Doctoral Dissertation

Understanding the Effectiveness of Augmented Reality for Training Individuals

Isidro III Mendoza Butaslac

Program of Information Science and Engineering
Graduate School of Science and Technology
Nara Institute of Science and Technology

Supervisor: Professor Hirokazu Kato
Interactive Media Design Lab. (Division of Information Science)

Submitted on January 31, 2023

A Doctoral Dissertation
submitted to Graduate School of Science and Technology,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of Engineering

Isidro III Mendoza Butaslac

Thesis Committee:

Supervisor	Hirokazu Kato (Professor, Division of Information Science)
	Kiyoshi Kiyokawa (Professor, Division of Information Science)
	Masayuki Kanbara (Associate Professor, Division of Information Science)
	Yuichiro Fujimoto (Assistant Professor, Division of Information Science)
	Taishi Sawabe (Assistant Professor, Division of Information Science)

Understanding the Effectiveness of Augmented Reality for Training Individuals[*]

Isidro III Mendoza Butaslac

Abstract

Computer-mediated training utilizing augmented reality (AR) is slowly being added and used together with traditional training methods as more technological barriers are addressed over the years of research and development. Although the use of AR for training has a lot of potential, its effectiveness is not as straightforward as there are a lot of concepts that are contested and poorly understood that warrants further investigation. The goal of this study is to understand the implementation of augmented reality training systems (ARTS) by exploring how and under which circumstances ARTS work for training individuals, in this case people involved in Physical/Occupational therapy and Surgical training. This study took a realist approach through the formulation of a program theory. This is done by the theory elicitation of context-mechanism-outcome (CMO) configurations, then tested through the confirmation of each of the respective hypotheses of these configurations. Empirical evidence and logical induction were both used to appraise the training effects of these CMO configurations. Specifically, empirical evidence involved user studies and expert feedback while logical induction involved related work generalizations and model conceptualizations. Finally, this program theory is generalized by identifying what facilitates or constrains the implementation of training with AR.

Keywords:

Augmented Reality, Training, Participatory Design, Realist Evaluation, Surgery, Physical and Occupational Therapy

---

[*] Doctoral Dissertation, Graduate School of Science and Technology, Nara Institute of Science and Technology, January 31, 2023.

Contents

List of Figures

List of Tables

1. Introduction

1.1 Background of Augmented Reality

1.1.1 Augmented Reality - Overview and Definitions

Around the 1990s, the term “augmented reality” was coined by Thomas Caudell and David Mizell as a technology that is “used to augment the visual field of the user with information necessary in the performance of the task” [42]. Thirty years later, their definition of AR resurfaced as the prevailing phrase used to characterize the modern computing paradigm, and it has changed how people and the world interact. The preliminary tests in the early 1990s have vividly showed that the pioneers acknowledged the technology’s promise and realized that AR could bring a lot of benefits when used in real-world scenarios. Classic examples of AR would be the work by Feiner et. al, KARMA, which was the first knowledge-driven AR application. Their system was able to do automatic inferences and give a sequence of instructions for repair and maintenance procedures appropriately, as seen in Figure 1.1 [66]. Another example is the work of Rekimoto and Nagao, NaviCam, which is considered to be the first true handheld AR display. It could detect color-coded markers based from the camera image taken from the video feed, rendering the information on a video see-through style of view just like in Figure 1.2, showing the description about Rembrandt [179].

AR can be considered a variant of virtual environments (VE), otherwise known as virtual reality (VR). VE/VR systems work by completely immersing the user in a digital world. In comparison, AR helps the user to see the real world and it allows virtual objects to be superimposed on or combined with the real world. In this sense, AR expands the reality instead of replacing it altogether. As a result, the user will assume that virtual and real objects can coexist in the same workspace. With this consideration, Azuma established three defining characteristics of AR systems[14]:

1. Combines real and virtual

2. Interactive in real time

3. Registered in 3-D

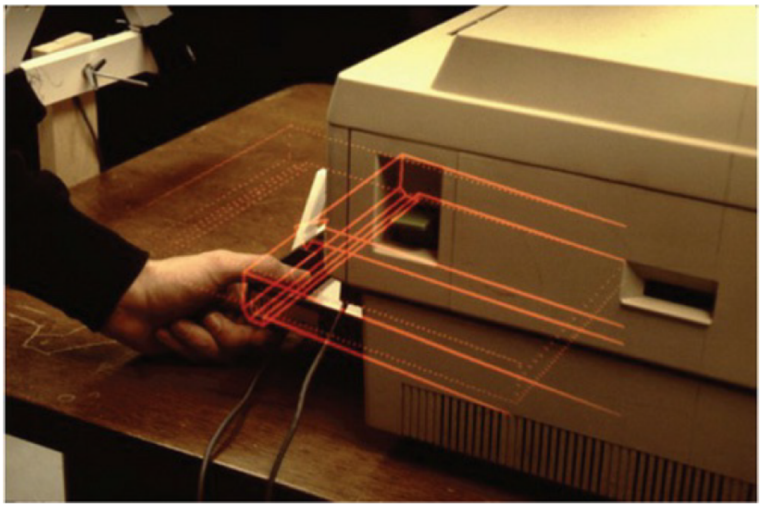


Figure 1.1: KARMA: Knowledge-based Augmented Reality for Maintenance Assistance

Figure 1.2: NaviCam: first true handheld AR display

The definition above by Azuma allows for other systems other than head-mounted displays (HMDs) to be considered as AR. The usage of AR technology techniques, such as computer vision and object recognition, can help the user interact and digitally manipulate the data surrounding the real world. With the help of AR, information about the environment can be overlaid and augmented in the user's real-world view. AR has the capability of improving the perception and interaction between a user and the real world. This is because virtual objects show details that the user cannot perceive directly with their senses. The extra information the virtual objects convey can help the users perform real-world tasks [14].

Owing to AR's multidisciplinary aspect, researchers from various backgrounds have come to define it a bit differently. As Santos et al. pointed out, what is considered to be AR is debatable depending on the quality of the implementation. They cite the example of "imitating the effect of AR... is simulated only by flashing relevant information on a screen. It does not employ any kind of tracking," which is still considered by some scholars as AR [188]. We have decided to include them because they also provide useful insights about the training methods and strategies, even if they are, in the strict definition, not AR. In this loosened definition of AR, we have found in our collected articles that authors tend to use the term Mixed Reality (MR), as long as it satisfies the elements of physical venue, virtual medium, and the user's interactive imagination [98]. Examples of training setups that fall into this category are works that use special input equipment such as fetoscopes [109] or phantom haptic interfaces [111, 212], displayed in a stationary device such as a large monitor.

1.1.2 Augmented Reality in Medicine

AR in the medical field is already a thoroughly researched area. For example, AR technology can be used to give surgeons "X-ray vision" to allow them to see the insides of the patient. To name a few of the classic literatures in this area, in Figure 1.3 we show in real-time 3D ultrasound visualization and display the fetus view inside the womb of an expecting mother, with the doctor wearing the HMD has the impression of using a "3D stethoscope" [207]. Another example is on the operation of biopsy of the breast through an ultrasound-guided needle. In

Figure 1.4 we see an example of guiding the insertion of the needle of a practice biopsy operation [208].

Figure 1.3: AR Visualization of fetus inside womb

There is already a plethora of technologies that are used in conjunction with the motor rehabilitation. To name one example that is similar to the prototype that will be discussed later, is a system that measures and manipulates pelvic motion during step training on a treadmill [100]. The system can be used in passive mode to monitor pelvic trajectories, either manually determined by a therapist or pre-recorded from unimpaired subjects, and then replay those trajectories with a non-linear algorithm for force regulation. The device also has the ability to record and repeat the pelvic movements that occur during normal walking. These kinds of technology offer numerous possibilities for gait preparation in stroke therapy thus removing the physical therapist's excessive repeated activity in a non-ergonomic role [144].

The trend of using AR to therapy is expressed in the rehabilitative clinical environment situation, however therapists still solely rely on the “clinical eye”, which is the medical professional's senses and deductive ability that is gained through years of experience practicing the field. The potential that we suggest in this work is to use AR to complement and improve clinical observation skills by providing additional information and data. This gives the therapist to not

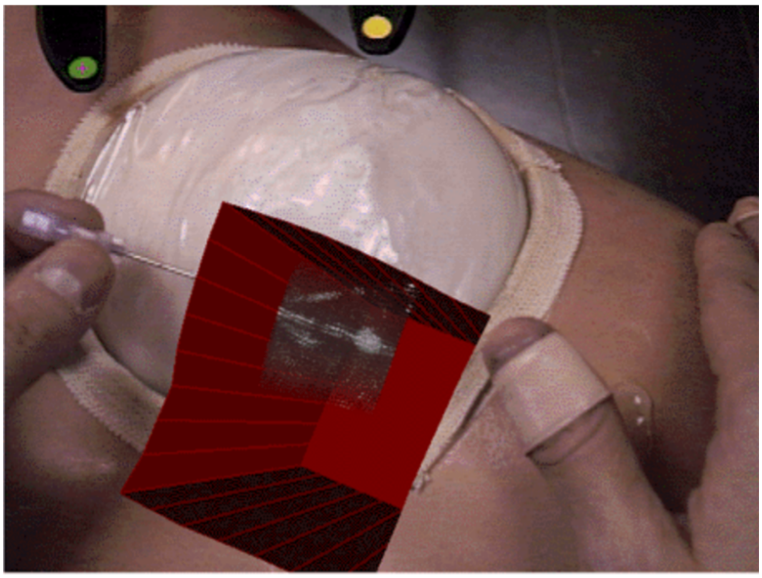


Figure 1.4: AR Visualization of needle during biopsy of breast

only be provided a clearer picture of the patient's performance but also provide appropriate objective means for assessment and classification. Some examples of data that can be used for this kind of application can be the motion of patient observed in 3D, their interaction in relation to the environment, and some measurable physical parameters about the patient.

Augmented Reality in Rehabilitation

On example of an AR system that is used in Rehabilitation is the AUSILIA project. This system is an apartment-wide project fitted with a technical infrastructure that allows the individual to be tracked and supported at the premise during his stay. It is an atmosphere in which the individual stays for a certain period of time, and his daily operations are controlled by an integrated system in the room full of sensors. In this way, the doctor will be able to monitor and evaluate the patient remotely, all important data such as interactions with the environment and also physiological parameters, as shown in Figure 1.5[43].

AUSILIA's main purpose is the design and development of cost-effective and efficient solutions based on a scalable architecture consisting mainly of commonly used hardware and software modules on the market or accessible as open source. The method applied allows for a scalable and up-do-date platform that reduces

Figure 1.5: Assisted Unit for Simulating Independent Living Activities, AUSILIA project in Italy

the cost of designing, developing and maintaining the laboratory thus creating affordable solutions for the end user. The design of the apartment was structured to make it a possibility to collect data in real time, and provide user experience feedback based on their interactions with the sensorized environment. The collection of this data provided many opportunities: Providing a platform for successful environment monitoring to reduce potential risks to the heal and safety of the patients (as these devices can be transferred to the patient's home for remote monitoring and telemedicine purposes, upon discharge from the hospital); Providing a resource for the evaluation of the advantages received by the patient with the network to guarantee that the medical and technical personnel are able to evaluate the best support approach for a given user. Assessing the level of confidence the individual retains when working with the services offered, by measuring his emotional and stress parameters; Populating the laboratory database for further analysis and research on the patient's behavior, such as detection of standard repetitive circumstances or trends, user profile descriptions, etc. Confidentiality will not be an issue because of the introduction of a data anonymization protocol that would ensure privacy. The technologies and techniques the monitoring system AUSILIA benefits from are for example video cameras, presence sensors, identification sensors, pressure and force sensors, detectors for sliding, falling, and localization toolkits [171].

The project focuses on the advantages of a new interaction methodology, where the scenario presented includes the appropriate parameters at the same moment, and the associated data is precise and contextualized for use of therapist as shown in Figure 1.6. The image shows the information of the patient where the patient shape and motion is visualized through skeletonization, and the information of the force applied to the floor. The right side image proposes one possibility of visualization for these kind of data, while the left side image shows the same set of data but in a detached and decontextualized manner [43].

Figure 1.6: Visualization of Patient Information in AUSILIA

From the AUSILIA project follows the work of Stocco where he introduces AR to the rehabilitative setting with the goal of improving the quality of the medical service they provide. He presented a way to augment the therapist's clinical eye with data such as skeletons, applied force visualizations, and an overview of the status of the patient as seen in Figure 1.7 [210]. My work extends from where he has left off, such as providing real and useful clinical applications to the data of skeletons, providing a stream of actual (not fake) sensor values be it environmental or physiological, and providing in-situ visualizations of the acquired data rather than doing a panel-view that acts like an interactive 2D monitor in 3D space.

Augmented Reality in Surgery

When compared to the use of 2-D pictures in image-guided surgeries, the integration of the surgical site with digital preoperative patient-specific images

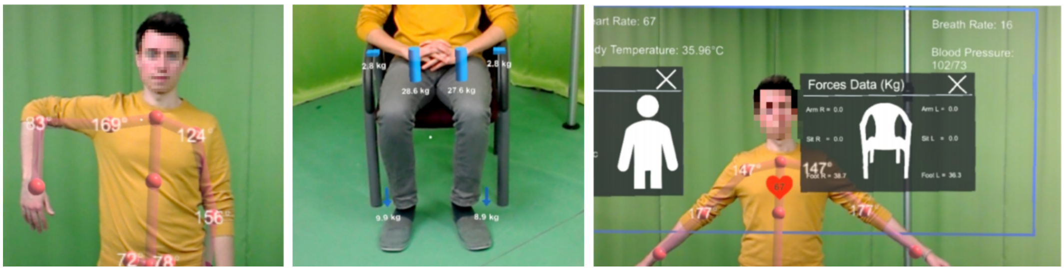


Figure 1.7: Previous Work: Improving Clinical Eye by Supplementing AR Visualizations to Therapists

enhances performance. Indeed, the current research indicates that surgeons are becoming increasingly interested in incorporating augmented reality (AR) into surgical operations, resulting in enhanced surgical safety and effectiveness. Due to limited organ movement and distortion, augmented reality is proven effective in the delivery of surgical operations in orthopedic surgery, neurosurgery, head and brain space, and hepatobiliary and pancreatic surgery, where tissue deformation and surgical precision remain technological challenges [222].

AR enables the projection of aligned CT, MRI, and ultrasound images onto the patient's body. This assists with orientation, anatomical delineation, awareness, and surgical approach development [168]. As shown in Figure 1.8, Pratt et al. conducted a case series in which they employed the Microsoft HoloLens AR to guide operational incisions during extremities reconstruction surgery by superimposing computed tomography angiography scan information of the subsurface vascular architecture on a patient's body. Without compromising environmental sterility due to the device's self-contained nature and ability to be operated by hand gestures and voice control, and with minimal procedure modifications, they concluded that the Microsoft HoloLens headset was a potent tool with the potential to reduce anaesthetic time and surgical morbidity. In addition, they found it beneficial to enhance training and provide remote assistance to the operating surgeon [175]. In another case study of AR-guided surgery by Gregory et al., the capacity to superimpose the patient's anatomy on their body and the ability to gain immediate feedback from colleagues streaming the process resulted in increased safety during shoulder replacement surgery [84].

