the virtual world. He also stands that involvement, interest, and emotion are aspects of an experience and do not influence the state of being present. However, they can be used to evaluate it. An example of this is comparing an emotional response in both the virtual environment and the similar real-world to determine if they are the same.

In order to measure the sense of presence of the user in a VE, two main methods are used (Schuemie et al., 2001):

> **Subjective measurements:** these measures are based on subjective ratings through post-experimental questionnaires, which allow users to express their feeling of presence in a VE. The advantage of using such metrics is that they give feedback on the user experience in the VE. This provides new opportunities to understand better the phenomena linked to presence. However, the questions are often dependent on a theory or definition of presence (Schuemie et al., 2001). This may limit the scope of the results of these questionnaires. Their validity is also continually questioned (Sanchez-Vives and Slater, 2005). They are also based on subjective feedback of the users that may be influenced by other factors. Nevertheless, this method is the most

commonly used in presence research. One example of questionnaires used to assess the sense of presence is the "Presence Questionnaire" (Witmer and Singer, 1998).

Objective measures: these measures can be divided into two categories:

- Behavioral measures: these measures include, for instance, reflex responses (e.g., avoiding an obstacle that arises suddenly in the VE) or the direction of gaze during the user interaction with the VE.
- Physiological measurements: these measures include, for instance, heart rate, skin temperature, and skin conductance, or even brain activity measured through electroencephalograms (EEG).

However, the use of these objective measurements remains limited and their validation involves comparisons with subjective measurements.

Action/psychological fidelity *Action fidelity* is defined "in terms of relations between performance in the simulator and performance in the simulated system" (Stoffregen et al., 2003). It exists when the performance in the simulator transfers to the real-world situation that is being simulated. A way of measuring the action fidelity is to measure the transfer of learning to the real task (when possible) or the variance in performance across trials.

As presented by Champney et al. (2014), *psychological fidelity* of a training simulation refers to the similarity in producing the essential underlying psychological process that takes place within the learner as experienced in the operational environment as well as the emotional cues present during the experience. It has the potential to enable the use of cost-effective, low fidelity simulations during training, which can maximize transfer in terms of retention and generalization (Kozlowski and DeShon, 2004a).

Interface and environment fidelity

Waller et al. (1998) defined simulator fidelity as "the extent to which the virtual environment experiences and interactions with it are indistinguishable from the participant's observations of and interactions with a real environment". This definition is based on a subjective evaluation of the simulator by users, which contemplates that there are three information domains in VE training: the real-world environment, the training environment, and the trainee's mental representation of the environment (Waller et al., 1998). In their work, the authors indicate that simulator fidelity can be decomposed into two psychological aspects of the simulation, the environmental fidelity, which mediates the mapping

between real-world and the training variables, and the interface fidelity, which mediates the mapping between the training environment and the user's mental representation of it.

They define *environmental fidelity* as the degree of resemblance between the training and the real-world variables of the environment. In addition, they define *interface fidelity* as the degree of similarity between the input and output devices functioning in the VE to the way in which the user would interact in the real world. Both definitions are presented as psychological concepts rather than objective correspondence between the VE and real-world.

Physical and functional fidelity

Pioneer simulator systems had focused on the *identical element principle* which corresponds to include the maximum common elements between the simulator and the operational environment (Champney et al., 2014), producing high levels of *physical fidelity*. In this approach, the equipment to use, the controls, the reactions and behaviors, and the simulators' look, feel and motion are built to be as realistic as possible.

An alternative to focusing on identical elements is to replicate the functions and behaviors from the real to the simulated environment. *Functional fidelity* defines the similarities in how the simulation behaves like the operational environment in reaction to tasks/actions (inputs) performed by the learner (who is engaged in the experience) (Champney et al., 2014).

However, this approach is technology-centered and does not consider the users' experience, i.e., how learners perceive and behave during training in the simulation compared to the real world.

Display, interaction and scenario/simulation fidelity

McMahan et al. propose to evaluate fidelity through three dimensions: *interaction fidelity*, *display fidelity* and *scenario fidelity* (Bowman and McMahan, 2007; McMahan, 2011; Ragan et al., 2015).

They define *display fidelity* as the objective degree of exactness with which real-world sensory stimuli are reproduced by a display system (i.e., the verisimilitude of the displayed output), commonly referred to as immersion (Bowman and McMahan, 2007; Ragan et al., 2015).

Moreover, *interaction fidelity* is defined as the objective degree of exactness with which real-world interactions are reproduced in an interactive system (i.e., the realism of the input devices and interpretation software) (Ragan et al., 2015). McMahan (2011) has

proposed the FIFA (*Framework for Interaction Fidelity Analysis*) framework to objectively classify interaction techniques in terms of interaction fidelity.

Finally, *Scenario fidelity* is defined as the objective degree of exactness with which behaviors, rules and object properties are reproduced in a simulation as compared to the real/intended experience (i.e., the realism of the simulated scenario and the associated model data) (Ragan et al., 2015). Nilsson (2016) rejoins this definition, but uses the word *simulation* instead of *scenario* fidelity. He defines *simulation fidelity* as the objective degree of exactness with which real-world physics and characteristics are reproduced in a simulation (i.e., the realism of the models on which the virtual world is generated, e.g., geometric, lighting, or physical models).

A more in-depth discussion of this model will be presented later in section 2.3

Buy-in/user acceptance

Another concept that appears in the literature review is the *buy-in* or user acceptance (Alexander et al., 2005), which refers to the degree to which a person recognizes that an experience or event is useful for training and may depend on the user identification with avatars and/or the whole simulated environment.

2.2.3 Synthesis

Table 2.1 summarizes our literature review on simulator fidelity. Since the appearance of the simulator fidelity concept, many researchers have contributed to this topic. However, only a few have proposed practical guidelines or frameworks to design and evaluate virtual reality simulators from a fidelity perspective.

In addition, most of the reviewed works are based upon controlled evaluations of fidelity, which means that they usually perform a direct comparison of similar systems while keeping under control one or more components of fidelity. This allows determining their effects and the possible interactions between them. Some other evaluations are performed comparing high-fidelity simulators with *true fidelity* (when this is possible), conveying that reality is the highest possible fidelity (see Table 2.1).

To conclude, simulator fidelity has recently gained importance in the design of virtual reality simulators. However, none of the existing classifications and proposed frameworks have been developed explicitly for the surgical domain.

Previous research has shown that interaction fidelity plays a central role in transferring knowledge from a virtual reality trainer to the real world (Chellali et al., 2016; Drews and Bakdash, 2013). It is affected by the ease of interaction and the level of user control of

the system (Hamblin, 2005). This suggests that building efficient virtual reality training systems depends on designing an effective interaction for these systems. However, as pointed out in chapter 1, there is currently a lack of guidelines to determine the levels of fidelity of interaction techniques during the design of VR surgical simulators. The previous review permits identifying existing models of simulator fidelity that can potentially be adapted to tackle this issue.

The model we have chosen to use in this thesis work presents the simulator fidelity as the compound of three dimensions: interaction, display and scenario fidelity. We chose this model because it is a detailed framework which proposes an interaction-design-based classification of several components of a VR simulator in terms of fidelity. Moreover, this framework was extensively evaluated through user studies. Indeed, some of the existing components of this model can be generally applied to compare objectively two setups, such as the FIFA (*Framework for Interaction Fidelity Analysis*) framework for interaction fidelity or the visual display framework. Applying this framework in the surgical domain will help us determine the fidelity components that are most important for the related simulator, and specifically, which is the level of fidelity expected for each one. More particularly, this framework will be useful to investigate the levels of fidelity of interaction techniques and their impact on user training performance. It will be combined with subjective measures, such as presence questionnaires, to assess the user's subjective experience and sense of presence. However, further reflections might be required for assessing the fidelity of all the aspects involved in the simulation.

In the next section, we discuss this model and its components more in-depth from an interaction design perspective.

Table 2.1 Review of simulator fidelity works

Authors	Description / Objectives	Application Domain	Task	Simulator	Fidelity Dimensions	Evaluation	Framework/ Formulation
Waller et al. (1998, 2001)	Examine the variables in transfer of spatial knowledge using fidelity, in spatial representations for VE training	training	navigation	VR	environment, interface	user studies	interface & environment fidelity definitions
Slater (2003)	Addresses the confounding of the term presence with several different distinct aspects of experience	NA	NA	NA	immersion, presence	NA	presence & immersion definitions
Maran and Glavin (2003)	Review of currently available simulators. The use of different levels of simulation in a continuum of training is discussed.	medical	NA	physical, manikin, computer-based	physical, psychological / functional	NA	–
Stoffregen et al. (2003)	Present simulator fidelity analysis	flight simulation	NA	VR	stimulus experiential action	NA	stimulus experiential action definitions
Kozlowski and DeShon (2004b)	Present a theoretically based strategy for training research and design that focuses on psychological fidelity	training	NA	NA	physical, psychological	NA	psychological fidelity approach
Alexander et al. (2005)	Examine the potential of technology developed for video games and MMPGs to be used in training	gaming & military training	NA	computer-based	physical, psychological, immersion, presence, buy-in	NA	–
Hamblin (2005)	Explore a variety of individual differences known to affect HCI and their potential effect on transfer of training from VEs	training	manually assembly	VR	environment, interface	controlled experiments	–

Johnsen et al. (2005)	Development of an immersive virtual patient system and discussions on the feedback from a pilot group of students	medical	diagnosis	VR	presence, immersion	pilot study	–
Fuchs (2006)	Present simulator fidelity definitions	NA	NA	VR	perceptive, psychological	NA	perceptive psychological fidelity definitions
Bowman and McMahan (2007)	Understand how the various components of immersion affect measurable user performance, understanding, and preference in a wide variety of VEs	military training & entertainment	visualization, manipulation, search, spatial understanding	VR	immersion, display	user studies	display fidelity framework
Lee and Kim (2008)	Description of a haptic teleoperation system of a mobile robot and experimental study results	robotics teleoperation	remote navigation	VR	presence, display	user studies	–
Backlund et al. (2009)	Extend the concept of game-based training systems by adding multiple screen view and a novel interaction mode	training & gaming	search	VR	immersion, physical, functional, psychological	questionnaires	–
Cornacchione Jr (2012)	Explores educational technology and management education by analyzing fidelity	gaming & education	NA	VR	physical, psychological, functional	comparison studies	–
Lee et al. (2013)	Investigate the validity of Mixed Reality Simulation studying visual realism effects on search tasks	NA	search	VR (AR)	display, interaction, scenario, simulation	user studies	–
Champney et al. (2008, 2014)	Present a human experience approach to simulator fidelity optimization, incorporating required cues for functional and psychological fidelity	training medicine	combat	VR	functional, psychological, physical	NA	SPOT taxonomy (physical) extended for functional and psychological fidelity

McMahan et al. (2010, 2012) McMahan (2011) Ragan et al. (2015, 2013)	Development and Evaluation of VR environments for Gaming and Military domains. Theoretical framework definition and application to those environments	gaming & military training locomotion, targeting spatial judgment	selection, manipulation, navigation,	VR	display, interaction, simulation	user studies	FIFA framework for interaction, display and scenario fidelity
Bertrand et al. (2015)	Describe and evaluate an immersive VR simulation employing bimanual interaction tasks	education & training	bimanual interaction	VR	interaction	user studies	application of FIFA
Benyahia et al. (2015)	Validate and improve the interface and environment fidelity of the system	surgery	biopsy	VR	interface, environment	user studies	–
Nilsson (2016)	Study of the factors influencing the degree of perceived naturalness of Walking-in-Place locomotion	NA	navigation	VR	display, interaction, simulation	NA	application of FIFA

2.3 User-interaction design approach for simulator fidelity

In order to better understand the fidelity dimensions of a simulator included in McMahan and Herrera (2016) model, we revisit here the human-system communication model presented in the previous chapter (see Figure 1.6).

Based on the flow of information presented on this model (see Figure 2.2), we can include the three components of fidelity presented by McMahan and Herrera (2016), which were partially explored in the literature for other domains, such as gaming and military training:

- Display fidelity
- Scenario fidelity
- Interaction fidelity

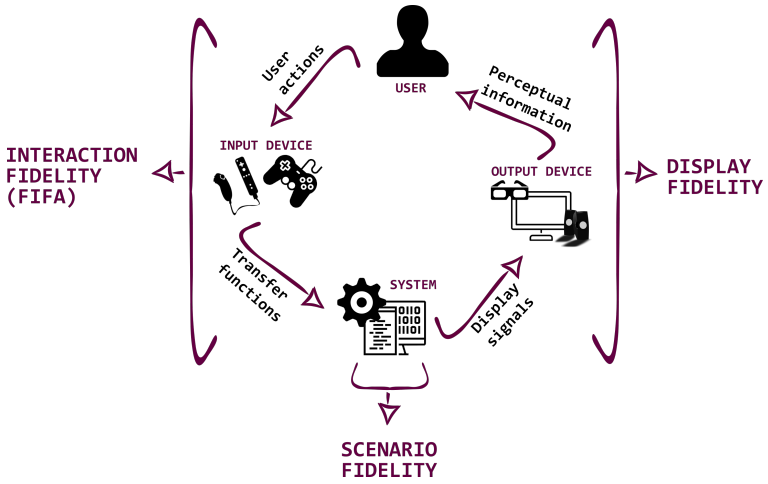


Fig. 2.2 Human-system communication design and FIFA components

2.3.1 Display fidelity

Display fidelity is related to the different stimuli a human being can perceive. Therefore, it includes visual display fidelity, as the stimuli perceived by human vision, haptic display fidelity as perceived by the kinesthetic human system, audio display fidelity as perceived by human hearing. An approach to characterize visual display fidelity is presented by Bowman and McMahan (2007). It considers components like the field of view, the display size and the resolution, the frame rate, and stereoscopy (see Table 2.2). An analogous

reasoning can be done for the rest of the dimensions: auditory (hearing), haptic (touch), olfactory (smell), and gustatory (taste), e.g., for haptic display fidelity, we can consider components like the force feedback degree of freedom, the force range, the torque range, the force resolution, and the workspace. However, only the visual display fidelity component has been discussed so far in the literature. Haptic fidelity, which is an important dimension for surgical simulation, was not explored. Since this was identified in chapter 1 as an essential component for a VR surgical simulator, an extension will be necessary in order to use this model for the analysis of this component.

Table 2.2 Visual display fidelity framework

Component	Definition
Stereoscopy	The display of different images to each eye to provide an additional depth cue
Field of View (FOV)	The size of the visual field (in degrees of visual angle) that can be viewed instantaneously by the user
Field of Regard (FOR)	The total size of the visual field (in degrees of visual angle) surrounding the user
Display Resolution	The total pixels displayed on the screen or surface
Display Size	The physical dimensions of the display screen or surface
Refresh Rate	How often the display draws provided rendered data
Frame Rate	How often rendered data is provided to the display

2.3.2 Scenario fidelity

Scenario fidelity has not been fully explored, and it embraces many components, which range from models, objects properties, physics, VE to 3D effects, specialized algorithms and artificial intelligence. Some of these components can be of crucial importance for some aspects of a VR surgical simulator if we intend to replicate the real-world experience. Therefore, this component will require a more in-depth analysis before being used. However, this component is considered as out of the scope of this thesis work.

2.3.3 Interaction fidelity

We will more particularly focus on this component in this work.

McMahan (2011) has proposed the framework of interaction fidelity analysis (FIFA), a theoretical framework for the study of the interactions fidelity aspects in the fields of gaming and military simulation. This framework has been used to objectively characterize and evaluate different components of the fidelity of the interaction identified by the author (McMahan et al., 2010, 2012). However, it has never been applied to the specific field of VR surgical simulation.

The *FIFA* framework decomposes interaction fidelity into three main components: *bio-mechanical Symmetry*, *control symmetry* and *system appropriateness*.

Table 2.3 FIFA framework (McMahan, 2011)

Biomechanical Symmetry	
The objective degree of exactness with which real world body movements for a task can be reproduced through the interaction to successfully complete the task. Basically, it is the amount of correspondence between the movements of the body realized during interaction, and those that would be made in the same real task	
Kinematic symmetry	The objective degree of exactness with which a real-world body motion for a task can be reproduced
Kinetic symmetry	The objective degree of exactness with which a force involved in a real-world movement can be reproduced
Anthropometric Symmetry	The objective degree of exactness with which the body segments involved in a real-world action can be matched by the body segments involved in an interaction that successfully completes the task
Control symmetry	
The objective degree of exactness with which control in a real-world task is provided through interaction. Basically, it is the amount of correspondence between the control provided by an interaction technique, and the possible control in the real world	
Dimensional Symmetry	The objective degree of exactness with which control dimensions in a real-world task are provided through interaction
Transfer Function Symmetry	The objective degree of exactness with which a real-world transfer function is reproduced through interaction
Termination Symmetry	The objective degree of exactness with which the termination of a real-world interaction is reproduced
System Appropriateness	
The other factors that characterize how suitable the system is for implementing a particular aspect of interaction	
Input Accuracy	The degree to which the registered values represent the "true" values
Input Precision	The degree to which repeated measurements in static conditions yield the same results
Latency	The temporal delay between user input and feedback generated by the system in response to it
Form Factor	The shape and "size" of the input device used

The biomechanical symmetry component describes the matching between the body movements made during the interaction in the simulator and those necessary to perform the same task in the real world. The control symmetry component represents the corre-

spondence between the control provided by an interaction technique and the possible control in the real world. Finally, the system appropriateness component is used to describe the other factors that characterize how suitable the system is for implementing a particular aspect of the interaction. See Table 2.3 for a complete definition of each of this components.

However, it is currently unclear whether this framework can be used to determine the appropriate levels of interaction fidelity in virtual surgical simulators. In the following section, we will present a concrete example of how this framework can be used in the specific case of a VR surgical simulator.

2.3.4 A practical example

As a way to better explain the concepts and fidelity components presented in this model (display, interaction and scenario fidelity), we present here an interpretation of each of them through examples of a VR simulator for tissue dissection using a scalpel.

Analysis of the display fidelity of the system

The display fidelity concerns the degree of realism of the used rendering technology. For visual display fidelity, we can opt for a low level of fidelity by providing a desktop monitor screen to visualize the VE in a 2D view, or a higher fidelity level by providing a HMD device, allowing the user to see the VE in 3D. In fact, both solutions can provide stereoscopy view (i.e., rendering images to each eye to provide depth perception), but the HMD provides higher degrees of FOV and FOR. Nonetheless, a higher-level of display fidelity does not guarantee better effectiveness of the system. Hence it would be necessary to test which of these components influence the users' performance in the simulator to decide upon the best solution to render the visual aspects of the simulation. For haptic display, we can choose to provide haptic rendering through a haptic desktop device. Depending on the characteristics of this device, we can provide different levels of fidelity. For instance, we can provide a high-fidelity solution by providing 6 DOF for force feedback at a 1KHz update rate, or just 3 DOF for positional forces. This analysis can be extended to include and characterize all the rendering modalities being used. This analysis will then permit to determine the level of display fidelity of the whole system. After that, we can then carry out comparative user studies and analyze each component's influence on user performance and user perception of the simulated stimulus to adequately choose the interface to use for each rendering modality that better suits the training objectives of the surgical simulator.

Analysis of the interaction fidelity of the system

Let us consider an interaction such as the manipulation of the virtual scalpel in this simulator. We can say that if the user utilizes a classic Xbox joystick to manipulate the scalpel (for example, by using the right stick to move in the X and Z axis and the up/down directional pad to move on the Y-axis), we have a lower level of interaction fidelity as compared to the user manipulation through the stylus of a haptic device (e.g., freely moving the stylus in a 3D space and matching its position to that of the virtual scalpel). This classification is made through the application of the interaction fidelity framework. In fact, the joystick interaction technique has a lower biomechanical symmetry since the movements done by the user in this interaction technique (i.e., fingers moving the stick and pressing buttons) do not correspond to the ones expected in the real-world task (i.e., moving the upper arm, forearm, hand and fingers to manipulate the scalpel). On the other hand, interaction through the haptic device has a higher biomechanical symmetry since the movements to manipulate the stylus can be identical to those of the real scalpel manipulation. Moreover, the transfer function places the haptic-based interaction technique as a more realistic one by providing a $1:1$ transfer of the position (i.e., one millimeter moved in the real space is translated to one millimeter of movement in the VE).

Analysis of scenario fidelity

We can also describe different scenario fidelity aspects of the VE. For instance, we can provide basic 3D models (e.g., rigid cubes and spheres) to represent the tissue and organs, or more realistic ones with complex forms that deform on contact with the instrument. In addition, we can provide realistic fluid models for blood and patient's respiratory movements. The deformation of the different patient's organs and tissues and how they respond to interactions with instruments are the focus of many research works. As mentioned earlier, this component is out of the scope of this thesis project.

Conclusions on simulator fidelity and existing models

We have presented so far an overview of the existing works and theories in the field of simulator fidelity. McMahan (2011) model that includes the FIFA framework was discussed as the most suitable approach to examine the fidelity of interactions to be included in the VR surgical simulators. However, this model suffers from some limitations. Indeed, it does not provide a framework for analyzing some components, such as the haptic display

fidelity, and it was not originally designed specifically for the surgical domain. Considering these limitations, our first contribution in this work was to extend this model to iteratively incorporate and evaluate the suitable components for developing virtual reality surgical simulators. We discuss this in the following section.

2.4 Towards a simulator fidelity framework for surgical VR simulators

Our literature review in chapter 1 suggests that the definition of effective interactions for virtual reality medical simulators may impact their usability and effectiveness for training purposes. We have subsequently identified three main interaction aspects of a virtual reality simulator that may play a role in the training of technical skills:

- **Changing the virtual camera's point of view** (*navigation*): In this thesis work, we have investigated this aspect for a VR biopsy simulator, where the main task to train on the system is the needle insertion inside a virtual body. The navigation around the operative field was considered as a secondary interactive task. The main objective was to determine whether the viewpoint changing technique's interaction fidelity, although considered a secondary task, impacts the user's primary task performance in the biopsy simulator. We have investigated this question by designing and comparing two navigation techniques used to change the user's viewpoint. These techniques were characterized in terms of their levels of interaction fidelity by applying the FIFA framework. They were designed following an iterative approach, which consisted of three steps: an in-depth analysis of real biopsy procedures with a focus on navigation tasks, the design of interaction techniques, and evaluation studies. In chapter 3, we describe the biopsy task analysis and then present the design and evaluation of the navigation techniques through two iterative experimental studies.

- **The avatar control and representation:** One important aspect observed during the analysis of real surgery procedures is that surgeons visualize their fingers when manipulating the surgical tools. Hand and finger movements are important for performing motor tasks. This raises the question of whether it is important for the trainees to see their hands and their movements when simulating the manipulation of tools in VR. Unfortunately, our review of exiting simulator fidelity models did not provide us with a framework for characterizing the degree of fidelity for users' self-avatar representation and control. Therefore, in the following section, we propose

and discuss a model for self-avatar representation fidelity, which will be used to characterize the avatar fidelity of the users' hands representation in a bimanual tool-based motor skills trainer. This is used then to determine the influence of the self-avatar fidelity on users' performance, learning outcomes, and transfer of skills to the real procedure in a VR trainer for a pick and place task simulating basic technical skills in surgery. This topic is investigated in chapter 4.

- **The haptic display fidelity:** No existing framework allows us to characterize haptic display fidelity in order to deduce what is required for haptic feedback in surgical simulators. The analysis of this topic is out of the scope of this thesis work. Nonetheless, we propose in the next section a haptic display fidelity framework that can be used in the future to objectively evaluate the user interfaces to use for haptic rendering, in a similar way to the visual display fidelity framework (Bowman and McMahan, 2007). By following the same approach, (analysis of the literature, design of a user study, implementation of a prototype, and evaluation of the research hypothesis), designers can determine, in the future, the appropriate levels of haptic display fidelity to include in their surgical VR simulators.

2.4.1 Self-avatar representation and control fidelity

As presented in chapter 1, the importance of the self-avatar representation and control has been investigated for some time, showing that there is a strong correlation between the feeling of presence and the degree of association of the users with their virtual representation and movements (i.e., sense of embodiment). In addition, some studies show that the users' performance in VR simulators can also benefit from having a self-avatar representation. However, the current FIFA framework does not include a tool to evaluate the self-avatar fidelity, and there is no framework proposition for characterizing it. For surgical VR simulators, the users use their hands to manipulate different tools while performing technical skills. Therefore, the self-avatar in this context concerns the users' hands.

In this section, we propose a framework for self-avatar fidelity of the users' hands, adapted from our research on avatar representation fidelity (Gamelin et al., 2020).

Three distinct components were identified for characterizing the self-avatar fidelity of virtual hands: the *visualization perspective fidelity*, the *kinematics fidelity*, and the *appearance fidelity*.

Visualization perspective fidelity The first characteristic concerning the self-avatar is the perspective from which the users visualize their virtual hands in the simulator (i.e., point of view). This dimension presents two levels:

- A *first-person perspective*, where the user visualizes the virtual hands co-located with their real hands. This level of point of view usually requires an immersive display for "hiding" the users' real hand (e.g., HMD).
- A *third-person perspective*, where the user visualizes the virtual hands at an external position, not co-located with their own hands, as if the virtual hands were the puppets of the real hands.

Intuitively, we can consider that a first-person perspective represents a higher fidelity metaphor for the users' point of view in self-avatar visualization. Previous works have investigated the difference between first- and third-person perspectives for a full-body avatar and their impact on the sense of ownership (Debarba et al., 2015; Fribourg et al., 2020; Gorisse et al., 2017). Additional studies are required to analyze this dimension for virtual hands.

Kinematics fidelity This dimension considers the dynamics of the self-avatar. It is defined as the degree of exactness with which the avatar replicates the movements of the represented user (Gamelin et al., 2020). In this line, we can analyze all the DOF of the hands' movements, i.e., hand position and rotation, as well as fingers' flexion and spreading.

Moreover, for training motor skills, we can also consider the level of relevance of the control capabilities with the real movement we want to train, in other words, the minimal control functions required in the real-world task. For example, we should provide the user with control capabilities for the hand position and rotation rather than individual fingers' movements for using a scalpel in a surgical procedure.

Finally, kinematics fidelity is also characterized by the synchronicity of the movement. For example, we can have low levels of synchronicity with virtual hands that "follow" the real ones by providing animated transitions between different goal posture states or high-fidelity ones with hands that produce real-time movements.

Appearance fidelity This dimension refers to the visual aspects of the self-avatar, regardless of the motion capabilities. We can define it through three different components: the *appropriateness of the morphology*, the *photorealism* and the *visual identity*. These

components were previously defined for full-body avatars in the context of collaborative virtual environments (Gamelin et al., 2020).

- The *appropriateness of the morphology* describes how close the selected morphology matches that of the real entity or how realistic the morphology is represented (Gamelin et al., 2020). Different degrees of fidelity can be established for the hands' morphology's appropriateness, ranging from the lowest level representation through a basic element (e.g., a sphere) to a full human-like morphology (i.e., five articulated fingers connected to a main body).

- The *photorealism* describes the realism of the representation as a generic human hand. An example of a low-level model could be a cartoonish hand, whereas a higher level contemplates, for instance, having a realistic rendering with tissue details, nails and hair.

- The *visual identity* concerns the individual characteristics of the users' hands, which allow them to recognize the virtual hands as their own. It can concern, among others, the complexity (e.g., thin or wider hands representation), skin colors, scars and wrinkles, tattoos, as well as jewelry worn.

Previous studies (as presented in chapter 1) have compared different degrees of morphology (Argelaguet et al., 2016; Giraud et al., 2013), visual realism (Lok et al., 2003) and visual identities (Jung et al., 2018) for virtual hands in VR simulators. Nonetheless, further studies are needed for the specific domain of surgical simulation in VR.

To summarize, we propose to extend the fidelity framework to contemplate the self-avatar representation and control fidelity in VR simulators, specifically for hands representation, which was an important aspect to consider during the design of VR surgical simulators (see chapter 1). Previous works have studied the effects of self-avatar fidelity in users' state of presence, ownership and agency for VR simulators. However, none of them focus on how the self-avatar's fidelity might impact surgical skills training. We investigate this issue for a tool-based motor task in chapter 4.

2.4.2 Haptic display fidelity

Humans interact with their surroundings through the five sensory channels: visual, auditory, olfactory, gustatory and haptic. The haptic channel enables manipulation and active exploration of the physical world, allowing the perception of some features like hardness, roughness (friction or smooth), texture, shape, and weight. Haptic interaction consists then of perception and manipulation (see Figure 2.3).

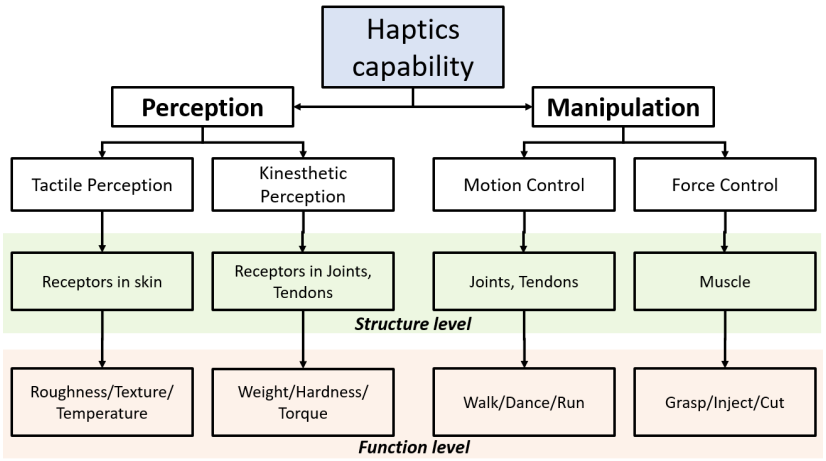


Fig. 2.3 Classification of the human haptic perception and manipulation capabilities (Wang et al., 2014)

As mentioned in chapter 1, a haptic interface is a feedback device able to produce a sensation to the user's skin or muscles, like the sense of touch, rigidity or weight. It can be classified into two main categories: force-feedback devices and tactile feedback devices (perception of pressure, vibration, and texture). Haptic rendering is "the process of computing the force required by contact with virtual objects based on the operator's motion measurements" (Hannaford and Okamura, 2008). The computational process of haptic rendering can be seen in Figure 2.4. Indeed, the realism that can be achieved for the sense of touch depends on the device's capabilities in terms of positional sensing and force feedback DOF. To characterize the haptic display fidelity, we focused on the interface rather than the underlying force models and algorithms used for the simulation, which relate to the scenario fidelity dimension.

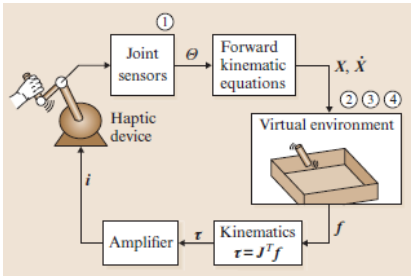


Fig. 2.4 Schematic diagram of the haptic rendering cycle (Hannaford and Okamura, 2008): Virtual object moves in the virtual environment according to the user's movements of the haptic device. Joint displacements are sensed in the device (1), processed through collision detection (2), surface points (3), and force calculation (4)

Haptic display fidelity framework

As was previously presented by (Bowman and McMahan, 2007), we can characterize the visual rendering output through objective components. In the same way, we propose to classify haptic rendering systems in terms of their characteristics. This way, we can enumerate several objective components that allow us to determine the degrees of realism we can have in terms of haptic feedback (see Table 2.4).

Table 2.4 Haptic display fidelity framework

Component	Definition
Sensed Degree of Freedom	The total degrees of freedom sensed for the device movement
Force Feedback Degree of Freedom	The total degrees of force feedback
Force Range	The upper and lower limits of force feedback
Torque Range	The upper and lower limits of the torque
Force resolution	Which is the resolution of the force allowed to be applied
Frame Rate	Or Update Rate. The number of times per second that a force is updated
Workable Workspace Size	The physical dimensions of the workable workspace for the device movement

Some of these characteristics might have a more significant influence on surgical training. We hypothesize that the DOF for force feedback is one of the most critical components to analyze. Indeed, depending on the surgical procedure intended to simulate, using only positional forces might be enough and has been used for several surgical trainers (e.g., needle insertion and MIS). Whether fully-actuated 6 DOF for force feedback devices (i.e., including torques) can significantly impact the users' performance and transfer of skills remains an open research question. In this line, Forsslund et al. (2013) studied the influence of the degree of freedom for force feedback on users' task performance for an instrument positioning task in virtual reality. They conclude that 6 DOF haptic rendering significantly improved task performance over 3 DOF haptic rendering. Moreover, other aspects, such as the workspace available and the frame rate provided for haptic rendering are worth being explored.

To summarize, the inclusion of the sense of touch in virtual environments is currently used to complete the user's perception. Nowadays, there are many devices with great capabilities to generate force feedback. By characterizing those capabilities (Table 2.4), we can then investigate which interface and properties are needed for a specific VR surgical simulator. Although this model's justified relevance for the design of VR surgical trainers, its validation requires several studies, which are considered out of this thesis's scope. Therefore, future works need to be conducted for its validation.

2.5 Conclusions

Our analysis in chapter 1 suggests that the definition of effective interactions for virtual reality medical simulators could have an impact on their usability and effectiveness for training purposes. From a technological point of view, there are three main general characteristics that we can extract from the virtual reality surgical simulators studied. First, a means for visualizing the working site is provided by a screen monitor or HMD, sometimes with a stereoscopic view, with one or simultaneous fixed views of the virtual environment, or means for changing the users' point of view in real-time. Second, the virtual working site can include tissue models, bone representations, or merely abstract objects, as well as the tools used to interact with them. Finally, they include the use of one or two physical tools, usually interfaced with haptic devices, which are manipulated by the users' hands. They serve to control the virtual tools and provide the necessary haptic feedback from their interaction with the other virtual objects.

From a user-centered point of view, surgical simulation can be seen as a traditional 3D interactive system, where interactions can be described through three essential tasks. First, the manipulation of tools, generating kinesthetic and force feedback from collisions of the tools with the environment elements. Second, the navigation, where the user moves his head to visualize the surgical site from different perspectives. Finally, we can add the avatar control as the user's control of his body during manipulation. In this particular case (surgical technical skills training), the avatar concerns the user's hands. The first task corresponds to the simulation's primary task, and the other two to secondary ones. In this thesis, we focus only on the two secondary tasks. Therefore, if we want to design an efficient simulator, we need to determine how realistic these tasks need to be simulated, which relates to simulator fidelity.

In this chapter, we have introduced and defined the main concepts related to simulator fidelity in VR. Indeed, the fidelity of a simulator is strongly associated with the transfer of skills to the real world. Hence it is important to base the design of surgical VR simulators on this concept. We have also presented an overview of existing fidelity models and research studies that contribute to this topic. The variety of domains in which this term is employed and the different nature of the simulators make it difficult to reach a classification that embraces all the dimensions found in the literature. Nonetheless, we have chosen one of the existent works, which already proposes a three-dimensional model to characterize different simulator interactions and components in terms of display, interaction, and scenario fidelity. However, it was never used for the surgical domain.

From the main interaction aspects identified above that could play a role in technical skills training, the navigation inside the VE and the manipulation tasks can be analyzed through the interaction fidelity framework dimension (FIFA). Indeed, the current chapter sheds light on the important role the interaction fidelity component plays in the transfer of knowledge from a virtual reality trainer to the real world. This suggests that higher fidelity interactions would improve the efficiency of VR surgical simulators for learning technical skills. However, it is not clear whether high interaction fidelity is necessary for all the interaction aspects. In this work, we do not focus on the manipulation task. We hypothesize that as being the main task of the simulation, it must be simulated in its highest level of fidelity in terms of interaction, including the same degrees of freedom for the tool movement and the same shape of the tool. Navigation, more particularly, how to change the point of view in the VR simulator, is explored in chapter 3. The focus is made on interaction techniques for supporting the needle insertion task in a VR biopsy simulator. We investigate the impact of interaction fidelity viewpoint changing techniques in users' performance in a designed VR biopsy simulator.

The second main aspect identified, which is how the environment is visually rendered to the user and the haptic feedback rendering, can be analyzed through the display fidelity dimension. The existing fidelity framework for this dimension focuses only on how to characterize the visual display fidelity. However, other stimuli, such as haptic display, are more relevant for surgical VR simulators. In this chapter, we proposed a framework to characterize the levels of haptic display fidelity, which contributes to the existing framework. In this work, we do not focus on studying the impact of haptic fidelity on training. Nonetheless, this framework extension could be useful in the future to determine which are the essential characteristics for haptic rendering in VR to guarantee effective training and transfer the skills to the real-world procedure. We do not focus on the scenario fidelity dimension either. Concerning this dimension for surgical simulators, we can move from lower levels of object models to more realistic, deformable ones. Moreover, higher expressions of scenario fidelity might include specialized algorithms for fluids and blood. As presented in this chapter, this dimension has not been thoroughly explored for VR simulation in general, nor for surgical simulators either.

Finally, the avatar control and representation in VR simulators are not included in the fidelity framework formulation we chose to focus on (interaction, display and scenario fidelity). Therefore, we extended this model with a self-avatar framework to characterize the users' virtual representation's degree of fidelity. The avatar control and representation is the second component we focus on in this thesis work. We study how the users' self-avatar representation's fidelity can influence the learning of a tool-based motor-skill task

in a VR simulator in chapter 4. For that purpose, we use the new model we have proposed in this chapter to characterize the level of avatar fidelity for users' hands in VR simulators.

To summarize, two components were selected from surgical VR simulators to study: viewpoint changing task and self-avatar representation and control. We study how some degrees of fidelity for these components can influence users' performance during surgical technical skills training, see chapter 3 and chapter 4 respectively for viewpoint change and avatar.

Chapter 3

Head movement

***

In the previous chapters, we have identified interaction fidelity as a main aspect to consider during VR surgical simulators' design for technical skills training. More particularly, the correct navigation inside the virtual environment remains unexplored for surgical simulators.

In this chapter, we will investigate the design of this interaction component for a virtual reality surgical simulator in order to improve its effectiveness. More particularly, we will focus on how the users can efficiently change their point of view during a needle insertion task in a virtual reality biopsy simulator.

After a detailed analysis of the navigation task during biopsy, we identified the key points of view used by the physicians when performing this task. Based on this analysis, two techniques to change the VE camera's point of view are proposed and compared using the state-of-the-art interaction fidelity framework (FIFA) (McMahan, 2011) to categorize their levels of interaction fidelity. The first technique is based on tracking the user's head, while the second one is based on touch interactions. They were evaluated through two user studies to determine which technique is more suitable for performing the needle insertion task, improving users' task performance, and which is preferred by the users.