Furthermore, Wachs showed that image-guided surgery, although allowing for

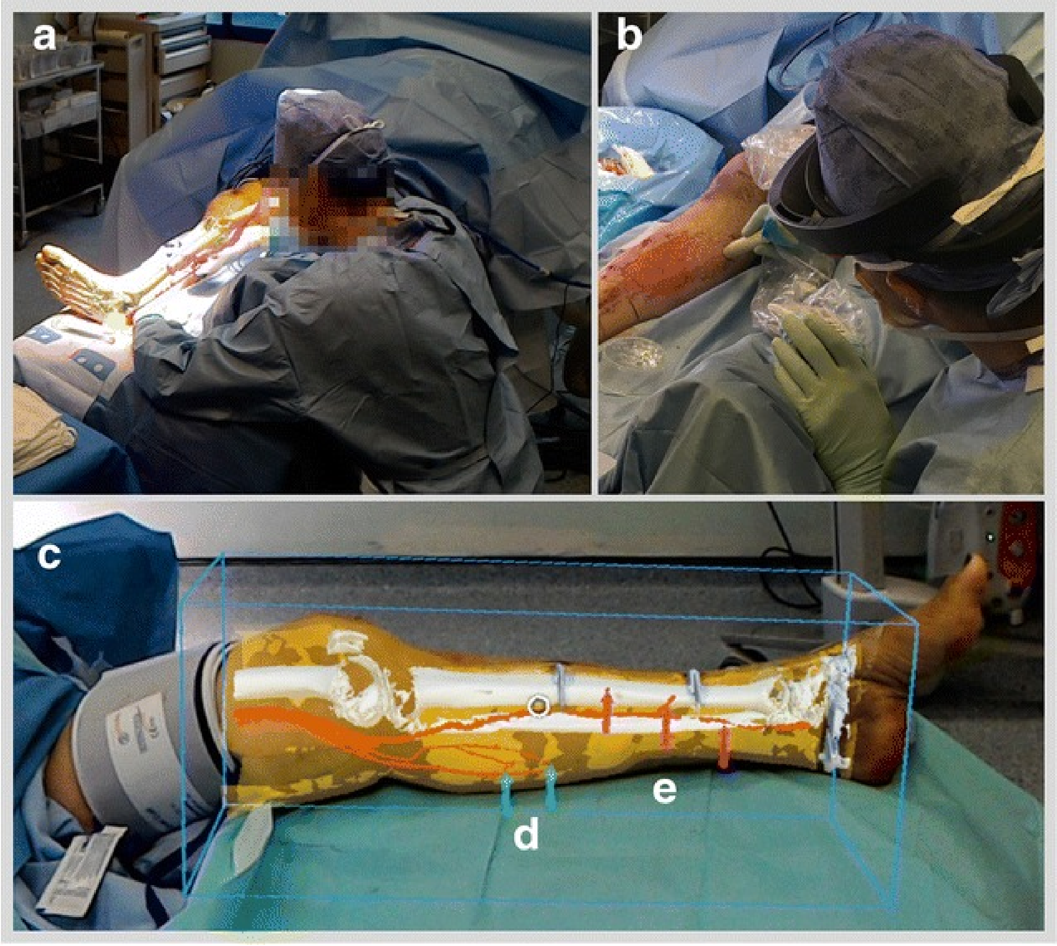


Figure 1.8: AR has potential for superimposing real-time, three-dimensional interior anatomy on the patient. (a) AR overlay of models seen from a distant HoloLens; (b) audible Doppler ultrasonography to check perforator position. (c) Superimpose of visualizations using a bounding box. [175]

enhanced safety and surgical performance, produces cognitive difficulties from the surgeon to the patient and the monitor, which adversely impacts cognitive and motor activities [224]. While thinking in situ about the location, overlays of data utilizing AR reduce eye movements away from the work space and increase surgical field attention as opposed to computer displays. In a simulated surgical work, Stewart and Billinghurst claim enhanced attention to the surgical field using a "through-the-lens" HMD in contrast to a peripheral display a regular computer monitor [209]. As a result, AR may minimize attention shifts and the time required to complete a task, resulting in an improved perception of performance by surgeons, due to the fact that superimposing data on artifacts reduces the cognitive separation between data and the artefact in hand, thereby facilitating better rationalization about information. This is corroborated by eye movement study demonstrating a loss of recall and spatial awareness when eyes are diverted from a work place [222].

AR also offers new opportunities for postoperative surgical navigation. Using several depth cameras, a computer-based system might record a surgical process and rebuild in 3D the changing circumstances occurring throughout the surgery. It permits a reexamination of the technique for training or assessment reasons. Cha et al. proposed a computer-based system for acquiring and designing 3D-plus-time data, allowing a user to retrospectively walk around the reconstruction of the procedure room while managing the recording of the medical operation using simple controls (e.g., play, pause, rewind, fast forward) [72].

1.2 Approach

1.2.1 Training Definition

The definition of training has evolved since the dawn of the industrial revolution. A literature review conducted by Somasundaram et al. indicated that one of the earliest definitions of training was given by Black in 1961: "imparting job knowledge to employees so that they can carry out orders smoothly, efficiently, and cooperatively" [27]. As we approach the 21^{st} century, training according to Goldstein and Ford is defined as the "systematic acquisition of skills, rules, concepts, or attitudes that results in improved performance in another environment"

[81].

It can be observed that such a common word used in everyday life has a plethora of interpretations. From this, Somasundaram et al. resolved this ambiguity by analyzing and synthesizing the available definitions of training and development, and listed the dependent variables, area of focus, and core elements of these definitions [204]. Based on the results from their review paper, Somasundaram et al. proposed three major categories for elucidating the purpose of training:

1. Develop or gain knowledge
2. Develop or gain skills
3. Improve performance

Aside from training systems, a common terminology that is also widely used in the development of technologies is support/assist systems. It is important to note that training systems and support/assist systems have the same goal of improving the performance of the end user, whether it be for the benefit of the user or some organization. The biggest difference here is that training systems also offer the user the opportunity to develop and gain knowledge and skills that he or she did not have before. Since support/assist systems do not offer these benefits, this may lead the user to develop a dependency toward the system such that when the system is removed, one’ s performance heavily drops.

1.2.2 Realist Approach

The premise of a realist approach is that the same intervention will not work everywhere and for everyone. The emphasis on this kind of approach is looking at “what works, for whom, under what circumstances and how” [232]. Realist approach focuses on the concerns of causality (the act of causing something) and attribution (the act of attributing something). In this field of study, Pawson and Tilley introduced the phrase “realist evaluation” [167].

All assessment methods are founded on philosophical ideas. Evaluation that is based on reality (a philosophical perspective in which the social world is viewed as real). Consequently, non-observable entities and processes, such as culture, class, and economic systems, may have a significant impact on whether or not an intervention is successful. In terms of the movement of people, resources, and data, social systems such as the family, schools, and economic institutions have dynamic borders. These social systems interact, hence system boundaries must be created for the assessment, despite the fact that these limits may not exist in reality. Interventions are inherently dynamic, open systems. These may interact with other social systems; hence, causality is not a straightforward, linear process. They may come from changes in and interactions between various social systems.

Realist techniques are suitable for assessing complex interventions, such as community-based public health programs with a greater possibility for learning. They are especially valuable for examining interventions with varied results in order to better understand how and why divergent outcomes arise. It is inappropriate when it is well known how, why, and where programs function, when the program is straightforward and one-size-fits-all, or when simply the net impact of the intervention is of interest.

Context, mechanisms, and outcomes are the three pillars of realist evaluation. The appraiser produces Context-Mechanism-Outcome (CMO) hypotheses, – i.e., hypotheses on which mechanisms are expected to work in specified contexts and the observed outcomes as they occur.

Context impacts whether an intervention's mechanisms function. For instance, results may differ based on economic, geographical, historical, social, and political conditions, as well as individuals' cultural beliefs. Variations within the intervention's intended demographic might also impact which techniques may be used. A Realist Evaluation hypothesizes which context elements will influence how and for whom an intervention will function, then collects data on those context elements.

Mechanisms are "···underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest." [10]. Due to the fact that mechanisms need the proper environment to function, any changes to the system might impact the causative chain. In social interventions, the mechanism is the cognitive or emotional reasoning of the target people in response to the intervention's asset, possibility, or limitation [167].

Short-, medium-, and long-term planned or unintentional outcomes may be produced by an intervention. Additionally, there may be several outcomes with varied value for various participants.

“In summary, realism holds that mechanisms matter because they generate outcomes, and that context matters because it changes... the processes by which an intervention produces an outcome. Both context and mechanism must therefore be systematically researched along with intervention and outcome. By implication, research or evaluation designs that strip away or ‘control for’ context with a view to exposing the ‘pure’ effect of the intervention limit our ability to understand how, when and for whom the intervention will be effective.” [231]

1.2.3 Design Philosophy

To quote an excerpt of Ehn 1989, as cited in Spinuzzi’s 2005 paper on the Methodology of Participatory Design (PD), “ [PD] attempts to steer a course ‘between tradition and transcendence’ - that is, between participants’ tacit knowledge and researchers’ more abstract, analytical knowledge. [63, 205]” The value of this kind of philosophy to design becomes more and more apparent as we head towards more distinct specializations of knowledge, for instance, surgical and rehabilitation expertise. The reason is that as we focus more on the tacit knowledge of this body of works, it becomes progressively challenging to verbalize and express these concepts to people of different fields. Tacit knowledge is akin to how jargons work, such that its convenience is its ability to encapsulate a complex concept into a single word/phrase. On the contrary, this same quality is its disadvantage as it becomes tough for outsiders to grasp. PD approach to research is therefore paramount when handling multidisciplinary projects.

More detailed definitions on the concepts about the Realist Approach and PD will be elucidated on section 3.

1.3 Thesis Overview

This section provides an overview of the research questions, my research contributions in an attempt to answer these questions, the hypotheses formulated to test the main research question based on the realist approach, and finally a short summary of the flow to describe this thesis.

1.3.1 Research Questions

Main Research Question:

How and under which circumstances does augmented reality work for training individuals?

Sub-research Questions:

RQ1 What are the different training mechanisms of augmented reality training systems?

RQ2 How effective can the prototype systems authored together with the co-designers realize the desired outcomes of training?

RQ3 What facilitates or constrains the implementation of training with augmented reality?

1.3.2 Research Contributions

The work described in this thesis makes the following contributions:

1. This study explored how ARTS work and under which circumstances by constructing theoretical frameworks of CMO configurations. This is done through the confirmation of the hypotheses formulated for each of the CMO configurations. (answers Main RQ)
2. This study identified the underlying mechanisms of ARTS through the synthesis of literature on computer-mediated education, learning, and training. (answers RQ1)
3. This study showed the training effects of these CMO configurations through both empirical evidence and logical induction. (answers RQ2)
4. This study identified what facilitates or constrains the implementation of training with AR, giving example scenarios in the medical setting. (answers RQ3)

1.3.3 CMO Hypotheses

CMO-1 The presentation of real-time visualizations of invisible current conditions about the patient can train the clinical eye of the therapist.

CMO-2 Providing real-time feedback that supports the understanding of exproprioceptive information can train the locomotion skill as performed in actual exercise.

CMO-3 Incorporating a valid and reliable shared AR experience into the usual rehabilitation workflow improves therapist's diagnosis and communication with the patient.

CMO-4 The use of patient's real CT data can trigger perceptions of deep immersion, which will result in improved learning, knowledge, and comfort with knowledge and skill performance.

CMO-5 Observing different perspectives of higher skilled operative performance can facilitate the transfer of tacit knowledge and skill to learners.

CMO-6 Progressive difficulty of AR presentation can promote deliberate practice, which will make the transfer of knowledge and skills from practice to actual performance smoother.

1.3.4 Thesis Outline

The work in this thesis is outlined as follows. In section 2, this study tries to capture the current state of AR training by conducting a systematic review. From this review, we gain an understanding on what are the training methods, strategies, and principles are used with AR, and these are used to describe the mechanisms part in our CMO configurations. In section 3, this thesis elucidates more on the methodology, philosophy, and design principles of this thesis, which is about the realist approach and participatory design. In section 4 this work tries to evaluate ARTS by empirical evidence while in section 5 tries to evaluate by logical induction. Each of the sub-sections in section 4 and section 5 tries to answer one hypothesis of the aforementioned CMO configurations, focusing on Physical and Occupational Therapy, and Surgery Training, respectively. Finally section 6 resummarizes everything and synthesize the answers found for each of the CMO configurations. These are then all generalized to make proposals for what makes an effective ARTS. Limitations and suggestions for future work are also discussed to provide the community interested in the development and implementation of ARTS hints of what are sensible research directions to move towards to.

2. Review of ARTS

2.1 Background of the Review

2.1.1 Scope of the Study

Human skills are acquired through training. Martin et al. defined training as, "the development and delivery of information that people will use after attending it" [142]. As defined, the acquisition of a skill often involves a personal or social necessity. This includes the activity to acquire surgical and diagnostic skills by medical students, the activity to learn how to assemble new products by factory workers, and rehabilitation by hemiplegic patients to use their upper limbs. The mechanisms of skill acquisition through training are closely related to cognitive processes and motor learning theory [70, 193]. Throughout the vast history of research, numerous "training methods" based on these theories have been proposed to effectively acquire skills.

Computer-mediated training uses a computer to aid human training from multiple perspectives, thereby increasing efficiency and lowering costs, such as minimizing training time and providing a more individualized training format [153, 15]. This paper focuses on computer-mediated training that utilizes augmented reality (AR) technology, which can superimpose information on the real world. We refer to systems that provide this type of training as augmented reality training systems (ARTS). Task support, which helps people to efficiently perform specific tasks, is one of the most common applications of AR. However, in the last couple of decades, it has also been gradually used for training. Although the use of AR for training has the potential to bring benefits, it is not straightforward. Many studies have reported its effectiveness in task support systems in helping people understand and perform certain tasks efficiently by adding the aid content using AR. The user can "rely" on the system in that case. AR task support systems are typically not designed for the user to stop using them in the future. By contrast, in many training contexts, the training system should be designed such that the user can perform the task without the system in the future.

An ideal ARTS should include special techniques and approaches (i.e., training methods) to achieve this goal. We adopted Martin et al.'s classification of training methods and analyzed how each of the existing training methods is extended by the proposed ARTS [142].