The chapter is structured as follows. First, we present a brief state of the art on biopsy VR simulators to lead to our research problem for the first part of this work. We then describe the two iterations carried out (preliminary and main study) to design the needle insertion trainer and interaction techniques. Finally, we present the user studies conducted to compare these interaction techniques and discuss the main results and conclusions.

3.1 Biopsy and virtual reality simulators

Biopsy consists of inserting a specific needle into the human body to reach a target tissue (e.g., a tumor). This procedure is usually performed with limited real-time visual feedback to guide the manipulation of the needle (Chellali et al., 2012). Therefore, the accuracy of this task requires high haptic and three-dimensional spatial abilities.

The needle insertion task has already been simulated through VR technologies with a focus on haptic feedback simulation. For example, Gerovich et al. (2004) have developed a needle insertion simulator where the user could see in 2D a sample of a four-layer tissue: skin, fat, muscles and bones. Using a haptic device, users could feel force feedback when the needle passes through the tissue layers. Shin et al. (2011) have presented a needle insertion simulator with a haptic device. However, no visual feedback experience was included in this system. Other simulators have also been developed for training different needle insertion procedures (Corrêa et al., 2014; Henshall et al., 2015; Sutherland et al., 2013; Torres et al., 2012). The reader can refer to Corrêa et al. (2018) recent literature review on needle insertion simulators. However, the absence of any widespread adoption of these systems suggests that more work is needed to improve their design.

3.2 Research problematic

The previous review and the state of the art on the navigation task in VE presented in chapter 1 suggests that the existing needle insertion simulators and studies focus on the design of the main operative task to train. However, little is known about the impact of secondary tasks, such as how to change the point of view, on the user's performance. In this chapter, we focus on how the user can interact inside the VE to find the best point of view during a needle insertion task in a VR simulator.

As discussed in chapter 1, the principle of orbiting around 3D objects (point of interest) is the most suitable for navigation in VR surgical trainers since it allows the trainee to visualize the 3D objects from different angles to extract information about the tools' position with respect to them. We have also identified two types of navigation techniques that vary in their levels of interaction fidelity: naturalistic and metaphoric techniques. However, no consensus was found to help designers choose the appropriate level of interaction fidelity of navigation techniques for these systems.

To summarize, secondary tasks such as navigation are not the main focus during VR surgical simulators' design, and although many techniques have been proposed for navigating in these systems, their impact on users' performance on main tasks remains

unknown. Besides, in chapter 2, simulator fidelity was pointed out as important to guarantee the transfer of knowledge to the real world. However, the influence of the interaction techniques' fidelity on users' performance for surgical simulators has not been fully investigated yet. In this chapter, we address this issue for a virtual reality biopsy simulator, where the main task is the needle insertion, and the secondary task is the navigation around the operative field (i.e., changing the point of view). The research problem we want to address in this chapter is whether the interaction fidelity of the viewpoint changing technique impacts the user's task performance in the biopsy simulator. We investigate this issue by designing and comparing two existent techniques used to change the user's viewpoint. These techniques have different levels of interaction fidelity and were designed following a user-centered approach with an in-depth analysis of real biopsy procedures. In the following sections, we describe the biopsy task analysis and then present the design and evaluation of these techniques through two iterative experimental studies.

3.3 Biopsy task description

The task analysis of the biopsy procedures presented here is based on a previous study conducted by the research team (Van Nguyen et al., 2015). The study is based on observations of videos from actual interventions and interviews with expert clinicians. Hereafter, we describe the main phases of the biopsy procedure and their requirements.

During a biopsy, clinicians must insert a needle accurately in the patient's body to reach target tissue such as a tumor. The biopsy procedure can be divided into three main phases. The first phase is the planning. This phase consists of identifying the tumor position inside the body by means of an imaging system (e.g., a computerized tomography (CT) scan), choosing the appropriate needle entry point on the body surface, and determining the trajectory that the needle must follow to both, reach the target and take into consideration safety constraints (e.g., avoiding vital organs or bones). The second phase is the needle insertion and manipulation. It consists of inserting the needle through the chosen entry point and following the defined path to reach the tumor and extract a sample tissue to analyze. To reduce the radiation exposure from the CT scan in this phase, the clinician does not have live visual feedback of the needle position. Only the CT scan images' spatial memorization from the planning phase and the haptic feedback provided by the needle in contact with the different tissues during insertion are used to guide the clinician. The third phase is the verification of the needle position inside the patient's body. This phase consists of performing another offline CT scan of the patient to determine whether the

needle insertion is being correctly performed and whether the needle tip reached the target. It is essential to mention that the number of CT scans should be minimized in order to reduce the patient's radiation exposure. If the needle is not correctly positioned, the clinician can take it out of the skin and then reinsert it to correct its orientation. The procedure ends when the needle tip is correctly positioned inside the target to perform the tissue sampling. The clinician then removes the needle with the sample tissue outside the body.

3.3.1 Analysis of viewpoint change during biopsy

To design our techniques, we have conducted here a more specific analysis related to navigation tasks during a biopsy. This allowed us to identify the key viewing angles used by the physicians when performing this procedure (see Figure 3.1). Performing a biopsy requires first to have a top viewing angle (by looking at the same direction as the needle insertion axis) in order to determine the entry point on the skin surface to position the needle on this entry point correctly. After the needle is positioned on the entry point and inserted inside the skin, the physician changes to an inclined viewing angle with small head movements (between 30 and 60 degrees relative to the needle insertion axes) in order to associate the position of the needle inside the tissue layers with the perceived haptic feedback. Finally, the clinician uses a close-up lateral viewing angle (almost perpendicular to the needle direction axis) to determine how deep the needle inside the patient's body is. All this information is updated regularly through CT scan images to ensure the proper execution of the task.

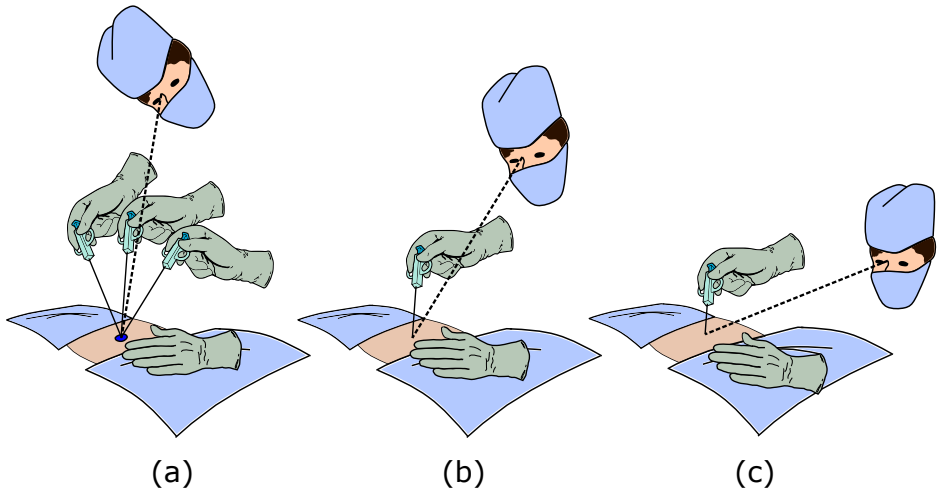


Fig. 3.1 The different points of view required to perform the biopsy task. (a) Top viewing angle to determine the entry point and position the needle on it. (b) Inclined viewing angle to associate the needle position with the haptic feedback. (c) Close-up lateral viewing angle to determine the needle depth into the skin

3.4 The design of the interaction techniques

Based on the previous analysis, two distinct techniques are proposed to change the point of view of the working scene in our biopsy simulator. The first technique is based on tracking the user's head, while the second one is based on touch interactions. Both techniques should cover a field of view similar to that used by the clinicians. These techniques correspond to existing metaphors usually used for navigation inside virtual environments but have different levels of interaction fidelity. Therefore, we have inspected them by applying McMahan's FIFA framework (McMahan, 2011) to classify them according to their levels of interaction fidelity.

Head-tracking technique

This technique allows the user to change the point of view by moving his head. As observed in the recorded videos, clinicians use this way of navigation to choose an appropriate point of view when manipulating the biopsy needle.

The technique consists of getting the user's head position through a tracking device and using this information to change the virtual camera position in the VE (see Figure 3.2). This technique offers the advantage of being intuitive because the camera in the VE replicates the natural movements performed by a user when exploring the real-world environment.

The implementation in our simulator uses Kooima's generalized perspective projection algorithm proposition (Kooima, 2008). This algorithm deals with the issue of motion parallax by calculating the appropriate projection matrix through the user's position with respect to the center of the screen.

Touch-based technique

In this technique, the non-dominant hand is used to set the appropriate point of view, while the dominant hand is used to manipulate the needle. We have designed this technique based on previous discussions with clinicians and observation of the real-world task. Indeed, clinicians use their non-dominant hand as a local frame of reference when performing the needle insertion task (see Figure 3.3). This can be related to the asymmetrical bimanual model (Guiard, 1987), which indicates that the non-dominant hand serves as a frame of reference for tasks performed using the dominant hand. This technique offers the advantage of a finer tuning of the point of view by allowing the user to have better control with smaller camera movements.

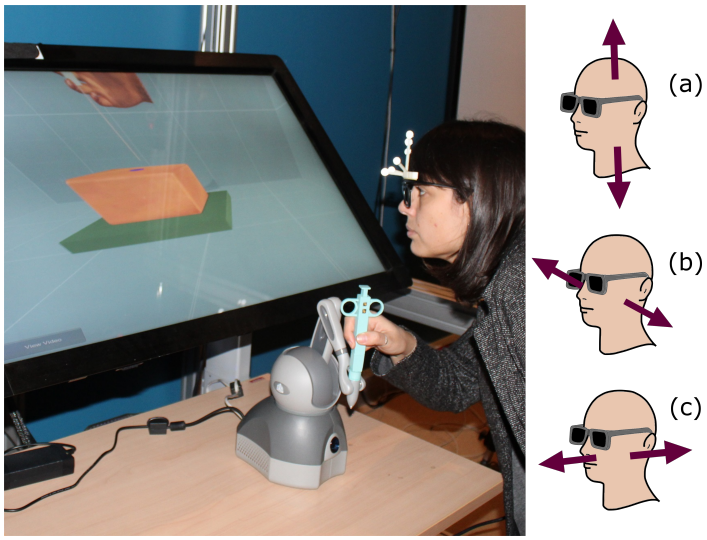


Fig. 3.2 (left) User interacting in the VR simulator using the head-tracking technique. (right) Head-Tracking movements: (a) Get up and down views (move the head up/down). (b) Get lateral views (move the head sideways). (c) Get closer to/far from the object (move the head ahead/backward)

For our implementation, two different gestures are used, dragging and pinch. The dragging gesture uses one finger to rotate the camera in the opposite direction of the movement, allowing the user to explore lateral and up/down views. The pinch gesture allows zooming into the scene for a close-up look (see Figure 3.4). Similar gestures were previously used by Fu et al. (2010) for navigating simulated astrophysical environments.

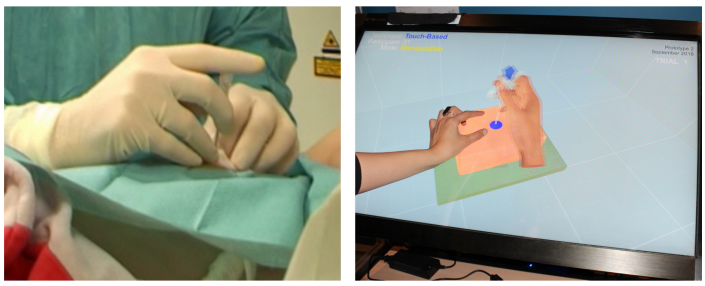


Fig. 3.3 Clinicians use their non-dominant hand as a local frame of reference. (left) Clinician performing a real biopsy. (right) User using the touch-based interaction technique

3.4.1 Analysis of interaction fidelity of the techniques

Concerning the design choices and fidelity levels, we controlled the two dimensions of the display and the scenario fidelity by providing a moderate visual (high resolution and stereoscopy, but moderate framerate and display size, and low FOV and FOR) and haptic (3 DOF force feedback, no torque, a max force of 3N, and a frame rate of 1 kHz) fidelity

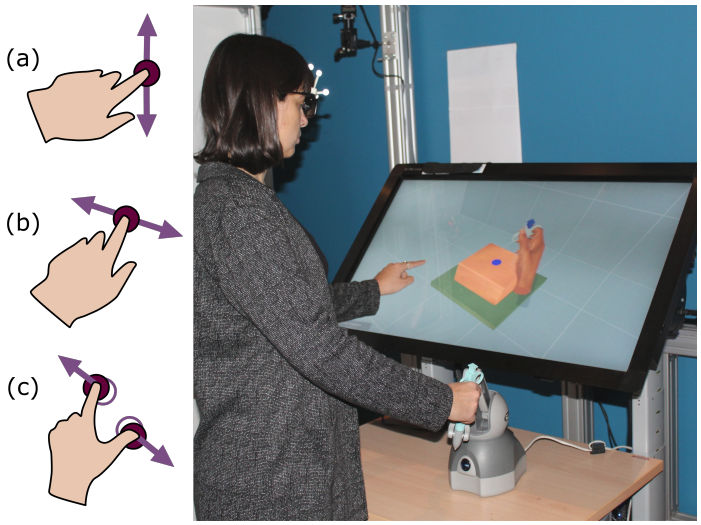


Fig. 3.4 (left) Touch-Based gestures: (a) Dragging gesture to rotate the scene (up-down). (b) Dragging gesture to inspect the skin from its lateral sides. (c) Pinch gesture to zoom the scene. (right) User interacting in the VR simulator using the touch-based technique

levels and a low scenario fidelity (a simplified task in a multi-layer tissue sample, with no representation of the patient's body).

For interaction fidelity, we controlled both manipulation and avatar fidelity. As mentioned in chapter 2, the manipulation task requires high levels of interaction fidelity (i.e., exact forces and movements, shape-alike interface). The avatar fidelity was designed with a low level of fidelity (static human-like hand with third-person perspective).

Finally, the evaluated component is the viewpoint changing task, for which we investigated the level of interaction fidelity. To characterize the two interaction techniques' fidelity levels, we have used the state-of-the-art interaction fidelity framework. The FIFA framework analyzes three dimensions of the interaction fidelity: biomechanical symmetry, control symmetry and system appropriateness (see subsection 2.3.3 for the definition of each dimension). Its application consists of determining whether an interaction technique matches the natural human interaction in the real world for each of these dimensions. In our case, we want to analyze which of the proposed techniques reproduces more reliably the natural human navigation during needle insertion.

Hereafter, we present for each dimension of the FIFA framework the associated level of fidelity attributed to each proposed technique. A summary of the analysis can be found in Figure 3.5.

FIFA analysis		Natural navigation	Head-tracking navigation	Touch-based navigation
Biomechanical symmetry	Kinematic symmetry	Move all body, particularly shoulders, neck and head to change head position	Move all body, particularly shoulders, neck and head to change head position	Move upper arm, forearm, hand and fingers to manipulate object
	Kinetic symmetry	Moderate muscle forces	Moderate muscle forces	Small muscle forces
	Anthropometric symmetry	Trunk, neck and head	Trunk, neck and head	Upper arm, forearm, hand and fingers
Control symmetry	Dimensional symmetry	Control over the position $x+y+z$	Control over the position $x+y+z$	Control over the position $x+y+z$
	Transfer function symmetry	1:1 position-to-position	1:1 position-to-position	1:N movement-to-deltaAngle-to-position; 1:N movement-to-position;
	Termination symmetry	Stop moving	Stop moving	Take off the finger from the screen
System appropriateness	Input accuracy	N/A	Head tracker	Touch detection
	Input precision	N/A	Head tracker	Touch detection
	Latency	N/A	Head tracker	Touch detection
	Form factor	N/A	N/A	Touch Screen

Reference		
Low	Moderate	High

Fig. 3.5 Interaction fidelity analysis of head-tracking and touch-based techniques using FIFA

Biomechanical symmetry: The head-tracking technique is considered a higher-fidelity interaction technique than touch-based technique regarding this component because the body movements involved and the forces applied are those of the real human navigation.

Control symmetry: Although it is possible to control the virtual camera position with both techniques, it is important to notice that for the head-tracking technique, the user's movements are directly mapped to the virtual camera movement. In contrast, finger movements must be converted into a delta angle to update the camera position. Moreover, the scaling performed with the zoom gesture must be translated to a displacement in order to finally change the camera position. Additionally, touch interaction has a lower level of termination symmetry when compared to the head-tracking. To analyze this, we considered that the termination symmetry in the natural navigation task corresponds to stopping the body (i.e., head) movement, which is precisely reproduced in the head-tracking technique.

System appropriateness: For this dimension, we consider that both techniques have minimal errors in accuracy and precision (the fingers detection and the head-tracking).

However, the gesture recognition has a higher latency and requires a touch input device, while the head-tracking is done naturally by attaching passive markers to the user's head. Therefore, the head-tracking technique presents a higher level of system appropriateness.

Nonetheless, although the touch-based technique is presented as a lower fidelity technique in terms of system appropriateness, it can be more advantageous in terms of movement stability. In fact, users can view the scene from one fixed viewpoint while manipulating the device from their natural viewpoint. On the other hand, head-tracking can result in a less stable point of view because any small movement of the user's head results in a camera movement (low input accuracy and tracking latency). This may impact the user's performance during needle manipulation.

As a whole, head-tracking is classified as a high-fidelity interaction technique and touch-based as a moderate-fidelity technique. Nevertheless, the touch-based technique presents some advantages that can lead to a better performance in the particular task of needle insertion using the VR trainer.

3.5 Viewpoint change fidelity, first study

3.5.1 Working hypotheses

After designing our techniques, we have conducted a first user experiment to compare them. Our main research question for this experiment was whether the fidelity of the viewpoint changing technique (secondary task) would impact the users' performance during needle insertion (main task) in the VR biopsy trainer. Besides this central question, we wanted to investigate whether the fidelity of this interaction technique will also influence the users' subjective experience through the sense of agency and the sense of ownership, as well as easiness to use and preference. We can interpret the sense of agency in this specific context as the users' impression to be able to control the camera's point of view, and the sense of ownership as the impression that the camera's point of view is the users' own eyes.

Therefore, we have defined three main working hypotheses for our experiment:

- H1) Users will perform the needle insertion task better using the touch-based technique.

  We think that the touch-based technique will allow users to be more accurate when performing the needle insertion task because it will grant them a more

stable point of view and let them carry out smaller camera movements. Consequently, this is expected to allow them to perform the overall task faster, with fewer trials and more accurately because they will reach the target much more easily. On the other hand, in the head-tracking technique, the users' head position will be continuously tracked, and the camera position updated. This may result in a less stable point of view, which will impact the accuracy of the needle manipulation and also increase the overall completion time and the number of insertion trials.

- H2) Users will have a higher sense of agency and ownership when using the head-tracking technique.

  We hypothesize that the head-tracking technique will more positively affect the sense of agency and the sense of ownership, compared to the touch-based technique, due to its higher biomechanical symmetry and a stronger control symmetry with natural navigation.

- H3) Users will find the head-tracking technique easier to use and prefer it over the touch-based technique.

  We hypothesize that the head-tracking technique will be easier to perform due to the similarity to a known task, and this will make users prefer the use of this technique over the touch-based one.

3.5.2 Apparatus

In order to evaluate the two proposed techniques, a prototype of a needle insertion simulator was developed. The implementation was done using Unity3D[1] (version 5.3.2) with C#. Unity3D was chosen as the development platform because it permits an easy integration of the various devices used, such as the haptic interface and the expansion modules for network communication and stereoscopic 3D view.

The VE consists of a virtual biopsy needle (modeled in Blender) and a rectangular object simulating a soft two-layer tissue (penetrable surface) placed on a table (not penetrable surface). A green sphere representing a tumor target was placed inside the tissue. In addition, a virtual rigid-hand holding the needle was added because it was reported to be useful as a spatial reference in needle insertion VR simulators (Van Nguyen et al., 2015) (see Figure 3.6g).

[1] https://unity.com

The virtual needle position and orientation are controlled by a physical interface, the Geomagic Touch[2] haptic device, providing 6-DoF for position (see Figure 3.6d) and 3 DOF of force feedback ($3.3 N mm^{-1}$ maximum stiffness force). The used force feedback model is based on a needle insertion inside a pig liver at a constant speed (Barbé et al., 2006) and was implemented by extending the C# wrapper for Geomagic Touch device developed by Kirurobo[3] (Kirurobo, 2014). Finally, to increase the system's interface fidelity, the haptic device's stylus was replaced by a 3D printed model of a biopsy needle-holder. This holder replicates the exact measures of a real biopsy needle-holder (see Figure 3.6e).

A 40 inches 3D monitor was used to display the VE (see Figure 3.6b). It was positioned with a 45° inclination with respect to the horizon, as suggested in previous studies (Van Nguyen et al., 2015). Finally, the VE was displayed with stereoscopic rendering because previous studies show that it provides depth cues that can help users to better understand scene details when precision is required (Boritz and Booth, 1997; Lee and Kim, 2008).

Head-tracking technique: The user's head position tracking was performed using the SmartTrack[4] camera system (see Figure 3.6a). This system consists of two infrared cameras for optical tracking, capable of detecting passive and active markers and operating at 60fps. Passive markers were attached to the stereoscopic 3D glasses (see Figure 3.6f). In order to communicate the position of the markers to the Unity3D application, the *ARTTrackSM*[5] library (developed internally by the team) was used (Davesne, 2015). Our implementation of the head-tracking technique is based on the generalized perspective projection method (Kooima, 2008).

The usable tracking space provided by the camera system was limited. It allowed the user to perform around 120° sideways and 110° upside-down exploration. This constraint was, however, sufficient to answer to the requirements of the biopsy navigation.

Touch-based technique: The touch-based technique was implemented using an interactive 3D multi-touch table (see Figure 3.6b) and by extending the *Touch Script*[6] framework developed for Unity3D and freely available in the asset store.

[2]https://www.3dsystems.com
[3]https://github.com/kirurobo/ManagedPhantom
[4]https://ar-tracking.com
[5]https://www.ibisc.univ-evry.fr/ fdavesne/tec/
[6]https://github.com/TouchScript/TouchScript

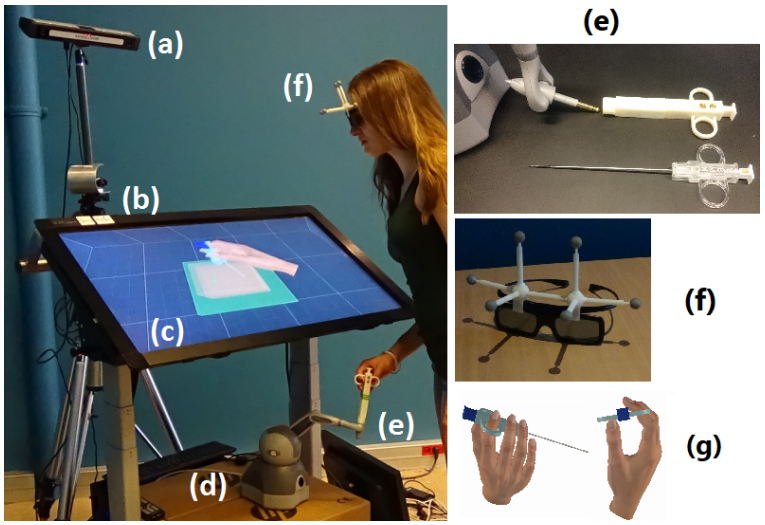


Fig. 3.6 Biopsy VR simulator implementation. (a) SmartTrack camera to track the user's position. (b) Touch input screen with 3D stereo capabilities. (c) Virtual scene of the needle insertion task. (d) Geomagic haptic device to implement force feedback. (e) 3D printed needle attached to the device. (f) 3D stereo glasses with passive markers to track the user's position. (g) Virtual hand

To be able to compare both techniques, the amplitude of the point of view using the touch-based technique was limited in order to have the same space as the head-tracking technique (i.e., 360° exploration with the touch-based technique was not allowed).

3.5.3 Participants

Fourteen participants (10 males and 4 females) from the University (students and staff) were enrolled in this study ($N = 14$). The mean age was $\mu = 26.50$ ($\sigma = 3.98$). Twelve of them were right-handed. They all had normal or corrected to normal vision, and six wore their correction glasses during the experiment. All of them have previous video game experience (including smartphone games), with four of them being regular players. Eight participants reported previous experience with 3D VEs, and seven participants used haptic devices before this experiment (mainly in demonstrations or previous user studies). All the participants were naive subjects with limited experience with needle insertion tasks. This is similar to what one can expect from future users of the system –novice medical students who start learning basic technical skills.

3.5.4 Experimental design

A within-subject design was used, with one independent factor (viewpoint changing technique) with two levels: the head-tracking (HT) and the touch-based (TB). Thus, all

participants performed the experimental task twice, once using each technique. The presentation order of the techniques (automatically assigned by the application) was counterbalanced to avoid any learning effect.

3.5.5 Experimental task

The experimental task consisted of a needle insertion inside a transparent tissue model. The goal was to position the needle tip in the center of a green target located inside the transparent tissue.

Before each trial, a red sphere was displayed in the center of the scene, indicating the starting point. The participants had to position the virtual needle in the center of this sphere to start the trial. This ensured that the haptic device was positioned at the exact same position at the beginning of each trial. Once the trial started, the green target was placed inside the tissue model. The participants were instructed to insert the needle inside the tissue to reach the target. They used the viewpoint changing technique to choose the entry point, verify the virtual needle's position with respect to the target, and confirm if the virtual needle reached the target's center. A red alert appeared if they touched the virtual table support, indicating that they have inserted the needle in too far. Once the participants considered that they reached the target's center, they had to validate the current trial by pressing the button included in the physical needle holder.

The participants were instructed to perform the task as quickly as possible (to evaluate the completion time) and to position the needle tip as accurately as possible in the center of the green target (to evaluate the accuracy).

3.5.6 Experimental procedure

The experiment started with signing the consent form and filling-in the demographics questionnaire. After that, the participants were presented with an instructional video explaining how to visualize the environment from different points of view by using each of the proposed techniques, how to use the different equipment (i.e., haptic device and validation button), and the task procedure to perform. The participants were then installed in the VR prototype area. They started the familiarization phase, where they learned how to manipulate the haptic device, how to use the physical holder's button, and they were able to experience the haptic feedback of the virtual needle inside the tissue. After that, the experimental session started for the first technique (automatically assigned by the simulator). It consisted of performing five trials of the needle insertion task. Finally, the participants were asked to fill in a questionnaire to subjectively evaluate the viewpoint

changing technique used. This procedure was repeated for the second technique. See Figure 3.7 for a summary of the experimental procedure.

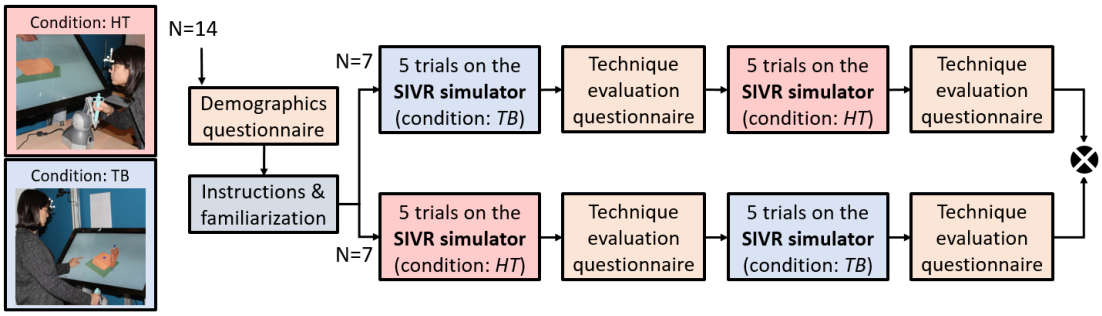


Fig. 3.7 The experimental procedure for the viewpoint change study

3.5.7 Data collection and analyses

Objective and subjective measurements were recorded for this experiment. The completion time and the distance between the needle tip position and the target's center were recorded for each trial to analyze the techniques' effectiveness. Completion time calculation started when the start red sphere was reached and ended when the physical button was pressed to validate the trial.

The subjective data consisted of responses to a questionnaire for each technique, using a 7-point Likert scale (see Appendix C). The questions included four different criteria: the realism, the possibility to navigate and manipulate inside the environment, the sense of ownership, and the sense of agency. Some of the questions (Q1-Q11) were extracted from the "State of Presence Questionnaire" (Witmer and Singer, 1998) and the rest of them (Q12-Q16) were inspired from questionnaires used in the literature (Hoyet et al., 2016).

Data analyses

All data analyses were performed using SPSS software version 21.0 (IBM Corp., Armonk, NY, USA) with the appropriate statistical tests. We have used a confidence level of 95% for all our statistical analyses. Therefore, a result is considered significant when $p < .05$.

First, the collected data was analyzed to determine whether parametric tests can be used. We have checked the data normality assumption through the Shapiro-Wilk test on the completion time data and the accuracy data. The results indicate that all data follows a normal distribution. Therefore, the paired samples t-test was used for all the variables data to compare the mean values.

The non-parametric Wilcoxon Signed Rank Sum test was used to compare the mean scores of the subjective questionnaire data (ordinal data).

3.5.8 Results

Objective measures

Completion time: The paired-samples t-test showed a significant main effect of the technique on the completion time ($t = 2.25$, $p = .044$) (see Figure 3.8). The participants performed the task significantly faster in the HT condition.

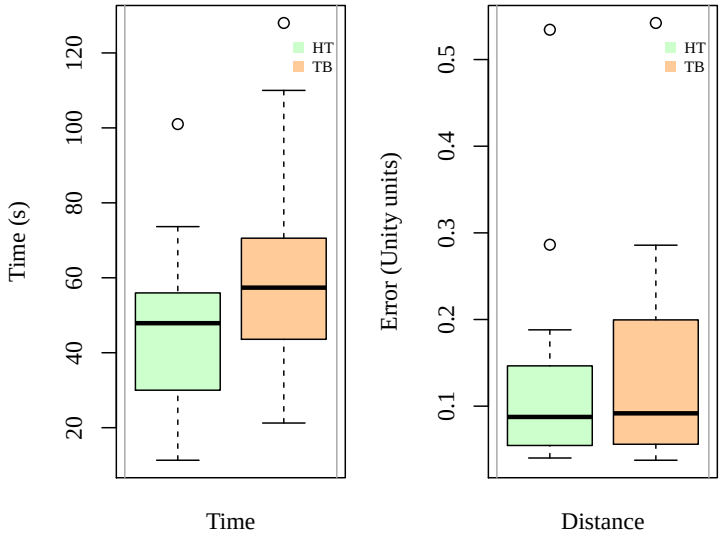


Fig. 3.8 The median value (with inter-quartile range –Q1-Q3) for time and accuracy measures in the VR prototype for the HT and TB techniques

Accuracy: Although participants were more accurate in the HT condition, the paired-sample t-test showed no significant effect of the technique on the accuracy ($t = 0.56$, $p = .584$) (see Figure 3.8).

For a more in-depth analysis, we have compared the accuracy according to each axis (X, Y, and Z) separately. A one-way ANOVA shows no significant effect of the axis ($F(2,39) = 2.38$, $p = .105$) on the accuracy for the HT condition. On the other hand, the one-way ANOVA shows a significant effect of the axis ($F(2,39) = 7.12$, $p = .002$) on the accuracy achieved with the TB condition. Post hoc comparisons with Bonferroni correction show that participants were more accurate on the X-axis (lateral) and on the

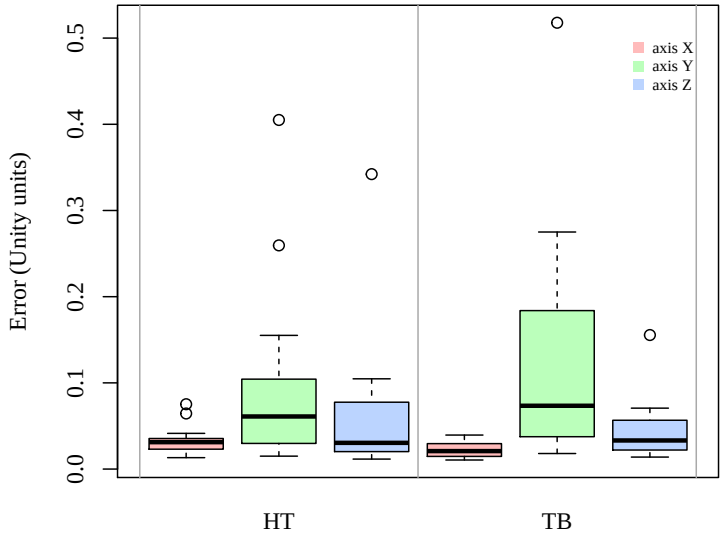


Fig. 3.9 The median value (with inter-quartile range –Q1-Q3) for the analysis of the accuracy measure by axis in the VR prototype for the HT and TB techniques

Z-axis (vertical) than on the Y-axis (depth) ($p = .003$, $p = .020$; respectively). No significant effect was found between X and Z axes (see Figure 3.9).

Subjective evaluation

Overall, 71% of the participants preferred the head-tracking technique over the touch-based technique, and 64% of them thought that it is easier to learn. Finally, half of them believe that they performed better with it. Regarding the questionnaire answers, although the responses show a higher score in favor of the HT technique for all the criteria (see Figure 3.10), the Wilcoxon non-parametric tests show no significant effect of the technique for any of the criteria (see Table 3.1).

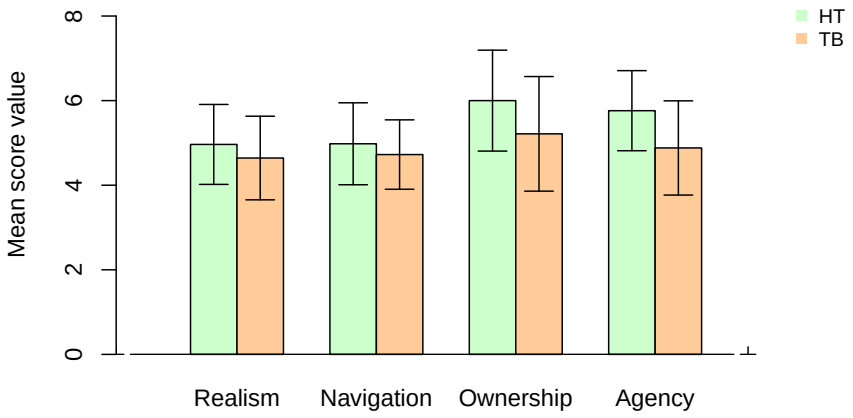


Fig. 3.10 Subjective measures results for user-study (first-iteration)

Table 3.1 Statistical results for Presence questionnaire

	Wilcoxon tests	
Criteria	Z	p
Realism	−0.63	.528
Possibility to Navigate/Manipulate	−1.18	.238
Sense of Ownership	−1.72	.085
Sense of Agency	−1.86	.063

$*p < 0.05$

3.5.9 Discussion

Our first user study results do not allow us to validate the hypothesis of participants having a better performance with the moderate fidelity technique (H1). Indeed, the results suggest that the higher fidelity technique for changing the point of view permits to perform the needle insertion task significantly faster than the moderate fidelity technique in our virtual reality simulator. These results are consistent with Chen et al. (2013) study, who compared a head-tracking navigation technique with a lower fidelity interaction based on joystick's control, and concluded that the head-tracking metaphor allowed participants

to achieve lower task completion times. The results show also that this technique allows users to perform the task more accurately, although no significant difference was found. A more in-depth analysis of the user's performance using the touch-based technique shows that users were more accurate on the lateral (X) and up (Z) positions of the needle compared to the depth (Y) position. This suggests that the top view of the tissue was not fully explored during target validation using this moderate interaction fidelity technique.

Not surprisingly, the head-tracking technique was reported to be easier to learn (64% of participants) and was preferred over the touch-based technique (71% of participants); therefore, H3 is validated. This can be attributed to the fact that this technique is based on the way users naturally navigate in the real world to inspect objects from different sides. Concisely, head-tracking technique has a higher level of bio-mechanical symmetry.

Besides, the head-tracking technique seems to generate a higher feeling of ownership and agency, although the differences were not significantly different compared to the moderate fidelity technique. These results can be attributed to the technique's fidelity level, particularly to the higher control symmetry level that the head-tracking technique has over the touch-based technique. Therefore, H2 is not validated, although the results lean towards head-tracking evoking a higher sense of agency and ownership when compared with the moderate technique.

Finally, general remarks were made by participants concerning the non-animated hand characteristics as well as its posture not reproducing participants one. Indeed, in this first study, participants were able to only choose between a left or a right hand, which had a medium-light skin tone. Improvements need to be made to increase participants association with their virtual hand (the research topic of hand representation and control was explored in the second part of this thesis project for a bimanual task, and it is presented in chapter 4).

The preliminary results presented above (Ricca and Chellali, 2016; Ricca et al., 2017) suggest better performance and preference by the users for the head-tracking technique, which is characterized as a high-fidelity interaction technique, compared to the touch-based technique, characterized as a moderate-fidelity interaction technique.

However, our observations during the experiment have shown that the participants did not fully use the touch-based technique to correctly verify the position of the needle during the insertion. Some of them kept the same viewpoint during the whole trial, which did not allow them to fully understand the 3D position of the needle with respect to the target. This may have impacted their performance using this technique. This issue can be attributed to an excessive simplification of the task design. In fact, the participants were able to insert

the needle and validate the trial without explicitly using the techniques to change the point the view. This does not reflect the actual biopsy procedure requirements in which the clinicians need to verify the needle position during the whole procedure. Subsequently, we have decided to perform a new design iteration of the simulator prototype by changing the task to better match the actual needle insertion task during biopsy and encourage participants to use the viewpoint changing techniques to perform the task correctly. This new iteration required us to come back to the task analysis phase in order to better design the biopsy procedure.

3.6 Viewpoint change fidelity, second study

3.6.1 New design choices

Based on the first study results, we have decided to modify the task implemented in the first prototype by including a visualization metaphor of a patient's scanner image during the biopsy procedure. By adding this metaphor that better reflects the procedure requirements, we were expecting the users to make use of the viewpoint changing techniques to perform the task with higher performance. This was also expected to improve the simulator's design by teaching the trainees how to better take into account the viewing angle during the needle insertion. Subsequently, two modes were included in the system. First, the verification mode simulates the clinician's need to visualize the offline scanner images to check the needle position without manipulating it. Second, the manipulation mode simulates the needle manipulation and insertion without visual feedback of the needle's position inside the body during its manipulation.

Based on these new design choices, a second version of the prototype was designed and developed. A second user study was carried out to investigate the original research question of whether there is an impact of the interaction fidelity of the viewpoint changing technique (secondary task) on the user's performance during the needle insertion task (main task) (see subsection 3.5.1).

3.6.2 Apparatus

Two changes were made to the initial needle insertion trainer. First, the two system modes (verification and manipulation) were added. A pedal was used as a user interface to switch between these two modes (see subsection 3.6.5 for details). Second, concerning the head-tracking technique, the SmartTrack camera system was replaced by the OptiTrack[7] camera system due to the original system's malfunctioning (see Figure 3.11a). This new camera system consists of six infrared cameras capable of detecting active and passive markers (like the ones mounted on the 3D glasses). A VRPN[8] client-server application was developed to communicate the user's head position to the Unity3D application.

[7] https://optitrack.com
[8] https://github.com/vrpn/vrpn

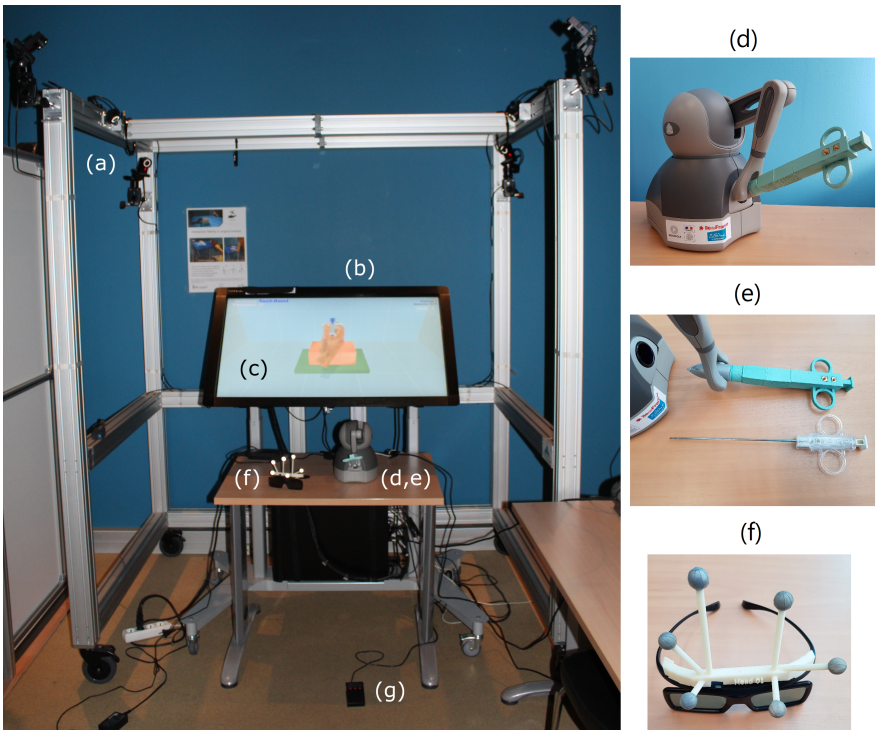


Fig. 3.11 Biopsy VR simulator implementation. (a) OptiTrack IR cameras to track the user's position. (b) Touch input screen with 3D stereo capabilities. (c) Virtual scene of the needle insertion task. (d) Geomagic haptic device to implement force feedback. (e) 3D printed needle attached to the device. (f) 3D stereo glasses with passive markers to track the user's position. (g) Pedal used to change modes

3.6.3 Participants

Twenty-one new participants (15 males and 6 females) from the University (students and staff) were enrolled in this study ($N = 21$). None of them have participated in the first study. The mean age was $\mu = 32.19$ ($\sigma = 9.13$). Seventeen of them were right-handed. They all had normal or corrected to normal vision, and six wore their correction glasses during the experiment. Seventeen of them have previous experience with video games (including smartphone games), with six of them regularly playing video games. Fifteen of them reported previous experience with 3D VEs, with only five of them having used haptic devices before this experiment (mainly in demonstrations or previous user studies).

The experimental protocol was approved by the institutional ethics committee of *Université Paris-Saclay* (CER Paris Saclay) prior to enrolling any human subject. Informed written consent was obtained from all the subjects involved in this study prior to their participation.

3.6.4 Experimental design

The same design as the first study was used in this study (see subsection 3.5.4). One male participant had to be excluded from the data set because he did not correctly follow the experimental instructions. This left 10 participants who started the experiment with the head-tracking based technique ($N = 10$) and 10 participants who started with the touch-based technique ($N = 10$).

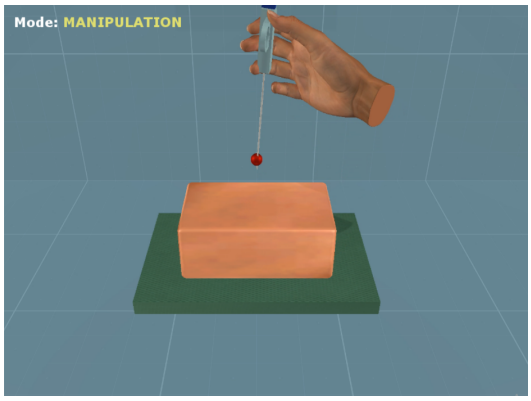
3.6.5 Experimental task

The changes made to the experimental task in this second experiment include the addition of the visualization metaphor of a patient's scanner image during a biopsy procedure.

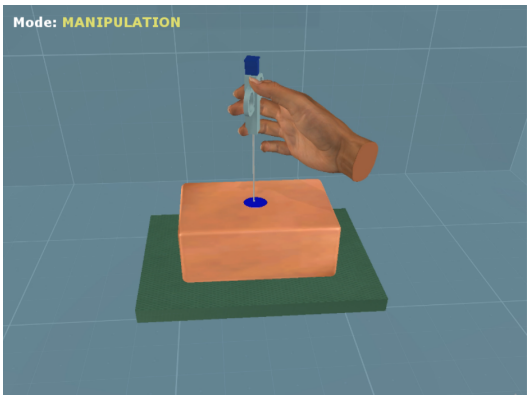
Once the trial started (after the red sphere position was reached with the virtual needle, see Figure 3.12a), the green target was placed inside the tissue model. This time, the tissue model was opaque, similar to the patient's body. Therefore, the participants were not allowed to see the target inside it. To simplify the task, the entry point, represented by a blue circle, was displayed on the top of the tissue (see Figure 3.12b). This replaces the initial planning phase, identified during our task analysis (see section 3.3). The participants were instructed to insert the needle inside the tissue following this entry point, allowing them to focus only on the needle manipulation and the verification phases.

Based on our task analysis, the system was designed to provide two different visualization modes. The first mode, called the "manipulation mode", allowed the participants to move and rotate the virtual needle as pleased and experience the collision forces with the tissue (see Figure 3.12c). Moreover, the participants were able to use the viewpoint changing technique to choose the desired point of view. The second mode, named the "verification mode", showed a frontal plane cut section of the tissue located at the target's middle position. In this mode, a truncated cone was rendered from the blue circle in the tissue's surface to the green target, and served as a guide to conduct the needle insertion (see Figure 3.12d). In this mode, the participants could not manipulate the virtual needle (i.e., the haptic device's position was constrained). They could only change the point of view by using the interaction technique. This mode simulates the verification phase using the CT scan images during the real-world procedure.

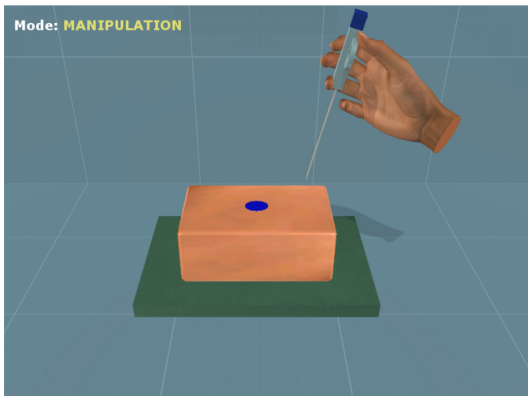
A priori, the participants did not know the green target's location, so they started each trial by switching to the verification mode to locate the target and determine the best needle orientation to reach it according to the proposed entry point. In addition, participants were instructed to use this mode to verify at any time the position of the virtual needle with respect to the target. The pedal was used to allow the participants



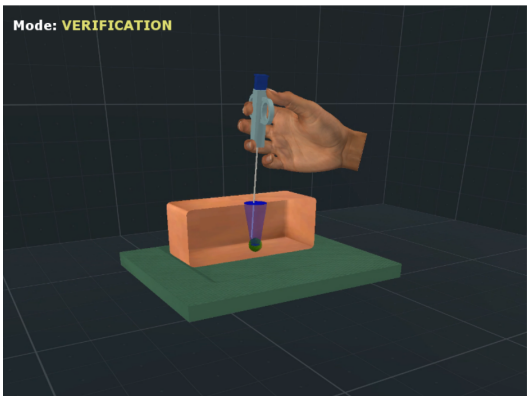
(a) Starting point determined by the red sphere.



(b) Blue circled guide where the insertion must be done.



(c) Manipulation mode where the participant can control the virtual needle.



(d) Verification mode where the participant can use the cone guide to verify the needle's position.

Fig. 3.12 Biopsy simulator task description

to switch between the two modes (see Figure 3.11g). In this case, the pedal needed to be pressed and hold to display the verification mode and released to come back to the manipulation mode. The pedal's use allowed them to switch between the two modes hands-free so that they could continue to use the touch-based technique using the non-dominant hand without interruption. Once the participants considered that they reached the target's center, they had to validate the current trial by pressing the needle holder's button in the verification mode.

The participants were instructed to perform the task as quickly as possible (to evaluate the completion time) and to position the needle tip as accurately as possible in the center of the green target (to evaluate the accuracy). Finally, they were asked to minimize the number of times they switched between the two visualization modes to evaluate users' ability to maximize the patient's safety (the patient's radiation exposure needs to be minimized).

3.6.6 Experimental procedure

The experiment started with signing the consent form and filling-in the demographics questionnaire. After that, the participants were presented with the simulator. They were able to choose the virtual hand characteristics: dominant hand (left/right), male/female model, and skin color (light, medium or dark). The simulator automatically assigned the first technique to test, for which a detailed tutorial was presented. It consisted of a series of steps in the form of short video and audio instructions on how to perform the needle insertion task, how to use the different devices, and how to use the viewpoint changing technique to update the virtual scene's point of view. This training phase was alternated with short interaction sequences for familiarization with the setup. All the steps could be repeated if necessary and permitted to progressively understand each step. The last training step consisted of performing a whole needle insertion procedure to guarantee that the participant had understood the experimental task correctly. Once the tutorial was finished, the experimental session started for the first technique. It consisted of performing five trials of the needle insertion task. Finally, the participants were asked to fill in a questionnaire to subjectively evaluate the viewpoint changing technique. This procedure was repeated for the second technique and finished with a comparative questionnaire and a global system usability questionnaire. See Figure 3.13 for a summary of the experimental procedure.

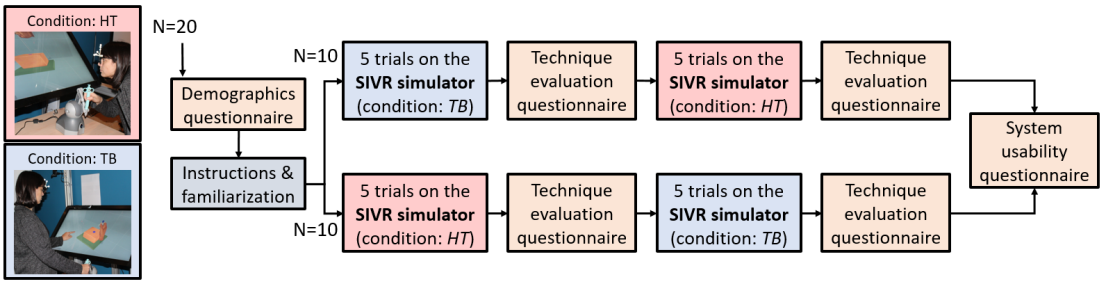


Fig. 3.13 The experimental procedure for the viewpoint change study

3.6.7 Data collection and analyses

Both objective and subjective measurements were recorded to compare the two techniques. The user performance was evaluated through the accuracy of the needle insertion, the completion time, the time spent in the verification and manipulation modes, the number of times the participant switched to the verification mode, and the number of errors committed. All the data was automatically recorded on a log file.

Time

The time counter for the task started once the virtual needle was placed on the center of the red sphere and ended once the users pressed the validation button on the haptic device, indicating that they considered that the needle tip was in the center of the green target. We have also calculated the time spent in the verification and the manipulation modes and the mean time spent during a single insertion trial. An insertion trial began when the user approached the needle to the tissue model and ended when the button was pressed for validation, or when the users removed the needle from the tissue to correct its orientation and re-try the insertion.

Accuracy

The distance between the virtual needle and the center of the green target (upon validation of the user) was used to measure the insertion's accuracy.

Changes between modes

A counter was incremented each time the user switched to the verification mode to register the participants' skill to perform the task by minimizing the patient's exposure to radiation.

Error rates

Errors were measured through the number of times the participants performed the insertion outside the blue circled guide and the total number of insertions required to finally reach the target.

Subjective data

The same comparative questionnaire (5-point Likert scale) used in the first study was employed to compare both techniques through four different criteria: realism, possibility to navigate and manipulate inside the environment, the sense of ownership, and the sense of agency.

Participants were also asked to indicate which technique was easier to use, had a better performance with, preferred for the whole task, and which one was specifically preferred for performing the three subtasks:

- finding the entry point;
- guiding the needle towards the target;

- checking the position of the needle's tip with respect to the target.

Finally, the "System Usability Questionnaire" (SUS) (Brooke, 1996) was used to obtain a general usability score of the VR biopsy system.

Data analyses

All data analyses were performed using R version 3.3.2 (R Core Team, 2017) on RStudio (RStudio Team, 2016, Boston, MA) with the appropriate statistical tests. We have used a confidence level of 95% for all our statistical analyses. Therefore, a result is considered significant when $p < .05$.

First, the collected data was analyzed to determine whether parametric tests can be used. We have checked the normality assumption of the data through the Shapiro-Wilk test on the completion time data, the accuracy data, the verification, manipulation and insertion times data, the number of switches to the verification mode, and the different error measures data.

The results indicate that all but the number of switches to the verification mode data follows a normal distribution. Therefore, the non-parametric Wilcoxon Signed Rank Sum test was used to compare the means for this dependent variable, and the paired samples t-test was used for all the normally distributed data measures. In addition, we have used Pearson's correlation test to analyze, for each participant, the correlation between his/her performances with each interaction technique. The non-parametric Wilcoxon Signed Rank Sum test was used to compare the mean scores of the subjective questionnaire data (ordinal data). Results for the objective measures are summarized in Table 3.2.

3.6.8 Results

Time

The paired-sample t-test shows a significant effect of the technique on the mean completion times for the task ($t = 3.95$, $p < .001$). Participants performed the task significantly faster in the TB condition (24.4% less time).

The paired-sample t-test shows a significant effect of the technique on the total time spent in the manipulation mode ($t = 4.03$, $p < .001$). Participants spent significantly less time in the TB condition (24.5% less time).

The paired-sample t-test shows a significant effect of the technique on the total time spent in the verification mode ($t = 2.25$, $p = .036$). Participants spent significantly less time in the TB condition (24.0% less time).

The paired-sample t-test shows no significant effect of the technique on the mean insertion time ($t = -0.60$, $p = .555$). In addition, the Pearson's correlation test between the two variables shows that the insertion time was highly correlated between the two techniques ($r = .715$, $p < .001$).

The results graph can be observed in Figure 3.14.

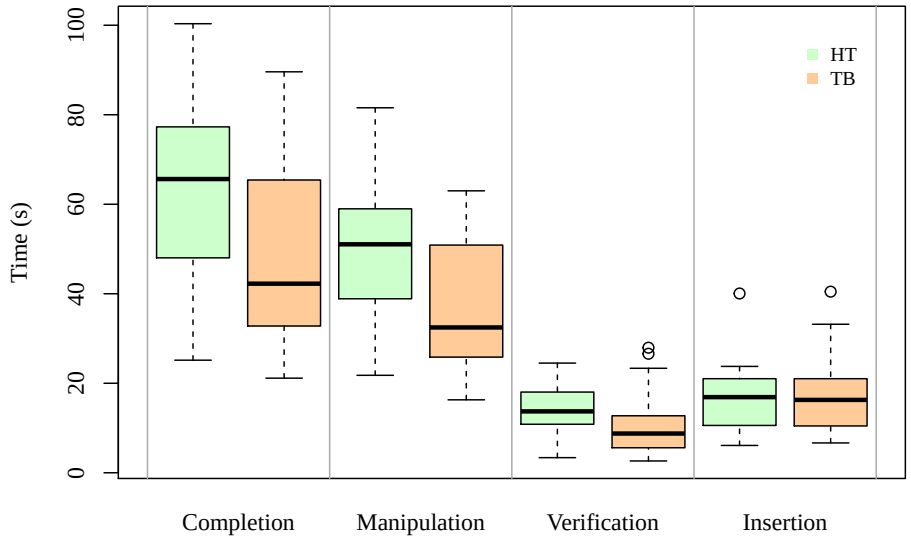


Fig. 3.14 The median value (with inter-quartile range –Q1-Q3) for time measures in the VR prototype for the HT and TB techniques

Accuracy

The paired-sample t-test shows no significant effect of the technique on the mean distance between the needle tip and the target center ($t = -0.62$, $p = .541$) (see Figure 3.15). In addition, the Pearson's correlation test shows a moderate positive correlation between the mean distance between the needle tip and the target in the two techniques ($r = .445$, $p = .049$).

Changes between modes

The non-parametric Wilcoxon test shows a significant effect of the technique on the number of switches to the verification mode performed by the participants ($Z = -2.49$, $p = .006$). Participants switched to the verification mode less often in the TB condition (29.5% less) (see Figure 3.15).

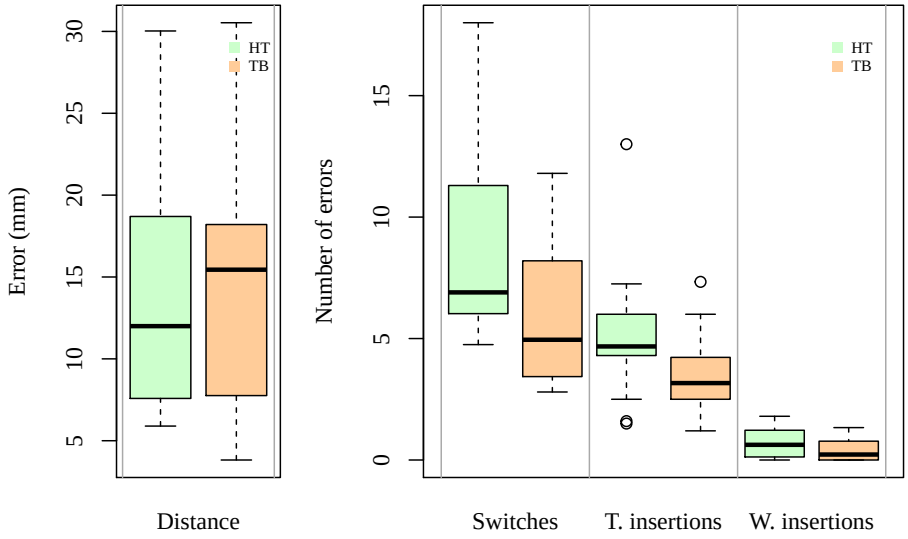


Fig. 3.15 The median value (with inter-quartile range –Q1-Q3) for accuracy, number of switches to verification mode, and error measures (total number of insertions and number of wrong insertions) in the VR prototype for the HT and TB techniques

Error rates

The paired-sample t-tests show a significant effect of the technique on the number of insertions performed outside the blue guide ($t = 2.61$, $p = .017$), and the total number of needle insertions required to reach the target ($t = 1.63$, $p < .001$). Participants performed the task with significantly fewer errors in the TB condition (43.1% less wrong insertions and 31.9% fewer insertions required) (see Figure 3.15).

Table 3.2 Descriptive and statistical analyses for the objective data

			t-test	
Measure	**HT** μ (σ)	**TB** μ (σ)	t	p
Completion Time	63.37 (19.72)	47.91 (19.71)	3.95	$< .001*$
Manipulation Time	49.38 (15.35)	37.28 (14.41)	4.03	$< .001*$
Verification Time	13.98 (5.50)	10.63 (7.32)	2.25	$.036*$
Insertion Time	16.72 (7.97)	17.60 (9.08)	-0.60	.555
Accuracy	13.50 (6.65)	14.51 (7.10)	-0.62	.541
Wrong Insertions	0.72 (0.59)	0.41 (0.42)	2.61	$.017*$
Total Insertions	5.11 (2.42)	3.48 (1.51)	1.63	$< .001*$

			Wilcoxon test	
Measure	**HT** μ (σ)	**TB** μ (σ)	Z	p
Mode Switches	8.66 (3.71)	6.11 (3.03)	−2.49	.006∗

$*p < .05$

Subjective data

The analysis of grouped questions through the non-parametric Wilcoxon tests show no significant effect of the technique on any of the criteria: the realism, the possibility to navigate and manipulate inside the environment, the sense of ownership, and the sense of agency.

On the other hand, the non-parametric Wilcoxon tests show a significant effect of the technique on the participant's mean scores for question Q5 (*To what extent were you able to actively explore the virtual environment?*), with the mean score being significantly higher in the TB condition ($Z = -2.54$, $p = .011$). No significant effects were found for the other questions.

In addition, the results show that 55%of the participants found the touch-based technique easier to use. There was no preference of the technique for the perceived performance and the preferred viewpoint changing technique. Regarding the preference of the technique to perform the individual subtasks, the results show that the touch-based technique was preferred to find the entry-point (65%) and to verify the final position of the needle with respect to the target (55%), but no preference was observed for the guidance subtask.

Finally, the SUS score reports a mean value of $\mu = 75.13$ ($\sigma = 9.75$), which stands for a grade B ("Good", percentile range $70-79$) on the usability scale.

3.6.9 Discussion

The results show that using the touch-based technique facilitated the completion of the needle insertion task on our VR simulator as compared to the head-tracking technique. This is observed through the decrease in the total time spent to perform the task, which was significantly lower in this technique. In addition, the participants also reduced the number of switches to the verification mode. This also reflects a better needle insertion performance with the touch-based technique requiring less often to check the needle position during manipulation. Moreover, the observation of the time spent in each mode shows that the use of this technique decreased the completion time for the manipulation

of the needle and the verification of its position subtasks. For the manipulation task, the participants also had a significantly lower number of wrong insertions (insertions performed outside of the blue guide) with this technique. This suggests that using the touch-based technique helped the participants insert the needle through the predefined entry point compared to the head-tracking technique. This is also supported by the comparative questionnaire on the preferred technique to perform this subtask (65% of the participants preferred this technique).

We have also analyzed the mean time spent to perform a single insertion using both techniques. The results show that participants spent a similar amount of time for each insertion with a significantly high correlation of performances between the two techniques. In addition, no difference in the subjective preference between the two techniques to perform this subtask. The same observation can be made for the verification phase. In fact, only a small difference (55% vs. 45%) in the subjective preference between the two techniques to perform this subtask is observed.

Considering these results as a whole, they suggest that the users' main issues using the head-tracking technique appeared while positioning the needle on the entry point before its penetration inside the tissue. This can be associated with a less stable point of view when using this technique, resulting in a less accurate needle positioning. Actually, this can be explained by the fact that when the users were trying to reach the entry point, they needed to move their heads to determine the best orientation of the needle, and at the same time keep their hand still to do the insertion, which sometimes required several trials. Indeed, the head-tracking technique is more demanding in terms of muscle forces and movements (biomechanical symmetry), which was previously found to be one of the reasons for poor performance in high-fidelity techniques (McMahan et al., 2012). On the other hand, the touch-based technique offered a stable viewpoint during the needle positioning towards the entry point. This resulted in a more successful and thus faster needle positioning and penetration through the entry point. In this case, they were able to choose and keep a top viewing angle on the tissue while moving their head to manipulate the needle correctly. Indeed, its design, based on Guiard (1987) principles for bimanual interaction, proved to have a beneficial impact on the dominant hand's precision, which is in line with previous results (Bertrand et al., 2015). After the needle penetrated inside the tissue and during the needle insertion towards the target, an inclined viewing angle was necessary. This was easy to get using both techniques. In this phase, the needle was guided by the haptic forces generated by the interactions between the tissue and the needle. This made the insertion easier. This can explain the fact that the needle insertion times were equivalent and positively correlated between the two techniques. This is also

observed in the verification phase because the users were not allowed to move the needle while checking its position. The increase in time in this phase can mainly be related to the increase of the insertions' number, each insertion requiring an additional amount of time to check whether the needle was correctly inserted. This was also associated with an increase in the number of switches to the verification mode. Finally, the techniques had no incidence on accuracy. In addition, the results show a moderate correlation in performance between the techniques with no significant difference in the mean values. This further confirms that both techniques offer an efficient interaction to check the needle tip's final position and verify whether it is in the target's center. Therefore, H1 is validated except for accuracy performance. Indeed, other works have shown that head-tracking metaphors are not always perceived as more efficient than touch interactions, and that a mixture between head movement interactions with manual ones offers a better compromise than head interaction alone (Spindler et al., 2012).

On the other hand, the subjective data analyses do not allow us to validate the hypothesis of the user's higher feeling of ownership and agency (H2) or preference for the technique (H3) when using the head-tracking technique. This can be explained in part by the users' lower performance with this technique, which may have impacted their subjective evaluation. This suggests that the participants considered that higher fidelity is less important in this case. The touch-based technique was better evaluated regarding actively exploring the VE. This suggests that, while both techniques offer exactly the same FOV amplitudes, the virtual camera's higher stability in the touch-based technique was more comfortable for the users and helped them better explore the environment. Although it was not objectively measured, a small delay can be observed for the head-tracking technique. This lag corresponds to the VRPN client-server communication, and it was not perceived as significant for the users (Q9: *To what extent were the delays between your actions from their consequences perceived?*, $Z = -1.04$, $p = .297$).

Many works have shown that high interaction fidelity techniques generally improve the users' performance and that middle-fidelity techniques tend to perform even worse than lower-fidelity ones (Bhargava et al., 2018; McMahan, 2011; McMahan et al., 2012; Nabiyouni, 2017; Nabiyouni et al., 2015). However, the results presented here show that high-fidelity interaction techniques are not always necessary for surgical simulators, and a well-designed moderate-fidelity technique can have a real advantage for training purposes. This highlights the fact that more research is needed to understand the impact of fidelity on VR simulators' design.

Finally, in terms of usability, the SUS questionnaire reported a good value (75%), which encourages the choices made for the design of this VR biopsy simulator.

3.7 Conclusions

Summary

Following an iterative design approach, we have designed two interaction techniques for viewpoint change for our virtual reality needle insertion simulator. The first technique is based on tracking the user's head position, while the second one is touch-based and relies on manipulating the user's point of view using the user's non-dominant hand. Our main objective was to investigate the impact of the interaction fidelity to perform the navigation task (secondary task) on the user performance for a needle insertion task (main task). In other words, whether higher interaction fidelity during navigation is required to improve user's performance. The two techniques were inspected using the FIFA framework to characterize their levels of interaction fidelity. This framework was originally designed for military applications and games, and we have used it for the first time in the specific field of surgical simulators. Based on this framework, the head-tracking technique was classified as a high interaction fidelity technique, while the touch-based technique was classified as a moderate interaction fidelity technique. The two techniques were then implemented in a virtual needle insertion simulator and compared in a user study.

Findings

The results show that the higher interaction fidelity technique (head-tracking) is more intuitive and easier to learn, globally preferred by the users, and improved participants needle insertion task performance when the simulated biopsy task was simplified (first study (Ricca and Chellali, 2016; Ricca et al., 2017)). On the other hand, when the task was more realistic (i.e., better contemplated the different sub-tasks required to complete a biopsy procedure), the moderate interaction fidelity technique (touch-based) offered a better task completion performance (second study (Ricca et al., 2021a)). This increase in performance was more particularly observed during the needle positioning before its insertion inside the tissue, which impacted the overall task completion performance. In addition, no impact of this technique was observed on the accuracy of the needle insertion.

As a whole, these results indicate that a high-fidelity viewpoint changing technique is not required to perform the needle insertion task. A moderate technique, which design rationale is based on observing the real procedures, permits users to reach similar accuracy performance as the higher fidelity technique but faster, with fewer errors and higher safety (i.e., minimizing patients' exposure to radiation). In addition, the touch-based technique

is easier to implement and permits to avoid additional calibration constraints associated with head-tracking technologies. Moreover, in a setup like the one presented in these studies, the head-tracking technique is more expensive due to requiring an external camera system to track the user's head position.

Our study also confirms that the design of interaction techniques to perform secondary tasks, such as navigation can influence the users' primary task performance in VR simulators. This further suggests that these tasks should be highly considered and carefully designed when implementing and evaluating such systems.

Limitations and perspectives

By observing the results of the designed viewpoint changing techniques in both studies, we noticed that the users' performance differs depending on the realism of the task simulated. The task realism, as presented in chapter 2, is related to the scenario fidelity. Therefore, it would be important to study and classify the scenario fidelity components for our simulator to determine which aspects influenced the users' performance and also if there is an interaction between the scenario and interaction fidelity components.

In addition, while the head-tracking technique was classified as a higher interaction fidelity technique, it is not yet fully realistic, which can lead to an uncanny valley effect (Mori et al., 2012). Since the touch-based technique's primary added-value relies on its usefulness to correctly position the needle to perform the insertion, we can also propose a combination between head movement interactions with hand movements to provide a better compromise.

Moreover, as for system improvements, participants commented that the actual position of the haptic device with respect to the monitor was not the optimal one. This may have lowered the interaction fidelity of the system as a whole. The main problem seems to come from the fact that the users visualize their real hands while manipulating the haptic device, and see the virtual needle's representation with a non-animated hand model attached in another position, hence presenting a mismatch between the visual and motor coordinate systems. In this current setup, users need to perform a mental translation between their motor and visual systems to carry out the task. To overcome this issue, some other possibilities can be explored. For instance, some researchers have used a mirror to see a co-located representation of the virtual needle and the user's real hand (Arsenault and Ware, 2000; Coles et al., 2011; Fortmeier et al., 2016; Lemole Jr. et al., 2007; Mastmeyer et al., 2014). This solution allows placing the haptic device in a closer position in comparison to the real-world task. However, there are no evaluation studies that show an advantage of this approach for training. Another possibility would be using

a head-mounted display, where the users look in the same direction as they manipulate the needle and switch to a more immersive virtual environment. However, this may raise other issues, such as dealing with the self-body (hand) perception during tool manipulation. Indeed, while using a head-mounted device, the users can visualize a synthetic environment all around them, but they do not have a direct visualization of their physical bodies, specifically, of their hands (used to manipulate the tools). We investigate the self-avatar representation (second design component) in chapter 4.

To summarize, in this first part of the thesis project, we were able to get insight into the impact of the interaction techniques' fidelity in a single experimental session. Our objective for this first part was to determine the impact of the fidelity of the interaction techniques for secondary tasks on users' primary task performance. Indeed, as revealed by the literature review on surgical simulators (see chapter 1), secondary tasks are not the main focus of those simulators' design, yet their inclusion and appropriate design, as shown by our two studies results, can lead to higher users' task performance in terms of completion time, and also improving the whole user experience in the VR simulator. Our first approach through these two studies was to apply the state-of-the-art interaction fidelity framework (FIFA) in the surgical domain and determine its suitability and limitations in the context of technical skills training. However, as mentioned in chapter 1, to determine the VR simulator's effectiveness, a real-world assessment of the learned skills must be performed. Although the biopsy task was a suitable choice due to requiring a certain degree of navigation (secondary task) to perform the needle insertion (primary task), performing a real-world evaluation is a strong limitation. Indeed, the needle insertion task simulated was relatively easy, limiting the possibility to carry out a longitudinal study with repeated training sessions. Moreover, transfer of knowledge to the real task is challenging to implement (in terms of objective assessment) and carries ethical and patients safety issues due to requiring the use of cadavers or patients. To better approach our main objective, which is to determine how the VR simulators' fidelity can influence technical skills learning, we chose to focus on another motor task to train, more challenging for users to learn, and approachable for us to assess the users' transfer of skills in a real-world setup. As presented in chapter 2, the second design component retained for studying is the fidelity of the self-avatar representation and control in VR surgical simulators. We explore this aspect in the next chapter.

Chapter 4

Hand movement

***

In the previous chapter, we have raised users' concerns about not fully associating the virtual non-animated hand represented in the VE with the way their real hands hold the physical tools and the movements performed with them. Indeed, the user's self-representation (avatar) in surgical VR simulators has not been thoroughly studied. The need to visualize our own representation inside the virtual environment and the required fidelity of this visualization are not yet determined to guarantee performance improvement or technical skills learning. In motor skills training, such as the assembling of mechanical parts, instrument handling, or tying knots in surgery, the user representation concerns primarily the user's hands.

In this second part of the thesis, we will focus our attention on the user's representation and control in immersive surgical simulators. More specifically, we want to determine the impact that the users' hand movement visualization might have on performance, training outcomes, and skill transfer of a tool-based motor task.

The chapter is structured as follows. First, we present a brief summary of the state of the art on hand visualization in different VR applications presented in chapter 2. We then describe the design and implementation of our pick, transfer and place VR simulator with hand visualization support. In order to determine the degree of self-avatar fidelity of the hand representation, we characterize its implementation by applying the proposed extension model of self-avatar fidelity for surgical VR simulators (see chapter 2).

Finally, we present the two user studies conducted, validation and longitudinal, which analyze the impact of hand representation in motor skills training, and discuss the main results and conclusions.

4.1 Hand representation in virtual environments

Several works have focused on studying the user's hand representation and its influence on the sense of presence and embodiment. Findings support the inclusion of virtual hands to enhance the user experience in a VR simulator. In addition, the virtual hands' appearance has been shown to have an impact on performance in immersive VR applications. However, it is important to understand the impact of including these hand representations on motor skills tasks.

In addition to visual feedback, haptic feedback is another aspect to consider when studying tool-based motor skills training. Although the visual feedback dominates how stimuli are perceived and drives the multimodal interaction (Hecht and Reiner, 2009; Posner et al., 1976), haptic feedback has also been shown to impact motor skills training (Lin et al., 2012). Research in the field has shown that motor learning performance is improved in terms of accuracy and trajectory optimization when haptic feedback is used Grant et al. (2020), which highlights the importance of including multimodal feedback in VR simulators that aim to train and transfer tool-based motor skills.

The objective of this second part of the project is to build on the existing literature, which shows a positive impact of hand presence on performance for direct object manipulation, and to go beyond by investigating the impact that the hand representation can have on motor tasks completion in immersive VR when the user manipulates a tool (indirect object-manipulation tasks). For that purpose, we have designed an immersive VR simulator where users could perform a tool-based pick and place task to train a motor skill. We have also conducted a first validation study to investigate the influence of hand visualization on users' performance and experience in this system. For that purpose, the users were asked to perform the task on the system with two conditions: presence or absence of the representation of their virtual hands. Furthermore, a longitudinal user study was also conducted on the developed VR simulator and a physical implementation of the same task. The purpose of this second study is to first evaluate if user's hand visualization impacts the learning process of the motor skill in VR, and if there is a transfer of knowledge from the VR simulator to the real-world task.

4.2 Validation study

Previous research suggests that adding a virtual hand will increase depth perception and give users additional spatial cues when manipulating tools (Van Nguyen et al., 2015). Moreover, visual feedback from arms was reported to be better for movement percep-

tion in tool-based tasks than the visualization of an isolated tool (Guerraz et al., 2018). Therefore, we expect the presence of the hands' representation to improve the users' task performance.

Besides, previous works have shown that the users' sense of presence increased with the self-avatar representation for tasks involving direct manipulation of objects (Lok et al., 2003). More particularly, users' higher sense of ownership is associated with avatars' higher appearance representations (Canales et al., 2019). In contrast, users' higher sense of agency is associated with the users' ability to reproduce more realistic movements with the virtual body (Argelaguet et al., 2016). Therefore, we aim to investigate whether these findings also apply when users utilize tools to manipulate objects.

Based on the previous observations, we have defined the following working hypotheses for our user study:

- H1) Users' performance (time, accuracy, and errors) would be improved when they have the real hand representation compared to the no hand representation condition.
- H2) The users will have a higher sense of presence with the hand visualization condition in a tool-based interaction task:
  - H2.1) Users in the real hand movement condition would feel a higher sense of ownership than users in the no hand representation condition.
  - H2.2) Users in the real hand movement condition would feel a higher sense of agency than users in the no hand representation condition.

4.2.1 Participants

Forty-one participants (26 males, 15 females) from the University (students and staff) enrolled in this study ($N = 41$). The mean age was $\mu = 35.68$ ($\sigma = 11.64$). Thirty-one of them were right-handed, 8 left-handed, and 2 ambidextrous. All of them have normal or corrected to normal vision, and 18 wore their correction glasses during the experiment. All of them have previous video game experience (including smartphone games), with 18 of them being regular players. Twenty-two participants reported previous experience with 3D VEs, and 14 participants used haptic devices before this experiment (mainly in demonstrations or previous user studies). Twenty-three participants have previously used HMDs, with only 2 of them using them regularly and for no longer than 30 minutes per session.

The institutional ethics committee of *Université Paris-Saclay* (CER Paris Saclay) approved the experimental protocol before enrolling any human subject. All the participants involved in the study gave informed written consent before their participation.

4.2.2 Experimental design

A within-subject design was used, with one independent factor with two levels: the presence of the animated virtual hand (VH) and the absence of the animated virtual hand (NH) (i.e., only the virtual tools visualization) in the VR simulator. All participants performed two sessions on the VR simulator (one for each condition). The presentation order of the condition was counterbalanced to avoid any learning effect. One female participant had to be excluded from the data analysis since she felt uncomfortable using the plastic handles in the VR prototype and could not finish the whole experiment. This left 20 participants who started the experiment with the VH condition ($N = 20$), and 20 participants who started with the NH condition ($N = 20$). The sample size was calculated beforehand to use a two-tailed t-test for matched pairs ($\alpha = 0.05$ and $\beta = 0.10$) to be able to detect a medium effect size ($d = 0.55$), which determined that at least 37 participants were required in this experiment to have an actual power of 0.90.

4.2.3 Experimental task

For this experiment, we have designed and developed a VR simulator for training a motor skill task. We have designed a simple tool-based pick, transfer, and place task (PT&P), in which the trainee had to pick a set of small virtual objects from a specific position using a first tool, transfer them from one tool to another, and place them in final position using the second tool.