Another important aspect of training is the evaluation. Measuring the effectiveness of ARTS training can be more complex than one of the AR task support systems. The reason for this is that it is not possible to directly assess the training effect from the training itself, and it is necessary to check the acquisition of the target skill by the trainee after training. Another reason is that most training requires continuity. The requirement for continuity not only increases the cost of the experiment, but it also creates the need to separate the skill improvement of the user by the continuous training from the skill improvement that is brought by the ARTS. We refer to the evaluation types of Barsom et al. that are widely used, especially for medical systems and applications [20]. As described in Section 2.2.2, this is a five-point framework for the validity of a system to be guaranteed in order for it to reach social implementation.

In this paper, we categorize ARTS technologies that are proposed in various fields (medical, rehabilitation, industrial, etc.) in terms of their training methods [142], their use in AR are summarized, and their trends and characteristics are discussed. In addition, five evaluation types [20] are used to identify trends in the evaluations adopted in the existing ARTS papers for each application field and training method. To the best of our knowledge, there is no paper that discusses ARTS in multiple fields from the viewpoint of the training method and evaluation type. The training methods and evaluation types emphasized in each field potentially differ, but these have not been sufficiently clarified. Furthermore, the intrinsic values of AR to training are difficult to ascertain by focusing on just one specific field; they only become apparent by looking at the similarities and differences across multiple fields. Our main objective is to discuss the intrinsic value of AR in training by enumerating these across disciplines. Our research questions are as follows:

R1 What are the training methods for implementing ARTS and how are these methods used in the different fields?

R2 What types of evaluations are used to assess the existing ARTS and how do they contribute to the validity of each training method?

R3 How is AR utilized for each training strategy?

2.1.2 Current State of AR Training

As a growing trend in technology, AR is already being used in a variety of fields. In the survey that we conducted, we identified three major fields in which AR is used: medicine, rehabilitation/exercise, and industry.

It is difficult to gain hands-on expertise executing treatments without danger in medicine. AR helps to overcome this barrier by letting medical students understand anatomy and perform procedures. Understanding the interior anatomy of the human being can be a challenging task as it is outside the everyday experience and requires more creative imagination. With the help of AR, the gap between these ideas and operating processes is simplified further [127, 60, 180].

Rehabilitation and exercise in AR can enable more individuals to obtain tailored training at home [171], and perhaps increase adherence and facilitate more regulated execution of physical training sessions [128]. AR in different forms has been used effectively in skill training for doctors/healthcare staff, and it is considered to be appropriate for instruction that demands strong motor skills and meticulous spatial movement [28].

AR technology significantly benefits industry by allowing the transfer of information from expert to novice to go from just conceptual knowledge to usable hands-on experience in a swift manner [160]. For example, HMDs are trending in industrial AR as they enable the delivery of real-time, step-by-step guidance and feedback from trainers during practice. This technology is ideal for industrial workers since it allows trainees to learn faster and practice more frequently [78].

We may have identified three major categories that are the most common applications for ARTS; however, its usefulness is not limited to only these aforementioned fields. For example, we identified research on ARTS that focuses on general work activities, autonomous driving, or some form of course education such as business ethics or musical instruction. Despite the fact that these are included, they are labeled as ”Other” for scoping reasons.

2.1.3 Distinction over Existing Surveys

In the last couple of years, AR has proliferated and a number of survey papers have emerged that describe its trends. One example is a survey that was conducted by Santos et al., in which they performed a meta-analysis of AR learning experiences (ARLE) [188]. They concluded in their study that “there is a need for valid and reliable questionnaires to measure constructs related to ARLEs to iteratively improve ARLE design” [188].

A few years later, in 2020, Bianchi et al. reported that the results of their Systematic Review verified that there is no standard protocol for evaluation, at least in the scope of the medical teaching–learning process content of AR [24]. This sentiment however can be generalized to the broader sense of AR training (i.e., across multiple fields) by doing a preliminary search in the IEEE Xplore search engine using the search string [augment* AND reality AND train* AND review] on all metadata.

There are Systematic Literature Reviews that consider the evaluation methods of AR applications in the educational scenario, however saying that these methods approach are quantitative, qualitative, mixed, or not specified may be too general of a classification to draw concrete conclusions about the training [161]. There are also Literature Reviews about AR usage performance that extracted learnability factors from Kolb’s Experiential Learning Theory, and validated these factors through surveying students and academicians [88]. This study by Hafsa et al. is a good example of linking the effectiveness of AR with relation to the user’s learning; however, this method of validation is quite costly timewise and moneywise; thus, a “survey” type of validation would prove rather difficult if it is applied for a generalized study scope.

Barsom et al. [20] also checked the validity of AR systems and proposed five validity stages through which ARTS should be assessed to complete a full validation process. However, their work limits the search to AR applications for the purpose of training or educating medical professionals. They focused more on how AR is feasible to the medical field at different stages of validation.

Barsom et al.’s work is different from ours as our motivation is to gain a general understanding of the trends of ARTS research through the categorization by training methods and evaluations. We then use this understanding of the trends to extract the utilization of AR that is useful for the design of training. Hopefully, this will serve as a reference to people who want to develop ARTS but do not know what strategies they should adopt and how they should evaluate it.

In summary, the following are the distinct characteristics of this work: We consider studies that have a high impact on the AR research community. We identify the common application fields in which ARTS is currently used. We categorized ARTS by its training methods and evaluation types.

It is also important to note that this is not the first survey paper to use the study impact as a consideration for the inclusion criteria. For example, Dey et al. [61] used Google Scholar to find the total number of citations of each paper to calculate the average citation count (ACC) per year since it was published. For this paper, we have utilized Google Scholar's h-5 index to indicate the impact of a publication in the last five years. The decision to adopt this strategy rather than the ACC is to address the issue of missing out on recently published research. The trade-off between this and Dey et al.'s method is further elucidated in Sections 2.3.3 and 2.3.4.

Another thing to take note is that the screening process for the papers does not include a limiting range for the date of publication. The search process includes the related studies until the day the search was performed. This is because this paper is the first to review ARTS that is not restricted to any application field and it summarizes them by their training methods and evaluation types.

2.2 Training Categorizations

2.2.1 Training Methods

In a study by Martin et al., they developed an exhaustive list of possible methods that can describe and encapsulate the different types of training systems. This was accomplished by documenting the strategies, techniques, and procedures that are associated with the core process of the sample training systems. From their study, they have determined 13 core training methods that are able to represent any kind of training. These include: a case study, games-based training, internship, job rotation, job shadowing, lecture, mentoring and apprenticeship, programmed instruction, role-modeling, role play, simulation, stimulus-based training, and team-training [142].

However, these results show the core methods of training in a general sense. Since one of the goals of this paper is to define the key characteristics of training systems that is specific only to AR, the authors have narrowed down these 13 core methods. Based on the results of the method that is described in Section 2.3, each of the 64 studies included for synthesis utilizes one of these training methods: simulation, programmed instruction, games-based training, job shadowing, and mentoring.

Table 2.1 provides a summary of how Martin et al. defined these five training methods. With regards to job shadowing and mentoring, one could possibly classify these two together as both deal with the handing-down of expert performance and skill to the novice. However, it is still more advantageous to distinguish the two as job shadowing shines in the presentation of the desired result, whereas mentoring fosters the mentor–mentee relationship to stimulate the skills and knowledge transfer. Limbu et al. defined usage of an approach similar to that of Martin's, where "demonstration of the task" corresponds to job shadowing, and the "modelling the task with task analysis" corresponds to the mentoring style of training [136].

2.2.2 Training Evaluations

In the most general sense, an evaluation is characterized as the process of judging the worth or value of something. In the nomenclature of research design, an evaluation is achievable by using the concepts of measuring constructs. According to Nelson, there are two important dimensions when considering evaluation measurement methods [155].

The first dimension is reliability, which simply refers to the consistency of a measurement. Although reliability is one important aspect in research design, this is not discussed as it is safe to assume that studies that are published in high impact conferences/journals assure the reliability of their results.

The second dimension, validity, is defined as the "extent to which the scores from a measure represent the variable they are intended to," as stated by Chiang et al. [49]. Barsom pointed out that a full validation process is needed for a training system to be ready for implementation to the real-life environment [20]. This full process compromises the face, content, construct, concurrent, and predictive validity. Barsom et al. uses the terminologies of the validity types to judge whether the evaluations used for augmented reality applications are sufficient for training and education. All of these are defined by Barsom et al. and

Table 2.1: Definition of training methods relevant to AR
(directly taken from Martin et al. [142])

Method	**Definition**
Simulation	Involves the use of a simulator where specific skills are developed through repeated practice with a multisensory experience of imitated conditions.
Programmed instruction	Involves the delivery of training through instruction that is delivered by a program via some electronic device without the presence of an instructor.
Games-based training	Trainees compete in a series of decision-making tasks which allows them to explore a variety of strategic alternatives and experience the consequences which affect training the other players, but with without risk to the individuals or the organization.
Job shadowing	Involves a trainee closely observing someone perform a specific job in the natural job environment for the purpose of witnessing first-hand the details of the job.
Mentoring	Involves a one-on-one partnership between a novice employee with a senior employee. Mentorship aims to provide support and guidance to less experienced employees.

are summarized in Table 2.2. Finally, there is also a need to check whether these evaluations are in relation to the user, as the primary goal of a training system is to increase the knowledge, skills, and abilities rather than the promotion of a tool [34].

2.2.3 Training Strategies

According to Salas et al., training can be divided into four stages: information, demonstration, practice, and feedback. These stages are applied in terms of the content that is to be learned (i.e., training strategies) [184]. As different contents are mainly learned and trained in each strategy, the manners in which AR is utilized in each strategy is expected to differ. Along with the categorization, we investigate how AR works and what benefits it can bring to each stage.

The main training strategies and the utilization methods of AR adopted in each study are shown in Table 2.3. Each color indicates the training method employed in each paper. With a few exceptions, it can be seen that AR is utilized in different training strategies, but for the same training method.

Information

The first strategy is to convey information to the trainees (i.e., the concepts, facts, and information they need to learn). To obtain complex skill acquisition such as surgical skills, it is necessary to learn sufficient information that is relevant to each step in the task prior to practice. This is "learning" and it is generally performed by using textbooks and instructional videos. The effects of ARTS on learning are discussed in detail in Santos et al.[188]. They cite the multimedia learning principle [143] and extend it to learning with an AR annotation. They state that "people learn better from annotated virtual words onto physical objects than from separate multimedia (e.g., illustrated manual) and physical objects," in terms of time contiguity. In terms of spatial contiguity, "people learn better when corresponding virtual words and physical objects are presented near rather than far from, each other on the screen." As an example of this, Sankaran et al. promoted the acquisition of necessary knowledge to prevent sepsis by displaying educational content near the relevant location in omnidirectional images [185].

Table 2.2: Validity types for augmented reality applications (ARA)
(directly taken from Barsom et al. [20])

Validity	**Definition**
Face	The degree of resemblance between an ARA and the educational construct as assessed by medical experts (referents) and novices (trainees).
Content	The degree to which the ARA content adequately covers the dimensions of the medical content it aims to educate (or is associated with) (“the truth, whole truth, and nothing but the truth”).
Construct	Inherent difference in outcome between experts and novices on outcome parameters relevant to the educational construct.
Concurrent	Concordance of subject outcome parameters using tie ARA compared to outcome parameters on an established instrument or method, believed to measure the same educational construct (preferably the golden standard training method).
Predictive	The degree of concordance of ARA outcome parameters and subjects’ performance on the educational construct it aims to resemble in reality.

Table 2.3: Training Strategies applied for each study.

Strategy	**Description and Study**	
Information	Display of educational content in relevant locations	
	🟦 [68, 185, 97]	
	🟥 [3, 196]	🟩 [152]
Demonstration	Presentation of ideal behavior	
	🟧 [8, 227, 137, 50]	🟩 [47, 117]
	🟪 [108, 197, 83]	🟦 [119]
Practice	Expanding the flexibility of training content by adding information to physical objects	
	🟥 [90, 29, 151, 234, 126]	
	🟩 [93, 47, 217, 65]	🟦 [9]
	Facilitation of the skill transfer through imitation of the actual	
	🟥 [109, 54, 18, 103, 3]	
	Task support by presenting relevant information	
	🟥 [151, 189, 31, 52, 67, 1, 212, 235]	
	🟦 [97, 99]	🟩 [241]
Feedback	Feedback on the performance, problems, and ways to improve	
	🟥 [29]	🟦 [238]
	🟧 [182]	
	Real-time feedback for motion compensation	
	🟩 [134, 32, 65, 198, 117, 62]	🟥 [151]
	Real-time visualization of invisible current conditions	
	🟩 [147]	

Color representation: 🟥 Simulation, 🟦 Programmed instruction, 🟩 Games-based, 🟪 Job shadowing, and 🟧 Mentoring.

Demonstration

The second training strategy is to demonstrate the desired behavior, cognition, and attitudes to trainees. According to Schmidt's schema theory, learned movements are not stored by individual concrete motor programs, but by abstracted schemas [193]. The construction of this schema requires a body schema (postural schema), which is a cognitive standard for intuition by knowing the current posture of the body, the positional relationships of the body parts, and how much each body part needs to be moved to perform a certain action. A body schema is unconscious, subjective, and it has body-centered spatial coordinates (i.e., first-person perspective) [159]. Therefore, it is expected that motor learning is more efficient to start with the first-person view "demonstration." AR-applied job shadowing (e.g., [108, 197]) is considered to be an effective method for acquiring body schematics because it allows for first-person observation of the (ideal) movements of the expert.

Practice

The third strategy is to create opportunities to practice the knowledge, skills, and abilities that need to be learned. The nature of the application of AR to the practice is divided into three categories (which may be applicable to different examples).

Practice-1: Extend the flexibility of training content by adding information to the physical object

Even if a physical body model is elaborately made, the functions and degrees of freedom it possesses are limited. On the other hand, AR can provide versatile training conditions by superimposing the simulated organs model or the surface texture on the physical model. For instance, the medical trainees would need to change the trajectories of cutting the skin according to the shape of the overlaid organs in surgery.

Practice-2: Facilitation of skill transfer by imitating real objects and their superimposed display

When the training environment deviates from the real one, the acquired skill may be difficult to directly apply to the real environment. This is attributed to the enormous cognitive load that is involved in filling in the difference. Therefore, it is generally desirable for the training environment to resemble the real environment. The visual difference between the actual body and physical model can be compensated by imitating the texture of the real environment and superimposing it by using AR. This is expected to facilitate the application of skills that are acquired during training to the real environment.

Practice-3: Assistance in task execution through the presentation of task-related information

AR can reduce the cognitive load that is required for performing a task or it enables the user to perform the task more accurately by displaying cues that are related to the task execution at the relevant location on the real object. Abhari et al. assisted novice surgeons by superimposing a 3D trajectory of the pre-planned instruments on a display [1]. This is equivalent to the AR task support, but care should be taken when using this strategy in a training context. This information is not provided in the real task. If users rely on them, this can potentially lead to failure of the task execution in the real environment. For instance, Hulin et al. revealed that in programming-by-demonstration for an arm robot, the user's performance is impaired if visual effects are applied in the training phase [99]. Hulin et al. concluded that this is because users have come to rely too heavily on visual information during training. It should be used in combination with an appropriate skill transfer, such as training with less auxiliary information after becoming familiar with the task.

Feedback

The fourth strategy is to give feedback to the trainee on how they are doing with respect to learning; consequently, it allows for remediation. Here, we define feedback as information provided to the user that is generated adaptively according to the user's actions or their results.

Feedback-1: Feedback on the performance, problems, and ways to improve

The simplest feedback is for the system to evaluate the trainee's behavior and its results. The feedback also provides the user the evaluation, the problem, and how to improve it. Along the time axis, this can be classified into real-time feedback and summary feedback. The real-time feedback is presented to the user almost immediately when the system finds a problem with the user's current state or behavior. The summary feedback is a method of presenting a summary of the overall evaluation and areas for improvement at the end of a section of the experiment or during a short break. According to Chollet et al., the former has a motivational maintenance effect, whereas the latter has a substantial effect on skill improvement in the context of public speaking training (note that the system in this paper is not ARTS) [51]. Zhao et al. proposed a CNN-based method for automatically evaluating the user performance in neonatal endotracheal intubation training [238]. In addition, based on this evaluation, the system generates and presents summary feedback, which is color-coded to indicate areas that need more practice.

Feedback-2: Real-time feedback for motion compensation

One way to effectively use AR with real-time feedback is to present the difference between the optimal position/posture and the current ones of the user's body or the grasped object. Although this is effective in that the correction content can be intuitively understood, it must be designed so it does not lead to excessive user dependence. Sigrist et al. proposed a feedback method for oar pedaling training that combines sonification (the process of turning information into sounds [94]) of the difference between the current and optimal movements with visual information by using AR [198]. The feedback is designed to disappear when the user reaches the optimal state. They revealed that by iterating the training of constantly adjusting one's own state so that the feedback disappears, the skills improved even in the absence of feedback.

Feedback-3: Real-time visualization of invisible current conditions

An example of special real-time feedback is biofeedback, which aims to control one's own internal state (e.g., calming down). This can be interpreted as replacing a skill whose acquisition process is unclear with a different task that is easier to perform: self-regulation of physiological index values (e.g., brain waves and heartbeat) that are presented visually in real-time to meet certain criteria. During this training, it is important to associate the display information with the control target. For example, an attempt has been made to strengthen this connection by using AR to display electroencephalogram (EEG) information on the head of a mirror image of oneself [147]. The authors stated that this can also be used to train users to learn how to coordinate brain activity in some areas of the brain.

2.2.4 Training Principles

Transfer appropriate processing

Schmidt and Bjork provided a summary of a number of studies on verbal and motor learning in which the variables that maximized training performance were distinct from those that promoted transfer or assured long-term retention [192]. Generalization and maintenance of abilities are strengthened when techniques involve “transfer suitable processing,” or cognitions learners must engage in to apply their training in the transfer environment. Such tactics make performance in training more difficult and varied; nonetheless, learners get a deeper understanding of fundamental laws and concepts. It is commonly known, for instance, that “drilling” (continuous repetition of stimulus-response pairings) improves quick learning in training. However, extensive research indicates that while this style of training allows quick skill acquisition, it is less likely than other forms of training to transfer to post-training contexts. Other factors found to facilitate the type of deep learning that leads to transfer include contextual interference during practice (e.g., embedding performance cues within “noise”), variability in practice conditions, withholding knowledge of results until trainees have completed multiple trials (i.e., not providing continuous feedback), and gradual removal of knowledge of results. Though we are only providing an overview of a vast research domain, the concept is straightforward: as trainees begin to master a skill, training and practice conditions should become more challenging, trainer support should decrease, and practice conditions should increasingly resemble transfer conditions. Salas points out that “practice opportunities should require trainees to engage in the same cognitive processes they will need to engage in when they return to work. Often, that will mean designing sufficient challenge into the training [184].”

Behavioral role modeling

A number of sensorimotor and social skills have been taught via behavioral role modeling [214]. This training method is based on Bandura's theory of social learning [19]. In particular, trainees acquire new abilities through seeing others do them. Initially, trainees are given a list of behaviors (skills) to be acquired. These learning goals are conveyed most effectively as guidelines [214]. Second, behavioral models exhibit the desired behaviors, often using audio and/or video material. The demonstrative aspect is most successful when both good and negative examples are presented, as opposed to solely positive examples [17]. Third, participants engage in practice of the intended behaviors. When learners construct certain situations themselves during training, practice chances are most productive. Finally, trainees get comments on their progress and encouragement to implement newly acquired abilities. Instructing learners to establish their own transfer objectives aids behavior modification in this respect [214]. Salas suggests "demonstrating effective workplace behaviors based on demonstrated behavioral modeling practices [184]."

Self-regulation

Organizations may design training to improve learning by encouraging self-regulation among learners. In the context of training, self-regulation refers to individual's thinking patterns that enable them to maintain intense concentration on learning by self-monitoring performance, comparing advancement to an end goal, and adjusting learning approach as necessary. Although self-regulation is commonly viewed as a characteristic or behavior of learners, recent research indicates that encouraging self-regulatory activity during training programs can increase trainees' focus [201] and enhance their comprehension [22]. In two trials, one of which was conducted online, Sitzmann et al. discovered that encouraging trainees to self-regulate led to rapid and lasting increases in their semantic knowledge [200]. Salas elucidates the importance of the condition to "engage learners in self-regulatory processes during training and to encourage them to reflect and adjust. Simple questions such as 'Are you learning what you need to learn?' or 'Would you be ready to take an exam on this material?' may be sufficient to affect trainee learning [184]."

2.3 Review Method

2.3.1 Search for Prototypes

A systematic literature search was conducted using the search string [(augment* OR mix*) AND reality AND train* AND system]. This was performed on January 11, 2021. In this study, the main search engine used was from the IEEE Xplore digital library, complemented with additional papers from PubMed and SciFinder. The search resulted in 985 articles from IEEE Xplore and 121 articles from PubMed and SciFinder.

2.3.2 Inclusion Criteria

To identify the trends of the training systems in the field of AR, we defined the following criteria that must be met for an article to be considered as part of the database used in this analysis.

1. The paper was submitted in a top AR/VR conference, an IEEE Transaction, or a conference/journal with a h-5 index higher than 20.
2. The number of pages is more than four.
3. The full research paper is publicly accessible.
4. The work can be classified to AR/MR.
5. The work claims to be classified as a training system
6. The system is evaluated at least once.

The purpose of the first three criteria is to identify the papers that have a big impact and influence over the latest trends in ARTS research. The purpose of the bottom three criteria is to determine whether the article really did work on ARTS that is aligned with our definition as described in the previous section.

2.3.3 Study Selection

Figure 2.1 presents the flow diagram of the articles included in the analysis, which follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[158]. In the process of identification, we gathered 985 articles from IEEE Xplore and 121 articles from other digital libraries. Afterwards, we removed duplicates that were present across each database. In the screening process, we filtered the 1,102 articles to follow inclusion criteria 1–3 and we were left with a total of 281 articles. For criterion 1, we used the conference/journal's h-5 index as displayed in Google Scholar metrics, not the individual articles themselves. The reasoning why the score of 20 was chosen is based on the original paper by Hirsch et al., where this score characterizes a successful scientific activity [96]. The caveats of using this metric are also explained in that paper; however, this number can give a rough estimate to the quality of the research.

In the eligibility process, the articles were read carefully and the authors confirmed whether they satisfied inclusion criteria 4–6. For criterion 4, we refer to the loosened definition of AR/MR described in Section 1.1.1. This includes training systems such as some haptic interface plus monitor setup, like the works of [109, 189, 4, 90, 29]. What is excluded are works that describe purely VR systems; however, comparative studies between AR and VR such as the work of Qin et al. [176] are included.

As for criterion 5, we examined whether the goal of the developed system is in fact for the improvement of a person's knowledge, skills, or abilities, aligned with the earlier definition of training in Section 1.2.1. Examples of false cases that passed the initial screening are machine learning-related studies, as they also provide hits for the keyword "training," although not for people but for neural networks.

Finally for criterion 6, we checked whether the study contains evaluation conducted on the users. As we want to check for the evaluation as a training system, this metric should be in relation to the user, not the system performance such as tracking speed or rendering time of virtual objects. After removing the articles that did not meet all of the inclusion criteria, 64 articles remained for the analysis and synthesis.

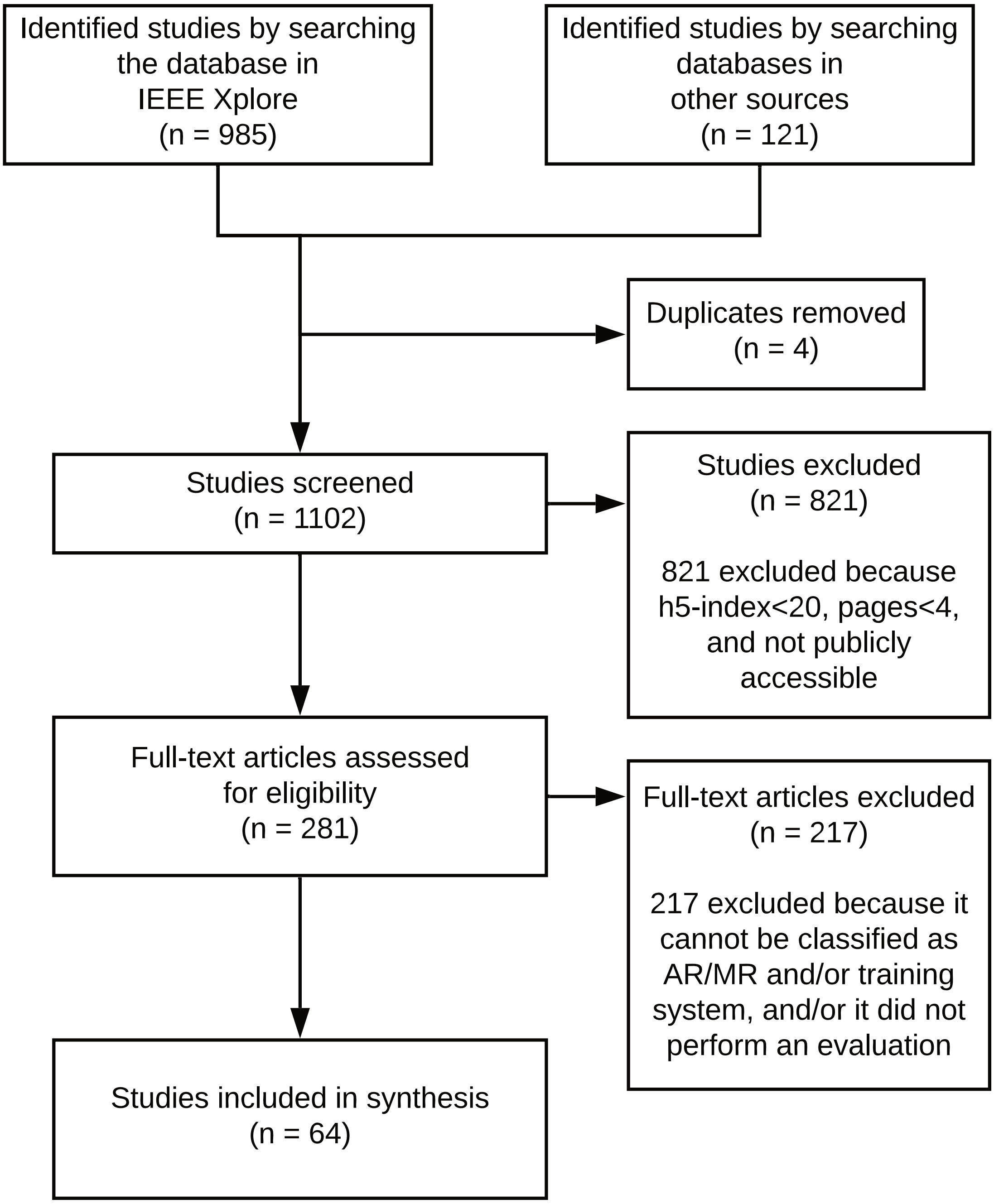


Figure 2.1: Flow diagram following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [158].

2.3.4 Limitations

One limitation is the selection of the database used. Our aim is to identify the current AR trends that encompass a wide variety of application fields while also having a huge impact on the AR community. The best candidate we identified that has a huge influence in the AR community is the IEEE Xplore database. To complement the IEEE Xplore results in order to have a wide variety of application samples, we used PubMed and SciFinder only.

Another important limitation of note is the use of the h-5 index. The use of this tool is important to determine the impact; however, this impact that we have measured is not the impact of the paper itself but from the conference/journal venue it was published. This opens the weakness of being vulnerable to missing out on important articles that were published in venues that were not as impactful. However, our decision to filter these by using the h-5 index rather than the actual paper citation itself is to overcome the problem of missing out on important papers that have recently been published; thus, having a low citation count. This trade-off has been considered while thinking about the inclusion criteria. There are also several impactful fields contributing to AR/VR that do not have journals with h-5 index greater than 20. This is currently an important limitation of the study. Hopefully, future authors of review/survey papers can propose better suggestions on how to quantify paper impact other than the two aforementioned methods.

Finally, we set the condition that the work should be considered as AR or MR; however, different authors have different terminologies and usage. In this paper, we discuss the definition of AR according to Azuma; however, we also includ works that are a little different from that definition but consider themselves as AR/MR systems. .

2.4 Results and Discussion

2.4.1 Overview

In this section, the results of the qualitative analysis are described. A list of all 64 papers that meet all the criteria in Section 2.3.2 is shown in Table 2.4. The table includes the reference number, application field, purpose, training method, eval-

Figure 2.2: Number of papers in each training method and application field.

uation type, AR device used, and user that was employed for the evaluation. As for the application field, the training for a medical skill (e.g., surgery, diagnosis) was accounted for in 35 papers. This was followed by rehabilitation/exercise (17 papers), industry (four papers), and other fields (eight papers). It is noteworthy that there is very little training in the industrial field in contrast to the extremely large amount of support systems that use AR.