This task was inspired by the peg and transfer task used in physical and VR simulators to train surgeons' bi-manual dexterity and hand-eye coordination skills in laparoscopic surgery (Chellali et al., 2016). However, in contrast with the laparoscopic surgery simulation, the trainees using our system could see the objects and tools in 3D and in the same direction where they manipulate them. In contrast, in the laparoscopic surgery simulation, they observe their actions on a 2D screen decoupled from their working space.

The experimental PT&P task was performed with two forceps tools and six small cubes. The goal was for the participant to grab each cube with one of the forceps, transfer it to the other forceps, and place it on a square target.

At the beginning of each task, the participant had to place the forceps at the initial position and wait for the countdown to start (see Figure 4.1a). This ensured that the tools were positioned at the same starting position for all the trials and all the participants.

The initial and target positions of each cube were numbered with the same number (see Figure 4.1b). Due to the limited workspace provided by the haptic devices used, both initial and final cubes positions were carefully chosen to enable full movement with the tools during the whole task. These positions remained the same for all trials and all participants. The cubes' manipulation order had to respect the number sequence, i.e., the participant should first pick the cube number 1 and place it in the target position number 1, continue with the cube number 2, and so on until placing the cube number 6.

To make the task more challenging for participants, each cube had a green sticker on one of its faces. Initially, the cubes were oriented with the sticker facing upwards. The participant had to place the cubes with the sticker facing downwards, regardless of the orientation (i.e., the side of the sticker must be in contact with the target once placed), and to align the cubes with the target squares (see Figure 4.1c). This forced the participant to rotate the cubes during their transfer.

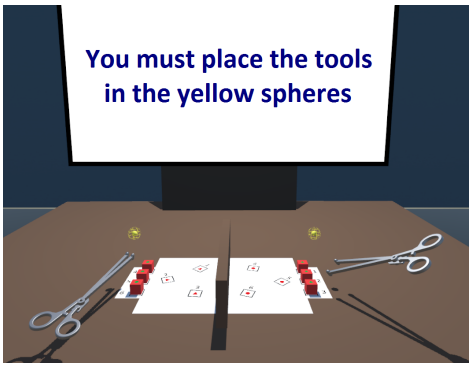
The working area was divided into two sides by a wall, with three cubes placed on each side. Each of the tools was used only to manipulate a cube on one side of the working area, which means that the right tool was used to pick and place cubes on the right side of the wall and the left tool on the left side. The only zone where both tools were authorized to manipulate a cube at the same time was the exchange zone, which was on top of the wall (see Figure 4.1d). They were able to do as many manipulations of the cube between the tools as necessary.

An example of how all these rules apply is:

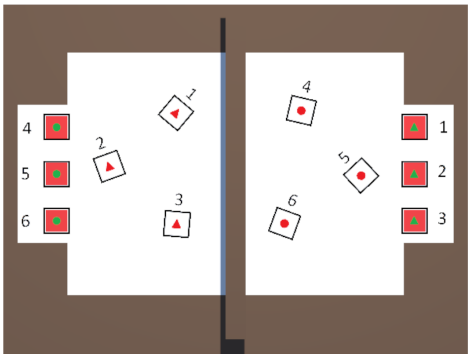
- pick cube 1 with the right tool,
- transfer it to the left tool in the transfer zone (taking into account that the green sticker must be placed facing downwards),
- place it with the left tool on target number 1.

Continue with cubes 2 and 3 following the same rule. Then, for cubes 4 to 6, participants had to use the left hand/tool for picking, and the right hand/tool for placing them.

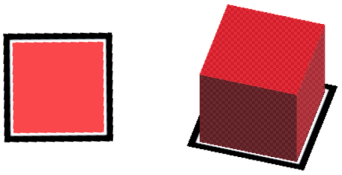
The participants were asked to move the cubes as quickly as possible and place them on the targets as precisely as possible. They were also instructed to avoid errors, which were tool-wall collisions and cube drops. When a cube was dropped, they were allowed to pick it up and continue. They had to use the appropriate tool (left or right) to pick it up, depending on which side of the wall the cube fell in. Finally, either tool could be used to pick it from the wall.



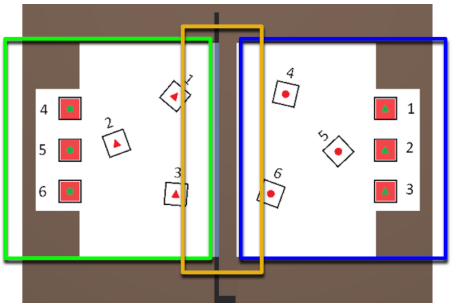
(a) The initial position of the forceps determined by the yellow spheres on top of cubes 2 and 5.



(b) The initial position of the cubes and target square numbering.



(c) The final position of a cube.



(d) The left and right tools zones (green and blue, respectively) and the exchange zone (yellow).

Fig. 4.1 Pick, transfer and place task description

4.2.4 Apparatus

VR prototype

The VR prototype consisted of two physical interfaces, an HTC VIVE[1] HMD providing a first-person perspective, and a pair of data gloves (see Figure 4.2). The physical interfaces were used to control two virtual forceps. Two Geomagic Touch[2] haptic devices were used, each providing 6 DOF for the position and 3 DOF for force feedback. The interfaces' styluses were removed, and 3D printed models of forceps handles were attached instead (see Appendix A for further information concerning the plastic holders). Each handle added one extra DOF by allowing to open and close the tool. The opening angle was obtained through a potentiometer installed inside each handle and connected to an Arduino Uno card[3].

The VE consisted of 6 small pickable 3D cubes (dimensions: 2 cm side-length, mass: 29 g) with a green sticker on one of the faces, two virtual forceps representing the tools, and

[1] https://www.vive.com
[2] https://www.3dsystems.com
[3] https://store.arduino.cc/arduino-uno-rev3

the user's hand representation. It also included a delimited working area (dimensions: l = 30.5 cm, w = 22 cm) divided by a wall and the cubes' initial and target positions specified. The initial cube positions were marked on each side of the wall. Both the initial and final positions of the cubes were numbered.

The PT&P VR simulator was designed as a client-server system. The server side was developed in C++ and used the chai3d[4] (Conti et al., 2003) framework for haptic force feedback simulation and ODE physics engine for collisions detection. The tools' interaction with the virtual cubes and wall was computed on this application. The cubes and tools' positions and orientations and the opening angle of the tools were continuously sent to the client-side using a UDP socket. The opening angle of the tools was obtained through serial communication with the Arduino. Haptic feedback was displayed when the virtual tools were in contact with the cubes, the wall, and the table surface, similar to what one can experience when manipulating real objects. The client-side was developed using Unity3D[5] (version 2018.3.6) with C#. It received the computed positions of the different components of the VE and rendered them in the HMD. It also communicated with the data gloves to obtain the user's fingers' movements and animated the virtual hands accordingly. The client and server were executed in two different computers (CPU: Intel i7, GPU: GeForce GTX 1060/1070, RAM: 16 GB), directly connected through a UTP cable. The application frame rate was 90 fps for visual rendering and 500-600 fps for haptic rendering.

One of the main challenges faced during this prototype's development was to match the virtual hands' positions with that of the virtual tools and the user's egocentric view of the VE. The hands' positions were obtained through the HTC trackers attached to each glove, providing the same frame of reference as the user's head (through the HTC VIVE HMD tracking). Also, a wood platform was built to position the two haptic devices. A third HTC tracker was placed on this fixed platform to track its position and locate it in the same reference frame as the HTC VIVE. The positions of the virtual tools (controlled by the haptic devices) were then associated with the corresponding position inside the VE.

Hand representation and control

In order to explore the influence of hand representation on performance for a motor task trainer, we chose to design a metaphor that provides a global moderate avatar fidelity. In this section, we first present the process of implementation of the virtual hand representation and control, and then we characterize it in terms of fidelity.

[4]https://www.chai3d.org
[5]https://unity.com

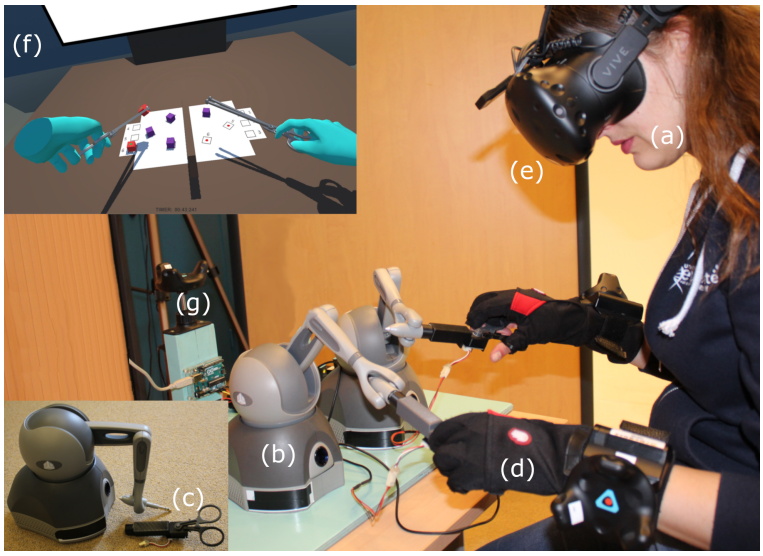


Fig. 4.2 The VR prototype for the PT&P task. (a) User interacting with the system (b) physical interfaces (c) 3D printed handle, forceps model (d) data gloves with HTC trackers (e) HTC VIVE HMD (f) virtual scene: working table, wall, cubes, forceps and hands representation (g) HTC tracker for the whole setup

In chapter 1 we have mentioned that there are two main approaches to capture the user's hand movements: through optical trackers, such as the Microsoft Kinect[6] or the LeapMotion[7], or by using inertial trackers or data gloves (many commercial solutions available).

We performed some tests with the LeapMotion device. It already provides easy integration to Unity3D applications through a well-documented API and examples. It can be attached to the HMD, facilitating the frame of reference association with the users' view. Moreover, with this device, it is possible to get, in addition to hand tracking, full finger-tracking, with better tracking results for the thumb, index and middle fingers. Nonetheless, the tracking was unstable once the user grabbed the forceps tool (used to perform the pick and place task), which led to a frustrating user experience. Therefore, we have decided to explore data gloves technologies.

The first tests were carried on with 5DT data gloves[8], which provide optical sensors to track fingers (only flexion of each finger), but without an integrated solution for positional tracking of the hand. Finally, we have decided to use the Noitom Hi5 VR Gloves[9], with a VIVE tracker attached to each of them to track the users' hand movements. These data gloves provide 9 DOF, 1 DOF for each finger (flexion), an extra DOF for the thumb finger

[6]https://developer.microsoft.com/en-us/windows/kinect/
[7]https://www.ultraleap.com
[8]https://5dt.com/5dt-data-glove-ultra/
[9]https://hi5vrglove.com

(they do not offer a full finger-tracking), and 3 DOF for hand rotation. Extra 3 DOF for the position of the hand were obtained using the HTC tracker attached. Although the Noitom VR gloves require calibration for each user, the procedure lasts only one minute.

The virtual hand representation consisted of a human-like hand with a neutral color. Since virtual hands' appearance may impact immersive VR applications, we chose to use a virtual hand model with a realistic shape and movements (obtained through the data gloves sensors) while controlling color and gender differences.

Self-avatar fidelity analysis: Concerning the fidelity levels, in this second part of the thesis project we have also controlled the two dimensions of the display and the scenario fidelity by providing a high visual (stereoscopy, and high resolution, framerate, FOV and FOR) and a moderate haptic (3 DOF force feedback, no torque, no grasping force, a max force of 3N, and a frame rate of 600 Hz) fidelity levels and a moderate scenario fidelity (small objects with high physical behaviors, no patient's body, complex task).

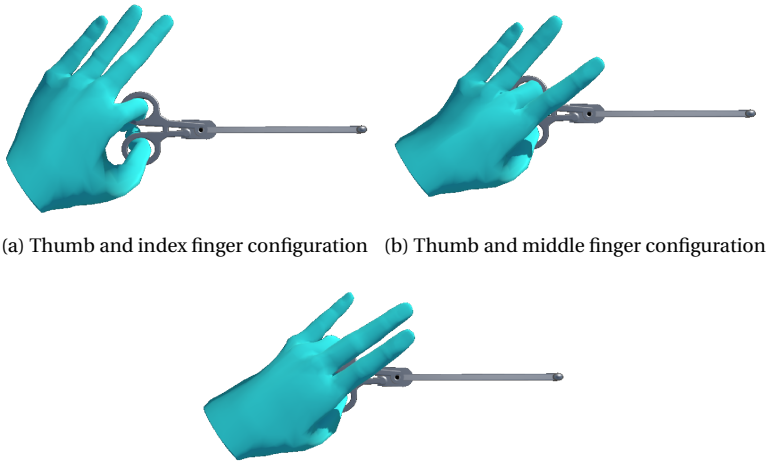
For interaction fidelity, we controlled both manipulation and navigation tasks, representing a high level of fidelity (physical handles and 6 DOF for virtual tool movement, plus 1 DOF for the tool's opening and closing, and head-tracking for the navigation).

Finally, the evaluated component is the avatar representation and control. In chapter 2 we have proposed a model for characterizing the self-avatar representation, i.e., the users' hands in VR surgical simulators. By applying this model to our designed virtual hand, we can determine that it has a high-level of visualization perspective fidelity since a first-person point of view is provided. Indeed, the users' virtual hands global positions were obtained through the HTC tracker controllers (attached to the data gloves), whose positions were in the same reference frame as the HTC VIVE HMD, hence being rendered in the "same" position where their real hands were (i.e., co-located). For the kinematics fidelity analysis, we characterize the hand control as a moderate-to-high fidelity approach. In terms of the movements reproduction, we provide 1 DOF for flexion of each finger not directly involved in the grasping of the tool (*free fingers*, commanded by the data gloves sensors), and an extra DOF for simulating the "pinch" movement carried on with the (other) two *main* fingers to open and close the tool (controlled by the potentiometer included in the physical plastic handles, see Appendix A). Indeed, after conducting some pilot tests, we observed that the initial arbitrary choice of using the index and thumb fingers to control the opening/closing of the tool was not preferred by all the participants in terms of comfort. Therefore, we have decided to propose three different configurations for this aspect: the thumb and index, the thumb and middle, and the thumb and ring fingers (see Figure 4.3). No spreading movement between the fingers was provided (limitations of

the data gloves interface). Moreover, the low precision technology used to track the users' hands global positions and rotations (the HTC VIVE tracking system) did not allow us to rely on this information to update these measures during the whole simulation. Indeed, although the haptic devices were rendered in their exact position with respect to the HMD (through a third HTC tracker installed in the rigid wooden platform), a mismatch was observed between the haptic forceps and virtual hand positions. This error might be related to several factors: the low-level precision of the HTC trackers, the experimental calculation of the haptic end effectors positions, and the fact that a unique hand size model was used for hand representation. To fix this mismatch issue, we chose to include two movements states:

- A *free-mode* where participants had full movement control of the virtual hand, i.e., maximum interface capabilities: 6 DOF for hand position and rotation plus 5 DOF for finger flexion and an extra DOF for thumb-index pinch.
- An *on-the-task mode* used when users were manipulating the tools. In this mode, the hands' global movements (6 DOF) were controlled by the haptic devices (6 DOF), main fingers by the potentiometer installed in the forceps interface, and free fingers by the data gloves sensors.

The participant's virtual hands were animated into pre-configured models of the hand attached to the tool to switch between the free to on-the-task modes. This transition phase corresponds to a lower level of synchronicity of the movements.



(a) Thumb and index finger configuration (b) Thumb and middle finger configuration

(c) Thumb and ring finger configuration

Fig. 4.3 The different configurations for hand-tool association

Finally, our hand representation was characterized as presenting a moderate appearance fidelity. By using a human-like hand model, we are representing a high level of appropriateness of the morphology. However, the photorealism of the hand representation and the visual identity are both characterized as low-level components. Indeed, realistic details are presented in these human-like hands, e.g., the nails, but we chose a neutral skin color to symbolize the use of gloves in surgical procedures, which renders a final version more cartoonish than realistic in terms of photorealism. In addition, the visual identity is very low for the hands' representation in this study. Neither the cosmetic aspects (e.g., participants' skin real texture), nor the dimensions and gender of the users' hands are personalized in our hands' model.

4.2.5 Experimental procedure

Before the experiment, participants were asked to read and sign the consent form and complete the demographics questionnaire. They were then asked to read the experimental instructions presenting the prototype and explaining how to use the different devices and how to perform the task. After that, they were moved to the VR prototype area. They were asked to put in the gloves and the HMD, grab the two forceps handles and sit comfortably. The instructions were shown on the virtual screen. For the VH condition, a calibration of the data gloves was performed before the familiarization session. They started then the familiarization session, which consisted of trying the system by performing a pick, transfer, and place of one cube (placed in the middle of the starting zone) on each side of the wall. At this point, the participants would feel comfortable with both the tools manipulation and task completion. After the familiarization phase, the experimental session started by performing three trials of the experimental task on the VR prototype for the first condition. To reduce cognitive load and allow participants to focus their attention on the motor task to perform, the cubes' location per trial and cubes' order to manipulate remained the same. At the end of the third trial, the participants were asked to subjectively evaluate their experience with the system in the current condition through the presence-state questionnaire. This procedure was repeated for the other condition. We counterbalanced the order of conditions presentation to control any learning effect (this was also verified through the proper statistical tests). Finally, the participants were asked to answer a system usability questionnaire for the VR simulator. The whole procedure lasted, on average, 45 minutes for each participant, with each condition taking between 10 to 15 minutes (VH condition was 2-3 minutes longer due to the data gloves calibration phase). See Figure 4.4 for a summary of the experimental procedure.

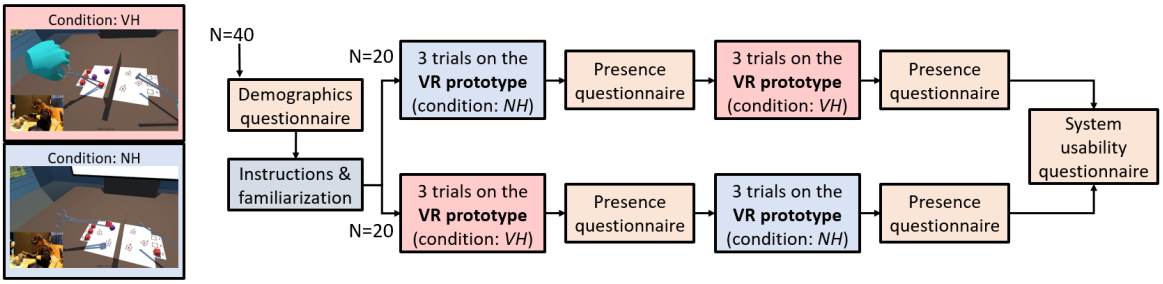


Fig. 4.4 The validation study procedure

4.2.6 Data collection and analyses

To compare the two VR conditions (VH, NH), objective and subjective measurements were recorded. The user's performance was evaluated through the accuracy of placing the cubes, the task completion time, and the number of errors during the task (the number of cubes dropped and the number of collisions of the tools with the wall). All the data was automatically recorded on a log file for each experimental condition.

Time

The time calculation of the task (total time) started once the tool touched the first cube, and ended when the last cube was placed. We have also calculated the time for the three subtasks: pick (pick time), transfer (transfer time), and place (place time). The pick subtask for a cube started when the tool touched it in its initial position and ended when the other tool touched the cube. This also corresponds to the beginning of the transfer subtask. The transfer subtask was determined by the exchange of the cube in the transfer zone at any moment in the task. Finally, the place subtask consisted of manipulating the cube with the placement tool outside of the transfer zone at any moment in the task. This means that the user could come back to the transfer subtask from the place subtask if necessary, i.e., to correct the cube's orientation.

Accuracy

The distance and the minimal rotation angle difference between each cube center and the corresponding target center were used to measure each cube placement accuracy. The task's accuracy was determined as the average of the six cubes' distance error (position distance) and rotation difference (rotation difference).

Errors rates

An error was counted for each time a cube was dropped during the task (total drops).

The second error measure corresponded to the number of times one of the tools touched the wall (total wall collisions).

Subjective data

The subjective data consisted of responses to a "Presence" questionnaire for each condition, using a 5-point Likert scale. The questions included eight different criteria: realism, possibility to act, quality of the interface, possibility to examine, self-evaluation of performance, haptic, ownership, and agency. Some of the questions (Q1-Q21) were extracted from the State of Presence Questionnaire (Witmer and Singer, 1998), and the rest of them (Q22-Q25) were inspired by questionnaires used in the literature (Hoyet et al., 2016). Also, participants were asked to rate which condition they preferred the most (VH or NH).

Finally, the "System Usability Questionnaire" (SUS) (Brooke, 1996) was used to obtain a general usability score of the VR system to improve the system in future iterations.

Both questionnaires can be found in Appendix C.

Data analyses

All data analyses were performed using R version 3.6.0 (R Core Team, 2019) on RStudio (RStudio Team, 2016, Boston, MA) with the appropriate statistical tests. We have used a confidence level of 95% for all our statistical analyses.

First, we have checked the normality assumption of the data through the Shapiro-Wilk test on the time, accuracy, and the two error measures for the VR prototype data. The results indicate that the total time, place time, position distance, and rotation difference follow a normal distribution. Therefore, the paired samples t-test was used for these variables data to compare the mean values for the two conditions in the VR prototype. The non-parametric Wilcoxon Signed Rank Sum test was used to compare the means for all dependent variables that are not normally distributed.

In addition, we have used the Pearson's correlation test (for variables following a normal distribution), and the Spearman's correlation test (for the other variables) to analyze, for each participant, the correlation between his/her performance on both VR conditions (VH and NH). This test was used to determine whether each participant achieved the same task performance in both conditions regardless of differences in personal abilities between participants. To further investigate the similarity in performance between both

conditions, equivalence tests (TOST (Lakens et al., 2018)) were also carried out for each dependent variable.

Finally, the non-parametric Wilcoxon Signed Rank Sum test was used to compare the mean scores of the "Presence" questionnaire data (ordinal data). The general mean score is provided for the "System Usability Questionnaire".

4.2.7 Results

Objective measures

The plots for all the performance measures for both conditions are detailed in Figure 4.5, and descriptive statistics in Table 4.1.

Time: The paired-sample t-test shows no significant effect of the condition on the mean total time ($t < 0.01$, $p = .999$) nor on the mean place time ($t = -1.29$, $p = .206$). The non-parametric Wilcoxon test shows no significant effect of the condition on the mean pick time ($Z = -0.61$, $p = .541$), nor on the mean transfer time ($Z = -0.36$, $p = .722$).

In addition, the Pearson's correlation test shows a moderate correlation between the mean total time ($r = .374$, $p = .018$), and a high correlation between the place time ($r = .618$, $p < .001$) in both conditions. The Spearman's rank correlation test shows a high correlation between the mean pick time ($r = .513$, $p = .001$), and a moderate correlation between the transfer time ($r = .469$, $p = .002$) in both conditions.

Furthermore, the equivalence test was significant for the total time ($t(39) = -1.89$, $p = .033$), pick time ($t(39) = -2.24$, $p = .016$), transfer time ($t(39) = -1.93$, $p = .030$), and place time ($t(39) = 2.17$, $p = .018$).

These results suggest that participants spent an equivalent amount of time to perform the task in both conditions, globally, and for each sub-task.

Accuracy: The paired-sample t-test shows no significant effect of the condition on the mean position distance ($t = 1.04$, $p = .305$), nor on the mean rotation difference ($t = 0.02$, $p = .983$).

In addition, the Pearson's correlation test shows a high correlation between the mean position distance in both conditions ($r = .672$, $p < .001$), and a high correlation between the mean rotation difference in both conditions ($r = .787$, $p < .001$).

Furthermore, the equivalence test was significant for both, the position distance ($t(39) = -3.36$, $p < .001$), and the rotation difference ($t(39) = -3.20$, $p = .001$).

We can observe that participants achieved the same accuracy in both conditions regardless of personal motor skills.

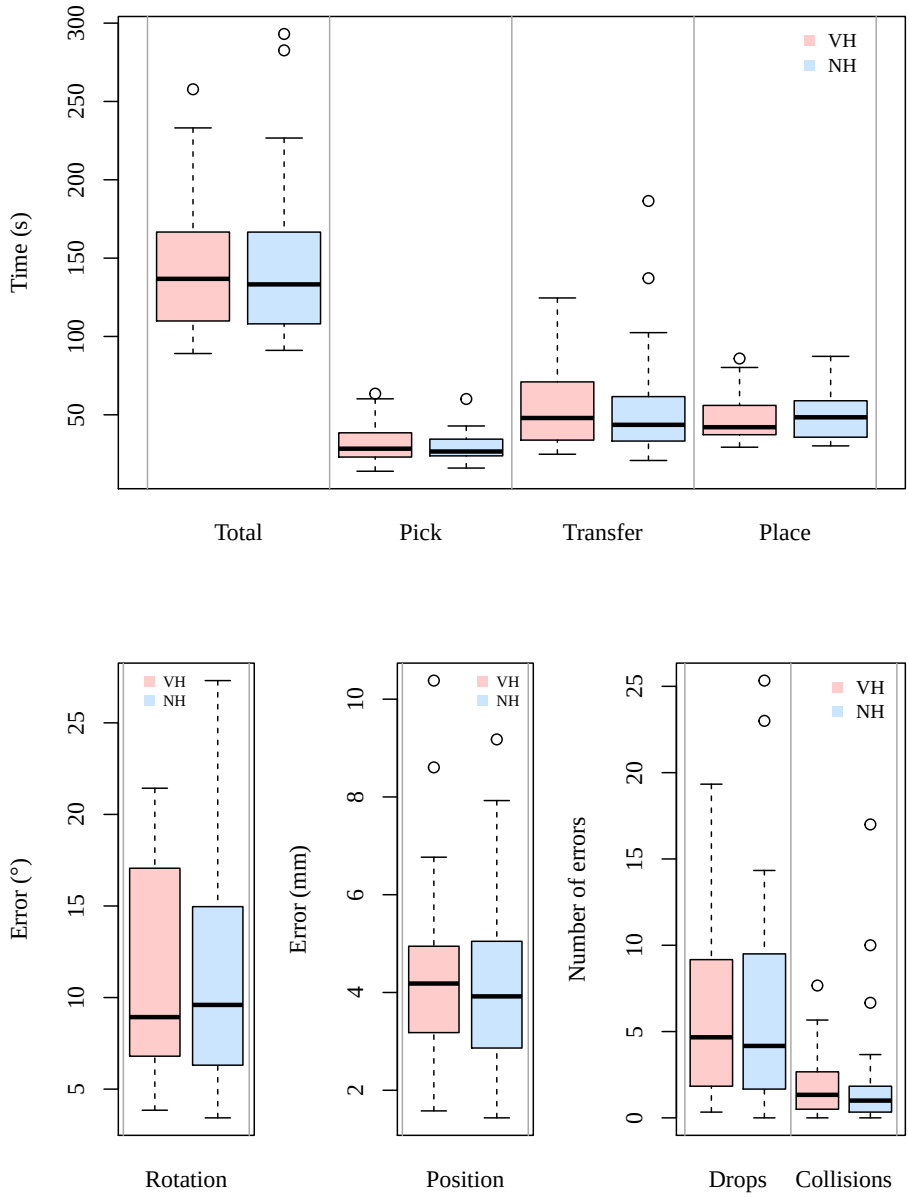


Fig. 4.5 The median value (with inter-quartile range –Q1-Q3) for all measures in the VR prototype for the VH and NH conditions. On the top: time measures; On the bottom: accuracy (rotation and distance), and errors (cubes drops and tool-wall collisions)

Errors rates: The non-parametric Wilcoxon test shows no significant effect of the condition on the mean total drops ($t = -0.20$, $p = 0.838$), nor on the mean total wall collisions ($t = -1.00$, $p = .317$).

In addition, the Spearman's rank correlation test shows a high correlation between the mean total drops ($r = .707$, $p < .001$), and a moderate correlation between the total wall collisions ($r = .426$, $p = .006$) in both conditions.

Furthermore, the equivalence test was significant for both error measures, ($t(39) = 2.21$, $p = .017$) and ($t(39) = 2.18$, $p = .018$) for total drops and total collisions, respectively.

These results show that error rates remained equivalent between both conditions too.

Table 4.1 Descriptive mean and standard deviation for the objective measures

Measure	**VH** μ (σ)	**NH** μ (σ)
Total time	144.09 (42.02)	144.08 (47.13)
Pick time	30.80 (11.47)	28.67 (8.37)
Transfer time	54.46 (26.32)	53.80 (32.25)
Place time	47.15 (14.28)	49.76 (14.99)
Rotation difference	10.92 (5.62)	10.90 (6.25)
Position distance	4.35 (1.78)	4.12 (1.80)
Total drops	5.83 (4.71)	6.33 (6.05)
Total collisions	1.81 (1.75)	1.83 (3.10)

Subjective evaluation

The analysis of grouped questions through the non-parametric Wilcoxon tests show no significant effect of the condition on any of the criteria: realism, the possibility to navigate and manipulate inside the environment, the possibility to examine, self-evaluation of the performance, haptic, and sense of ownership. However, a significant effect of the condition on the sense of agency can be observed ($Z = -2.29$, $p = .022$), with the mean score being significantly higher in the VH condition. Results are summarized in Table 4.2.

On the other hand, the non-parametric Wilcoxon tests show a significant effect of the condition on the participant's mean scores for questions Q8 (*"Were you able to anticipate what would happen next in response to the actions that you performed?"*) ($Z = -1.97$, $p = .049$), Q14 (*"How much delay did you experience between your actions and expected outcomes?"*) ($Z = -2.00$, $p = 0.045$), and Q24 (*"I felt that I was losing control of my hand when the virtual (hand/tool) was not responding correctly."*) ($Z = -2.32$, $p = .020$), with the

mean scores being significantly higher in the VH condition for Q8 and Q24, and the NH condition for Q14. No significant effects were found for the other questions.

Regarding the preference of the visualization of the hand, 37.5% of the participants preferred the visualization of their hands, 32.5% preferred to visualize only the virtual tools, and 30% felt no difference between the two options.

Finally, the SUS score reports a mean value of $\mu = 74.25$ ($\sigma = 13.61$), which stands for a grade B ("Good", percentile range 70 – 79) on the usability scale.

Table 4.2 Results of the grouped questions criteria mean comparison

			Wilcoxon test	
Criteria	**VH** μ (σ)	**NH** μ (σ)	Z	p
Realism	3.65 (0.60)	3.67 (0.59)	−0.21	.837
Possibility act	3.75 (0.53)	3.73 (0.64)	−0.18	.854
Quality of interface	2.47 (0.66)	2.48 (0.71)	−0.19	.854
Possibility examine	4.08 (0.63)	4.03 (0.61)	−0.15	.885
Self performance	3.65 (0.66)	3.77 (0.58)	−0.82	.415
Haptic	3.64 (0.86)	3.60 (0.71)	−0.43	.667
Ownership	3.58 (1.03)	3.60 (1.06)	−0.10	.918
Agency	3.17 (0.59)	2.91 (0.86)	−2.29	.022∗

$*p < .05$

4.2.8 Discussion

The results show that although the greatest number of participants preferred to visualize their virtual hands (37.5% of the participants), participants achieved the same performance for both conditions (presence and absence of the hands). Indeed, there was no significant difference between the two conditions for any objective measure (time, accuracy, and error measures). Moreover, the results show that all of these measures present a significant high to moderate correlation between the two experimental conditions and equivalent results. This does not allow us to validate our hypothesis on the participants' performance for the hand visualization (H2). These results contradict previous research, which has shown that the visualization of one's upper limb movements during a manual task could be beneficial for the training of motor skills (Castro-Alonso et al., 2014). A possible explanation for this finding may rely on the fact that the users were holding the physical handlers during the whole experimental session for our pick and place task.

This might have generated a form of embodiment beyond the body itself, allowing the participants to include the tools as if they were part of their bodies (Maravita and Iriki, 2004) as if the hands were elongated to the extremity of the tools. This phenomenon is coherent with the "ready-to-hand" (*zuhanden*) concept introduced by Martin Heidegger (Heidegger, 1927). By "ready-to-hand" Heidegger describes a tool that has become invisible, receding to the background of the work, and we are no longer conscious of it. This, in turn, allows our concentration to be focused on the work as opposed to the tools that we use to perform the work. In our case, the participants were more focused on the objects to be manipulated and the task rather than their hands. These results suggest that the user's animated hand's visualization is not necessary to perform a simple tool-based motor task in a VR simulator. This is an important finding for designing immersive VR motor skills simulators since obtaining a high-fidelity hand animation may be technically challenging and expensive. Furthermore, hand tracking devices, such as data gloves or optical cameras, add a small delay during the simulation. For our simulator, a delay was indeed perceived by the users in the visual hand condition (Q14: *How much delay did you experience between your actions and expected outcomes?*) compared to the tools-only visualization.

On the other hand, the subjective data analyses show that the users feel a higher sense of agency when visualizing their hands' real movements in the VE. This finding is coherent with previous results on existing VR simulators, including fully animated hands for motor skill trainers (Mohler et al., 2010; Ossmy and Mukamel, 2017). This is also supported by responses to question 8 (Q8: *Were you able to anticipate what would happen next in response to the actions that you performed?*) of the presence questionnaire where participants in the hand condition felt more confident controlling the tools and events in the simulator. Furthermore, the fact that participants experienced a higher sense of agency with the hands' real movements representation reinforces the importance of kinematics fidelity of the avatar to enhance the sense of embodiment.

On the contrary, the sense of ownership score was not significantly different between the two conditions. As supported by previous research, visual fidelity correlates with the sense of ownership. In our case, we have used human-like hands, but have kept a neutral skin color (that could be similar to wearing surgical gloves). In order to increase the sense of ownership, personalized gender, size, and skin-colored hands could be proposed to the users, similar to what was done for the second iteration of the needle insertion user study (see chapter 3). Therefore, H1 is validated, except for the sense of ownership.

Finally, in terms of usability, the SUS questionnaire reported a good value (74%), which encourages the choices made for this VR pick and transfer simulator's design.

4.2.9 Conclusions of the validation study

Through this study, we have investigated the impact of self-avatar visualization on a VR simulator for tool-based motor skills training.

We have designed and developed a VR simulator for a pick, transfer, and place task. Our main objective for this part was to investigate the impact of the users' hand visualization on task performance and user experience. In other words, whether having animated hands reproducing the users' movements allow users to achieve better performance than only visualizing the tools. A user experiment was designed and conducted to compare the users' performance for the two conditions (hand visualization and tools visualization) in the VR simulator. Results show that although users prefer to visualize their animated hands, they achieved the same performance on both VR conditions. These results suggest that the presence of the animated user's hand representation is not necessary when performing a tool-based motor skill task in a VR simulator. These findings have practical implications for the design of VR simulators for training tool-based motor skills tasks. Adding users' hand representation may require cumbersome and expensive additional devices. Therefore, to reduce the costs of VR trainers, designers can benefit from this study's findings to build cost-effective and less constraining VR simulators.

In this user study, we have investigated the impact of hand visualization on user's performance during only one experimental session. However, the impact of hand visualization on the long-term training of motor skills in immersive VR simulators and the transfer of these skills to the real world remains unanswered. These issues were investigated in the last part of this thesis project, through the following study.

4.3 Longitudinal study

After analyzing the impact of having a virtual hand representation in the VE on participants' performance in our motor skills VR simulator, we have decided to investigate whether these results would also apply for long-term motor skills training in immersive VR and if the learned skills transfer to a real-world setup. For this purpose, a new user study was designed. Our virtual reality prototype was used to train participants for a short period of time. Two training groups were formed. The first group was trained with the virtual hand representation, while the second group was trained without it (tools visualization only). In addition, a physical prototype was developed, implementing the same pick, transfer and place task, to study the transfer of skills learned in VR to a real-world task.

The study follows a pre- post- retention-test design, with two training groups and a control group. Our research questions for this work are:

- Would training groups significantly improve their performance in the VR simulators when compared to the control group?
- Would the experimental condition in the VR simulator impact the performance of the two training groups differently?
- Would the training in the VR simulator improve the trainees' performance in the physical prototype?

Therefore, we hypothesize that:

- H1) Participants in the training groups will improve their performance (i.e., time, accuracy, and errors) in the VR simulator when compared to the control group.

  We hypothesize that the training groups will benefit from the training sessions to automatize the motor task execution, expecting an improvement from the pre-test to the post-test. In addition, if the training were consequent enough, we would expect retention of this training after a short pause, which means that the performance at the post-test will be no significantly different from that at the retention-test. On the other hand, we can expect a small improvement for the control group between the three evaluation phases (pre-, post- and retention-test). However, we think that the training groups' improvement will be significantly greater than the control group's improvement.

- H2) Participants who trained under the virtual hand representation will improve their performance in the VR simulator compared to participants who trained in the no hand condition.

In the first study, we did not observe any differences in performance for a single experimental session. Nonetheless, we think that users can improve differently under the two conditions after a longer exposure time to the simulator. In particular, we hypothesize that the hand-visualization training condition will more positively affect the users' performance by providing an additional spatial frame of reference. Therefore, the group trained under the hand-visualization condition is expected to reach a higher performance level in the post-test than the group trained under the no-hand visualization condition

- H3) Participants who trained in the VR simulator will improve their performance in the physical prototype as compared to the control group.

  As mentioned before, both the VR simulator and the physical prototype implement the same tool-based pick and place task. Therefore, we expect that the motor skill learned in the VR prototype will transfer to the physical setup for both training groups.

4.3.1 Participants

Previous studies show that a minimum of 6 participants per group is required for this type of study (Linsk et al., 2018; Zhang et al., 2013). Therefore, 30 participants from the University (students and staff) were enrolled in this study to prevent tiredness due to the study's length. A total of 26 participants (16 males, 10 females) were retained for the study ($N = 26$). The mean age was $\mu = 37.00$ ($\sigma = 10.87$). Nineteen of them were right-handed, six left-handed (23%), and one ambidextrous. All of them had normal or corrected to normal vision, and twelve wore their correction glasses during the experiment. All of them had previous experience with video games (including smartphone games), with thirteen of them playing video games regularly. Thirteen participants reported a previous experience with 3D virtual environments (VEs), and nine participants used haptic devices before this experiment (mainly in demonstrations or previous user studies). Thirteen participants have previously used head-mounted displays (HMDs), with only two of them regularly using them, and for no more than 30 minutes per session.

The experimental protocol was approved by the institutional ethics committee of *Université Paris-Saclay* (CER Paris Saclay) prior to enrolling any human subject. An informed written consent was obtained from all the subjects involved in this study prior to their participation.

Participants were divided into three groups: the control group (CG) and two training groups (HT and TT). In Table 4.3 we summarize the characteristics of each group.

Table 4.3 Demographics summary by group

		CG	HT	TT	Total
Participants	N	9	9	8	26
Gender	Female	3	3	4	10
	Male	6	6	4	16
Dominant hand	Left	3	2	1	6
	Right	5	7	7	19
	Ambidexter	1	0	0	1
Age	< 30	4	3	4	11
	30 - 45	2	3	2	7
	> 45	3	3	2	8
Glasses	Worn during experiment	5	2	5	12
Video games	Play regularly	5	5	3	13
3D VEs	Previous experience	3	6	4	13
HMD	Previous experience	3	6	4	13
Haptic devices	Previous experience	2	4	3	9

4.3.2 Experimental design

The experiment followed a three-way mixed-design with one between-subjects factor, the group, with three levels: control group (CG), hand-training group (HT) and tool-training group (TT); and two within-subjects factors, the phase, with four levels: the pre-test (PRE), the training (TRAIN), the post-test (POST) and the retention-test (RET), and two conditions for the VR prototype, the presence of the animated virtual hands (VH) and the absence of them (NH).

All participants ($N = 26$) performed the PRE, POST and RET phases. Each of these phases consisted of a single experimental session (75 minutes mean duration) where participants performed three different tests of the pick, transfer and place task (the same as that performed in the previous experiment). The first test was performed in the physical prototype (real-world task simulator) and the two others on the VR prototype (one for each condition). The order of presentation of the two conditions in the VR prototype was counterbalanced for each group (CG, HT and TT). All the participants started with the physical prototype test for all three phases.

The training phase consisted of ten training sessions on the VR prototype to be performed in two successive weeks (30 minutes mean duration each). The VR prototype

condition used in these sessions depended on the participants' group. Only HT and TT groups participated in the training phase. For the HT group, the condition used in all the ten sessions was the VH condition, and for the TT group, the NH condition.

4.3.3 Experimental task

The experimental task for the longitudinal study was the same as that used for the validation study (PT&P) (see subsection 4.2.3).

4.3.4 Apparatus

Two prototypes were developed for this study, a physical and a VR simulator, both implementing the same task (PT&P).

VR prototype: No changes were performed to the VRP with respect to the validation study (see subsection 4.2.4).

Physical prototype: It consisted of 6 small plastic cubes (dimensions: 2 cm side-length, mass: 29 g) with a green sticker on one of the faces, two forceps tools, two gloves, and a working table. The working table consisted of a scanner with the final targets' position of the cubes (represented by squares), and a wooden wall (dimensions: l = 30 cm, w = 1 cm, h = 18 cm) placed in the middle of it. The initial cube positions were specified on each side of the scanner. Both the initial and final positions of the cubes were numbered. A computer display was placed in front of the participants to instruct them before the task, with a camera on top of it to record the whole session (see Figure 4.6).

Passive markers were attached to each forceps tool and the working table to configure three tracking rigid-bodies. An OptiTrack[10] camera system, which consisted of six infrared cameras, was capable of detecting the passive markers and communicating their position and orientation through a VRPN client-server application. Tools position and orientation data were collected but are not included in the analysis of the current results.

4.3.5 Experimental procedure

All the participants (N = 26) started with PRE phase (day 0). Before the experiment, participants were asked to read and sign the consent form and complete the demographics questionnaire. They were then asked to read the experimental instructions presenting

[10] https://optitrack.com

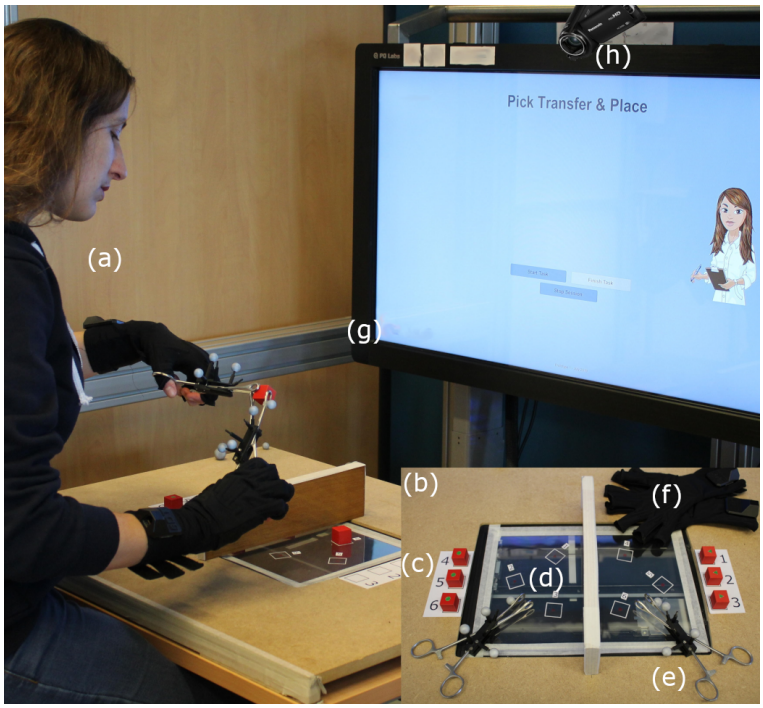


Fig. 4.6 The physical prototype for the PT&P task. (a) User interacting with the system (b) working table composed of the scanner and the wall (c) cubes (d) cubes' targets (e) forceps (f) gloves (g) screen for the instructions (h) recording camera

the prototypes and explaining how to use the different devices in both setups and how to perform the task.

After that, they were installed in the physical prototype area. They were instructed to sit facing the instructions screen, put on the gloves and grab the two forceps. They started then the familiarization session, which consisted of trying the system by performing a pick, transfer and place of one cube on each side of the wall. At this point, the participants would feel comfortable with both the tools' manipulation and the task completion. After the familiarization, the experimental session started by performing three trials of the experimental task on the physical prototype to establish their baseline performance on this system (real-world task). Prior to each trial, the experimenter placed the six cubes in their initial positions, and at the end of each trial the scanner was used to capture an image with the six cubes' final positions. After finishing the experimental session on the physical prototype, the participants switched to the VR prototype area. They were asked to put in the gloves and the HMD. For the VH condition, a calibration of the data gloves was performed before the familiarization session. The same steps were then followed for the VR prototype: familiarization session and experimental session (3 trials of the experimental task) on the VR prototype for the first condition. In these tests, the

instructions were shown in the virtual screen displayed inside the HMD. This procedure was repeated for the second condition. Finally, the participants were asked to answer a system usability questionnaire for the VR simulator.

This procedure was repeated in the POST (day 0 + 2 weeks) and RET (day 0 + 3 weeks) phases for all the participants. At the end of the RET phase, participants were asked to answer a comparative questionnaire between the physical and the VR setups.

The training (HT and TT) groups ($N = 17$) participated in the training phase, which consisted of ten training sessions of 30 minutes each, during two weeks. The first training session took place the day after the PRE phase (day 0 + 1 day), and the last one the day before the POST phase. In each training session, the participants were installed in the VR prototype area. They were asked to put in the gloves and the HMD. For participants in the HT group, a calibration of the data gloves was performed for each session. The participants were then asked to perform the experimental task ten times or the maximum number of trials possible in 20 minutes. They were allowed to rest their arms and remove the HMD between each trial if they felt tired or uncomfortable. At the end of the training phase (training session number 10), participants answered a post-training questionnaire.

See Figure 4.7 for the experimental procedure summary.

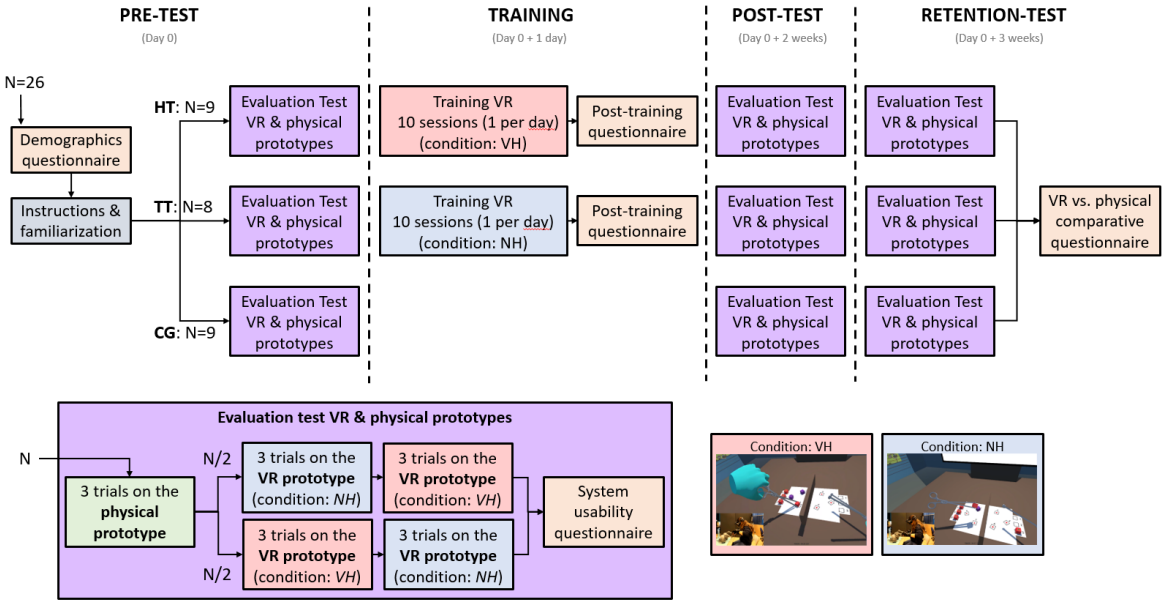


Fig. 4.7 Experimental protocol for the longitudinal study for the control group (CG), and the two training groups (HT: hand-training, and TT: tool-training)

4.3.6 Data collection and analyses

Both objective and subjective measurements were recorded.

The objective data corresponded to the user's performance, which was evaluated through the same measures used in the validation study (see subsection 4.2.6): the accuracy of placing the cubes, the task completion time and the number of errors during the task (the number of cubes dropped and the number of collisions of the tools with the wall). For the virtual prototype, all the data was automatically recorded on a log file. For the physical prototype, measures were obtained through the analysis of the video recordings (for time and errors) and the scanner images of the cubes' final positions (for accuracy, described in Appendix B). For this study analysis, we focus only on the total time, both accuracy measures (position and rotation), and the errors of dropping a cube.

The subjective data consisted of responses to three questionnaires. All of them used a 5-point Likert scale: the "System Usability Questionnaire" already described for the validation study, a comparative questionnaire for the two prototypes (VR vs. physical prototypes), and a post-training questionnaire. The comparative questionnaire evaluated the perceived similarity between the two setups in terms of the tools and cubes' weight, the shape and comfort of the tools, the force feedback and gripping force, the realism of the manipulation, and the similarity of the two situations as a whole. The post-training questionnaire evaluated the perceived performance, involvement, contribution of each session, quantity of training sessions, duration of each session, learning contribution, and interference of training with other daily tasks. The post-training statistical analysis is detailed in Appendix D.

Finally, participants were asked to comment on their experience with the system during the debriefing session (end of the experiment), see Appendix D. All the questionnaires can be found in Appendix C.

Data analyses

All data analyses were performed using SPSS software version 21.0 (IBM Corp., Armonk, NY, USA) with the appropriate statistical tests. We have used a confidence level of 95% for all our statistical analyses. Therefore, a result is considered significant when $p < .05$.

The analysis of the objective data is divided into three stages. Firstly, we used a three-way mixed-design ANOVA on the pre-, post- and retention-test participants' performance on the VR prototype (with two test conditions –VH and NH), Tukey's HSD post hoc test, as well as pairwise comparisons with Bonferroni correction. Secondly, we analyzed the learning curve for each training group (HT and TT) to better understand the learning

process. Finally, we determined whether the training in the VR prototype led to a transfer of skills to the physical prototype by using a two-way mixed-design ANOVA on the pre-, post- and retention-test participants' performance on the physical prototype, Tukey's HSD post hoc test, and pairwise comparisons with Bonferroni correction.

Outliers were removed from the data and replaced with the group means before analysis. For the VR prototype, two data points on the total time, three on the rotation difference and total drops, and four on the position distance were considered outliers. For the physical prototype, three data points on the rotation difference measure were considered as outliers. Furthermore, residuals normal distribution for each cell of the study was validated through visual inspection of Q-Q plots, and homogeneity of variances through Levene's test. For the two- and three-way interaction effects, Mauchly's test of sphericity was calculated, and Huynh-Feldt correction was applied when sphericity was not assumed (ϵ is reported).

For the subjective data analysis, a two-way mixed ANOVA was used to analyze the mean differences in the "System Usability" questionnaire participants' global score. Finally, we present a descriptive analysis of the comparative questionnaire (VR vs. physical prototypes).

4.3.7 VR prototype results

The first step in the data analysis aims to determine whether users' performance in the VR prototype is different between groups through time. For this purpose, we analyze whether participants' performance on the VR prototype changes through time (pre-, post-, retention-test) for the different groups and conditions. The mean and standard deviation of each measure is summarized in Table 4.4 Results of the three-way mixed-design ANOVA (three-way and two-way interaction effects, plus main effects) are summarized in Table 4.5. In addition, one-way ANOVAs were performed for the PRE phase measurements to compare groups' initial performance (baseline) for both conditions. This aimed to determine whether all groups had a comparable performance at the beginning of the experiment. Moreover, to assess the training impact, independent-samples t-tests were applied to compare groups performance between PRE and POST phases for each condition separately, and paired-samples t-test were employed to compare for each group participants' performance between the two conditions.

Table 4.4 Mean and standard deviation (μ (σ)) for all time measures in the VR prototype

Condition	Group	Total time (μ (σ))	Position (μ (σ))	Rotation (μ (σ))	Drops (μ (σ))
PRE-TEST					
VH	CONTROL GROUP	142.65 (41.13)	3.99 (1.04)	9.99 (5.80)	5.63 (3.12)
	HAND TRAINING	138.45 (42.84)	4.35 (1.77)	12.17 (5.98)	7.04 (7.00)
	TOOL TRAINING	120.88 (19.37)	4.07 (0.78)	9.74 (3.85)	5.50 (2.44)
NH	CONTROL GROUP	122.68 (14.42)	4.39 (2.34)	9.40 (4.79)	5.22 (4.43)
	HAND TRAINING	142.08 (14.42)	4.51 (2.12)	12.56 (8.48)	6.59 (5.55)
	TOOL TRAINING	144.56 (44.04)	4.42 (1.19)	10.28 (5.16)	5.95 (2.29)
POST-TEST					
VH	CONTROL GROUP	104.41 (9.02)	3.78 (1.28)	8.74 (3.20)	4.12 (2.22)
	HAND TRAINING	74.82 (15.10)	1.88 (0.80)	3.26 (1.00)	1.11 (1.21)
	TOOL TRAINING	72.28 (12.35)	1.86 (0.60)	5.58 (3.34)	1.42 (1.32)
NH	CONTROL GROUP	115.93 (22.69)	4.59 (1.36)	9.61 (4.91)	5.19 (3.47)
	HAND TRAINING	75.43 (11.45)	2.54 (1.24)	3.65 (1.20)	1.24 (2.07)
	TOOL TRAINING	76.20 (11.17)	2.56 (0.98)	5.15 (2.37)	1.23 (0.56)
RETENTION-TEST					
VH	CONTROL GROUP	96.95 (16.40)	3.17 (0.91)	7.50 (6.00)	4.41 (3.62)
	HAND TRAINING	73.29 (10.79)	2.12 (1.19)	4.83 (5.18)	0.93 (1.33)
	TOOL TRAINING	69.25 (12.57)	2.52 (1.68)	6.03 (4.31)	1.13 (0.85)
NH	CONTROL GROUP	106.53 (12.84)	3.06 (1.09)	8.15 (5.40)	3.83 (2.05)
	HAND TRAINING	70.78 (7.96)	2.23 (0.95)	3.52 (1.43)	0.52 (0.63)
	TOOL TRAINING	71.31 (11.62)	2.69 (1.64)	5.73 (3.17)	1.44 (1.31)

Baseline

Participants' baseline performance (PRE) in the VR prototype is summarized in Table 4.4. One-way ANOVAs indicate that groups in both conditions had a similar performance at the beginning of the study ($p > .05$), indicating that all the participants had a similar performance level at the beginning of the experiment.

Performance

Hereafter we present the results of the three-way mixed-design ANOVA for each objective measure considered (see Table 4.5 for interaction and main effects, and Figure 4.8 for the measures plots). Only significant results for each measure are reported below.

Table 4.5 Results of the two-way mixed design ANOVAs for each of the dependent variables in the VR prototype. Interaction and main effects of group and phase are presented

Measure	**Phase × Condition × Group**	**Phase × Condition**	$\epsilon^{\diamond}$
Total time	$F(2.7,31.2) = 3.01$, $p = .050$, $\eta^2 = .207$	$F(1.4,31.2) = 0.93$, $p = .836$ $\eta^2 = .004$	.678
Position	$F(3.7,42.1) = 0.14$, $p = .960$, $\eta^2 = .012$	$F(1.8,42.1) = 2.41$, $p = .106$, $\eta^2 = .091$	.914
Rotation	$F(4,46) = 0.73$, $p = .576$, $\eta^2 = .060$	$F(2,46) = 0.26$, $p = .773$, $\eta^2 = .011$	
Drops	$F(4,46) = 0.30$, $p = .879$, $\eta^2 = .025$	$F(2,46) = 0.21$, $p = .808$, $\eta^2 = .009$	

Measure	**Group × Phase**	**Phase**	$\epsilon^{\star}$
Total time	$F(2.5,28.2) = 5.28$, $p = .008*$, $\eta^2 = .314$	$F(1.2,28.2) = 86.76$, $p < .001*$, $\eta^2 = .790$	.614
Position	$F(4,46) = 5.60$, $p = .001*$, $\eta^2 = .328$	$F(2,46) = 32.66$, $p < .001*$, $\eta^2 = .587$	
Rotation	$F(4,46) = 4.16$, $p = .008*$, $\eta^2 = .266$	$F(2,46) = 18.26$, $p < .001*$, $\eta^2 = .443$	
Drops	$F(2.9,33.2) = 3.42$, $p = .030*$, $\eta^2 = .229$	$F(1.4,33.2) = 22.90$, $p < .001*$, $\eta^2 = .499$	.721

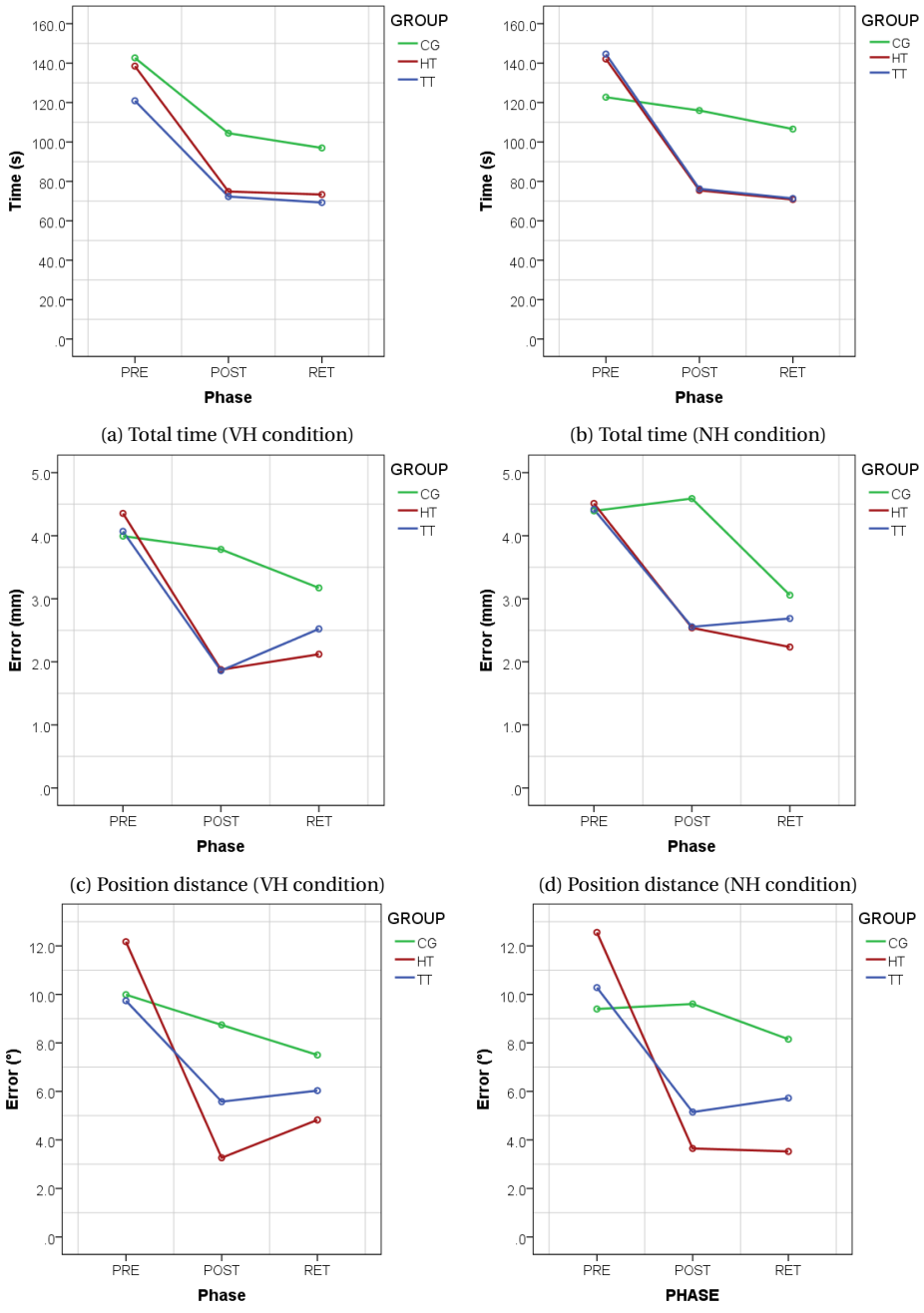
Measure	**Condition × Group**	**Condition**
Total time	$F(2,23) = 0.81$, $p = .458$, $\eta^2 = .066$	$F(2,23) = 1.11$, $p = .302$, $\eta^2 = .046$
Position	$F(2,23) = 0.36$, $p = .965$, $\eta^2 = .003$	$F(2,23) = 6.62$, $p = .017*$, $\eta^2 = .223$
Rotation	$F(2,23) = 0.16$, $p = .851$, $\eta^2 = .014$	$F(2,23) < 0.01$, $p = .952$, $\eta^2 < .001$
Drops	$F(2,23) = 0.20$, $p = .823$, $\eta^2 = .017$	$F(2,23) < 0.01$, $p = .979$, $\eta^2 < .001$

Measure	**Group**
Total time	$F(2,23) = 6.38$, $p = .006*$, $\eta^2 = .357$
Position	$F(2,23) = 1.98$, $p = .162$, $\eta^2 = .147$
Rotation	$F(2,23) = 1.04$, $p = .369$, $\eta^2 = .083$
Drops	$F(2,23) = 2.90$, $p = .075$, $\eta^2 = .201$

$*\ p < .05$

$\diamond$ *sphericity not assumed for the phase × condition factor, Huynh-Feldt correction used*

$\star$ *sphericity not assumed for the phase factor, Huynh-Feldt correction used*



(a) Total time (VH condition)

(b) Total time (NH condition)

(c) Position distance (VH condition)

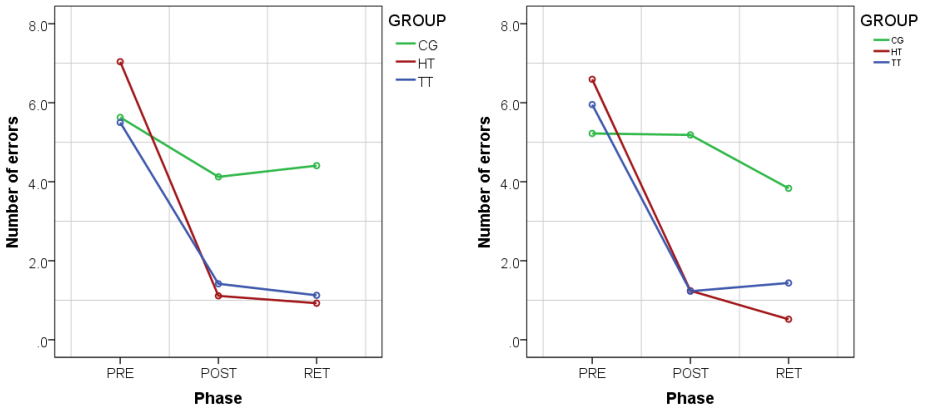
(d) Position distance (NH condition)

(e) Rotation difference (VH condition)

(f) Rotation difference (NH condition)

Fig. 4.8 Plot images of the different performance measures for each group at each phase for each condition for the virtual reality prototype

Time: For the total time, no significant three-way interaction effect between the phase, the condition and the group was found. In addition, no significant two-way interaction effects were found, neither between phase and condition nor between group and condition. A significant two-way interaction effect between phase and group was found



(g) Total drops (VH condition)

(h) Total collisions (NH condition)

Fig. 4.8 Plot images of the different performance measures for each group at each phase for each condition for the virtual reality prototype (*Cont.*)

($F(2.5,28.2) = 5.28$, $p = .008$, $\eta^2 = .314$, $\epsilon = .614$). Finally, a significant main effect of phase ($F(1.2,28.2) = 86.76$, $p < .001$, $\eta^2 = .790$, $\epsilon = .614$) and a significant main effect of group ($F(2,23) = 6.38$, $p = .006$, $\eta^2 = .357$) were found. No significant main effect of the condition was observed. Tuckey's HSD post hoc test showed a significant difference between control group and both, hand-training ($p = .024$) and tool-training groups ($p = .009$). Pairwise comparisons with Bonferroni correction showed that total time at PRE was statistically significantly different from the total time at POST and RET phases ($p < .001$), and POST was statistically significantly different from RET phase ($p = .011$).

Accuracy: No significant three-way interaction effect was found for any of the accuracy measures. In addition, there were no significant two-way interaction effects between phase and condition, nor between group and condition for these measures. There was a significant two-way interaction effect of group and phase for both, the position distance ($F(4,46) = 5.60$, $p = .001$, $\eta^2 = .328$), and the rotation difference ($F(4,46) = 4.16$, $p = .008$, $\eta^2 = .266$) accuracy measures. A significant main effect of phase was also found for both measures ($F(2,46) = 32.66$, $p < .001$, $\eta^2 = .587$, and $F(2,46) = 18.26$, $p < .001$, $\eta^2 = .443$; respectively for position and rotation). No significant main effect of group was found for these measures, and a significant main effect of the condition was found only for the position distance measure ($F(2,23) = 6.62$, $p = .017$, $\eta^2 = .223$). Pairwise comparisons with Bonferroni correction showed that both accuracy measures performance at PRE were statistically significantly different from those at POST phase ($p < .001$), suggesting learning, and no significant difference was found for these measures between POST and RET phases ($p < .746$, $p = 1.00$; respectively for position and rotation), suggesting no skill decay.

Errors: No significant three-way interaction effect was found for the cubes dropped. Moreover, there were no significant two-way interaction effects between phase and condition, nor between group and condition. A significant two-way interaction effect between group and phase was observed ($F(2.9, 33.2) = 3.42$, $p = .030$, $\eta^2 = .229$). There was a significant main effect of the phase ($F(1.4, 33.2) = 22.90$, $p < .001$, $\eta^2 = .499$). No significant main effect of the condition was found, neither a group main effect. Pairwise comparisons with Bonferroni correction showed that PRE and POST were statistically significantly different for the total cubes dropped ($p < .001$), suggesting learning, and no significant difference was found between POST and RET error ($p = .900$), indicating no skill decay one week after the training period.

To summarize, the results indicate that participants improved their task completion time as an effect of the training phase, with the training groups achieving better performance than the control group. We can also observe that participants in the control group have a higher and more accelerated improvement for time when evaluated in the VH condition. It can also be observed that participants, in general, continue to learn from post- to retention-test, suggesting that total time can still be improved. Participants' accuracy was measured as the distance error and angle difference between the placed cube and the target square. For both measures, we can observe that participants improved between pre- and post-test, suggesting a learning effect, while keeping almost the same performance between post- and retention-test, indicating that they were able to retain the learned skills after training stopped. Finally, the same results can be observed for the total cubes dropped, where participants improved significantly from pre- to post-test, and maintained their performance in the retention-test. As a whole, these results suggest that training in VR (both groups) improved the participants' performance as compared to the control group.

Learning

The change in performance from PRE to POST as an indicator of learning was analyzed for each evaluation condition separately by comparing each group 2-by-2. In addition, for each group, we compared this change in performance between the two conditions.

VH condition: Results from individual unpaired t-tests showed that the change in performance for the hand-training group compared to the control group was significantly different for both accuracy measures: the position distance ($t = -3.05$, $p = .008$), and the rotation difference ($t = -3.08$, $p = .007$). The change in performance for the tool-training

group compared to the control group was significantly different for the position distance measure ($t = -3.41$, $p = .004$). No other differences were found.

NH condition: Results from individual unpaired t-tests showed that the change in performance for the hand-training group compared to the control group was significantly different for the total time ($p = -4.00$, $p = .001$), the position distance ($t = -2.93$, $p = .010$), the rotation difference ($t = -3.44$, $p = .007$), and the total drops ($t = -2.16$, $p = .046$) measures. The change in performance for the tool-training group compared to the control group was significantly different for the total time ($t = -3.53$, $p = .003$), the position distance ($t = -3.33$, $p = .007$), the rotation difference ($t = -2.74$, $p = .015$), and the total drops ($t = -2.25$, $p = .040$) measures. No differences were found between the two training groups (HT and TT).

VH vs. NH condition Results from individual paired-samples t-tests showed that the change in performance between the two conditions was not significantly different for any of the four objective measures for the hand-training group, neither for the tool-training group, nor for the control group.

To summarize, for the VR prototype with the VH condition (virtual hand representation), participants trained under that condition significantly improved their accuracy compared to the control group, but not the total time and the errors made, although having better results than the control group. This can be explained by the fact that the VR prototype with the VH condition allowed the CG to significantly improve their time in only two sessions. The other training group (TT), which trained without the virtual hand representation, also improved their accuracy in the VR prototype with the VH condition compared to the control group, suggesting transfer from NH to VH condition. Both training groups significantly improved their performance (for all measures) in the VR prototype with the NH condition when compared to the control group. This also shows transfer from the VH training condition to the NH condition. Finally, the improvement from pre- to post-test was not different between the two conditions nor for the training groups, neither for the control group. This suggests that users' performance improvement in the VR prototype is not affected by the condition (virtual hand visualization vs. tools only visualization).

4.3.8 Learning curves results

In addition to participants' performance in the evaluation phases (pre-, post-, retention-test), we wanted to analyze the learning curve (ten training sessions) between the two groups (HT and TT) across time. For that purpose, we carried on a series of two-way mixed-design ANOVAs (training groups × training sessions) on each of the dependent measures registered for the VR prototype.

Two-way interaction of group and training sessions and main effects are presented in Table 4.6. In addition, Figure 4.9 presents for each measure the evolution of performance for each training group through time.

The two-way mixed-design ANOVA results show that no main effect of group was observed for any of the four measures, and that a main effect of the training session was presented in all of them. No other statistically significant results were found.

In summary, these results show that participants in the HT group (who trained visualizing their virtual hand representation) learned in the same way as participants in the TT group (who trained visualizing only the virtual tools). Both groups improved over time (training sessions).

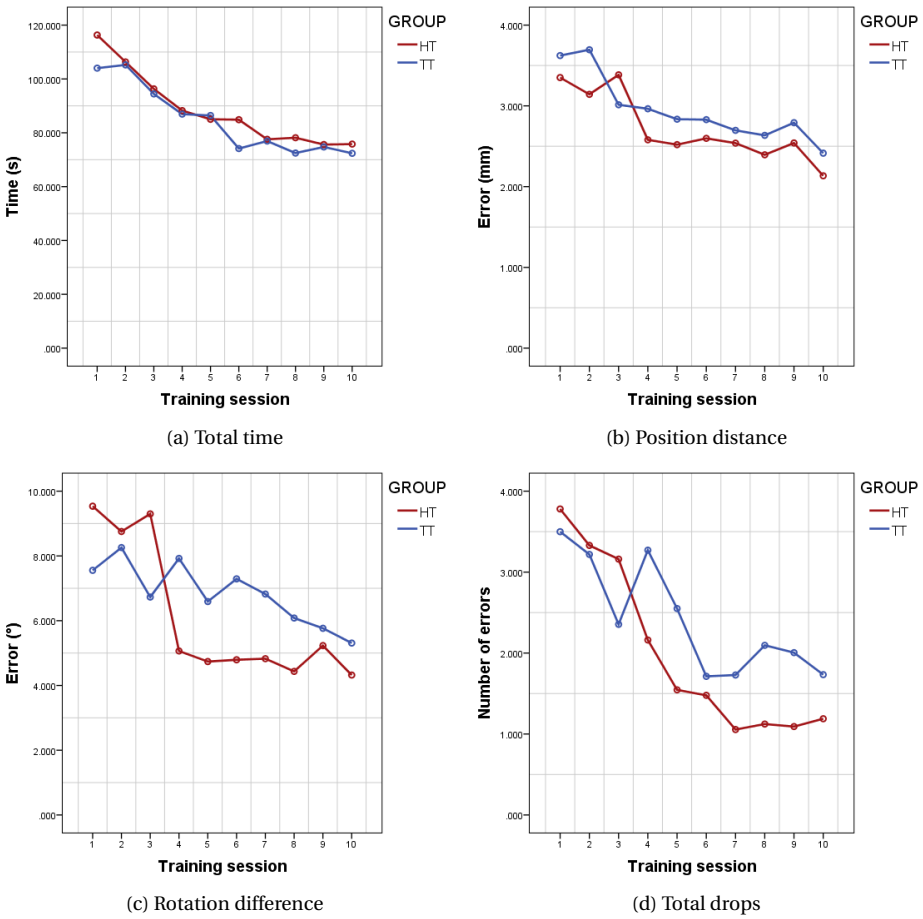
Table 4.6 Results of the two-way mixed design ANOVAs for each of the dependent variables in the VR prototype learning curves. Interaction and main effects of group and training sessions are presented

Measure	**Group**	**Training session**
Total time	$F(1,15) = 0.22,\ p = .648,\ \eta^2 = .014$	$F(3.0,45.0) = 21.60,\ p < .001*,\ \eta^2 = .590$
Position	$F(1,15) = 0.21,\ p = .655,\ \eta^2 = .014$	$F(5.9,88.8) = 9.41,\ p < .001*,\ \eta^2 = .385$
Rotation	$F(1,15) = 0.29,\ p = .597,\ \eta^2 = .019$	$F(2.1,31.5) = 6.43,\ p = .004*,\ \eta^2 = .300$
Drops	$F(1,15) = 0.42,\ p = .525,\ \eta^2 = .027$	$F(4.6,69.6) = 6.22,\ p < .001*,\ \eta^2 = .293$

Measure	**Group × Training session**	$\epsilon^\star$
Total time	$F(3.0,45.0) = 0.65,\ p = .589,\ \eta^2 = .041$	.333
Position	$F(5.9,88.8) = 0.83,\ p = .550,\ \eta^2 = .052$	.658
Rotation	$F(2.1,31.5) = 2.91,\ p = .067,\ \eta^2 = .162$	.234
Drops	$F(4.6,69.6) = 0.95,\ p = .452,\ \eta^2 = .059$	.516

$*\ p < .05$

$\star$ *sphericity not assumed for the training session factor, Huynh-Feldt correction used*



(a) Total time

(b) Position distance

(c) Rotation difference

(d) Total drops

Fig. 4.9 Plot images of the different learning curves for each measure for each training group for the virtual reality prototype

4.3.9 Additional findings: skills transfer results

In addition to the analysis of the VR prototype measures, we were interested in investigating whether participants' performance in the physical prototype was affected by VR training. For this purpose, we have investigated whether participants' performance on the physical prototype significantly changed through time (pre-, post-, retention-test) for the training groups and control group. The means and standard deviations for each measure are summarized in Table 4.7 Results of the two-way mixed-design ANOVA (interaction and main effects of phase and group) are summarized in Table 4.8. In addition, one-way ANOVAs were performed for the PRE phase measurements to compare groups' initial performance (baseline), and independent-samples t-tests were applied to compare groups' performance between PRE and POST phases (training impact).

Table 4.7 Mean and standard deviation (μ (σ)) for all time measures in the physical prototype

Group by Phase	**Total time** (μ (σ))	**Position** (μ (σ))	**Rotation** (μ (σ))	**Drops** (μ (σ))
PRE-TEST				
CONTROL GROUP	50.70 (20.15)	2.26 (0.83)	4.30 (1.34)	0.04 (0.11)
HAND TRAINING	49.48 (10.54)	2.33 (0.93)	4.52 (1.37)	0.07 (0.15)
TOOL TRAINING	39.43 (6.05)	2.07 (0.74)	3.21 (1.04)	0.08 (0.15)
POST-TEST				
CONTROL GROUP	43.73 (11.59)	2.21 (0.79)	3.91 (1.04)	0.07 (0.15)
HAND TRAINING	46.46 (5.18)	1.87 (0.66)	2.48 (1.08)	0.44 (0.41)
TOOL TRAINING	39.72 (7.65)	1.94 (1.08)	2.69 (0.79)	0.54 (0.67)
RETENTION-TEST				
CONTROL GROUP	37.46 (8.87)	2.19 (0.83)	4.05 (1.17)	0.07 (0.15)
HAND TRAINING	42.23 (6.78)	1.87 (0.79)	2.66 (1.12)	0.33 (0.53)
TOOL TRAINING	39.06 (6.54)	1.56 (0.39)	2.59 (0.97)	0.50 (0.31)

Baseline

Participants' baseline performance (PRE) in the physical prototype are summarized in Table 4.7. One-way ANOVAs indicate that groups had a similar performance ($p > .05$). In conclusion, we can say that participants from all the groups reported the same level of performance at the beginning of the experiment.

Performance

Hereafter we present the results of the two-way mixed-design ANOVA for each objective measure considered (see Table 4.8 for interaction and main effects, and Figure 4.10 for the measures plots). The significant results are detailed for each measure below.