2.4.2 Implications to Training Methods

Figure 2.2 shows the number of papers for each training method in each field. In descending incidence order: simulation (24 papers), games-based training (19 papers), programmed instruction (13 papers), job shadowing (four papers), and mentoring (four papers). The majority of the simulation studies (23 papers) were used in the medical field. Training dolls differ in many ways from the

Table 2.4: Summary of ARTS works by publication venue, field of application, purpose, training method, evaluation types, test users, and AR device used.
* A=Face, B=Content, C=Construct, D=Concurrent, E=Predictive

Ref.	**Venue**	**Field**	**Purpose**	**Training Method**	**Evaluation Type***					**User**	**Device**
					A	B	C	D	E		
[151]	JMIR	Medical	Digital Rectal Examination	simulation	✓	✓	-	-	-	actual	HMD
[54]	J. Healthc. Eng.	Medical	Orthopaedic Open Surgery	simulation	✓	-	-	-	-	mixed	HMD
[109]	IJCARS	Medical	Fetal Minimally Invasive Surgery	simulation	✓	✓	-	-	-	actual	Stationary
[18]	Heliyon	Medical	Cardiopulmonary Resuscitation Training	simulation	✓	-	-	-	-	actual	HMD
[176]	J. Healthc. Eng.	Medical	Peg Transfer Training	simulation	✓	✓	✓	-	-	actual	HMD
[103]	JMIR	Medical	Basic Life Support Training	simulation	✓	-	-	-	-	mixed	HMD
[189]	Anesth. Analg.	Medical	Anesthesiology Training	simulation	-	✓	-	-	✓	mixed	Stationary
[3]	Clin. Simul. Nurs.	Medical	Anatomy Nursing Skills	simulation	✓	✓	-	✓	-	actual	Handheld
[234]	Healthc. Technol. Lett.	Medical	Endoscopic Third Ventriculostomy Surgery	simulation	-	-	✓	✓	-	actual	HMD
[4]	Neurosurgery	Medical	Neurological Surgery	simulation	-	-	-	✓	-	alternative	Stationary
[90]	BMC Bioinform.	Medical	Artificial Cervical Disc Replacement Surgery	simulation	✓	-	-	-	-	actual	Stationary
[31]	Turk. Neurosurg.	Medical	Spinal Surgery	simulation	-	-	-	✓	-	actual	Handheld
[141]	MBEC	Medical	Ultrasound Guided Needle Insertion	simulation	✓	✓	✓	-	-	alternative	Stationary
[29]	World J. Surg.	Medical	Laparoscopic Surgery	simulation	✓	✓	✓	-	-	mixed	Stationary
[52]	BJUI	Medical	Urethrovesical Anastomosis Surgery	simulation	✓	-	-	✓	-	actual	Stationary
[67]	IJMRCAS	Medical	Soft Tissue Surgery	simulation	✓	-	-	-	-	actual	Stationary
[1]	IEEE TBME	Medical	Surgical Interventions Planning	simulation	-	-	✓	-	✓	mixed	HMD
[131]	IEEE TVCG	Medical	Micro-CT Analysis	simulation	✓	-	-	✓	-	alternative	HMD, Projected
[111]	IEEE TOH	Medical	Tumor Probing	simulation	-	-	-	✓	-	alternative	Stationary
[212]	IEEE TBME	Medical	Spinal Anesthesia Procedures	simulation	✓	✓	✓	-	-	mixed	Stationary
[235]	IEEE TBME	Medical	Needle Placement for Facet Joint Injections	simulation	-	-	-	-	✓	actual	Stationary
[196]	IEEE THMS	Medical	Anatomy Training	simulation	✓	✓	-	✓	-	actual	HMD
[79]	CCECE	Medical	Neurological Surgery	simulation	-	-	-	✓	-	alternative	HMD, Stationary
[68]	BMC Med. Educ.	Medical	Health Science Education	programmed	✓	-	-	-	-	actual	HMD
[194]	BMC Med. Educ.	Medical	Bladder Catheter Placement	programmed	✓	-	-	✓	✓	actual	HMD
[25]	Biomed. Eng. Online	Medical	Electrocardiogram device operation	programmed	-	-	-	✓	-	actual	HMD
[238]	ISMAR	Medical	Neonatal Endotracheal Intubation	programmed	-	-	-	✓	-	actual	HMD
[185]	IEEE VR	Medical	Sepsis Prevention Education	programmed	✓	✓	-	-	-	actual	HMD
[108]	World J. Urol.	Medical	Surgical Training	job shadowing	✓	✓	-	-	-	actual	Stationary
[177]	ISMAR	Medical	Anesthesia Education	job shadowing	✓	✓	-	-	✓	actual	HMD
[197]	ISMAR	Medical	Surgical Interventions Planning	job shadowing	-	-	-	✓	-	actual	HMD
[8]	Surgery	Medical	Surgical Instruction	mentoring	-	-	-	✓	-	actual	Handheld
[182]	NPJ Digit. Med.	Medical	Surgical Instruction	mentoring	-	-	✓	✓	-	actual	HMD
[227]	Sensors	Medical	Point of Care Ultrasound Training	mentoring	✓	-	-	✓	-	actual	HMD
[137]	IEEE VR	Medical	Number Matching Task/Austere Surgery	mentoring	✓	✓	-	✓	-	actual	HMD
[93]	JMU	Rehab	Gait and Balance Rehabilitation	games-based	✓	-	-	✓	-	actual	HMD
[47]	IJERPH	Rehab	Exergames	games-based	✓	-	-	-	-	actual	Stationary
[217]	Behav. Res. Methods	Rehab	Mirror and Imagery Training	games-based	-	-	-	-	✓	alternative	HMD
[134]	JNER	Rehab	Upper Limb Stroke Rehabilitation	games-based	-	-	-	-	✓	actual	Stationary
[32]	PLoS One	Rehab	Gait Rehabilitation	games-based	✓	-	-	-	-	alternative	HMD
[48]	JESF	Rehab	Tai-Chi Training	games-based	-	-	-	-	✓	actual	Stationary
[101]	ARM	Rehab	Balance and Mobility Rehabilitation	games-based	-	-	-	-	✓	actual	Stationary
[65]	JRRD	Rehab	Parkinson disease Gait Rehabilitation	games-based	-	-	-	-	✓	actual	HMD
[198]	Exp. Brain Res.	Rehab	Trunk-arm Rowing	games-based	✓	-	-	-	✓	alternative	Projected
[163]	Cogn. Behav. Neurol.	Rehab	Mild Cognitive Impairment Rehabilitation	games-based	-	-	-	-	✓	actual	HMD
[117]	IEEE VR	Rehab	Tai Chi Chuan Learning	games-based	✓	✓	-	-	-	actual	HMD
[21]	IEEE VR	Rehab	Eye-Hand Coordination Training	games-based	✓	✓	-	✓	✓	actual	HMD, Stationary
[126]	IEEE Access	Rehab	ADHD Children Treatment	games-based	✓	✓	-	-	✓	actual	HMD
[115]	IEEE TVCG	Rehab	Exercise for Reducing Obesity	games-based	✓	✓	-	-	✓	actual	Stationary
[62]	IEEE TNSRE	Rehab	Hemiparesis Stroke Rehabilitation	games-based	-	-	-	-	✓	actual	Stationary
[241]	IEEE/RSJ IROS	Rehab	Wheelchair Assistance	games-based	✓	-	-	-	✓	alternative	HMD
[50]	Biomed Res. Int.	Rehab	Chopsticks Telerehabilitation	mentoring	-	✓	-	-	✓	alternative	HMD
[220]	ISMAR	Industrial	Scenario-Based Training Authoring	programmed	✓	✓	-	✓	-	alternative	HMD, Stationary
[97]	IEEE Access	Industrial	Assembly Instruction	programmed	✓	-	-	✓	-	actual	Stationary
[99]	IEEE RO-MAN	Industrial	Transfer Task for Exploration Training	programmed	-	-	-	-	✓	mixed	Stationary
[83]	ISMAR	Industrial	Origami and Building Blocks	job shadowing	✓	✓	-	✓	-	alternative	Stationary
[206]	ACM/IEEE HRI	Other	Autonomous Driving	simulation	-	-	-	✓	✓	actual	HMD
[190]	EAIT	Other	Business Ethics	programmed	-	-	-	-	✓	actual	Handheld
[119]	Front. Psychol.	Other	Musical Instruction	programmed	-	-	-	-	✓	actual	Projected
[9]	IEEE ToE	Other	Remote Laboratory Education	programmed	✓	✓	✓	-	-	actual	Stationary
[147]	IEEE VR	Other	Neurofeedback Training	games-based	✓	-	-	✓	-	actual	Stationary
[239]	IEEE Access	Other	Hands-on Experiential Learning	games-based	✓	✓	-	✓	-	actual	HMD, Stationary
[152]	IEEE Access	Other	Spatial Memory Learning	games-based	✓	✓	-	✓	-	actual	Handheld
[133]	ISMAR	Other	General Work Activities	job shadowing	✓	✓	-	✓	-	actual	HMD

actual human body, which makes the learning transfer difficult. AR enables the appearance of a training doll to be similar to that of the actual human body by superimposing the patient's body textures and organ models on it. This is most likely the main reason why simulations are widely used in this field. On the other hand, the other fields, where imitation of the real environment is not comparatively necessary, employ different strategies.

Games-based training (19 papers) was the second most adopted, and 74% (14 papers) of these papers were in the field of rehabilitation/exercise. The majority of users in this field are people who have lost or are losing certain functions of their bodies owing to disease or injury. Rehabilitation often requires long-term continuity, although existing rehabilitation programs are often monotonous repetitions of tasks [2, 125]. One of the biggest problems is the difficulty in maintaining patient motivation [71]. Games-based training is an effective means of incorporating a mechanism for continuity into these monotonous structures through gamification. It is easy to understand that users in the rehabilitation field would find it helpful if the training system adopted the games-based method for the tedious tasks they typically have to endure.

Programmed instruction (13 papers) has been equally adopted in other fields, except rehabilitation. This may be because the normal rehabilitation process does not include general knowledge acquisition through classroom lectures.

2.4.3 Implications to Evaluation

As described in Section 2.2.2, we classified the evaluation into five types, while referring to Barsom et al. [20]. Figure 2.3 shows the number of papers of each evaluation type in each field. Note that some papers used more than one type of evaluation. The number of papers that adopted face evaluation is the largest (39 papers), followed by concurrent (27 papers), content (24 papers), and predictive (21 papers), whereas construct (eight papers) is the smallest by far. Face evaluation is performed to assess the degree of resemblance between training with the system and the educational construct with a questionnaire or a small interview. Owing to the implementation ease, the number of adoptions is large. The concurrent evaluation is a comparison with the existing training methods. The wide use of this evaluation is also understandable because it is already common practice for

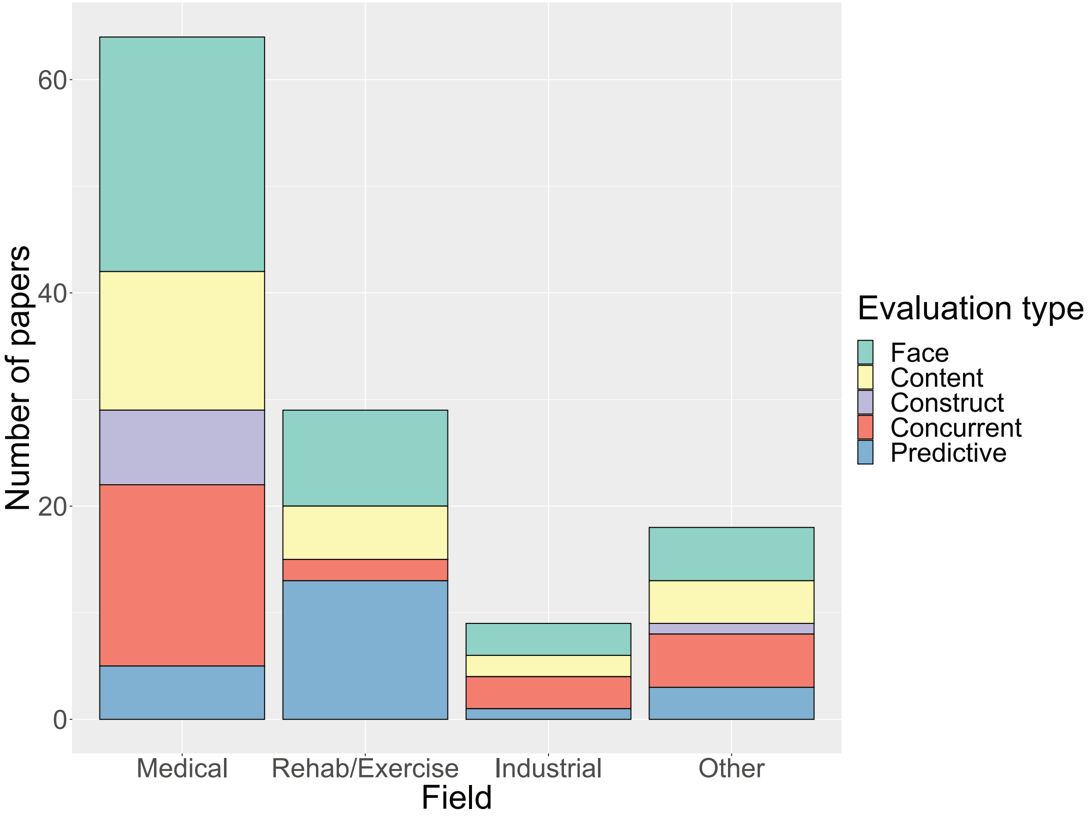


Figure 2.3: Number of papers in each evaluation type and application field.

studies to do comparisons of their proposed methods against that of the golden standards. Note that 21 studies conducted the predictive evaluation, which is a significant amount. Predictive evaluation is the most important type because it examines direct training effects (i.e., actual skill acquisition). However, it is also expensive because it requires training with the system and the evaluation to be carried out in separate steps. Thus, we initially presumed that the papers that included the predictive evaluation would be quite limited, although the results are contrasting. The construct evaluation examines the differences in the outcomes between two types of subjects with different skill levels when using the system, i.e., ascertain if the system is reflected with some skill that is possessed by the expert. The lack of its adoption can be interpreted as self-evident, or it is due to the high cost of the evaluation (two different groups of subjects—experts and novices—are required).

Looking at the differences among the fields, the adoption rate of the predictive evaluation is relatively high in rehabilitation. As mentioned earlier, games-based training is mainly used in the rehabilitation field. This method tends to transform the skills that need to be acquired by incorporating game elements. Therefore, it is understandable that many papers confirmed the impact of the gamified training content relative to the acquisition of the intended skills to be learned.

In addition, we analyzed the number of different types of assessments made in each paper. Although 43% (28/64 papers) had only one type of evaluation, more than half of the papers had a combination of two (18 papers), three (17 papers), and four (one paper) types of evaluations. There were no papers with five types of evaluations. Barsom et al. stated that the new training system can be considered for social implementation only when all these five types of validity are guaranteed [20].

2.4.4 Implications to the Users in the Evaluation

In addition, we analyzed the users that were used for the evaluation in each field. The users were divided into “actual” users (i.e., people who actually need to be trained, e.g., students in medical school in the case of training for surgical skills) and “alternative” users (i.e., people who do not need to train the target skill). In addition, papers that use both users are classified as “mixed”. Figure 2.4 shows the number of papers in each of the three categories. In some fields, it is very expensive to collect actual users of ARTS as test subjects, so it was assumed that many alternative users were reluctantly used. However, contrary to this, more than two-thirds of the papers conducted evaluations with actual users.

2.4.5 Implications to the Types of AR Device Used

We also looked at the type of AR device used for each study. For the classification of these devices, we adopted the categorization of AR displays according to distance from eye to display from Schmalstieg et al.’s book, “Augmented Reality: Principles and Practice” (i.e., HMD, Handheld, Stationary, Projected) [191]. However, this does not consider the quality of the AR device; for example, HoloLens and Google Cardboard (in which a smartphone is used as a VST device) are both considered HMD.