Table 4.8 Results of the two-way mixed design ANOVAs for each of the dependent variables in the physical prototype. Interaction and main effects of group and phase are presented

Measure	**Group**	**Phase**
Total time	$F(2,23) = 1.24$, $p = .309$, $\eta^2 = .097$	$F(1.8,40.2) = 7.31$, $p = .003*$, $\eta^2 = .241$
Position	$F(2,23) = 0.68$, $p = .517$, $\eta^2 = .056$	$F(2,46) = 2.28$, $p = .113$, $\eta^2 = .090$
Rotation	$F(2,23) = 5.16$, $p = .014*$, $\eta^2 = .310$	$F(2,46) = 9.07$, $p < .001*$, $\eta^2 = .283$
Drops	$F(2,23) = 4.43$, $p = .024*$, $\eta^2 = .278$	$F(2,46) = 6.14$, $p = .004*$, $\eta^2 = .211$

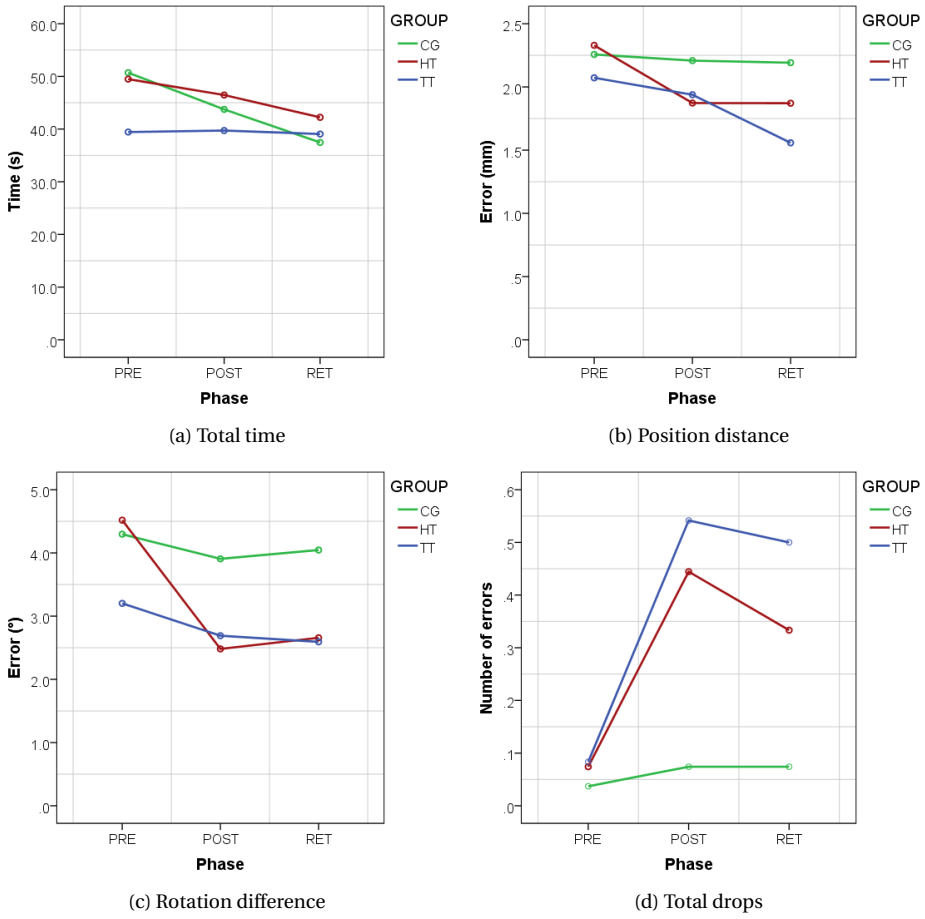
Measure	**Group × Phase**	$\epsilon^\star$
Total time	$F(3.5,40.2) = 2.06$, $p = .112$, $\eta^2 = .152$	.874
Position	$F(4,46) = 0.64$, $p = .639$, $\eta^2 = .053$	
Rotation	$F(4,46) = 2.75$, $p = .039*$, $\eta^2 = .193$	
Drops	$F(4,46) = 1.28$, $p = .292$, $\eta^2 = .100$	

$*\, p < .05$

$\star$ *sphericity not assumed for the phase factor, Huynh-Feldt correction used*

Time: There was a significant main effect of phase on the total time ($F(1.8,40.2) = 7.31$, $p = .003$, $\eta^2 = .241$, $\epsilon = .874$), but no difference as a function of group. No interaction between the two factors was observed. Pairwise comparisons with Bonferroni correction showed that total time did not improve between PRE and POST phases ($p = .314$), but it significantly improved between PRE and RET ($p = .012$) and POST and RET ($p = .020$).

Accuracy: Neither a phase effect, nor a group effect was found for the position distance accuracy measure. No interaction between the two factors was observed for this measure either. There was a significant main effect of the phase on the rotation difference ($F(2,46) = 9.07$, $p < .001$, $\eta^2 = .283$), as well as a group main effect on this measure ($F(2,23) = 5.16$, $p = .014$, $\eta^2 = .310$). Moreover, an interaction effect between phase and group was also observed ($F(4,46) = 2.75$, $p = .039$, $\eta^2 = .193$). Tukey's HSD post hoc test showed no difference between the two training groups ($p = .604$), nor between the CG and HT groups ($p = .91$), and a significant difference between CG and TT ($p = .013$). Pairwise comparisons with Bonferroni correction showed that PRE and POST were statistically significantly different ($p = .004$), suggesting learning, and no significant difference was



(a) Total time

(b) Position distance

(c) Rotation difference

(d) Total drops

Fig. 4.10 Plot images of the different performance measures for each group at each phase for the physical prototype

found between POST and RET rotation ($p = 1.00$), indicating no skill decay with one-week time after the training period.

Errors: There was a significant main effect of the phase on the total number of dropped cubes ($F(2,46) = 6.14$, $p = .004$, $\eta^2 = .211$), a group main effect ($F(2,23) = 4.43$, $p = .024$, $\eta^2 = .278$), and an interaction between the two factors ($F(4,46) = 1.28$, $p = .292$, $\eta^2 = .100$). Tukey's HSD post hoc test showed no difference between the two training groups ($p = .687$) or between the CG and HT groups ($p = .113$), and a significant difference was found between CG and TT ($p = .023$). Pairwise comparisons with Bonferroni correction showed that PRE and POST were statistically significantly different ($p = .012$), showing a negative learning effect on errors, but no significant difference between POST and RET rotation was found ($p = 1.00$).

In conclusion, the results indicate that participants' total time was improved, but not as an effect of the training phase in the VR prototype, since the training groups' completion time is not different from the control group at the end of the experiment. In addition, participants continue to improve between POST and RET phases, suggesting that they did not reach their best task completion time yet (see Figure 4.10a). Participants' accuracy, measured as the angle difference between the placed cubes and the target squares, improved between PRE and POST, with the training groups achieving better results than the control group. This suggests a learning effect due to the training phase in the VR simulator. In addition, no difference is reported between POST and RET, indicating that they retained the learned skill after a short time period without exposure (see Figure 4.10c). On the other hand, although we see that training groups improved their placement accuracy, measured as the distance error between the placed cubes and the target squares, compared to the control group, no statistically significant differences were found (see Figure 4.10b). Finally, we can see an increment in the number of errors between PRE and POST phases for both training groups, suggesting a negative effect of training in the VR prototype. In addition, this negative learning is retained after one week of pause (see Figure 4.10d).

Learning

The change in performance from PRE to POST as an indicator of learning was analyzed. Results from individual unpaired t-tests showed that the change in performance for the HT group was significantly different from that of the CG only for the rotation difference accuracy measure ($t = -2.29$, $p = .036$), and drops error measure ($t = 2.14$, $p = .048$). The rotation difference measures were found significantly different between HT and TT ($t = 2.22$, $p = .043$). No other differences were found.

In conclusion, after comparing the improvement (performance at PRE minus performance at POST) between each group, we confirm that training groups performance improvement (in the physical prototype) with respect to the control group is low, meaning that only a small transfer of skills is observed after training in the VR prototype.

4.3.10 Subjective data results

System Usability

In addition to the objective measures, we investigated the participants' subjective evaluation of the VR prototype in terms of usability. More particularly, we were interested in exploring whether the perceived system usability was the same across the three groups and across time.

The results of the two-way mixed-design ANOVA on the global SUS scores showed that there was no significant main effect of the phase, but a main group effect ($F(2,23) = 10.90$, $p < .001$). There was also an interaction effect between the two factors ($F(3.5,39.9) = 3.37$, $p = .022$, $\epsilon = .867$). Tukey's HSD post hoc test showed no difference between the two training groups ($p = .326$), a significant difference was observed between CG and both HT ($p < .001$) and TT ($p = .017$) groups. No other differences were found (see Table 4.9 for the mean SUS scores for each group among phases).

These results highlight the fact that participants in the training groups subjectively perceived the VR prototype as more usable at the end of the experiment as compared to the beginning of the experiment. In addition, their evaluation of usability (general score) was significantly higher than the control group.

Table 4.9 System usability global score for the VR prototype (group × phase)

Group by Phase	Score μ (σ)	Grade	Percentile range
PRE-TEST			
CONTROL GROUP	70.28 (14.06)	C	41 – 59
HAND TRAINING	80.28 (10.49)	A−	85 – 89
TOOL TRAINING	71.56 (15.64)	C+	60 – 64
POST-TEST			
CONTROL GROUP	64.72 (7.75)	C−	35 – 40
HAND TRAINING	86.94 (6.47)	A+	96 – 100
TOOL TRAINING	83.13 (11.16)	A	90 – 95
RETENTION-TEST			
CONTROL GROUP	65.56 (12.49)	C	41 – 59
HAND TRAINING	90.28 (6.78)	A+	96 – 100
TOOL TRAINING	84.06 (12.67)	A+	96 – 100

Virtual reality vs. physical prototypes

In addition to analyzing participants' performance in the physical prototype, we have used a comparative questionnaire between the two prototypes (see Table 4.11 for the

questions used) to determine which are the main perceived differences, and identify the potential improvement directions of our VR prototype.

The mean scores and standard deviations for each question are summarized in Table 4.10. In general, the results suggest that improvements are still to be made, notably regarding the force feedback in the VR simulation.

Table 4.10 VR vs. physical prototype individual questions statistical results

Question ID	**CG** (μ (σ))	**HT** (μ (σ))	**TT** (μ (σ))	**All participants** (μ (σ))
VRvsP-Q1	3.33 (0.71)	4.56 (0.53)	4.25 (0.89)	4.04 (0.87)
VRvsP-Q2	2.56 (0.88)	3.67 (0.87)	4.00 (0.76)	3.38 (1.02)
VRvsP-Q3	2.67 (0.87)	3.67 (1.23)	3.25 (1.28)	3.19 (1.17)
VRvsP-Q4	3.22 (1.09)	3.89 (0.60)	3.50 (0.93)	3.54 (0.91)
VRvsP-Q5	2.89 (1.17)	4.22 (0.44)	3.50 (1.07)	3.54 (1.07)
VRvsP-Q6	3.22 (1.09)	3.67 (0.87)	3.75 (1.58)	3.54 (1.17)
VRvsP-Q7	3.00 (0.71)	4.00 (0.50)	3.50 (0.93)	3.50 (0.81)
VRvsP-Q8	2.56 (0.53)	3.67 (0.50)	3.00 (1.07)	3.08 (0.85)
VRvsP-Q9	3.00 (0.71)	4.11 (0.60)	3.63 (0.52)	3.58 (0.76)

Table 4.11 VR vs. physical prototype questionnaire

ID	**Question**
VRvsP-Q1	How would you rate the feeling that your real hand was controlling virtual tools?
VRvsP-Q2	How did you find the resemblance of the weight of virtual tools compared to real tools?
VRvsP-Q3	How did you find the resemblance of the weight of the virtual cubes compared to the real cubes?
VRvsP-Q4	How would you describe the grip of virtual tools compared to real tools?
VRvsP-Q5	How would you rate the comfort of using virtual tools compared to real tools?
VRvsP-Q6	How did you find the shape of the handles used to manipulate the virtual tools?
VRvsP-Q7	How would you rate the resemblance of the sensations felt between the virtual and real manipulations?
VRvsP-Q8	How would you describe the force feedback during the interaction in VR compared to the real world?
VRvsP-Q9	How did you perceive the two situations (real and virtual) that have just been presented to you?

Answers interpretation: 1 = Not at all realistic / similar, 5 = extremely realistic / similar

4.3.11 Discussion

VR prototype for learning

The results of the longitudinal study show that participants improved their task completion time, accuracy (measured as the distance error and angle difference between the placed cube and the target square), and errors (measured as the total cubes dropped) as an effect of the training phase, with the training groups achieving a better whole performance than the control group. These results imply learning in the VR prototype and allow us to validate H1. On the one hand, the training groups reduced between pre- and post-test 40% to 48% the general mean task completion time, 42% to 56% the distance accuracy on placement, 42% to 73% the rotation difference, and 74% to 84% the total number of cubes dropped. On the other hand, the control group reported lower improvements than the training groups, less than 5% reduction of performance between pre- and post-test for all the VR prototype measures under the no hand condition, and between 5% and 27% for the virtual hand condition. In addition, no differences were reported between the hand-training and tool-training as a function of group.

Moreover, for both accuracy measures (rotation and position) and the total number of errors, the results also show that participants kept almost the same performance between post- and retention-test, indicating that they were able to retain the learned skills after training stopped. These results reinforce the learning effectiveness of the VR simulator for these performance measures. However, the task completion time was significantly different between the post- and retention-test, which means that participants, in general, continue to improve their speed after training. They reduced 2% to 8% the total task completion time between these phases (post and retention), suggesting that there is still a margin for improvement in terms of the total task completion time before reaching a plateau.

As a whole, these results suggest that the training of the pick and place task in our VR simulator (both groups) improved the participants' performance as compared to the control group (H1 validated for all the accuracy measures).

Hand visualization training condition

The results show that for the VR prototype with the virtual hand condition, participants that trained under that condition (hand-training group) significantly improved their accuracy as compared to the control group, but not the total time and the errors made, although having better results than participants in that group (control group). This can

be explained by the fact that the VR prototype with the virtual hand condition allowed the control group to significantly improve their task completion time in only two sessions. Indeed, participants in the control group had a greater and more accelerated improvement for task completion time when evaluated in the virtual hand condition, 25% time reduction for virtual hand condition compared to 6% for the no hand condition. Participants in the other training group (tool-training), which trained without the virtual hand representation, also improved their accuracy in the VR prototype with the virtual hand condition compared to the control group, suggesting a small transfer of learning for accuracy from the no-hand to the virtual hand condition.

Moreover, both training groups significantly improved their performance (for all measures) in the VR prototype with the no hand condition compared to the control group. This also shows a learning transfer in the other direction, from the virtual hand training condition to the no hand condition, for all the performance measures.

On the other hand, the results also show that the participants in the hand-training group were more accurate than the participants in the tool-training group in aligning the cubes during placement (rotation difference accuracy), with a greater improvement (between pre- and post-test) of 70% and 73% for the hand-training group, compared to 43% and 49% for the tool-training group, respectively for the evaluation in the VR prototype under the virtual hand and no hand conditions. However, no main effect of the condition was found in the three-way ANOVA for this measure.

Additionally, when considering the comparison of learning (pre minus post) between the two conditions analyzed separately for each individual group, the improvement was not different between the two conditions, nor for the training groups, neither for the control group. These results suggest that the users' performance improvement in the VR prototype is not affected by the condition (virtual hand visualization vs. tools only visualization). It does not allow us to validate our second hypothesis (H2).

Finally, the learning curve results show that participants learned in the same way for both training groups, significantly improving along time (training sessions) with no significant differences as a function of the condition. Nonetheless, a better and more accelerated improvement is visualized for both accuracy measures, where the hand-training group presented the best results.

As a whole, these results suggest that training in VR is not affected by the users' hand representation and control. Indeed, only small (no significant) improvements concerning the rotation difference accuracy measure in favor of having the hands representation were found. In line with the first study, this further confirms that self-avatar representation is

not essential for a tool-based manipulation task when having a moderate-fidelity hand representation.

Finally, the SUS score reports a mean value of $\mu = 79.81$ ($\sigma = 15.10$) at the end of the user experiment, which stands for a grade A- ("Good", percentile range 85 – 89) on the usability scale. Moreover, participants in the training groups subjectively perceived the VR simulator as more usable at the end of the experiment as compared to the beginning, and their scores were significantly higher than the control group. This indicates that the users' perceived system usability improved after a certain amount of training, suggesting that they started adopting the system.

Real-world performance after training

The results for the physical prototype show that participants' task completion time was improved. However, this improvement is not associated with the training in the VR prototype since the control group improved after the three trials as well.

On the other hand, an improvement in participants' position accuracy can be observed for both training groups, suggesting that the skills learned on the VR simulator are being transferred to the physical simulator, although this improvement was not significant.

In addition, participants' rotation accuracy improved between pre- and post-test, with the training groups achieving better results than the control group, and maintained between post- and retention-test, indicating retention of the learned skills after a short time without exposure. The improvement between pre- and post-test suggests a learning effect due to the VR simulator training phase, although learning (pre-test−post-test) was significant only for the hand-training group. Nonetheless, we might not that easily conclude that there was an effect of the training condition on these results, i.e., that only the visualization of the hand transferred the skills, because the participants in the tool-training group were extremely accurate since the beginning of the experiment, and further studies might be done to reach a conclusive decision.

Finally, the results show an increase in the number of errors between the pre-test and the post-test phases for both training groups compared to the control group, suggesting a negative effect of training in the VR prototype. In addition, this negative effect is retained after one week of pause. This finding can be related to the lack of gripping force in the VR simulator. Indeed, participants in the training groups got accustomed to the perceptive absence of forceps gripping force during the two-week training period. The plastic handles only provided sensing capabilities of realistically opening/closing the virtual tool in the VR simulator, but participants did not perceive a force effect when grabbing the cube. In addition, the answers to the comparative questionnaire for VR vs. physical prototypes

support this hypothesis. Most of the questions assessing the realism of the VR simulator to the real-world setup reported participants' scores between 3 and 3.5 (average quality), and in particular for force feedback modality. This issue is in line with the conclusions of chapter 2 regarding the need for a more in-depth study of surgical VR simulators' haptic fidelity characteristics.

As a whole, these results suggest that training in VR transfers to the real-world setup in a low portion. This allows us to validate our H3 hypothesis only for the accuracy performance, more particularly for the rotation difference measure. However, further exploration is required to determine whether this improvement in accuracy is related to the users' hands representation.

To summarize, our VR simulator has been shown to be effective for training a VR tool-based pick, transfer and place motor task (H1 was validated). However, the results do not support the need for the users to visualize their hands while interacting with this system (H2 was not validated). Further analysis of the learning curves and the application of correlation and equivalence tests can more vigorously support these conclusions. In addition, the motor task simulated in this work is similar to the tasks being simulated in existing surgical trainers (see chapter 1), which requires the operator to manipulate physical tools while interacting with virtual objects. Therefore, we can generalize these results on self-avatar to other surgical VR simulators. Moreover, our results show that only a small transfer of skills to the real-world setup is observed. Notably, the accuracy of aligning the cube with the target square (rotation difference) was improved in the physical prototype after training during two weeks in the VR simulator (H3 was partially validated). On the other hand, the number of cubes dropped (errors) was negatively affected by the physical interface's current limitations in our VR simulator. Finally, the system's usability reports a good score, encouraging some of the choices made in our simulator. However, other aspects, such as the force feedback, need to be appropriately addressed to build a more effective system. For that purpose, the proposed framework for evaluating the haptic display fidelity will be helpful in future studies.

4.4 Conclusions

Summary

In the second part of this thesis work, we have explored how the fidelity of secondary aspects of the simulation, such as the self-user representation, might influence the training of a primary motor task and the transfer to the real-world setup. Our main objective was to investigate the impact of self-avatar representation and control on a tool-based manipulation task. To answer this question, we have first proposed in chapter 2 a self-avatar fidelity model to characterize the users' representation and control of their virtual hand. The analysis of the implemented self-avatar, based on this model, indicates a high to moderate fidelity representation. We have also presented the design and implementation of two prototypes, an immersive VR simulator and a physical prototype, of a tool-based pick, transfer and place task. We have also presented a validation study and a longitudinal study to determine the impact that the user's hand movement visualization can have on the user's perception, performance, training outcomes, and skills transfer.

Findings

The first study results indicate that the users reached an equivalent task performance regardless of the hand representation condition. This suggests that a high fidelity of the user's hand representation is not necessary to perform the tool-based pick, transfer and place task (Ricca et al., 2020).

In addition, the longitudinal study shows that users improve their performance after a two-week training period in the VR simulator for both training conditions (visualization/absence of the virtual hand) when compared to the control group (Ricca et al., 2021b). This confirms that hand representation fidelity does not impact the learning curve of the users' technical skills in the VR system. Finally, although the system obtained a good usability score (participants subjective evaluation), only a marginal transfer of knowledge to the real-world task is observed. This reveals a need to continue improving the VR simulator's user interface, particularly the haptic feedback component, to ensure better skills transfer to the real world.

Limitations and perspectives

One limitation of our VR prototype is related to the plastic handles connected to the haptic interfaces. Some participants mentioned that they were uncomfortable using it,

which decreased the system's usability. Besides, several handles had to be made because they were fragile and broke after a couple of experimental sessions. To overcome these ergonomics and comfort issues, we have already improved the handles' design using a 3D scan of an ergonomic scissors handle with curved edges and a more robust design (see Appendix A). Nonetheless, solutions are to be explored on how to provide the gripping forces through them. Indeed, for this study, the gripping forces were reported as crucial for learning how to perform the tool-based bimanual task correctly. A future step would be to study the added value of using gripping force rendering for manipulation motor tasks in surgical VR simulators.

As for system improvements, the registration of accuracy measures in the physical prototype could be automatized (see Appendix B). We have used a scanner combined with image processing algorithms on the resulting images to capture the cubes' final placement in our setup. Although this solution is accurate in calculating the cubes' position and rotation differences (with respect to the targets), it is not robust to the users' subsequent movements (after placement), through which they can accidentally move a cube already placed. In addition, it requires a break in between trials to allow the scanner to capture the image. An attractive solution would be to use magnets inside the cubes to continuously track their positions and orientation (Taylor et al., 2019).

Moreover, another improvement for our VR simulator is to include a training debriefing report or use a feedback metaphor during simulation (Jeanne et al., 2017; Zhu et al., 2020) in order to inform the users about their performance. Indeed, several participants (second study) felt the need to know their improvement during the two-week training phase.

Finally, it would be interesting for the VR surgical simulator to add a new training group, which would train in the physical simulator. Indeed, this can help us better understand and characterize the proficiency measures required to be an expert in the simulation.

Conclusion

***

Our work on simulator fidelity and 3D interactions aims to improve the design of virtual reality surgical simulators. This conclusion summarizes the results of this research and opens up new perspectives.

Summary

The objective of this thesis work was to understand the simulator fidelity approach to design virtual reality surgical simulators to guarantee effective training and transfer of technical skills to the real-world procedure.

To investigate the impact of fidelity on surgical simulators design, we have adopted an **iterative design methodology**, which consisted of studying existing simulators from a human-computer interaction perspective and focusing on how simulator fidelity can contribute to the design of virtual reality surgical simulators. We have first studied existing surgical simulators to analyze the tasks to be trained and the interactions and artifacts necessary to accomplish effective training and transfer to the real world. The approach consists of identifying particular use-cases about a component of the simulator in order to design the appropriate interaction techniques to:

1. perform them in VR,
2. characterize their level of simulator fidelity,
3. evaluate them through user studies,
4. extract design guidelines that can be then generalized to other surgical simulators.

The literature review allowed us to understand the main concepts related to our object of study: *virtual reality simulators for learning technical surgical skills.* Two main types of surgical simulators are currently being used for technical surgical training: part-task

trainers and virtual reality simulators. On the one hand, part-task trainers are resource-intensive and do not provide an objective evaluation. On the other hand, virtual reality systems lack formal validation, limiting their adoption in current training curricula.

Based on the observation that current surgical virtual reality systems do not fully meet users' requirements, and their design is usually technology-driven, the objective of this work was **to make recommendations and guidelines for the design of virtual reality surgical simulators from an interaction design and simulator fidelity points of view**. Our analysis suggests that the definition of effective interactions for virtual reality medical simulators would impact their usability and effectiveness for training purposes. We concluded from the study of existing surgical simulators that besides the main task to simulate in surgical VR systems, there are other interactions and factors worth being analyzed for certain procedures that are not the main focus of current systems, e.g., **changing the point of view of the surgical workspace** and their impact on users' performance on the main tasks, **the user's self-avatar representation and control** in surgical simulators (i.e., the users' hands representation) and their impact on learning outcomes, and the **haptic feedback** impact on users' performance and transfer of technical skills. The common inherent question of the design of these aspects is the level of fidelity of the interactions and artifacts to include in the simulators. Indeed, the fidelity of a simulator is strongly associated with the transfer of skills to the real world, which is the main goal of surgical simulators (i.e., training on them must impact performance in the real procedure). Therefore, it is important to explore the design of surgical virtual reality simulators on this concept.

Previous research about simulator fidelity does not point out what levels of fidelity a designer can use to include these components (viewpoint changing task, self-avatar representation, haptic feedback) in a VR surgical simulator. It is unclear whether high interaction fidelity is necessary for all the interaction aspects or if lower fidelity approaches are enough. The state-of-the-art presented in this thesis work led us to some approaches to define simulator fidelity and analyze some of the dimensions it can be decomposed. However, we identified a lack of guidelines on how to fully characterize them and their influence on human performance, user experience, learning outcomes, and the transfer of skills learned through simulation to the real-world setup.

Through this thesis work we grow the knowledge about simulator fidelity dimensions, by extending the theoretical model, and by studying two components for surgical virtual reality simulators: viewpoint changing task and self-avatar representation and control. Through several user experiments we studied how the fidelity of the interactions and the

fidelity of the user's representation in the virtual environment can impact technical skills performance and training.

Contributions

In this section we summarize the main contributions of this thesis project.

→ Extension of the simulator fidelity model

An extended analysis of simulator fidelity is presented in chapter 2. Several dimensions and approaches on defining this concept were presented and discussed. McMahan (2011) and Bowman and McMahan (2007) simulator fidelity model that includes the FIFA framework was discussed as being the most suitable approach to be used to examine the fidelity of the interactions to be included in virtual reality surgical simulators. However, this model suffers from some limitations: it does not provide a framework for analyzing some components such as the haptic display fidelity and the self-avatar fidelity, which were identified as important for surgical simulators. Considering these limitations, our first contribution in this thesis work was to extend this model to iteratively incorporate and evaluate the suitable components for developing virtual reality surgical simulators. In this context, a haptic display fidelity model was proposed, based on Bowman and McMahan (2007) contributions for visual display fidelity; and a self-avatar representation and control fidelity model was defined, based on our previous work contributions for full-body avatar fidelity Gamelin et al. (2020). In the first part of this thesis project, we employed the existent interaction fidelity model (FIFA) for the first time in the surgical domain for viewpoint changing techniques analysis. Moreover, our proposed self-avatar representation and control fidelity model was employed in the second part of this project for a tool-based task.

→ The design of viewpoint changing techniques in virtual reality surgical simulators

In order to enrich the simulator fidelity model, we have chosen to focus on the viewpoint changing interaction fidelity, for which we have designed two interaction techniques with two levels of fidelity and which were compared through user studies. The results (of two iteration studies) show that the touch-based viewpoint changing technique improves the users' task completion performance during needle insertion while maintaining a similar

level of needle manipulation accuracy as compared to the head-tracking technique. These results suggest that high interaction fidelity is not always necessary when designing navigation techniques for surgical trainers. This study also highlights the importance of designing appropriate interactions for secondary tasks because they can influence the user's primary task performance in VR simulators.

→ The impact of avatar control and representation in immersive virtual reality motor skills trainers

We have studied how the fidelity of the users' self-avatar representation can influence a tool-based motor-skill task in an immersive VR simulator. For this purpose, we have used the new proposed model for self-avatar fidelity to characterize the level of avatar fidelity for users' hands in VR simulators. A validation study and a longitudinal study were conducted to compare the users' learning outcomes with and without visualizing their virtual hands. The results of the first study indicate that the users reached an equivalent task performance regardless the hand representation condition in a single experimental session. This suggests that high fidelity user's hand representation is not necessary to perform a tool-based motor task. In addition, the longitudinal study results indicate, first, that users improve their performance after a two-week training period, as compared to a control group, in a VR motor-skills trainer regardless of visualizing their hands representation during training; and second, that the hand representation does not impact the transfer of motor skills to the real-world. These results show that hand representation fidelity does not impact, neither the learning curve, nor the transfer of motor skills in VR systems for tool-based tasks.

Design recommendations

This thesis work permits to propose some recommendations for the design of virtual reality surgical simulators from an interaction and simulator fidelity points of view. We describe some recommendations to take into account when designing virtual reality surgical simulators for technical training. They are related to the two aspects studied in this thesis work: viewpoint changing techniques and user representation and control.

→ To support navigation tasks in virtual reality needle insertion trainers, designers should consider several aspects:

1. The realism of the surgical task can change how participants interact with the virtual reality system, therefore, scenario fidelity aspects are worth being explored during the design of this type of simulators.
2. The haptic feedback is required for simulating the needle guidance inside the virtual environment sub-task.
3. Stereoscopic view is required for simulating the needle tissue sampling sub-task.
4. Touch-based navigation metaphors should be considered for changing the point of view of the VE for fine motor skills main manipulation tasks.
5. The use of shadows of the instruments facilitates precision tasks.

→ To support self-avatar representation and control for tool-based motor skills training, designers should consider several aspects:

1. Gripping forces are to be included for tool-based tasks that require handling objects (e.g., pick and place), since its absence can lead to negative learning outcomes.
2. Including high fidelity representation and control of hand avatars is not necessary to improve user performance and training on VR technical skills simulators.
3. Hand avatars should be included only to improve the user experience through increasing ownership and agency.

Limitations and and short-term perspectives

The difficulties faced and limits of the scope of our contributions, as well as the research perspectives are discussed hereafter.

→ For our first study on the needle insertion simulator, the actual position of the haptic device with respect to the monitor was not optimal, which might have lowered the interaction fidelity of the system as a whole. The main problem seems to come from the fact that the users visualize their real hands manipulating the haptic device, and they see in addition the representation of the virtual needle in another position, hence presenting a mismatch between the visual and motor coordinate systems. The use of a mirror to see a co-located representation of the virtual needle and the user's real hand, or the use of HMD devices might fix this issue. This will increase

the visualization perspective fidelity component proposed in our model of hand avatar fidelity.

→ For our tool-based motor skills VR trainer, the plastic handles connected to the haptic interfaces lacked of gripping force feedback, which resulted in bad training by participants. In addition, ergonomic, fragility and comfort issues were raised during the two experimental studies. To overcome these issues, we have already worked on an improved model based on the 3D scan of an ergonomic scissors handle with curved edges and a more robust design. We think that providing comfortable interfaces might allow the users to fully concentrate on the task to perform and not being distracted by them.

→ One of the limitations of the user studies for navigation techniques is that transfer validation was not assessed. Indeed, a longitudinal study is required to verify that the skills trained in the VR simulator are positively reflected in the real-world procedure. As previously stated in this work, for some surgical procedures transfer assessment is complex and may present ethical and safety issues.

→ A limitation that prevents further generalization of our results on self-avatar representation to other motor skills and VR simulators can be attributed to the multisensory integration. Indeed, further studies are required to understand the contribution of each component: the visual feedback that originates from the tool and the hand, the functional characteristics of the tool (shape and dimensions) obtained through real tool manipulation, and the force feedback perceived through the physical device (tool). Therefore, it would be interesting to study in the future the impact of hand representation in the absence of haptic feedback, the use of purely virtual tools instead of physical ones, and the comparison between generic VR controllers and alike physical and virtual tools.

→ The two studies of viewpoint changing techniques in needle insertion simulation differ in the performance results of the distinct fidelity interaction metaphors used, suggesting that depending on the realism of the task, high or moderate fidelity navigation techniques perform differently. Therefore, it would be important to study and classify the scenario fidelity components for our simulators to determine which aspects have influenced the users' performance, and also if there is an interaction between the scenario and interaction fidelity components. Moreover, a combination of high and moderate interaction fidelity metaphors could be also explored.

→ Scenario fidelity dimension is not fully characterized, and there are currently no models to classify and characterize the fidelity of the elements associated with it. Further studies are required to enrich the simulator fidelity framework and investigate the possible impacts it may have on surgical VR simulators.

→ The influence of haptic fidelity on user's performance remains an open issue. Although the relevance of the model proposed for considering the haptic rendering capabilities in the design of VR surgical trainers, user studies are required to determine the impact of each component in technical skills learning and transfer.

Long-term perspectives

The long term perspective of this work is to reach a more complete fidelity framework for surgical simulators. In this work, simulator fidelity was characterized through three main dimensions: the display fidelity, the interaction fidelity and the scenario fidelity. Contributions were made through the proposition of individual models for haptic display fidelity and self-avatar representation and control fidelity. These models refine the whole fidelity framework allowing the characterization of two individual components inherent to surgical virtual reality simulators. Nonetheless, much work is needed to reach an structured framework able to fully characterize the fidelity of each aspect of the simulation separately, to determine its impact in training and transfer of technical surgical skills, but also to establish the interaction between the different fidelity dimensions. In this context, the scenario fidelity dimension remains unexplored. For surgical technical training, aspects related to the realism of the task, the physics algorithms and visual models of objects and tools might have an impact on how users assimilate training in VR. Moreover, haptic display fidelity plays a crucial role for motor-skills training. Several works have shown that practitioners performance improve on haptics-enabled simulators. Nonetheless, more studies are required to determine which aspects of the haptic display fidelity model are required at a high degree of realism. Finally, the main objective of virtual reality surgical simulators is to guarantee that the training acquired in virtual reality can be transferred to the real-world setup, therefore, the fidelity framework should also take into account the psychological level of learning, which can add a new dimension concerning the constructive process of learning and how it can be adapted and integrated in virtual reality. At the end, the fidelity framework should be used to better choose the fidelity degree of each component that favors training and renders a suitable learning experience to the user.

References

***

Abolhassani, N., Patel, R., and Moallem, M. (2007). Needle insertion into soft tissue: A survey. *Medical engineering & physics*, 29(4):413–431.

Abrahamson, S., Denson, J., and Wolf, R. (2004). Effectiveness of a simulator in training anesthesiology residents. *Quality and Safety in Health Care*, 13(5):395–397.

Agha, R. A., Fowler, A. J., and Sevdalis, N. (2015). The role of non-technical skills in surgery. *Annals of Medicine and Surgery*, 4(4):422–427.

Akhtar, K. S. N., Chen, A., Standfield, N. J., and Gupte, C. M. (2014). The role of simulation in developing surgical skills. *Current Reviews in Musculoskeletal Medicine*, 7(2):155–160.

Alexander, A. L., Brunyé, T., Sidman, J., and Weil, S. A. (2005). From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games. *DARWARS Training Impact Group*, 5:1–14.

Arbogast, P. and Rosen, J. (2012). Simulation in Plastic Surgery Training: Past, Present and Future. In Agullo, F., editor, *Current Concepts in Plastic Surgery*, chapter 12, pages 235–256. InTech.

Argelaguet, F. and Andujar, C. (2013). A survey of 3d object selection techniques for virtual environments. *Computers & Graphics*, 37(3):121–136.

Argelaguet, F., Hoyet, L., Trico, M., and Lécuyer, A. (2016). The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE Conference on Virtual Reality (VR)*, pages 3–10. IEEE.

Arikatla, V. S., Sankaranarayanan, G., Ahn, W., Chellali, A., De, S., Cao, C. G. L., Hwabejire, J., DeMoya, M., Schwaitzberg, S., and Jones, D. B. (2013). Face and construct validation of a virtual peg transfer simulator. *Surgical endoscopy*, 27(5):1721–1729.

Arora, A., Khemani, S., Tolley, N., Singh, A., Budge, J., Varela, D. A. D. V., Francis, H. W., Darzi, A., and Bhatti, N. I. (2012). Face and content validation of a virtual reality temporal bone simulator. *Otolaryngology–Head and Neck Surgery*, 146(3):497–503.

Arsenault, R. and Ware, C. (2000). Eye-hand co-ordination with force feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, pages 408–414. ACM.

Ayodeji, I., Schijven, M., Jakimowicz, J., and Greve, J. (2007). Face validation of the simbionix lap mentor virtual reality training module and its applicability in the surgical curriculum. *Surgical endoscopy*, 21(9):1641–1649.

Backlund, P., Engström, H., Gustavsson, M., Johannesson, M., Lebram, M., and Sjörs, E. (2009). Sidh: a game-based architecture for a training simulator. *International Journal of Computer Games Technology*, 2009.

Balcombe, J. (2004). Medical training using simulation: toward fewer animals and safer patients. *Alternatives to Laboratory Animals*, 32(1_suppl):553–560.

Barbé, L., Bayle, B., and de Mathelin, M. (2006). Bilateral controllers for teleoperated percutaneous interventions: evaluation and improvements. In *Proceedings of the American Control Conference*, ACC '06, pages 3209–3214. IEEE.

Basdogan, C., De, S., Kim, J., Muniyandi, M., Kim, H., and Srinivasan, M. A. (2004). Haptics in minimally invasive surgical simulation and training. *IEEE computer graphics and applications*, 24(2):56–64.

Batmaz, A. U., de Mathelin, M., and Dresp-Langley, B. (2018). Effects of image size and structural complexity on time and precision of hand movements in head mounted virtual reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 167–174. IEEE.

Baum, D. R., Smith, D. A., Hirshfeld, S. F., Klein, G. A., and Swezey, R. W. (1982). Specification of training simulator fidelity: A research plan. Technical report, DTIC Document.

Beaubien, J. M. and Baker, D. P. (2004). The use of simulation for training teamwork skills in health care: how low can you go? *BMJ Quality & Safety*, 13(1):51–56.

Benyahia, S., Van Nguyen, D., Chellali, A., and Otmane, S. (2015). Designing the user interface of a virtual needle insertion trainer. In *Proceedings of the 27th Conference on l'Interaction Homme-Machine*, page 18. ACM.

Bertrand, J., Brickler, D., Babu, S., Madathil, K., Zelaya, M., Wang, T., Wagner, J., Gramopadhye, A., and Luo, J. (2015). The role of dimensional symmetry on bimanual psychomotor skills education in immersive virtual environments. In *2015 IEEE Conference on Virtual Reality (VR)*, pages 3–10. IEEE.

Beyer, L., De Troyer, J., Mancini, J., Bladou, F., Berdah, S. V., and Karsenty, G. (2011). Impact of laparoscopy simulator training on the technical skills of future surgeons in the operating room: a prospective study. *The American Journal of Surgery*, 202(3):265–272.

Bhargava, A., Bertrand, J. W., Gramopadhye, A. K., Madathil, K. C., and Babu, S. V. (2018). Evaluating multiple levels of an interaction fidelity continuum on performance and learning in near-field training simulations. *IEEE transactions on visualization and computer graphics*, 24(4):1418–1427.

Biocca, F., Levy, M., and Lawrence, J. (1997). Communication in the age of virtual reality. *Psyccritiques*, 42(2).