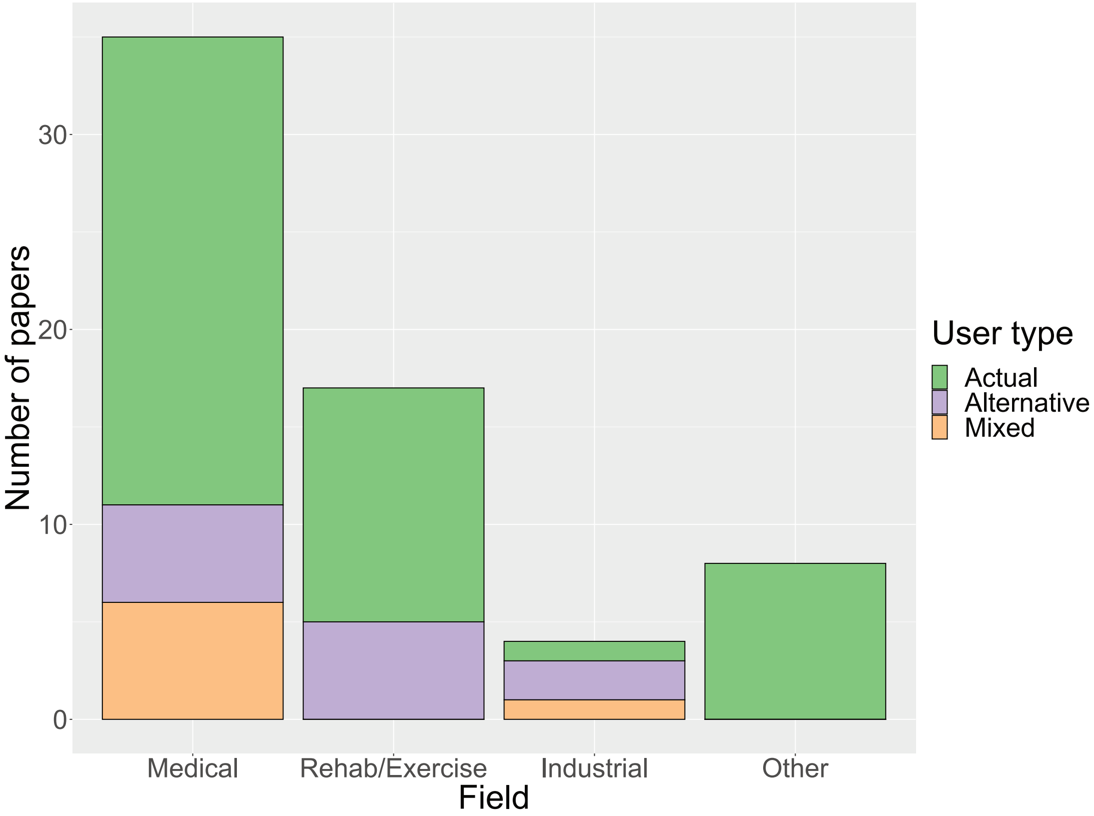


Figure 2.4: Number of papers for each user type for the evaluation.

In descending order, the distribution for the AR device used are as follows: HMD (34 studies), Stationary (27 studies), Handheld (five studies), and Projected (three studies). HMD and Stationary devices are the popular choices for implementing ARTS. When looking at the purpose for which these ARTS are used, they usually involve some type of activity that require the user's hands/body. Taking this point into consideration, it is logical that handheld displays are not optimal for training systems in general. Projected displays such as CAVE or other spatial AR displays are also rarely used as there are no additional training benefits they can offer compared with the World space alternative (i.e., Stationary displays) which are easier to implement. For the distribution of the type of AR device with regards to the training method used or the application field, no trends can be drawn as it is distributed evenly.

2.5 Narrative Synthesis for the Training Methods

In Section 2.4, we summarized the results in terms of the training method, evaluation type, users, and device used in the evaluation. However, from these results, we determined that the current implementation of ARTS has heterogeneity in regard to the study design; thus, statistical pooling cannot be performed. Because a meta-analysis is still difficult to accomplish in the current stage of ARTS, the most logical approach is to conduct a narrative synthesis. This section describes each of the training methods that are effective for ARTS by presenting the representative studies that constitute the essence of each method. We also look at how each method is evaluated by narrating some exemplars for each evaluation type.

2.5.1 Simulation

Simulation training simply consists of training under a simulated environment and the goal is to develop specific skills. The most obvious advantage of simulation training is that it can provide a risk-free environment that can be considered to be very risky if it is performed in a real-life environment. Another benefit with simulation training is that it offers the trainee the opportunity to do rehearsals and practice the process repeatedly [142]. The use of AR/VR simulation to supplement traditional teaching in surgery skills for example, is a better alternative ethically compared to cadaver dissections [60, 180]. In terms of the realism and didactic value, Botden's study [29] ascertained that the ProMIS AR was the better simulator for practicing laparoscopic skills in comparison with the LapSim VR simulator.

Implementation - Simulation

When considering how training in simulators is implemented and how the skills in this process are accumulated, it is important to discuss the concepts of skills generalization and skills transfer. Gallagher states that "skills generalization refers to the training situation where the trainee learns fundamental skills that are crucial to completion of the actual operative task or procedure. Skills transfer refers to a training modality that directly emulates the task to be performed in vivo or in the testing condition" [74]. Each of these has its own advantages.

For example, skills generalization can greatly boost the rate of learning while not having to invest substantially in the cost of the technology (in this case, the realism of the simulator). Meanwhile, skills transfer provides a more realistic approach and it generates a venue for rehearsal and practice.

Ingrassia et al. [103] presented a nice example of skills transfer training. They developed a system that is able to realistically reproduce the scenario of defibrillation training. The goal was to provide inexperienced trainees with an environment for self-instruction training to perform the CPR procedure, as shown in Figure 2.5. In this environment the trainees are able to do natural gestures and body movements, as if they were in the actual environment.

Another example of skills transfer training is provided by Muangpoon et al. [151]. The goal of the system was to provide clinicians and medical students an environment to learn, teach, and practice digital rectal examination (DRE). In this system, they are able to visualize a virtual hand that is overlaid on their real hand and the internal organs are overlaid on a benchtop model. Because the technique and manner of operation is vital for DRE, they have tracked the movement of the hand together with the amount of applied pressure.

For the skills generalization training, Qin et al. [176] offer a good example. They utilized a multiple platform simulator system (AR, VR, etc.) for the peg transfer task, as illustrated in Figure 2.6. Instead of recreating the whole surgical process itself, this type of training focuses more on reproducing the specific surgical maneuvers to gain the target skill.

Aside from looking at the perspective of the task itself, there are also training systems that deal with the planning of the operation. Abhari et al. [1] proposed a system that is able to facilitate training for the planning of a neurosurgical procedure. They overlaid patient-specific data onto a mannequin head to aid the planning of the operation. Their results show that the performance index of the non-clinicians significantly improved when they used their system. Furthermore, the performance time of the clinicians was significantly faster in comparison to using conventional planning environments. The reason for these improvements is that the participants were able to develop spatial reasoning ability, which cannot be gained in traditional methods.

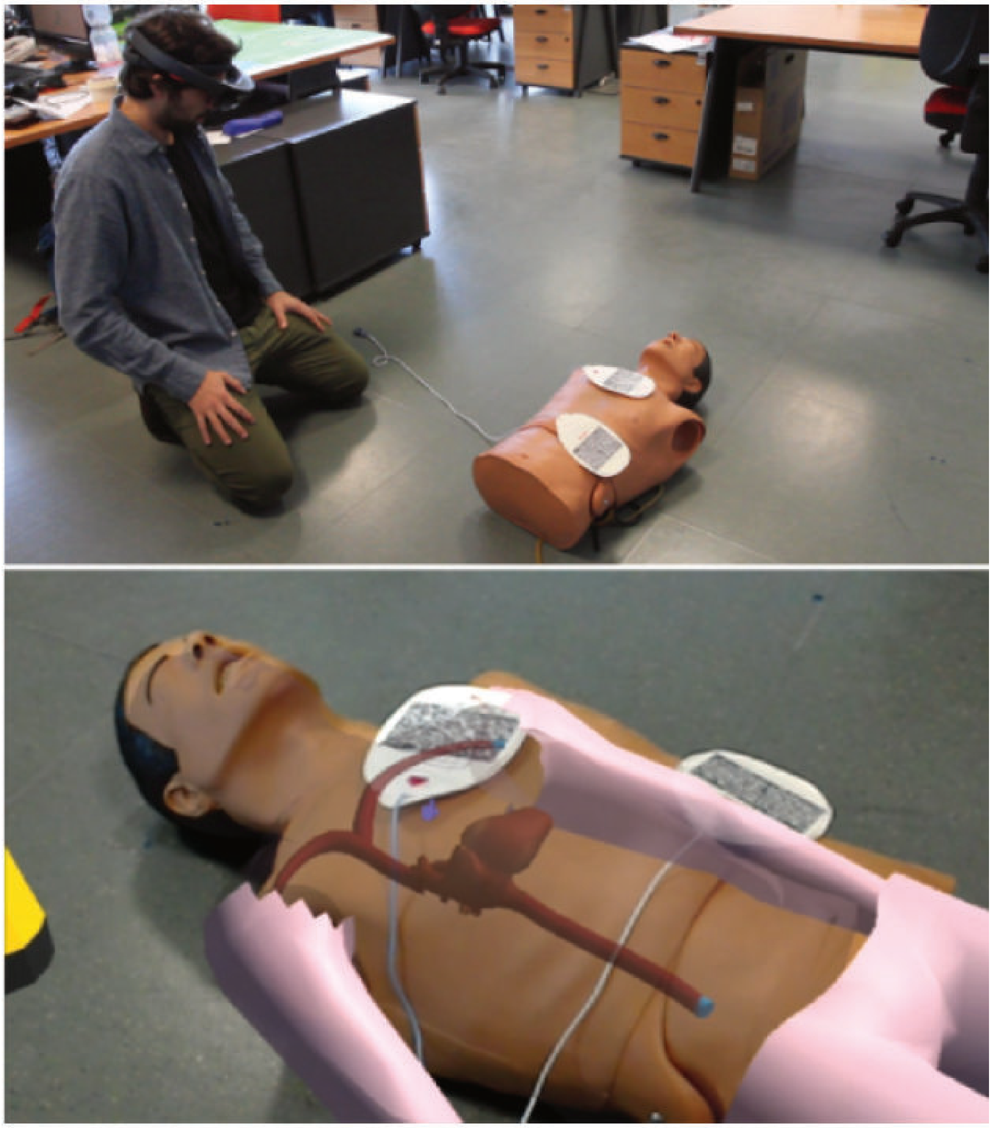


Figure 2.5: Environment for defibrillation training [103].

Evaluation - Simulation

- Face and content evaluation example:

Javaux et al. [109] evaluated the face and content by requesting surgeons that performed fetal minimally invasive surgery to complete certain objectives using a developed simulator. The skills targeted included basic fetoscopic and procedural skills for laser surgery. A five-point Likert scale was used to assess

Figure 2.6: Peg transfer simulator on box, AR, and VR platforms [176].

the face evaluation, which contained aspects such as the realism of the trainer, the body wall phantom, and the rendering of the environment (e.g., image quality, light propagation, and depth perception). Similarly, the content evaluation was accomplished by measuring the training capacities (e.g., scope handling and lasering) and the usefulness of each task.

- Construct evaluation example:

Qin et al. [176] conducted a Null-Hypothesis Significance Testing (NHST) between novice thoracic surgeons with four years of post-graduate experience and experts with 12 years of experience. They also considered the experience of the participant in regard to the box trainers, VR games, and HMDs. Meanwhile, Botden et al. [29] acquired 30 novices, 30 intermediates, and 30 experts of laparoscopy and the demographics included interns, surgical residents, and surgeons. The laparoscopic experience of each group was decided based on experience, such as the number of times they performed suturing in a clinical setting, or the number of times they participated in the surgical procedure, either as a spectator, a camera handler, or an assistant.

- Concurrent evaluation example:

An example of a concurrent evaluation is presented by Aebersold et al. [3]. The main goal of their study was to investigate whether the students had a better understanding and learning of nursing skills such as placing nasogastric tubes. Students were randomly appointed to either the control group (usual training method) or the AR group, which received the training module of anatomy simulation by using an iPad. There were 34 and 35 participants in each group, respectively. The nursing skills were assessed through a 17-item checklist scoring system and an NHST was carried out between these two groups. In another study, Chowriappa et al. [52] explored the effects of using an AR-based training module for robot-assisted urethrovesical anastomosis. In the user study, 52 participants were randomized to be either in the hands-on surgical training (HoST) group or the control group. They applied five global evaluative assessment of robotic skills (GEARS) to establish a comparison between the two groups and they determined the results that were based on the computed p-value.

- Predictive evaluation example:

In an investigation by Sappenfield et al. [189], a MR simulator was used to train anesthesiology residents to achieve the goal of improving supraclavicular access to the subclavian vein. When considering the evaluation metrics, they used an automated scoring system to objectively score the time, success, and errors/complications of the participant's performance. When considering the design of their study, they randomized the residents into two groups. The first group was exposed to real-time 3D visualizations in the first trial but not in the second trial. The second group was not exposed to 3D visualizations in the first and second trials, but they were later asked to experience the playback of the second trial's 3D visualizations before carrying out the third trial without the visualization guide. The comparison of scores between the subsequent trials and the differences between the groups that experienced real-time visualization, no visualization, and delayed visualization point toward the evidence of this study's predictive evaluation.

Discussion - Simulation

In this training method, it can be argued that the simulator's quality of realism affects the enhancement of the learning transfer. For this reason, 15 out of the 24 total simulation training papers that evaluated their system used the face evaluation type. Studies also usually perform content with the face because these evaluations used the same kinds of questionnaires such as five-point Likert scales. In the skills transfer training example, Muangpoon et al. [151] assessed the DRE clinicians and medical students with face and content evaluation. However, no evaluation metrics were performed for the other evaluation types. For the skills transfer training, the most important part to consider during the evaluation is the realism and quality of the simulation; thus, studies tend to evaluate the face and content.

In the investigation by Qin et al. [176], novices and experts were recruited for the peg transfer skill, and they evaluated the face, content, and construct. The lack of scenario information for skills generalization (i.e., patient/mannequin operation) puts more emphasis on the manner of doing the task. Because the difference in the expert's technique and novice's technique is emphasized, the value of performing a construct evaluation increases. Furthermore, Abhari et al. [1] performed the construct evaluation, which makes sense since the construct is arguably the most important when considering the planning phase for surgical procedures. In the simulation training method, there are 10 studies that performed a concurrent evaluation whereas only four had a predictive evaluation. It is notable that a lot of studies have adapted the practice of comparing their results to the golden standard (concurrent evaluation). However, we suggest that more studies should also investigate the effects of their simulators not only in the AR environment, but also how well these acquired skills from the simulators can translate toward real implementation (predictive evaluation).