Blanke, O., Slater, M., and Serino, A. (2015). Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron*, 88(1):145–166.

Bloom, B. S. (1986). Automaticity:" the hands and feet of genius.". *Educational leadership*, 43(5):70–77.

Boritz, J. and Booth, K. S. (1997). A study of interactive 3d point location in a computer simulated virtual environment. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '97, pages 181–187. ACM.

Botvinick, M. and Cohen, J. (1998). Rubber hands 'feel'touch that eyes see. *Nature*, 391(6669):756–756.

Bowman, D. A., Kruijff, E., LaViola Jr, J. J., and Poupyrev, I. (2004). *3D user interfaces: theory and practice*. Addison-Wesley, 1st. edition.

Bowman, D. A. and McMahan, R. P. (2007). Virtual Reality: How Much Immersion Is Enough? *Computer*, 40(7):36–43.

Bradley, P. (2006). The history of simulation in medical education and possible future directions. *Medical Education*, 40(3):254–262.

Brooke, J. (1996). SUS: a "quick and dirty" usability scale. *Usability Evaluation in Industry*, 189(194):4–7.

Brunelli, R. (2009). *Template matching techniques in computer vision: theory and practice*. John Wiley & Sons.

Buckley, C., Nugent, E., Ryan, D., and Neary, P. (2012). Virtual reality – a new era in surgical training. In Eichenberg, C., editor, *Virtual reality in psychological, medical and pedagogical applications*, chapter 7, pages 139–166. InTech.

Canales, R., Normoyle, A., Sun, Y., Ye, Y., Luca, M. D., and Jörg, S. (2019). Virtual grasping feedback and virtual hand ownership. In *ACM Symposium on Applied Perception 2019*, SAP'19, pages 1–9. ACM.

Cao, C. G. L., MacKenzie, C. L., and Payandeh, S. (1996). Task and Motion Analysis in Endoscopic Surgery. In *5th Annual Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 583–590. ASME.

Carlson, P., Peters, A., Gilbert, S. B., Vance, J. M., and Luse, A. (2015). Virtual training: Learning transfer of assembly tasks. *IEEE transactions on visualization and computer graphics*, 21(6):770–782.

Castro-Alonso, J. C., Ayres, P., and Paas, F. (2014). Dynamic visualisations and motor skills. In Huang, W., editor, *Handbook of human centric visualization*, chapter 6, pages 551–580. Springer.

Champney, R. K., Carroll, M., Milham, L., and Hale, K. (2008). Sensory-perceptual objective task (spot) taxonomy:a task analysis tool. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 52, pages 1929–1933. SAGE Publications.

Champney, R. K., Carroll, M., and Surpris, G. (2014). A Human Experience Approach to Optimizing Simulator Fidelity. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58(1):2355–2359.

Chellali, A., Ahn, W., Sankaranarayanan, G., Flinn, J., Schwaitzberg, S. D., Jones, D. B., De, S., and Cao, C. G. (2015). Preliminary evaluation of the pattern cutting and the ligating loop virtual laparoscopic trainers. *Surgical endoscopy*, 29(4):815–821.

Chellali, A., Dumas, C., and Milleville-Pennel, I. (2012). Haptic communication to support biopsy procedures learning in virtual environments. *Presence: Teleoperators and Virtual Environments*, 21(4):470–489.

Chellali, A., Mentis, H., Miller, A., Ahn, W., Arikatla, V. S., Sankaranarayanan, G., Suvranu, D., Schwaitzberg, S. D., and Cao, C. G. L. (2016). Achieving interface and environment fidelity in the Virtual Basic Laparoscopic Surgical Trainer. *International Journal of Human-Computer Studies*, 96:22–37.

Chen, W., Plancoulaine, A., Férey, N., Touraine, D., Nelson, J., and Bourdot, P. (2013). 6dof navigation in virtual worlds: comparison of joystick-based and head-controlled paradigms. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, pages 111–114.

Christie, M. and Olivier, P. (2009). Camera control in computer graphics: models, techniques and applications. In *ACM SIGGRAPH ASIA 2009 Courses*, pages 1–197. ACM.

Coles, T. R., Meglan, D., and John, N. W. (2011). The role of haptics in medical training simulators: a survey of the state of the art. *IEEE Transactions on Haptics*, 4(1):51–66.

Conti, F., Barbagli, F., Balaniuk, R., Halg, M., Lu, C., Morris, D., Sentis, L., Warren, J., Khatib, O., and Salisbury, K. (2003). The chai libraries. In *Proceedings of Eurohaptics 2003*, pages 496–500.

Cooper, J. and Taqueti, V. (2004). A brief history of the development of mannequin simulators for clinical education and training. *Quality and Safety in Health Care*, 13(suppl 1):i11–i18.

Cornacchione Jr, E. B. (2012). Fidelity and game-based technology in management education. *BAR-Brazilian Administration Review*, 9(2):147–167.

Corrêa, C. G., dos Santos Nunes, F. d. L., and Tori, R. (2014). Virtual reality-based system for training in dental anesthesia. In *Virtual, Augmented and Mixed Reality. Applications of Virtual and Augmented Reality (VAMR 2014)*, Lecture Notes in Computer Science, pages 267–276. Springer.

Corrêa, C. G., Nunes, F. L., Ranzini, E., Nakamura, R., and Tori, R. (2018). Haptic interaction for needle insertion training in medical applications: The state-of-the-art. *Medical Engineering & Physics*, 63:6–25.

Dang, T., Annaswamy, T. M., and Srinivasan, M. A. (2001). Development and evaluation of an epidural injection simulator with force feedback for medical training. *Studies in health technology and informatics*, pages 97–102.

Dargar, S., Kennedy, R., Lai, W., Arikatla, V., and De, S. (2015). Towards immersive virtual reality (ivr): a route to surgical expertise. *Journal of computational surgery*, 2(1):2.

Davesne, F. (2015). ARTTrackSM. https://www.ibisc.univ-evry.fr/ fdavesne/tec/. [Accessed: 1 March 2020].

De Araújo, B. R., Casiez, G., Jorge, J. A., and Hachet, M. (2013). Mockup builder: 3d modeling on and above the surface. *Computers & Graphics*, 37(3):165–178.

Debarba, H. G., Molla, E., Herbelin, B., and Boulic, R. (2015). Characterizing embodied interaction in first and third person perspective viewpoints. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 67–72. IEEE.

Drews, F. A. and Bakdash, J. Z. (2013). Simulation training in health care. *Reviews of Human Factors and Ergonomics*, 8(1):191–234.

Duffy, A., Hogle, N., McCarthy, H., Lew, J., Egan, A., Christos, P., and Fowler, D. (2005). Construct validity for the lapsim laparoscopic surgical simulator. *Surgical Endoscopy and Other Interventional Techniques*, 19(3):401–405.

Duncan, J. (2006). Fidelity versus cost and its effect on modeling and simulation. Technical report, DTIC Document.

Edelmann, J., Schilling, A., and Fleck, S. (2009). The DabR-a multitouch system for intuitive 3d scene navigation. In *2009 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video*, pages 1–4. IEEE.

Ehrsson, H. H. (2009). How many arms make a pair? perceptual illusion of having an additional limb. *Perception*, 38(2):310–312.

Eriksen, J. and Grantcharov, T. (2005). Objective assessment of laparoscopic skills using a virtual reality stimulator. *Surgical Endoscopy and Other Interventional Techniques*, 19(9):1216–1219.

Escobar-Castillejos, D., Noguez, J., Neri, L., Magana, A., and Benes, B. (2016). A Review of Simulators with Haptic Devices for Medical Training. *Journal of Medical Systems*, 40(4):104–126.

Fecso, A. B., Szasz, P., Kerezov, G., and Grantcharov, T. P. (2017). The Effect of Technical Performance on Patient Outcomes in Surgery: A Systematic Review. *Annals of Surgery*, 265(3):492–501.

Forsslund, J., Chan, S., Selesnick, J., Salisbury, K., Silva, R. G., and Blevins, N. H. (2013). The effect of haptic degrees of freedom on task performance in virtual surgical environments. *Studies in Health Technology and Informatics*, 184:129–135.

Forsslund, J., Sallnas, E.-L., and Palmerius, K.-J. (2009). A user-centered designed for implementation of bone surgery simulations. In *World Haptics 2009-Third Joint Euro-Haptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 391–392. IEEE.

Forsslund, J., Sallnäs Pysander, E.-L., and Lundin Palmerius, K. (2011). Design of Perceptualization Applications in Medicine. In *First workshop on Engineering Interactive Computing Systems for Medicine and Health Care (EICS4Med)*, pages 42–47. EICS4Med.

Fortmeier, D., Mastmeyer, A., Schröder, J., and Handels, H. (2016). A virtual reality system for PTCD simulation using direct visuo-haptic rendering of partially segmented image data. *IEEE Journal of Biomedical and Health Informatics*, 20(1):355–366.

Fribourg, R., Argelaguet, F., Lécuyer, A., and Hoyet, L. (2020). Avatar and sense of embodiment: Studying the relative preference between appearance, control and point of view. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):2062–2072.

Fried, G. M., Feldman, L. S., Vassiliou, M. C., Fraser, S. A., Stanbridge, D., Ghitulescu, G., and Andrew, C. G. (2004). Proving the value of simulation in laparoscopic surgery. *Annals of Surgery*, 240(3):518.

Fu, C.-W., Goh, W.-B., and Ng, J. A. (2010). Multi-touch techniques for exploring large-scale 3d astrophysical simulations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2213–2222. ACM.

Fu, Y., Cavuoto, L., Qi, D., Panneerselvam, K., Arikatla, V. S., Enquobahrie, A., De, S., and Schwaitzberg, S. D. (2019). Characterizing the learning curve of a virtual intracorporeal suturing simulator vblast-ss©. *Surgical Endoscopy*, pages 1–10.

Fuchs, P. (2006). *Le traité de la réalité virtuelle*, volume 2. Presses des MINES.

Gallagher, A. G., Richie, K., McClure, N., and McGuigan, J. (2001). Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality. *World journal of surgery*, 25(11):1478–1483.

Gamelin, G., Chellali, A., Cheikh, S., Ricca, A., Dumas, C., and Otmane, S. (2020). Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments. *Personal and Ubiquitous Computing*.

Gerathewohl, S. J. (1969). *Fidelity of simulation and transfer of training: a review of the problem*. US Department of Transportation, Federal Aviation Administration, Office of ....

Gerovich, O., Marayong, P., and Okamura, A. M. (2004). The effect of visual and haptic feedback on computer-assisted needle insertion. *Computer Aided Surgery*, 9(6):243–249.

Giraud, T., Paljic, A., and Leroy, L. (2013). "it"+"i": Virtual embodiments as hybrid experiences.

Goguey, A., Nancel, M., Casiez, G., and Vogel, D. (2016). The performance and preference of different fingers and chords for pointing, dragging, and object transformation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4250–4261.

González-Franco, M., Peck, T. C., Rodríguez-Fornells, A., and Slater, M. (2014). A threat to a virtual hand elicits motor cortex activation. *Experimental Brain Research*, 232(3):875–887.

Goodman, A. J., Melson, J., Aslanian, H. R., Bhutani, M. S., Krishnan, K., Lichtenstein, D. R., Navaneethan, U., Pannala, R., Parsi, M. A., Schulman, A. R., et al. (2019). Endoscopic simulators. *Gastrointestinal endoscopy*, 90(1):1–12.

Gorisse, G., Christmann, O., Amato, E. A., and Richir, S. (2017). First-and third-person perspectives in immersive virtual environments: Presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, 4:33.

Grant, B. L., Yielder, P. C., Patrick, T. A., Kapralos, B., Williams-Bell, M., and Murphy, B. A. (2020). Audiohaptic feedback enhances motor performance in a low-fidelity simulated drilling task. *Brain sciences*, 10(1):21.

Grenvik, A. and Schaefer, J. (2004). From resusci-anne to sim-man: the evolution of simulators in medicine. *Critical care medicine*, 32(2):S56–S57.

Guerraz, M., Breen, A., Pollidoro, L., Luyat, M., and Kavounoudias, A. (2018). Contribution of visual motion cues from a held tool to kinesthesia. *Neuroscience*, 388:11–22.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action. *Journal of Motor Behavior*, 19(4):486–517.

Hachet, M., Bossavit, B., Cohé, A., and de la Rivière, J.-B. (2011). Toucheo: multitouch and stereo combined in a seamless workspace. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 587–592.

Halsted, W. S. (1904). The training of the surgeon. *Bull Johns Hopkins Hospital*, 15:267–275.

Hamblin, C. J. (2005). *Transfer of training from virtual reality environments*. Phd, Wichita State University.

Hancock, M., Carpendale, S., and Cockburn, A. (2007). Shallow-depth 3d interaction: design and evaluation of one-, two-and three-touch techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1147–1156.

Hand, C. (1997). A survey of 3d interaction techniques. *Computer Graphics Forum*, 16(5):269–281.

Hannaford, B. and Okamura, A. M. (2008). Haptics. In *Springer Handbook of Robotics*, pages 719–739. Springer.

Harris, C. and Stephens, M. (1988). A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference*.

Hays, R. T. (1980). Simulator fidelity: A concept paper. Technical report, DTIC Document.

Healthcare, C. (2018). *LapVR*. Available at: https://caehealthcare.com/surgical-simulation/lapvr. [Accessed: 10 June 2018].

Hecht, D. and Reiner, M. (2009). Sensory dominance in combinations of audio, visual and haptic stimuli. *Experimental brain research*, 193(2):307–314.

Heidegger, M. (1927). *Sein und Zeit (Being and Time), trans. J. Macquarrie, John and Robinson, Edward*. Harper & Row, 7th. edition.

Henshall, G., Pop, S. R., Edwards, M. R., ap Cenydd, L., and John, N. W. (2015). Towards a high fidelity simulation of the kidney biopsy procedure. In *2015 IEEE Conference on Virtual Reality (VR)*, pages 191–192. IEEE.

Houde, S. (1992). Iterative design of an interface for easy 3-d direct manipulation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 135–142.

Hoyet, L., Argelaguet, F., Nicole, C., and Lécuyer, A. (2016). "Wow! I have six fingers!": would you accept structural changes of your hand in VR? *Frontiers in Robotics and AI*, 3(27):1–12.

Hurstel, A. and Bechmann, D. (2019). Approach for intuitive and touchless interaction in the operating room. *J—Multidisciplinary Scientific Journal*, 2(1):50–64.

Jankowski, J. and Hachet, M. (2015). Advances in interaction with 3d environments. *Computer Graphics Forum*, 34(1):152–190.

Jarc, A. M. and Curet, M. J. (2017). Viewpoint matters: objective performance metrics for surgeon endoscope control during robot-assisted surgery. *Surgical Endoscopy*, 31(3):1192–1202.

Jeanne, F., Thouvenin, I., and Lenglet, A. (2017). A study on improving performance in gesture training through visual guidance based on learners' errors. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pages 1–10.

Jiamu, C. (2001). The great importance of the distinction between declarative and procedural knowledge. *Análise Psicológica*, 19(4):559–566.

Johnsen, K., Dickerson, R., Raij, A., Lok, B., Jackson, J., Shin, M., Hernandez, J., Stevens, A., and Lind, D. S. (2005). Experiences in using immersive virtual characters to educate medical communication skills. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pages 179–186. IEEE.

JORF (2017). Arrêté du 27 novembre 2017 modifiant l'arrêté du 12 avril 2017 relatif à l'organisation du troisième cycle des études de médecine et arrêté du 21 avril 2017 relatif aux connaissances, aux compétences et aux maquettes de formation des diplômes d'études spécialisées et fixant la liste de ces diplômes et des options et formations spécialisées transversales du troisième cycle des études de médecine. *Journal Officiel de la République Française*.

Jung, S., Bruder, G., Wisniewski, P. J., Sandor, C., and Hughes, C. E. (2018). Over my hand: Using a personalized hand in vr to improve object size estimation, body ownership, and presence. In *Proceedings of the Symposium on Spatial User Interaction*, SUI'18, pages 60–68. ACM.

Khan, A., Komalo, B., Stam, J., Fitzmaurice, G., and Kurtenbach, G. (2005). Hovercam: interactive 3d navigation for proximal object inspection. In *Proceedings of the 2005 Symposium on Interactive 3D Graphics and Games*, pages 73–80. ACM.

Kilteni, K., Groten, R., and Slater, M. (2012a). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387.

Kilteni, K., Normand, J.-M., Sanchez-Vives, M. V., and Slater, M. (2012b). Extending body space in immersive virtual reality: a very long arm illusion. *PloS one*, 7(7):1–15.

Kim, H. K., Rattner, D. W., and Srinivasan, M. A. (2003). The role of simulation fidelity in laparoscopic surgical training. In *International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, Lecture Notes in Computer Science, pages 1–8. Springer-Verlag.

Kirurobo (2014). A C# (.NET) wrapper for Sensable PHANTOM Device. Available at: https://github.com/kirurobo/ManagedPhantom. [Accessed: 1 March 2020].

Kneebone Roger (2003). Simulation in surgical training: educational issues and practical implications. *Medical Education*, 37(3):267–277.

Kohn, L. T., Corrigan, J. M., and Donaldson, M. S. (2000). *To err is human: building a safer health system*. National Academies Press.

Kooima, R. L. (2008). Generalized perspective projection. Technical report, School of Electrical Engineering and Computer Science, Louisiana State University.

Kozlowski, S. W. and DeShon, R. P. (2004a). A psychological fidelity approach to simulation-based training: Theory, research and principles.

Kozlowski, S. W. and DeShon, R. P. (2004b). A psychological fidelity approach to simulation-based training: Theory, research and principles.

Kulik, A., Kunert, A., Keil, M., and Froehlich, B. (2018). RST 3D: A comprehensive gesture set for multitouch 3d navigation. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 363–370. IEEE.

Lab, T. P. (2018). *Vtopia | Virtual Reality to Medical Realities*. Available at: https://lab.troisprime.com/realisations/. [Accessed: 10 June 2018].

Lakens, D., Scheel, A. M., and Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2):259–269.

Lee, C., Rincon, G. A., Meyer, G., Höllerer, T., and Bowman, D. A. (2013). The effects of visual realism on search tasks in mixed reality simulation. *IEEE transactions on visualization and computer graphics*, 19(4):547–556.

Lee, S. and Kim, G. J. (2008). Effects of haptic feedback, stereoscopy, and image resolution on performance and presence in remote navigation. *International Journal of Human-Computer Studies*, 66(10):701–717.

Lemole Jr., G. M., Banerjee, P. P., Luciano, C., Neckrysh, S., and Charbel, F. T. (2007). Virtual reality in neurosurgical education: part-task ventriculostomy simulation with dynamic visual and haptic feedback. *Neurosurgery*, 61(1):142–149.

Lin, C.-L., Shaw, F.-Z., Young, K.-Y., Lin, C.-T., and Jung, T.-P. (2012). Eeg correlates of haptic feedback in a visuomotor tracking task. *NeuroImage*, 60(4):2258–2273.

Lin, L. and Jörg, S. (2016). Need a hand? how appearance affects the virtual hand illusion. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '16, pages 69–76. Association for Computing Machinery.

Linsk, A. M., Monden, K. R., Sankaranarayanan, G., Ahn, W., Jones, D. B., De, S., Schwaitzberg, S. D., and Cao, C. G. (2018). Validation of the vblast pattern cutting task: a learning curve study. *Surgical endoscopy*, 32(4):1990–2002.

Liu, D., Blickensderfer, E. L., Macchiarella, N. D., and Vincenzi, D. A. (2008). Transfer of training. In Hancock, P. A., Vincenzi, D. A., Wise, J. A., and Mouloua, M., editors, *Human Factors in Simulation and Training*, chapter 3, pages 49–60. CRC Press.

Lok, B., Naik, S., Whitton, M., and Brooks, F. P. (2003). Effects of handling real objects and self-avatar fidelity on cognitive task performance and sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 12(6):615–628.

Maguire, M. (2001). Methods to support human-centred design. *International Journal of Human-Computer Studies*, 55(4):587–634.

Mapes, D. P. and Moshell, J. M. (1995). A two-handed interface for object manipulation in virtual environments. *Presence: Teleoperators & Virtual Environments*, 4(4):403–416.

Maran, N. J. and Glavin, R. J. (2003). Low-to high-fidelity simulation –a continuum of medical education? *Medical Education*, 37:22–28.

Maravita, A. and Iriki, A. (2004). Tools for the body (schema). *Trends in cognitive sciences*, 8(2):79–86.

Marchal, D., Moerman, C., Casiez, G., and Roussel, N. (2013). Designing intuitive multitouch 3d navigation techniques. In *Human-Computer Interaction – INTERACT 2013*, Lecture Notes in Computer Science, pages 19–36. Springer, Berlin, Heidelberg.

Martinet, A., Casiez, G., and Grisoni, L. (2010). The design and evaluation of 3d positioning techniques for multi-touch displays. In *2010 IEEE symposium on 3D user interfaces (3DUI)*, pages 115–118. IEEE.

Mastmeyer, A., Hecht, T., Fortmeier, D., and Handels, H. (2014). Ray-casting based evaluation framework for haptic force feedback during percutaneous transhepatic catheter drainage punctures. *International Journal of Computer Assisted Radiology and Surgery*, 9(3):421–431.

McElhinney, B., Hart, R., Karthigasu, K., and Beard, A. (2012). Virtual Reality Simulation: A Valuable Adjunct to Surgical Training. In Eichenberg, C., editor, *Virtual Reality in Psychological, Medical and Pedagogical Applications*, chapter 8, pages 167–184. InTech.

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., and Scalese, R. J. (2010). A critical review of simulation-based medical education research: 2003–2009. *Medical Education*, 44(1):50–63.

McMahan, R. P. (2011). *Exploring the effects of higher-fidelity display and interaction for virtual reality games*. Phd, Virginia Polytechnic Institute and State University.

McMahan, R. P., Alon, A. J. D., Lazem, S., Beaton, R. J., Machaj, D., Schaefer, M., Silva., M. G., Leal, A., Hagan, R., and Bowman, D. A. (2010). Evaluating natural interaction techniques in video games. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 11–15. IEEE.

McMahan, R. P., Bowman, D. A., Zielinski, D. J., and Brady, R. B. (2012). Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):626–633.

McMahan, R. P. and Herrera, N. S. (2016). AFFECT: Altered-Fidelity Framework for Enhancing Cognition and Training. *Frontiers in ICT*, 3(29):1–15.

Medical, L. (2018). *SimMan*. Available at: https://laerdal.com/fr/products/simulation-training/emergency-care-trauma/simman-als/ [Accessed: 10 June 2018].

Medical-X (2018). *Lap-X VR*. Available at: https://www.medical-x.com/products/ [Accessed: 10 June 2018].

Mendes, D., Caputo, F. M., Giachetti, A., Ferreira, A., and Jorge, J. (2019). A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer graphics forum*, volume 38, pages 21–45. Wiley Online Library.

Mohler, B. J., Creem-Regehr, S. H., Thompson, W. B., and Bülthoff, H. H. (2010). The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment. *Presence: Teleoperators and Virtual Environments*, 19(3):230–242.

Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100.

Murphy-Chutorian, E. and Trivedi, M. M. (2009). Head pose estimation in computer vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4):607–626.

Nabiyouni, M. (2017). *How Does Interaction Fidelity Influence User Experience in VR Locomotion?* Phd, Virginia Polytechnic Institute and State University.

Nabiyouni, M., Saktheeswaran, A., Bowman, D. A., and Karanth, A. (2015). Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 3–10. IEEE.

Nilsson, N. C. (2016). *Walking Without Moving*. PhD thesis, University of Aalborg.

Ogawa, N., Narumi, T., and Hirose, M. (2019). Virtual hand realism affects object size perception in body-based scaling. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 519–528. IEEE.

Ortega, M. (2013). 3D object position using automatic viewpoint transitions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 193–196. ACM.

Ortega, M., Stuerzlinger, W., and Scheurich, D. (2015). SHOCam: a 3d orbiting algorithm. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, pages 119–128. ACM.

Orzech, N., Palter, V. N., Reznick, R. K., Aggarwal, R., and Grantcharov, T. P. (2012). A Comparison of 2 Ex Vivo Training Curricula for Advanced Laparoscopic Skills: A Randomized Controlled Trial. *Annals of Surgery*, 255(5):833–839.

Ossmy, O. and Mukamel, R. (2017). Short term motor-skill acquisition improves with size of self-controlled virtual hands. *PloS one*, 12(1).

Palter, V. N. (2011). Comprehensive training curricula for minimally invasive surgery. *Journal of graduate medical education*, 3(3):293–298.

Panait, L., Akkary, E., Bell, R. L., Roberts, K. E., Dudrick, S. J., and Duffy, A. J. (2009). The role of haptic feedback in laparoscopic simulation training. *Journal of Surgical Research*, 156(2):312–316.

Pearse, R. M., Moreno, R. P., Bauer, P., Pelosi, P., Metnitz, P., Spies, C., Vallet, B., Vincent, J.-L., Hoeft, A., and Rhodes, A. (2012). Mortality after surgery in Europe: a 7 day cohort study. *The Lancet*, 380(9847):1059–1065.

Pitzul, K. B., Grantcharov, T. P., and Okrainec, A. (2012). Validation of three virtual reality fundamentals of laparoscopic surgery (fls) modules. *Medicine Meets Virtual Reality 19 NextMed*, 173:349.

Posner, M. I., Nissen, M. J., and Klein, R. M. (1976). Visual dominance: an information-processing account of its origins and significance. *Psychological review*, 83(2):157.

Pulijala, Y., Ma, M., Pears, M., Peebles, D., and Ayoub, A. (2018). Effectiveness of immersive virtual reality in surgical training—a randomized control trial. *Journal of Oral and Maxillofacial Surgery*, 76(5):1065–1072.

Ragan, E. D., Bowman, D. A., Kopper, R., Stinson, C., Scerbo, S., and McMahan, R. P. (2015). 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*, 21(7):794–807.

Ragan, E. D., Kopper, R., Schuchardt, P., and Bowman, D. A. (2013). Studying the effects of stereo, head tracking, and field of regard on a small-scale spatial judgment task. *Visualization and Computer Graphics, IEEE Transactions on*, 19(5):886–896.

Reich, O., Noll, M., Gratzke, C., Bachmann, A., Waidelich, R., Seitz, M., Schlenker, B., Baumgartner, R., Hofstetter, A., and Stief, C. G. (2006). High-level virtual reality simulator for endourologic procedures of lower urinary tract. *Urology*, 67(6):1144–1148.

Reznick, R. K. (1993). Teaching and testing technical skills. *The American journal of surgery*, 165(3):358–361.

Ricca, A. and Chellali, A. (2016). Interaction Fidelity in Virtual Simulators: Two Navigation Techniques for a Virtual Biopsy Trainer. In *Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine*, IHM '16, pages 278–284. ACM.

Ricca, A., Chellali, A., and Otmane, S. (2017). Study of interaction fidelity for two viewpoint changing techniques in a virtual biopsy trainer. In *2017 Conference on IEEE Virtual Reality (VR)*, pages 227–228. IEEE.

Ricca, A., Chellali, A., and Otmane, S. (2020). Influence of hand visualization on tool-based motor skills training in an immersive VR simulator. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, ISMAR 2020, pages 260–268. IEEE.

Ricca, A., Chellali, A., and Otmane, S. (2021a). Comparing touch-based and head-tracking navigation techniques in a virtual reality biopsy simulator. *Virtual Reality*, 25(1):191–208.

Ricca, A., Chellali, A., and Otmane, S. (2021b). The influence of hand visualization in tool-based motor-skills training, a longitudinal study. In *28th IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE.

Rodriguez-Paz, J. M., Kennedy, M., Salas, E., Wu, A. W., Sexton, J. B., Hunt, E. A., and Pronovost, P. J. (2009). Beyond "see one, do one, teach one": toward a different training paradigm. *BMJ Quality & Safety*, 18(1):63–68.

Salisbury Jr, J. Kenneth, J. (1999). Making Graphics Physically Tangible. *Communications of the ACM*, 42(8):74–81.

Sanchez-Vives, M. V. and Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6(4):332–339.

Sankaranarayanan, G., Li, B., Manser, K., Jones, S. B., Jones, D. B., Schwaitzberg, S., Cao, C. G. L., and De, S. (2016). Face and construct validation of a next generation virtual reality (Gen2-VR©) surgical simulator. *Surgical Endoscopy*, 30(3):979–985.

Satava, R. M. (1993). Virtual reality surgical simulator. *Surgical endoscopy*, 7(3):203–205.

Satava, R. M. (2001). Accomplishments and challenges of surgical simulation. *Surgical Endoscopy*, 15(3):232–241.

Satava, R. M., Cuschieri, A., and Hamdorf, J. (2003). Metrics for objective assessment. *Surgical endoscopy*, 17(2):220.

Schijven, M. and Jakimowicz, J. (2003). Construct validity. *Surgical Endoscopy and Other Interventional Techniques*, 17(5):803–810.

Schuemie, M., van der Straaten, P., Krijn, M., and van der Mast, C. (2001). Research on presence in vr: a survey. *Cyberpsychology and Behavior*, 4(2):183–202.

Schwind, V., Knierim, P., Chuang, L., and Henze, N. (2017a). "where's pinky?" the effects of a reduced number of fingers in virtual reality. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '17, pages 507–515. Association for Computing Machinery.

Schwind, V., Knierim, P., Tasci, C., Franczak, P., Haas, N., and Henze, N. (2017b). "these are not my hands!" effect of gender on the perception of avatar hands in virtual reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 1577–1582. Association for Computing Machinery.

Science, S. (2018). *LapSim*. Available at: https://surgicalscience.com/systems/lapsim/. [Accessed: 10 June 2018].

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'brien, M. K., Bansal, V. K., Andersen, D. K., and Satava, R. M. (2002). Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, 236(4):458.

Sherman, V., Feldman, L., Stanbridge, D., Kazmi, R., and Fried, G. (2005). Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. *Surgical Endoscopy and Other Interventional Techniques*, 19(5):678–682.

Shin, S., Park, W., Cho, H., Park, S., and Kim, L. (2011). Needle insertion simulator with haptic feedback. In *Human-Computer Interaction. Interaction Techniques and Environments (HCI 2011)*, Lecture Notes in Computer Science, pages 119–124. Springer, Berlin, Heidelberg.

Simbionix (2018). *LAP Mentor*. Available at: http://simbionix.com/simulators/lap-mentor/. [Accessed: 10 June 2018].

SimforHealth (2018). *Tailor-made simulation*. Available at: https://simforhealth.fr/en/our-solutions/. [Accessed: 10 June 2018].

Simulations, M. M. (2018). *Gas-Man*. Available at: https://www.gasmanweb.com [Accessed: 10 June 2018].

Skinner, A., Auner, G., Meadors, M., and Sebrechts, M. M. (2013). Ambidexterity in laparoscopic surgical skills training. In *MMVR*, volume 2013, pages 412–416.

Slater, M. (2003). A note on presence terminology. *Presence connect*, 3(3):1–5.

Spindler, M., Büschel, W., and Dachselt, R. (2012). Use your head: tangible windows for 3d information spaces in a tabletop environment. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, pages 245–254. ACM.

Sportillo, D., Avveduto, G., Tecchia, F., and Carrozzino, M. (2015). Training in vr: a preliminary study on learning assembly/disassembly sequences. In *International Conference on Augmented and Virtual Reality*, AVR 2015, pages 332–343. Springer.

Stoffregen, T. A., Bardy, B. G., Smart, L., and Pagulayan, R. (2003). On the nature and evaluation of fidelity in virtual environments. In Hettinger, L. J. and Haas, M. W., editors, *Virtual and adaptive environments: Applications, implications, and human performance issues*, chapter 6, pages 111–128. Lawrence Erlbaum Associates.

Ström, P., Hedman, L., Särnå, L., Kjellin, A., Wredmark, T., and Felländer-Tsai, L. (2006). Early exposure to haptic feedback enhances performance in surgical simulator training: a prospective randomized crossover study in surgical residents. *Surgical endoscopy and other interventional techniques*, 20(9):1383–1388.

Sutherland, C., Hashtrudi-Zaad, K., Sellens, R., Abolmaesumi, P., and Mousavi, P. (2013). An augmented reality haptic training simulator for spinal needle procedures. *IEEE Transactions on Biomedical Engineering*, 60(11):3009–3018.

Sutherland, I. E. (1968). A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, AFIPS '68 (Fall, part I), pages 757–764. ACM.

Tang, C.-Y., Chin, W., Chui, Y.-P., Poon, W.-S., and Heng, P.-A. (2007). A virtual reality-based surgical simulation system for virtual neuroendoscopy. In *2007 IEEE International Conference on Integration Technology*, ICIT '07, pages 253–258. IEEE.

Taylor, C. R., Abramson, H. G., and Herr, H. M. (2019). Low-latency tracking of multiple permanent magnets. *IEEE Sensors Journal*, 19(23):11458–11468.

Torres, R. S., Bíscaro, H. H., Araújo, L. V., and Nunes, F. L. S. (2012). ViMeTGame: A serious games for virtual medical training of breast biopsy. *SBC Journal on 3D Interactive Systems*, 3(3):12–22.

Ullrich, S. and Kuhlen, T. (2012). Haptic Palpation for Medical Simulation in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):617–625.

Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., and Brooks Jr, F. P. (1999). Walking> walking-in-place> flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 359–364. ACM Press/Addison-Wesley Publishing Co.

Valdivieso, R. F. and Zorn, K. C. (2014). Surgery: Urological laparoscopic training—practice makes perfect. *Nature Reviews Urology*, 11(3):138.

Van Dongen, K., Tournoij, E., Van der Zee, D., Schijven, M., and Broeders, I. (2007). Construct validity of the lapsim: can the lapsim virtual reality simulator distinguish between novices and experts? *Surgical endoscopy*, 21(8):1413–1417.

Van Nguyen, D., Lakhal, S. B., and Chellali, A. (2015). Preliminary evaluation of a virtual needle insertion training system. In *2015 IEEE Conference on Virtual Reality (VR)*, pages 247–248. IEEE.

Våpenstad, C. and Buzink, S. N. (2013). Procedural virtual reality simulation in minimally invasive surgery. *Surgical endoscopy*, 27(2):364–377.

Våpenstad, C., Hofstad, E. F., Bø, L. E., Chmarra, M. K., Kuhry, E., Johnsen, G., Mårvik, R., and Langø, T. (2013). Limitations of haptic feedback devices on construct validity of the lapsim® virtual reality simulator. *Surgical endoscopy*, 27(4):1386–1396.

Villard, P.-F., Vidal, F. P., Ap Cenydd, L., Holbrey, R., Pisharody, S., Johnson, S., Bulpitt, A., John, N. W., Bello, F., and Gould, D. (2014). Interventional radiology virtual simulator for liver biopsy. *International journal of computer assisted radiology and surgery*, 9(2):255–267.

Vozenilek, J., Huff, J. S., Reznek, M., and Gordon, J. A. (2004). See One, Do One, Teach One: Advanced Technology in Medical Education. *Academic Emergency Medicine*, 11(11):1149–1154.

VRmagic (2018). *EYESI*. Available at: https://www.vrmagic.com/medical-simulators/eyesi-surgical [Accessed: 10 June 2018].

Waller, D., Hunt, E., and Knapp, D. (1998). The Transfer of Spatial Knowledge in Virtual Environment Training. *Presence: Teleoperators and Virtual Environments*, 7(2):129–143.

Waller, D., Knapp, D., and Hunt, E. (2001). Spatial representations of virtual mazes: The role of visual fidelity and individual differences. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1):147–158.

Wang, D., Xiao, J., and Zhang, Y. (2014). *Haptic Rendering for Simulation of Fine Manipulation*. Springer.

Weiser, T. G., Regenbogen, S. E., Thompson, K. D., Haynes, A. B., Lipsitz, S. R., Berry, W. R., and Gawande, A. A. (2008). An estimation of the global volume of surgery: a modelling strategy based on available data. *The Lancet*, 372(9633):139–144.

Westebring – van der Putten, E. P., Goossens, R. H. M., Jakimowicz, J. J., and Dankelman, J. (2008). Haptics in minimally invasive surgery – a review. *Minimally Invasive Therapy \& Allied Technologies*, 17(1):3–16.

Wijewickrema, S., Ma, X., Piromchai, P., Briggs, R., Bailey, J., Kennedy, G., and O'Leary, S. (2018). Providing automated real-time technical feedback for virtual reality based surgical training: is the simpler the better? In *International Conference on Artificial Intelligence in Education*, pages 584–598. Springer.

Wijewickrema, S., Piromchai, P., Zhou, Y., Ioannou, I., Bailey, J., Kennedy, G., and O'Leary, S. (2015). Developing effective automated feedback in temporal bone surgery simulation. *Otolaryngology–Head and Neck Surgery*, 152(6):1082–1088.

Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence*, 7(3):225–240.

Wu, S., Ricca, A., Chellali, A., Otmane, S., and Moreau, G. (2016). Two uni-manual manipulation techniques with constraints for 3d modeling applications on tabletpcs. In *Actes de la 28ième conference francophone sur l'Interaction Homme-Machine*, IHM '16, pages 8–18. ACM.

Wu, S., Ricca, A., Chellali, A., Otmane, S., and Moreau, G. (2017). Classic3d and single3d: Two unimanual techniques for constrained 3d manipulations on tablet pcs. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 168–171. IEEE.

Zhang, A., Hünerbein, M., Dai, Y., Schlag, P. M., and Beller, S. (2008). Construct validity testing of a laparoscopic surgery simulator (lap mentor®). *Surgical endoscopy*, 22(6):1440–1444.

Zhang, L., Sankaranarayanan, G., Arikatla, V. S., Ahn, W., Grosdemouge, C., Rideout, J. M., Epstein, S. K., De, S., Schwaitzberg, S. D., Jones, D. B., and Cao, C. G. (2013). Characterizing the learning curve of the vblast-pt(©) (virtual basic laparoscopic skill trainer). *Surgical endoscopy*, 27(10):3603–3615.

Zhou, M., Tse, S., Derevianko, A., Jones, D., Schwaitzberg, S., and Cao, C. (2012). Effect of haptic feedback in laparoscopic surgery skill acquisition. *Surgical endoscopy*, 26(4):1128–1134.

Zhu, B., Kaber, D., Zahabi, M., and Ma, W. (2020). Effects of feedback type and modality on motor skill learning and retention. *Behaviour & Information Technology*, 39(4):431–442.