2.5.2 Programmed Instruction

The use of programmed instruction with AR provides a lot of benefits in comparison to using more traditional approaches such as manuals and instruction books. For instance, programmed instruction can offer more flexibility when instructions and training procedures are updated to more recent techniques. Similar to the simulation training method, programmed instruction also enables the user to repeatedly practice and rehearse the specific skill [199, 150]. The unique point of programmed instruction over the other training methods is its potential to take advantage of multisensory features such as sound, text, and animations [142].

Implementation - Programmed Instruction

There are many ways of presenting instruction, but when deliberating about learning theory and instructional design, one of the more crucial aspects that need to be considered is the instruction sequencing [55]. Many studies have suggested that the sequence and arrangement of these learning activities impact how the information and knowledge is being handled [80, 138, 219]. According to theories of instructional design, one approach is to follow a simple-to-complex sequence [55]. There are also other design approaches such as simply conforming to the traditional methods (e.g., how kids learn the alphabet from A to Z). We refer to this delivery as the sequential instruction approach. The presentation of sequential instruction can also come in two forms: pre-recorded or real-time.

An example of a pre-recorded sequential instruction is presented in Sankaran et al. [185]. They used a 360-degree video recording session to enable the viewpoint of what it is like to experience the real clinical practices for sepsis prevention. To accelerate the training of medical students, the students were exposed to the recorded simulated environment along with the augmented information such as the patient graphical data or the pop-up information, as shown in Figure 2.7. At certain checkpoints in the program, the students were asked to answer a pop-quiz to assess their learning progress.

Furthermore, an example of a real-time sequential instruction is provided by Zhao et al. [238]. They developed an effective ARTS that is able to give a comprehensive understanding of the endotracheal intubation procedure by providing real-time instruction and evaluation. This was accomplished by overlaying 3D

Figure 2.7: Pop-up information about sepsis condition [185].

see-through visualizations of data such as the depth, penetration, and time to the manikin. In addition, the visualization of the instructions for each phase of the procedure was displayed to the user that was wearing the HMD.

By contrast, the other approach of the instruction delivery is the non-sequential instruction. In line with the component display theory of instructional design that was presented by Merrill, learner control is considered to be a very important component [148]. Learner control is “the idea that learners can select their own instructional strategies in terms of content and presentation components” [55]. In regard to this approach, the most essential point is the fact that students are able to have more control so they can learn at their own pace.

The study by Marquez et al. [9] presents the concept of non-sequential instruction. This work proposed the idea of the implementation of an augmented remote laboratory (ARL). Inside the ARL, students were free to experience and explore the environment and do laboratory activities that are analogous to traditional laboratory classes.

Evaluation - Programmed Instruction

- Face and content evaluation example:

Sankaran et al. [185] carried out face and content evaluation by using the system usability scale. Sankaran et al. asked 28 novice students for a score that had a scale ranging from one (strongly disagree) to five (strongly agree). The face evaluation questions contained statements such as “I needed to learn a lot of things before I could get going with this system.” Content evaluation statements included, “I thought the system was easy to use,” or, “I found the various functions in the system were well integrated.”

- Construct evaluation example:

In addition, Marquez et al. [9] performed construct evaluation using the results of the questionnaire and compared the scores between the evaluations from the teachers and the evaluations from the students.

- Concurrent evaluation example:

Schöb et al. [194] developed an MR system teaching tool that can provide students with instructions while learning a new practical task. Schöb et al. recruited 164 medical students to experience the bladder catheter placement. In this study, 107 students were assigned to the control group, and received instructions only from the instructor. The 57 remaining students received instructions from the MR guidance system while using HoloLens. Both groups were asked to take a standardized, non-timed objective structured clinical examination and were assessed by their learning outcomes. The control and study group were compared by performing an NHST.

- Predictive evaluation example:

Sari et al. [190] used AR techniques to teach ethics and moral imagination. They recruited 142 students that were taking a business ethics course. The study utilized the 3x2 experiment method, which consists of three training modes and two time periods. The three training modes comprised AR-based, paper-based, and no training. As evidence of the predictive evaluation, the authors performed an analysis of variance (ANOVA) to calculate the F-value and p-value for between-subjects and within-subjects. They also conducted a post hoc analysis test of the training methods to determine if the groups differed in terms of the moral imagination.

Discussion - Programmed Instruction

This section introduces two programmed instruction approaches inspired from the concepts of instructional design, sequential instruction, and non-sequential instruction. As most studies used the sequential approach, it is difficult to determine the differences between their evaluations. However, the low adoption number does not discredit the effectiveness of the non-sequential design of instruction. The non-sequential design shines when thinking about individually matching the pace of each student's personal growth.

As discussed in Section 2.4, when looking at the relationship between the programmed instruction and application field, it can be observed that programmed instruction is not utilized in the rehabilitation and exercise area. This is because programmed instruction is focused more on the development of knowledge rather than the development of skills. For the purposes of rehabilitation and exercise, the training needed require systems that allow for the forming of an ability and capability to do something, whether physical or psychological. On the other hand, programmed instruction is very effective when it is used for the endowment of knowledge. For this reason, most of the studies categorized to programmed instruction (except one) have their target user as the students that will actually use that specific knowledge (e.g., medical students [185, 9, 194], and business ethics students [190]).

2.5.3 Games-based Training

The culture of games has certainly risen in influence, especially for today's younger generation. It is no surprise that the use of game-like mechanics and gamification techniques can motivate users to engage in activities that might otherwise be less interesting [183, 30, 6]. The clear advantage of games-based training over other training methods is in its competitive nature, because it instills motivation that leads toward learning. The downside of this method though is that too many elements and components are integrated into the game. This may be difficult to clearly determine and guarantee which parts contributed to learning [142]. When comparing other games-based methods, AR games-based training stands out because it provides a nice balance between immersion in the game and being rooted in the real world. The latter is especially important when considering if the users have physical or psychological disorders.

Implementation - Games-based Training

From the database of papers we collected, games-based training is usually targeted for rehabilitation and exercise as demonstrated in an overwhelming 16 out of 19 papers. In terms of providing solutions for therapeutical treatments, games-based training systems can either address the physical or psychological problems of its target users.

One case of using games to treat physical problems is the work that is presented by Johnsen et al. [115]. In this study, Johnsen et al. created a MR system that allows obese kids to interact with a virtual pet. To trigger interactions with these pets, the kids need to input physical activities. Growth of the pet is proportional to the physical activity progress of the kid; thus, the virtual pet becomes a strong motivator for promoting health. Another example is the work by Trojan et al. [217], who proposed an AR hand training system that uses the mirror image approach. They used techniques such as finger flexing, hand posture fitting, and a “snake” video game to train the motor skills by hand-mirroring, as illustrated in Figure 2.8.

Other studies have also dealt with psychological disabilities. For example, Park et al. [163] examined the effects of their MR training system on patients that have mild cognitive impairments such as Alzheimer’ s disease. They designed games that recreated day-to-day activities to target cognitive functions such as selective attention, the visual/verbal working memory, and problem solving. One more case of this type is a study conducted by Kim et al. [126]. They developed an MR eye-contact game that is able to treat children with attention deficit hyperactivity disorder (ADHD). The game utilizes face recognition techniques to confirm the child’s attention, trigger the treatment, and develop interpersonal skills.

Evaluation - Games-based Training

- Face and content evaluation example:

Batmaz et al. [21] investigated the differences between the effectiveness of AR, VR, and the conventional 2D touchscreens to train professional athletes by using the eye-hand coordination reaction test. This study accomplished the face

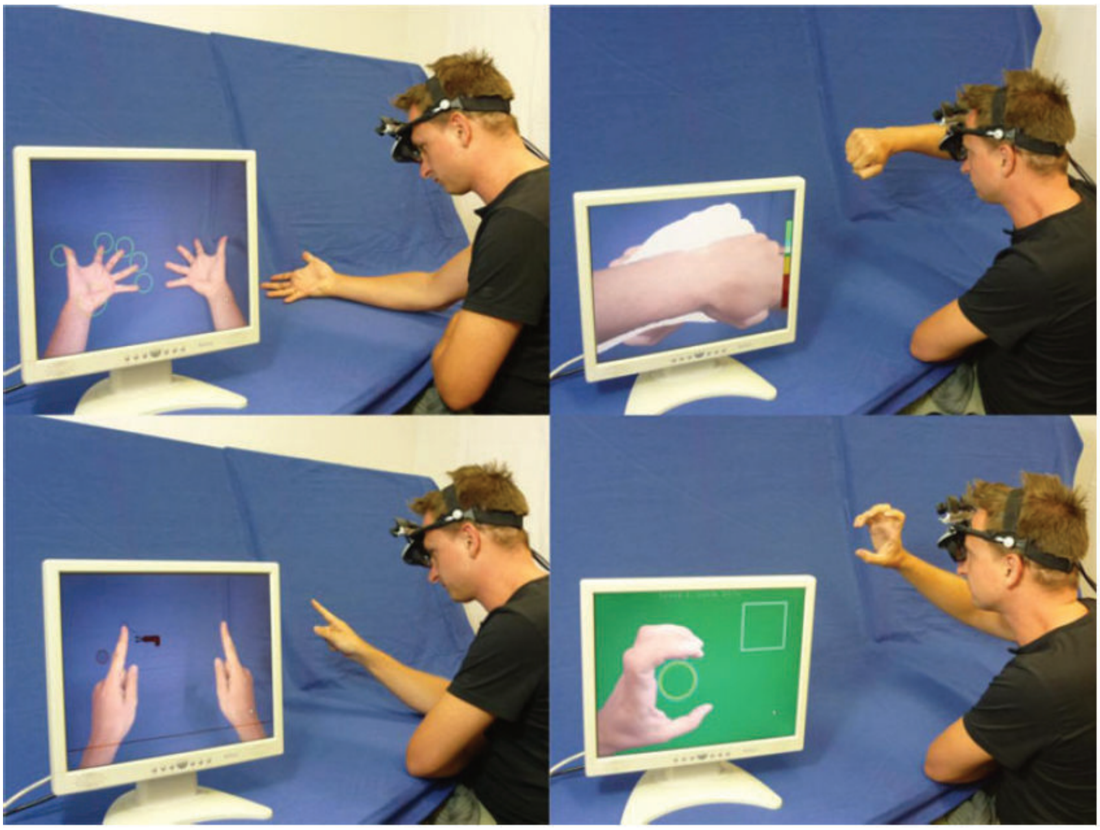


Figure 2.8: Example of activities for hand-mirroring ARTS (monitor view for demonstration purposes only. In the actual implementation, the users experienced these activities through an HMD) [217].

and content evaluation by collecting subjective opinions of 15 participants from their local university by using a seven-point Likert scale. Statements such as “increased sense of reality” and “having better perception of depth” indicate face and content, respectively.

- Concurrent and predictive evaluation example:

Moreover, Batmaz et al. [21] analyzed the effects of the time, error rate, and throughput while considering the experimental condition (AR, VR, or 2D screen). Aside from this, they also analyzed the effects for the haptic feedback and environments. They compared the effects by computing the F-value and p-value for each of the conditions, which suggests concurrent evaluation. Furthermore, they analyzed the performance improvement of the participants across multiple repetitions by using AR and VR in comparison to a 2D screen, which is indicative of a predictive evaluation.

Discussion - Games-based Training

Most cases for games-based training are targeted at rehabilitation and exercise. This suggests that games are a great avenue for developing or recovering physical and psychological skills. The reasoning for this is that tasks that are directed toward these skills are usually tedious, monotonous, and repetitive. Games are excellent at solving this problem because it can stimulate motivation and maintain user involvement.

One thing to note is that for games-based training, no studies have conducted a construct evaluation. This is self-evident as there is no user that has a high skill or a low skill for a physical or psychological disability. However, what matters is the confirmation if there is the transfer of training effects to the real context in terms of whether they acquired or recovered the targeted skill. In this situation, the predictive evaluation is of great importance, which is implemented in many studies (12 out of 19 papers).

2.5.4 Job Shadowing

In companies and enterprises, job shadowing refers to on-the-job training of a trainee that involves observing the model employee that is performing their usual work. When applying AR in training, job shadowing is the instance when a novice diligently observes the performance of an expert by using techniques such as sharing first-person views. The main advantage of job shadowing over the previous training methods is the ability of the novice to experience the perspective of the expert, which gives a broader outlook to the development of skills and techniques of professionals.

Implementation - Job Shadowing

Job shadowing in AR in practice can be done directly or indirectly. Direct job shadowing refers to the show-by-example approach where the expert representation (e.g., virtual hand) demonstrates the necessary techniques to the novice. On the other hand, indirect job shadowing refers to the use of tools such as an-

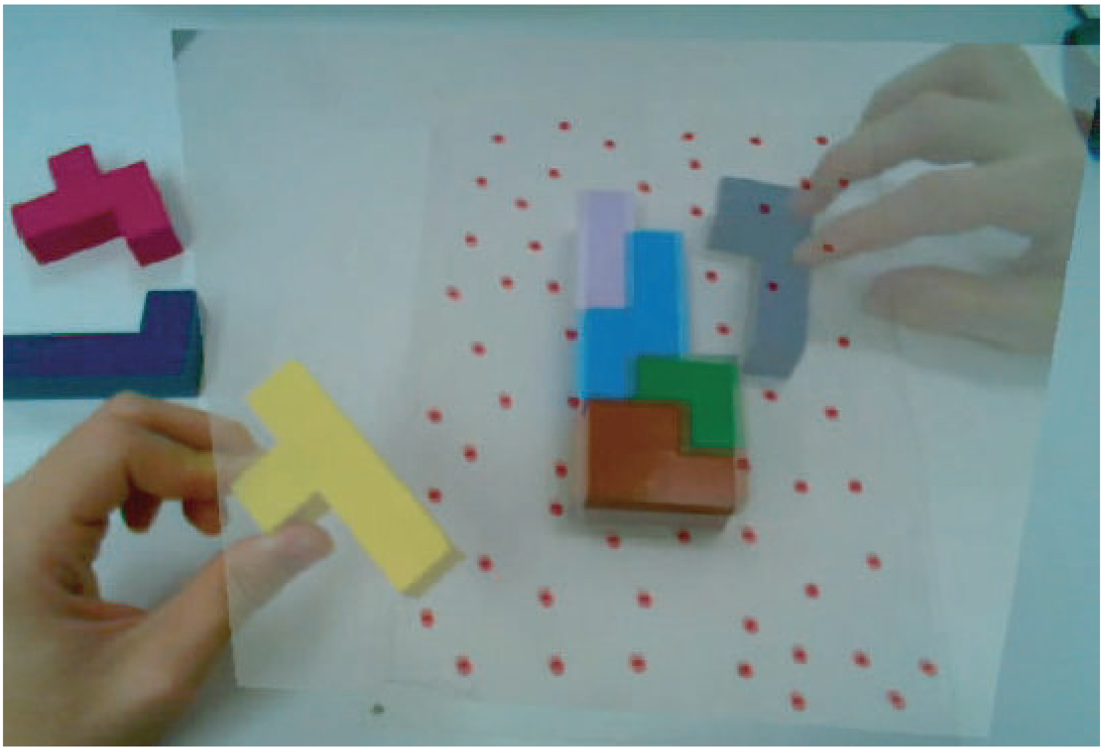


Figure 2.9: Instructor's hand overlay seen through an HMD [83].

notations (e.g., circles and pointers) to guide the novice to perform the proper techniques. Programmed instruction and indirect job shadowing may seem very similar at first glance, but they are slightly different in a sense that the former focuses more on developing the knowledge while the latter focuses on forming the skill.