Appendix A

Plastic handles and haptic integration

***

Several design iterations were made to build the 3D-printed plastic handles (used in the hand movement studies) based on the real metallic forceps (see fig. A.1). The initial design (see fig. A.2a) consisted of a main body connected to the haptic device and two handles, one fixed and solidary to the main body (printed separately to facilitate assembly), and a movable one (1 DoF). It was provided with a potentiometer which integrated axis served to both, guide the movable handle and register the angle difference. A custom-made circuit was added to scale and limit the measures to the range of movement needed (a maximum opening angle of 45°), and elastic bands were added to increase resistance on the opening. It also contained the two buttons already present in the stylus, which were only used during the development and debugging of the application, but were removed in the final solution. Two main issues were identified with this model. First, the potentiometer played the role of sensor and mechanical guide at the same time, which made the tool fragile and measures unstable after some use. Second, the size of the holes where fingers are placed were too small for most of the participants. Indeed, although they were designed to replicate the same size of the real tools' holes, the gloves used in the VRP had sensors placed in the fingers' extremities which did not allowed participants to comfortably grab the tools.

Another solution was explored, to have two movable handles so as to better reproduce the exact forceps' functioning. A group of four students (Master 1, Design specialty) worked on this prototype proposing an initial solution (see fig. A.2b). However, this solution, although more realistic in their movements, added more complications for the integration to the software application. Indeed, the main body moved independently to the handles and the point where haptic forces were rendered was not controllable.

Fig. A.1 The real metallic forceps

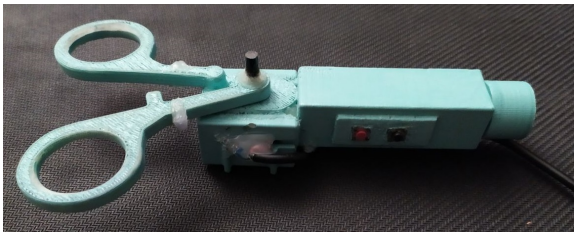
Nonetheless, the potentiometer (*MakeBlock*) they chose was a more reliable solution than the first prototype (guidance was not relied on the potentiometer).

The final prototype model (see fig. A.2c) was designed based on the first prototype. We have used a single body which was connected to the haptic device and contained the fixed handle. A *MakeBlock* angular potentiometer module was used, with an external metallic axis to guide the opening handle, leaving the sensor only responsible for measuring the angle variation. This allowed us to have a more robust and durable solution.

Another design iteration was performed after the user studies to improve the ergonomics and comfort of the handles. A 3D scanner (*EinScan-SE, Shinning 3D*) was used on a pair of real scissors (with a proper ergonomic design), and the result data was readjusted through several iterations to obtain a solid 3D object of the scissor handles (see fig. A.2d). As presented before in the results discussion, there are still improvements to make on this interface, notably, adding the gripping forces.

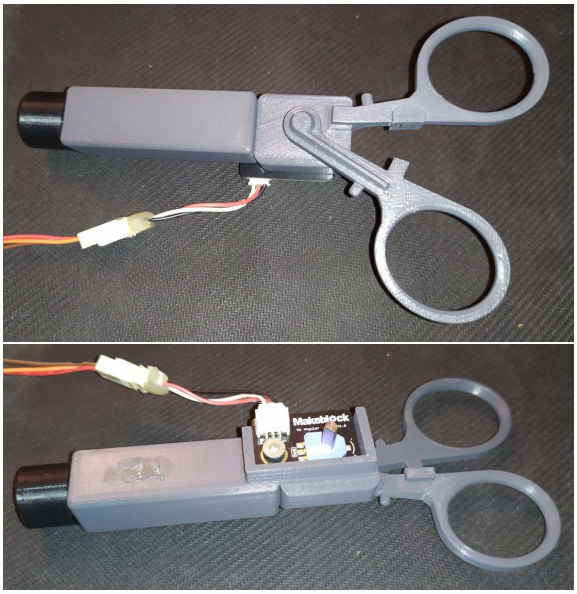
Integration

The chai3d framework was used for the haptic force feedback simulation. This open-source framework is specifically developed for computers haptics, visualization and real-time simulation, and it already supports several commercially-available haptic devices (*Geomagic Touch* included). We have used the finger-proxy algorithm for polygonal objects and the examples for multi-point contact to develop the forceps haptics simulation. As mentioned before, we have developed a distributed application, separating the haptics simulation from the visual rendering in the HMD. In the haptics simulation application, each virtual forceps was initially developed as two single contact points, each one of them implementing the finger-proxy algorithm. The position of each point was calculated with the haptics global position and orientation. One of this contact points corresponded to the

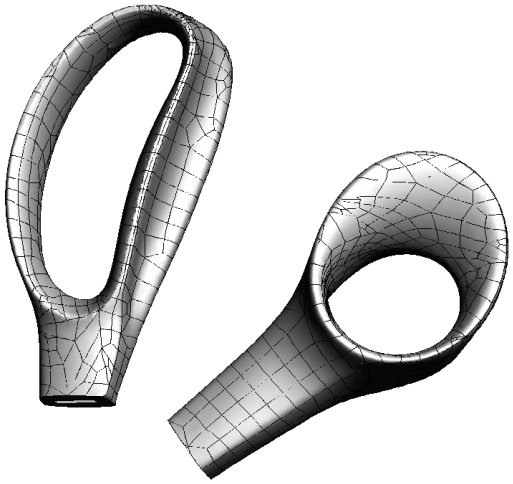


(a) First 3D printed forceps

(b) Development project (M1 students)



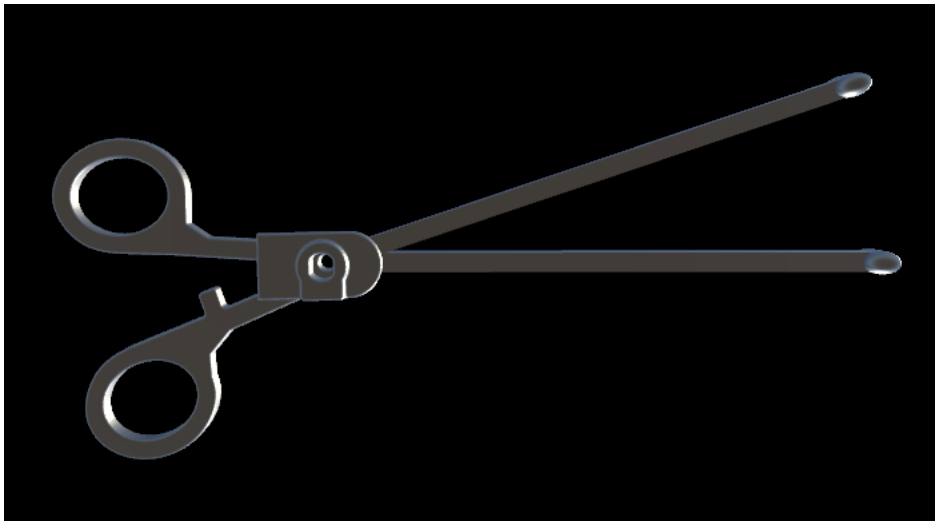
(c) Version used for the user studies



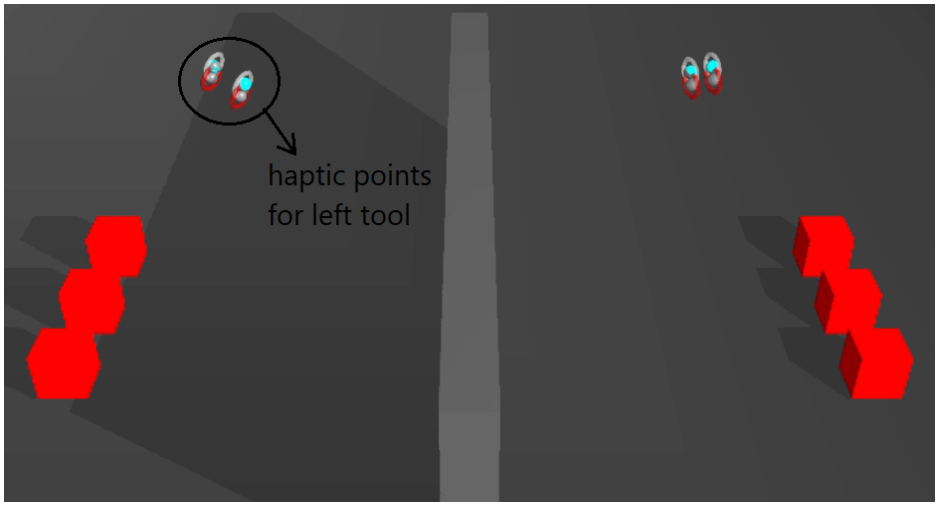
(d) Ergonomic handles designed for future iterations

Fig. A.2 The plastic handles design

fixed plastic handle extremity and the other to the movable handle. The position of this last contact point was adjusted with the opening angle obtained from the potentiometer through a serial communication with the Arduino UNO board. Each contact point behaved independently and allowed the user to “feel” the force interactions with the different static (wall and table surface) and dynamic (cubes) objects in the virtual scene (see fig. A.3). A “meta-proxy” structure was developed, which contained the constraints of the rigid objects, i.e., re-calculation of the contact points’ position and correction of the opening angle once the tool was closed and grabbing the cube. This allowed us to communicate the final information to the visual rendering application. The visual rendering application included the same 3D models of the forceps handles that were used for the printed ones.



(a) Virtual forceps representation



(b) Chai3d representation of the forcpes

Fig. A.3 The virtual representation of the forceps

Appendix B

Accuracy calculation for the physical prototype

***

For the physical prototype, measures were obtained through the analysis of the video recordings (for time and errors) and the scanner images of the cubes' final positions (for accuracy). In this section, we describe the procedure for calculating accuracy measures.

The distance and the minimal rotation angle difference between the center of each cube and the center of the corresponding target were used to measure the accuracy for the placement of each cube. The accuracy of the task was determined as the average of the six cube's distance error (position distance measure), and rotation difference (rotation distance measure).

The scanner, placed under the wooden wall, was used to capture an image of the final position of the cubes (2550 px × 3500 px). An OpenCV (CV) python application was developed to detect the position and orientation of the red cubes with respect to the black square targets.

First, a median filter (*kernel width* = 7) was used in the raw image (see fig. B.2a). After that, six regions of interest (ROIs) were determined (see fig. B.2c), one for each target, to locally identify the red square (cube position). For the squares detection, the CV template matching algorithm was used with the squared difference comparison method (Brunelli, 2009). The image used as pattern for the template matching algorithm was a red square with the green sticker (see fig. B.3a). For each ROI previously identified, the template matching algorithm was executed 90 times, one for each rotation of the cube (from 0° to 89°, with a step of 1°). The minimal result was retained, from which the rotation angle and center position was logged. This procedure was also performed on an image without the red squares (cubes) (see fig. B.2b), and with a different patch (see fig. B.3b), to determine

the targets' (black squares) position and orientation, and hence calculate the red squares' positions and rotations with respect to them.

Finally, CV Harris corners' detection algorithm (Harris and Stephens, 1988) was used to obtain the corners of the rotated patch chosen as solution, and draw them on the raw image (see fig. B.2d). In order to accomplish this task, first, a new image was generated from the patch rotated at the angle found as best solution, and then CV Harris algorithm was applied in this image. The resulted corners positions were then drawn on the original image taking into account the center position found and the ROI image position. See fig. B.1 for a graphical summary of the whole procedure. This last step was used to subjectively determine the accuracy of the result and correct the wrongly identified corners. 1.5% of the cubes were wrongly identified by the algorithm (subjective evaluation), which were manually corrected. Most of those errors were related to the bottom cube, were the CV template matching algorithm did not reach all the possible solutions (an enlarged image can solve this issue).

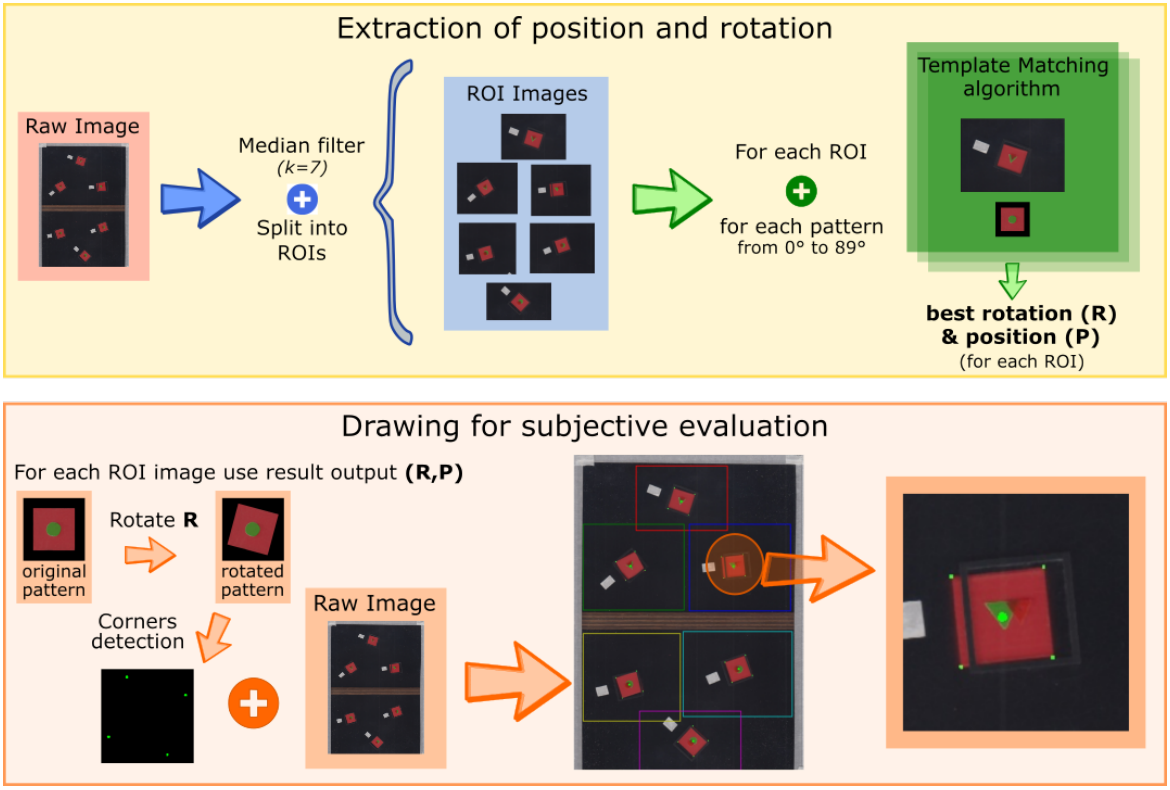
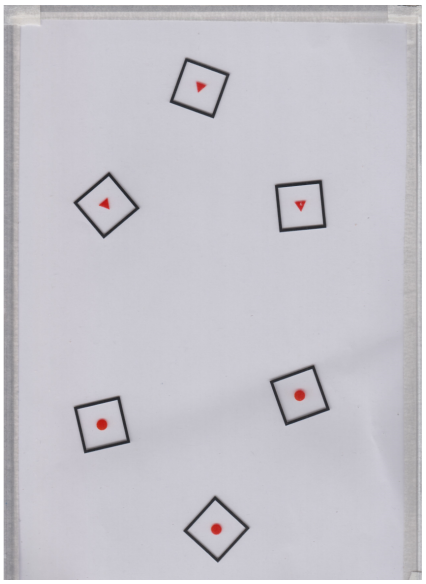
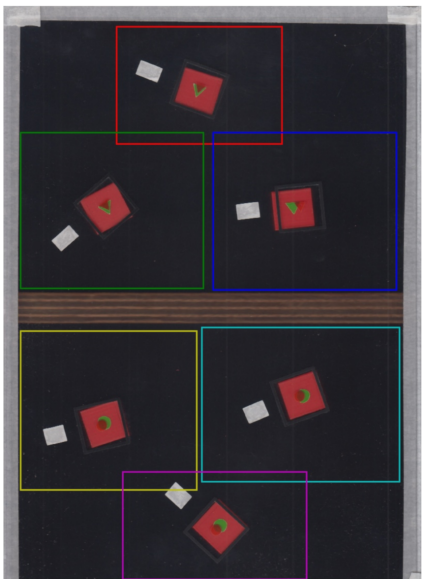


Fig. B.1 Summary of the cubes detection algorithm

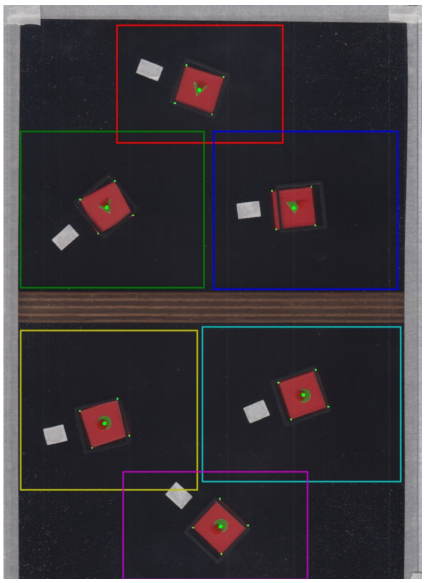
(a) Raw image used for cubes position and rotation extraction



(b) Raw image used for targets position and rotation extraction

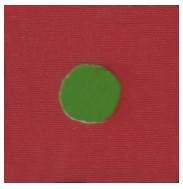


(c) Median filter and regions of interest



(d) Corners and center visualization

Fig. B.2 Example of the cubes' position detection



(a) Cube's detection pattern

(b) Target's detection pattern

Fig. B.3 Pattern images used in the template matching algorithm

Appendix C

User studies questionnaires

***

Head movement study

For both studies (navigation techniques evaluation) a demographics questionnaire was presented at the beginning of the experiment; a questionnaire concerning the users' state of presence was presented at the end of each technique's evaluation; a comparative questionnaire between both techniques was presented at the end of the second technique's evaluation. A system usability questionnaire was presented at the end of the experiment, only for the main study (second iteration).

Demographics questionnaire

1. Age _____
2. Gender
   ○ Female ○ Male
3. Dominant hand
   ○ Right ○ Left ○ Ambidextrous
4. Do you wear prescription glasses?
   ○ Yes ○ No
5. If so, are you wearing them during the experiment?
   ○ Yes ○ No
6. Have you ever used haptic feedback devices?

	1	2	3	4	5	
Never	○	○	○	○	○	Everyday

7. If so, in what occasion was it? _____

8. Do you have an experience in the manipulation/insertion of medical needles?

	1	2	3	4	5	
Never	○	○	○	○	○	Everyday

9. Have you ever used 3D virtual environments?

	1	2	3	4	5	
Never	○	○	○	○	○	Everyday

10. If so, in what occasion was it? _____

11. How often do you play 3D video games?

	1	2	3	4	5	
Never	○	○	○	○	○	Everyday

Presence questionnaire

Describe your navigation experience inside the 3D environment by checking the appropriate item on a 5-point scale, in agreement with the content of the question and descriptive labels. Please be aware of the full scale when you register your answers, especially when intermediate levels are involved. Answer the questions independently of each other and in the order in which they appear. Do not skip questions and do not return to a previous question to change your answer.

1. To what extent the mechanisms to change your point of view in the environment seemed natural?

	1	2	3	4	5	
Extremely artificial	○	○	○	○	○	Completely natural

2. To what extent the sense of movement within the virtual environment was it realistic?

	1	2	3	4	5	
Not realistic	○	○	○	○	○	Very realistic

3. To what extent your senses were fooled by the realism of the movement of objects in the space?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very misled

4. To what extent the experiences that you had in the virtual environment resembled those of a real environment?

	1	2	3	4	5	
Nothing at all	○	○	○	○	○	Very much alike

5. To what extent were you able to actively explore the virtual environment?

	1	2	3	4	5	
Nothing at all	○	○	○	○	○	Completely

6. How close were you able to examine the objects in the scene?

	1	2	3	4	5	
Very far	○	○	○	○	○	Very close

7. To what extent were you able to examine the objects in the scene from different angles?

	1	2	3	4	5	
Nothing at all	○	○	○	○	○	Completely

8. Were you able to anticipate the effects of the movement that you performed?

	1	2	3	4	5	
Nothing at all	○	○	○	○	○	Completely

9. To what extent were the delays between your actions from their consequences perceived?

	1	2	3	4	5	
No delay at all	○	○	○	○	○	Long delay

10. In terms of interactions and movements in the virtual environment, to what extent did you feel competent at the end of the experiment?

	1	2	3	4	5	
Not competent	○	○	○	○	○	Very competent

11. To what extent the control of your point of view interfered with the execution of the tasks required?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Greatly interfered

12. I felt that the movements within the scene corresponded to my own movements.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

13. I felt that the point of view of the scene corresponded to someone else's viewpoint.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

14. I felt as if the virtual representation of the scene moved just as I wanted, as if it obeyed my will.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

15. I expected the virtual representation of the camera to react in the same way that my own eyes.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

16. I felt able to interact with the environment in the way that I wanted.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

Comparative questionnaire between the navigation techniques

1. What do you think is the easiest technique to use?

   ○ Head-Tracking ○ Touch-Based

2. With what technique did you carry out your best performance?

   ○ Head-Tracking ○ Touch-Based

3. What is the technique better adapted to perform the sub-task “Find the entry point into the tissue”?

   ○ Head-Tracking ○ Touch-Based

4. What is the technique better adapted to perform the sub-task “Inserting the needle inside the tissue”?

   ○ Head-Tracking ○ Touch-Based

5. What is the technique better adapted to perform the sub-task “check the achievement of the target”?

   ○ Head-Tracking ○ Touch-Based

6. What technique do you prefer the most?

   ○ Head-Tracking ○ Touch-Based

Hand representation study

For both studies (hand representation evaluation) a demographics questionnaire was presented at the beginning of the experiment, and a system usability questionnaire was presented at the end of the experiment (retention-test). For the validation study, a questionnaire concerning the users' state of presence was presented at the end of each experimental condition. For the longitudinal study, a post-training questionnaire was presented for the two training groups at the end of the training phase, and a comparative questionnaire (virtual reality vs. physical prototypes) was presented after the system usability questionnaire at the retention-test for all the groups (control and training groups).

Demographics questionnaire

1. Participant number _____
2. Group
   - ○ Training group 1
   - ○ Training group 2
   - ○ Control group
   - ○ Single evaluation group

General information

1. Age _____
2. Gender
   - ○ Female
   - ○ Male
3. Dominant hand
   - ○ Right
   - ○ Left
   - ○ Ambidextrous
4. Do you wear prescription glasses?
   - ○ Yes
   - ○ No
5. If so, are you wearing them during the experiment?
   - ○ Yes
   - ○ No

Video games experience

1. Have you ever played video games?

   ○ Yes ○ No

2. If so, how many years since you have started playing video games? _____

3. If so, how often do you play video games?

   ○ Everyday

   ○ Once a week

   ○ Once a month

   ○ Once a year

   ○ Once every few years

Experience with virtual reality

1. Have you ever used 3D virtual environments?

   ○ Yes ○ No

2. If so, how often do you use 3D virtual environments?

   ○ Everyday

   ○ Once a week

   ○ Once a month

   ○ Once a year

   ○ Once every few years

3. If so, in which context do you use 3D virtual environments? _____

4. Have you ever used virtual reality head-mounted displays?

   ○ Yes ○ No

5. If so, for how long do you wear it?

   ○ Less than 10 minutes

   ○ Between 10 and 30 minutes

   ○ More than 30 minutes

6. If so, in which context do you use virtual reality head-mounted displays? _____
7. Have you ever used haptic feedback devices?

   ○ Yes ○ No

8. If so, when was the last time? _____
9. If so, in which occasion was it? _____

Presence questionnaire

Describe your navigation experience inside the 3D environment by checking the appropriate item on a 5-point scale, in agreement with the content of the question and descriptive labels. Please be aware of the full scale when you register your answers, especially when intermediate levels are involved. Answer the questions independently of each other and in the order in which they appear. Do not skip questions and do not return to a previous question to change your answer.

1. How much were you able to control the events?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Completely

2. How reactive (sensitive) was the environment to your actions?

	1	2	3	4	5	
Not responsive	○	○	○	○	○	Completely responsive

3. How natural did your interactions with the environment seem?

	1	2	3	4	5	
Extremely artificial	○	○	○	○	○	Completely natural

4. How much did the visual aspects of the environment involve you?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Completely

5. How natural was the mechanism which controlled movement through the environment?

	1	2	3	4	5	
Extremely artificial	○	○	○	○	○	Completely natural

6. How compelling was your sense of objects moving through space?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very compelling

7. How much did your experiences in the virtual environment seem consistent with your real world experiences?

	1	2	3	4	5	
Not consistent	○	○	○	○	○	Very consistent

8. Were you able to anticipate what would happen next in response to the actions that you performed?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Completely

9. How completely were you able to actively survey or search the environment using vision?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Completely

10. How compelling was your sense of moving around inside the virtual environment?

	1	2	3	4	5	
Not compelling	○	○	○	○	○	Very compelling

11. How closely were you able to examine objects?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Very closely

12. How well could you examine objects from multiple viewpoints?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Extensively

13. How involved were you in the virtual environment experience?

	1	2	3	4	5	
Not involved	○	○	○	○	○	Completely involved

14. How much delay did you experience between your actions and expected outcomes?

	1	2	3	4	5	
No delays	○	○	○	○	○	Long delays

15. How quickly did you adjust to the virtual environment experience?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Less than one minute

16. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

	1	2	3	4	5	
Not proficient	○	○	○	○	○	Very proficient

17. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Prevented task performance

18. How much did the control devices interfere with the performance of assigned tasks or with other activities?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Interfered greatly

19. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Completely

20. How well could you actively survey or search the virtual environment using touch?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Completely

21. How well could you move or manipulate objects in the virtual environment?

	1	2	3	4	5	
Not at all	○	○	○	○	○	Extensively

22. I felt that the virtual (hand was / tool represented) my hand.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

23. I felt that the virtual (hand was part / tool was the continuity) of my body.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

24. I felt that I was losing control of my hand when the virtual (hand / tool) was not responding correctly.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

25. I felt as if the virtual representation of the (hand / tool) moved as I wanted it as if it obeyed my will.

	1	2	3	4	5	
Completely disagree	○	○	○	○	○	Completely agree

26. In general I think I prefer to use the system:

	1	2	3	4	5	
With the virtual hand	○	○	○	○	○	Without the virtual hand

Post-Training questionnaire

In this questionnaire you can express your feelings after the end of the training phase (10 sessions).

1. After the 10 training sessions I felt very efficient at doing the task.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

2. I felt equally involved in each training session.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

3. I felt that each training session contributed to my improvement.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

4. I felt that the training phase had more sessions than necessary to be comfortable at doing the task.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

5. I felt that the training phase had fewer sessions than necessary to be comfortable at doing the task.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

6. I felt that the time for each training session was not enough to get involved in the task.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

7. I felt that the time allotted for each training session was longer than necessary to get involved in the task.

	1	2	3	4	5	
Strongly disagree	○	○	○	○	○	Strongly agree

8. I felt that understanding how the system worked took me too long compared to the time to learn the task.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

9. I felt that the training sessions interfered with my daily tasks.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

10. Do you think that there were missing functionalities in the system during the training phase?

○ Yes ○ No

11. If so, which functionalities would you add to the system? _____

12. You can write here other comments and suggestions related to the training phase: _____

VR vs physical prototypes comparative questionnaire

To answer this questionnaire, you should consider the virtual reality simulator as a whole (regardless of the two conditions tested).

1. How would you rate the feeling that your real hand was controlling virtual tools?

Not at all realistic	1	2	3	4	5	Very realistic
	○	○	○	○	○	

2. How did you find the resemblance of the weight of virtual tools compared to real tools?

Not at all similar	1	2	3	4	5	Very similar
	○	○	○	○	○	

3. How did you find the resemblance of the weight of the virtual cubes compared to the real cubes?

Not at all similar	1	2	3	4	5	Very similar
	○	○	○	○	○	

4. How would you describe the grip of virtual tools compared to real tools?

Not at all similar	1	2	3	4	5	Very similar
	○	○	○	○	○	

5. How would you rate the comfort of using virtual tools compared to real tools?

Not at all similar	1	2	3	4	5	Very similar
	○	○	○	○	○	

6. How did you find the shape of the handles used to manipulate the virtual tools?

Not at all similar	1	2	3	4	5	Very similar
	○	○	○	○	○	

7. In general how would you rate the degree of resemblance of the sensations felt between the virtual and real manipulations?

Not at all resembling	1	2	3	4	5	Very resembling
	○	○	○	○	○	

8. In general how would you describe the force feedback during the interaction in the virtual world compared to the real world?

Not at all resembling	1	2	3	4	5	Very resembling
	○	○	○	○	○	

9. How did you perceive the two situations (real and virtual) that have just been presented to you?

Not at all identical	1	2	3	4	5	Very identical
	○	○	○	○	○	

10. You can write here other comments and suggestions related to the comparison between the virtual reality and the physical prototypes: _____

System usability questionnaire

1. I think that I would like to use this system frequently.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

2. I thought the system was easy to use.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

3. I found the system unnecessarily complex.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

4. I think that I would need the support of a technical person to be able to use this system.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

5. I found the various functions in this system were well integrated.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

6. I thought there was too much inconsistency in this system.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

7. I would imagine that most people would learn to use this system very quickly.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

8. I found the system very cumbersome to use.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

9. I felt very confident using the system.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

10. I need to learn a lot before I can use this system.

Strongly disagree	1	2	3	4	5	Strongly agree
	○	○	○	○	○	

Appendix D

Additional results analysis

***

General participants' remarks

Participants expressed some general remarks at the end of the experiment in an open question. Hereafter we summarize them.

- Ergonomic:
  - Sitting position is uncomfortable and tiring.
  - Plastic holders in the VRP are uncomfortable and hurt after a long-time use.
  - Plastic holders finger holes in the VRP were too big and tools slipped.
  - HMD with glasses on is uncomfortable.
- Realism:
  - Cubes' movement is difficult to predict due to having only two contact points instead of planar contact in the VRP.
  - The two contact points made pick sub-task harder in the VRP.
  - Differences were found between left and right tools for the pick sub-task in the VRP, left tool was perceived as less accurate.
  - Cubes' size was perceived as smaller in the VRP.
  - Cubes and tools weights felt stronger in the VRP.
  - Vibrations were experienced when increasing speed in the VRP.
  - Haptic devices working-space is reduced for tool manipulation in the VRP (translation and rotation movements).
  - Haptic devices block on some configurations in the VRP (robots' articulations).

  - Place sub-task was easier in the VRP due to the absence of cubes sliding on the planar target surface (scanner in the PP).
  - Audio may have an impact on immersion.

- Hand visualization in the VRP:
  - The task required to better focus on the tools and cubes positions rather than the virtual hands.
  - Not having hands allowed to have a better focus on the task.
  - Virtual hands did not have an added value.

Post-training questionnaire:

Distributions of the questions scores for both groups were similar, as assessed by visual inspection.

Results of the Mann-Whitney U test showed that only score for $PTQ5$ was statistically significantly higher in HT group (group that trained in the VH condition) ($\mu = 2.4$) than in TT group (group trained in te NH condition) ($\mu = 1.33$), $U = 29.5$, $z = -2.37$, $p = .023$, using an exact sampling distribution for U. No significant differences were found for the other questions.

Table D.1 Results of the post-training individual questions.

			Mann-Whitney U		
Question	**VH** μ (σ)	**NH** μ (σ)	U	z	p
PTQ1	4.3 (0.48)	4.33 (0.71)	68.5	0.17	.880
PTQ2	4.8 (0.42)	4.78 (0.67)	69.0	0.23	.880
PTQ3	4.6 (0.70)	4.33 (1.00)	63.0	−0.20	.880
PTQ4	2.3 (1.49)	3.11 (1.27)	75.5	0.60	.566
PTQ5	2.4 (1.17)	1.33 (0.71)	29.5	−2.37	.023∗
PTQ6	1.1 (0.32)	1.22 (0.67)	62.0	−0.42	.833
PTQ7	2.9 (1.85)	2.00 (1.12)	54.0	−0.76	.487
PTQ8	1.4 (0.52)	1.67 (0.87)	79.0	0.88	.449
PTQ9	1.6 (1.07)	1.44 (0.73)	73.5	0.53	.651

$*p < .05$

Author's publications

***

International journals with peer review

1. **Ricca, Aylen**; Chellali, Amine and Otmane, Samir (2021). *Comparing touch-based and head-tracking navigation techniques in a virtual reality biopsy simulator.* Virtual Reality, 25:1, pp. 191-208. Springer Verlag. (10.1007/s10055-020-00445-7, hal-02861408) ***(Q1)***

2. Gamelin, Guillaume; Chellali, Amine; Cheikh, Samia; **Ricca, Aylen**; Dumas, Cedric and Otmane, Samir (2020). *The partner's avatar kinematic fidelity improves spatial communication in immersive collaborative virtual environments.* Personal and Ubiquitous Computing, *In press.* Springer Verlag. (10.1007/s00779-020-01431-1, hal-02898350) ***(Q2)***

International conferences with peer review

1. **Ricca, Aylen**; Chellali, Amine and Otmane, Samir (2021, March). *The influence of hand visualization in tool-based motor-skills training, a longitudinal study.* In the proceedings of the IEEE Virtual Reality 2021 Conference, Lisbon, Portugal, pp 1–8. IEEE. (hal-03132374) ***(Rang A)***

2. **Ricca, Aylen**; Chellali, Amine and Otmane, Samir (2020, November). *Influence of hand visualization on tool-based motor skills training in an immersive VR simulator.* In 19th IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2020), Recife, Brazil, pp. 260–268. IEEE. (10.1109/ISMAR50242.2020.00049, hal-02921192) ***(Rang A+)***

3. **Ricca, Aylen**; Chellali, Amine and Otmane, Samir (2018, April). *Study of self-avatar's influence on motor skills training in immersive virtual environments.* In the proceed-

ings of the 20th Virtual Reality International Conference (VRIC 2018), Laval, France, pp 1–4. ACM. (10.1145/3234253.3234304, hal-01968252)

4. **Ricca, Aylen**; Chellali, Amine and Otmane, Samir (2017, March). *Study of interaction fidelity for two viewpoint changing techniques in a virtual biopsy trainer.* In the proceedings of the IEEE Virtual Reality 2017 Conference, Los Angeles, CA, USA, pp. 227–228. IEEE. (10.1109/VR.2017.7892259, hal-01524543) ***(Rang A, poster presentation)***

National conferences with peer review

1. **Ricca, Aylen** and Chellali, Amine (2016, October). *Interaction fidelity in virtual simulators: two navigation techniques for a virtual biopsy trainer.* In the proceedings of Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine (IHM'16), Fribourg, Switzerland, pp. 278–284. ACM. (10.1145/3004107.30041 39, hal-01384326)

Communications to congresses, symposiums, seminars or workshops without proceedings

1. Gamelin, Guillaume; **Ricca, Aylen**; Chellali, Amine; Dumas, Cédric and Otmane, Samir (2018, October). *Effets de la forme de l'avatar du partenaire distant sur l'interaction spatiale collaborative dans un environnement virtuel.* Journées de la Réalité Virtuelle 2018. Evry, France.

2. **Ricca, Aylen**; Chellali, Amine; and Otmane, Samir (2018, April). *Influence de la main virtuelle sur une tâche d'apprentissage moteur dans un environnement virtuel immersif.* In Seminar of the Research Group Computer Graphics and Virtual Reality (GDR IGRV) –*Entraîner ses compétences motrices en réalité virtuelle.* Rennes, France.

3. **Ricca, Aylen**; Chellali, Amine; and Otmane, Samir (2017, November). *Training surgical skills in virtual reality simulators.* In colloque Evry Sciences et Innovation (ESI) – Les sciences pour l'autonomie de la personne. Evry, France. (*Second prize best poster*).

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre: Effets de la fidélité dans les simulateurs de réalité virtuelle sur l'apprentissage des compétences techniques en chirurgie

Mots clés: Interaction Homme Machine, Réalité Virtuelle, Conception Centrée-Utilisateurs, Simulation Chirurgicale, Fidélité des Simulateurs, Formation des Compétences Motrices

Résumé: La réforme des études médicales en France réserve une place importante à la simulation dans la formation des professionnels de santé. Dans ce contexte, les simulateurs en réalité virtuelle peuvent jouer un rôle important. Cependant, la fidélité de ces systèmes reste une question ouverte avec un manque de lignes directrices pour déterminer les niveaux appropriés de leur fidélité pour une formation efficace. Dans cette thèse, nous nous intéressons à la fidélité des simulateurs immersifs pour l'apprentissage des compétences techniques en chirurgie. Plus particulièrement, nous cherchons à comprendre comment la fidélité des interactions et artefacts pourrait affecter l'efficacité de certaines tâches effectuées dans le simulateur. Sur le plan théorique, nous proposons une articulation des notions de fidélité des simulateurs à travers un modèle basé sur les différents composants de la fidélité identifiés dans notre revue de littérature, et des différentes interfaces et interactions utilisées dans les simulateurs virtuels pour la formation des compétences techniques. Deux études expérimentales ont été réalisées pour explorer les facteurs de la fidélité du simulateur associés à une tâche de navigation et à une tâche de manipulation d'un outil et leurs impacts sur l'apprentissage de deux gestes techniques en réalité virtuelle. Les résultats montrent que des techniques d'interaction et artefacts du simulateur avec une fidélité modérée peuvent supporter la réalisation de tâches secondaires pour une formation efficace. Les concepteurs des simulateurs virtuels peuvent s'appuyer sur les recommandations issues de nos travaux afin d'éviter d'inclure certains dispositifs et composants de la fidélité qui peuvent être encombrants et coûteux sans avoir un réel impact sur l'efficacité de ces simulateurs.

Title: Effects of fidelity in virtual reality simulators on learning technical surgical skills

Keywords: Human Computer Interaction, Virtual Reality, User Centered-Design, Surgical Simulation, Simulator Fidelity, Motor Skills Training

Abstract: The reform of medical studies in France attributes an important role to simulation in healthcare professionals' training. In this context, virtual reality simulators can be very useful. However, the fidelity of these systems remains an open question with a lack of guidelines for determining their appropriate levels of fidelity to support effective training. In this thesis, we are interested in the fidelity of immersive simulators for the training of technical skills in surgery. More particularly, we investigate how the fidelity of interactions and artifacts to perform certain tasks in the simulator could affect the efficiency of these systems. From a theoretical perspective, we propose an articulation of the concepts of simulator fidelity through a model based on the various fidelity components identified in our literature review, and the various interfaces and interactions used in virtual simulators for the training of technical skills. Two experimental studies were conducted to explore the factors of simulator fidelity associated with a navigation task and a tool handling task and their impacts on the learning of two technical skills in virtual reality. The results show that the simulator's interaction techniques and artifacts with a moderate fidelity can support the completion of secondary tasks for effective training. The designers of virtual simulators can rely on the recommendations resulting from our work in order to avoid including certain devices and fidelity components, which can be cumbersome and expensive without a real impact on the efficiency of these simulators.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France