Goto et al. [83] exemplified the concept of direct job shadowing. This system provides visual guidance to the novice by overlaying in the first-person view an example of the correct way of performing the task. This is illustrated by the phantom representation of the instructor's hand as shown in Figure 2.9. In this study, Goto et al. addressed the problem of visual confusion of the real work environment from the visual guide by using techniques such as changing the transparency and enhancing the contours.

Conversely, Lee et al. [133] demonstrated the concept of indirect job shadowing. They developed a prototype system that is able to capture and share first-person view annotations to share the instructions. These shared instructions are then used by the novice to obtain an understanding of how to do things. Lee et al. used circles and arrows to guide the novice in performing general work activities (e.g., setting up presentation facilities in a seminar room).

Evaluation - Job Shadowing

- Face and content evaluation example:

Goto et al. [83] tested the face evaluation by using a five-point Likert scale for the subjective evaluation of an instructional video in terms if it is easy or difficult to follow. Goto et al. tested the content evaluation again by using a five-point Likert scale for the subjective evaluation of the visual effects (e.g., shift in the AR view, alteration of the transparency, etc.) to determine if it is suitable or not suitable for the specific task.

- Concurrent evaluation example:

Going back, Lee et al. [133] performed a user study with 18 participants from university students and staff and asked them to rate the user experience by using a seven-point Likert scale. Lee et al. allowed the user to undergo two conditions: video only and video with spatial cues. Lee et al. compared the results between the two conditions and did tests (e.g., paired t-test, Wilcoxon Signed Rank test, etc.) for the task completion time, error, and angular difference.

- Predictive evaluation example:

Quarles et al. [177] developed a system where the user can review the past training experience that is overlaid to the current experience of the user. Quarles et al. recruited 19 students and three educators to be a part of the user study. The evaluation metrics that Quarles et al. used included a fault test score and a confidence test score. As evidence of the predictive evaluation, Quarles et al. performed a comparison of the scores before and after using the after action review from the previously mentioned evaluation metrics.

Discussion - Job Shadowing

In the survey of papers that were gathered, only five studies contributed to the category of job shadowing. It is difficult to determine the trends with an insufficient sample size; however, this suggests that researchers still have a lot of scope to cover in this area that has not been investigated yet. Because job shadowing has the ability to facilitate the demonstration of an ideal execution and implementation, it can be considered as a great subject that should be focused for future ARTS. Currently, we have determined that there are two approaches to job shadowing: direct and indirect.

2.5.5 Mentoring

Mentoring is the process where the mentor provides guidance and support to the apprentice. The goal of mentoring is for the apprentice to grow in terms of developing the skill or experience so that they can do the work on their own. The advantage of AR mentoring, just like AR job shadowing, is the ability to facilitate the sharing of each other's views. In terms of the role of the actors, mentoring and job shadowing are similar in that experts share their knowledge and skills to the novice. The main, yet subtle, difference between these two training methods lies in the relationship between the experts and novice. Although job shadowing novices observe the expert performing, mentoring adds the element of the expert being more teacher-esque to the novice.

Implementation - Mentoring

Just like job shadowing, implementation of AR mentoring systems can also be classified as direct and indirect mentoring. Examples of the implementations for direct and indirect mentoring are akin to the strategies that are used in job shadowing (virtual hand overlay or annotations). However, the element of a two-way communication between the novice and expert, along with the condition of it being met in real-time, must be satisfied.

An example of direct mentoring is the work that was done by Wang et al. [227]. They created a telemedicine system that is capable of sharing hand pointing gestures through the use of Leap Motion and HoloLens. In this system, the mentor is able to see the first-person view of the trainee and the trainee is able to see the serialized hand data of the mentor. Above these, they are also able to communicate through a real-time audio stream. Another example is the work by Chinthammit et al. [50]. They developed Ghostman, a system in which the therapist and the patient are able to stream each other the ghost image of their hands. This is used to exchange the perspective that can be observed in the augmented environment. They designed the system in a way that the therapist

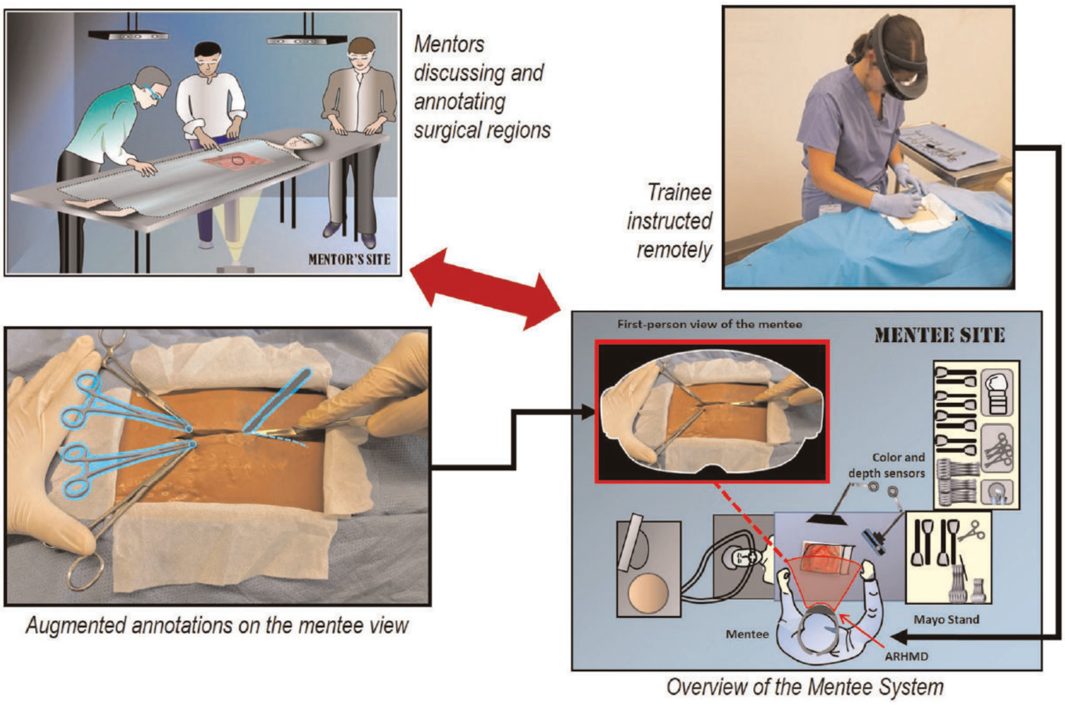


Figure 2.10: Telementoring system workflow [182].

is able to deliver instructions remotely to the patient in terms of how to properly execute the motor skills, and in this case, handling chopsticks.

By contrast, Rojas et al. [182] demonstrated the idea of indirect mentoring. They developed a telementoring system for cricothyroidotomies training, in which expert surgeons can share annotations and audio guidance to the onsite medical trainees, as summarized in Figure 2.10. To accomplish this, first, a stabilized first-person view of the working environment is streamed to the experts. Then, in a seamless manner, it provides instructions to the trainee through the use of annotations and 3D model augmentations projected through the AR HMD.

Evaluation - Mentoring

- Face evaluation example:

As described above in Section 2.5.5, Wang et al. [227] conducted a study that is an example of direct mentoring. For the face evaluation, they collected the trainee's and mentor's opinions about the system and asked them to choose answers such as "the technology was easy to setup and use" or "the technology was overly complex" on a five-point Likert scale.

- Content evaluation example:

The Ghostman work by Chinthammit et al. [50] provides good examples of statements that relate to the content evaluation. In this preliminary user study, they recruited 12 participants with no physical deficiencies. A five-point Likert scale of questions regarding the instructions or the overall training program that contributed to the learning of using chopsticks provided evidence for the content evaluation.

- Construct evaluation example:

As mentioned earlier, the work by Rojas et al. [182] is considered to be an indirect mentoring example that performed a construct evaluation. This is evident by comparing the scores between the different experienced subgroups, namely the low first responder experience (participants with fewer than 10 years of experience) and low cricothyroidotomy experience (participants with less than seven years of experience performing cricothyroidotomy as part of their training).

- Concurrent evaluation example:

Going back to the work by Wang et al. [227], recall that they assessed the concurrent evaluation by comparing the results of the HoloLens versus the full telemedicine setup (traditional method) by using evaluation metrics. The evaluation metrics include the global rating scale, trainee and mentor opinion, completion time, mental effort, and task difficulty ratings.

- Predictive evaluation example:

In the study conducted by Chinthammit et al. [50], the user experience questionnaire comprised a 2 (group) x 4 (test) mixed design ANOVA for the total skill error and the task completion time. The two groups refer to those who used Ghostman or the face-to-face (traditional instruction) method. The four tests contained statistics on the pretest, posttest, as well as 24 hours and seven days during the study.

Discussion - Mentoring

Similar to job shadowing, mentoring only has five studies allocated to this category. It is also difficult to contribute opinions regarding the trends of this method. Although the two may function in a similar manner, unlike job shadowing, which focuses on the demonstration of the ideal result, mentoring focuses more on the teacher–student relationship and how the teacher can guide the student to the ideal result. However, both methods have the potential to flourish the growth of the student/trainee if it is used correctly.

2.6 Conclusion

2.6.1 Overall Discussion

Training plays a crucial role in today's society as a means of developing competent and productive people in the workforce. With the goal of improving human performance, training serves as the medium for creating opportunities of nurturing and flourishing knowledge, skills, and abilities. There are many forms in which training can be successfully carried out. One such delivery that is proving to be effective is computer-mediated training, with an emphasis on AR training.

To answer R1 (What are the training methods for implementing ARTS and how are these methods used in the different fields?), we adopted Martin et al.'s categorization of training methods, and determined five methods for which AR is effective. Simulation training is particularly valuable when targeting operations that are very dangerous when they are executed in a real-life environment. Programmed instruction proves its usefulness when considering the efficacy of imparting substantial information and knowledge. Games-based training displays its advantage by inspiring motivation from users to induce learning. Job shadowing excels in the demonstration of an ideal performance, whereas mentoring facilitates the mentor-student relationship for effective knowledge and skill handover. We also found trends of these training methods across the different fields, such as high simulation studies for medical, and high games-based and no programmed instruction studies for rehabilitation.

To answer R2 (What types of evaluations are used to assess the existing ARTS and how do they contribute to the validity of each training method?), we described Barsom et al.'s validity types and used them as reference to standardize the evaluation of ARTS. For a training system to be ready for implementation in a real-life environment, Barsom et al. suggested that a "full validation process" needs to be completed. Only when the face, content, construct, concurrent, and predictive aspects of the system are evaluated, this can be considered as the ideal ARTS that can be used and deployed in real practice. Our survey results show that no study has achieved the full validation process; however, we have identified which evaluation types are considered to be more prioritized depending on the training method and strategy that was used.

To answer R3 (How is AR utilized for each training strategy?), we looked at the trends between the training methods and evaluations across the different fields. From these generalizations, we were able to discuss how AR is used based on Salas et al.'s strategy for training and gain an understanding of AR's utilization in relation to the different training methods. The first strategy was to convey the information, which was performed mainly by placing the educational content in the relevant locations. We found that this strategy was used by methods that needed to build up on the basic knowledge and foundation of the learner. The second strategy was to demonstrate the desired behavior, cognition, and attitudes so that learners can imagine the ideal result of their training. This strategy was used by methods that already established the basics, and have learners who want to go to the next step of understanding, seeing from the point-of-view of the expert (i.e., job shadowing and mentoring). The third strategy was to create opportunities to practice the knowledge, skills, and abilities learnt. This strategy was where simulation training methods were most applied as simulation studies tried to create rehearsal stages akin to that of the real environment or work scenario. The fourth strategy was to give feedback to the relevant things that aid the users in learning. Many of them presented information that intuitively guided the trainee's current body movements to the appropriate ones. This strategy was applied more to games-based training because an appropriate feedback presentation led to the gamification of training.

2.6.2 Findings and Suggestions

Before reading each full-text article, we presumed that only a limited number of studies would perform a predictive evaluation because it has a high evaluation cost owing to the need for the assessment to be realized in several steps. However, we were pleasantly surprised to find quite a few studies that went out of their way and evaluated their training system in an intricate manner. This is an important point as a predictive evaluation reflects the training effects gained from the system to the actual practice itself, which proves skills acquisition. Similarly, we noticed that more than two-thirds of the studies conducted their evaluations using actual users. Recruiting actual ARTS users as test subjects can sometimes be expensive; hence, the usage of alternative users can be the more practical option. Although not always the case, alternative users sometimes cannot fully reflect assessment results that interpret the performance to actual practice. From these two points, one can argue that the current evaluation design of ARTS is leaning in a positive direction.

The results of our survey indicated that only a handful of studies are focused on job shadowing and mentoring training methods. However, we recommend future ARTS researchers to focus more on this area because there are still numerous concepts and ideas to potentially uncover, which further proves its effectiveness in the application of training. Presently, we have only identified five out of 13 of Martin et al.'s training methods that are effective with AR according to the results of our study. However, this does not mean that the remaining methods are not applicable with AR technology. This merely suggests that no studies have implemented these methods yet. There is still much potential for AR to be implemented by using other methods, maybe even new methods that are not listed in Martin et al.'s core training methods.

We also noticed studies that explicitly performed Barsom et al.'s validity types usually come from the medical field. Barsom et al.'s validity types is becoming (if not already) a standard procedure in the medical field. Although the roots of the test validity concept are from the research design in behavioral and social sciences, adopting this concept in engineering has merits. This is particularly true for studies that handle people as subjects and performance indicators (e.g., ARTS). From this perspective, we can recommend that future ARTS research adopt Barsom et al.'s validity as an option for standardizing evaluations of training systems, regardless of the application field. When the standardization of the procedure for doing evaluations is realized, the next step is to strive for Barsom et al.'s full validation process. It is indisputable that proving the novelty and validity of an idea is paramount in the research community. However, it can also be pointed out that the implementation in actual practice is of great significance.

To quote the famous entrepreneur Scott Belsky, “it’ s not about ideas, it’ s about making ideas happen.” The true value of a training system becomes clear once it is assessed with a full validation; hence, we recommend that future ARTS not only focus on lab performance (i.e., implementation situated in ideal conditions), but also extend investigations in non-ideal scenarios and check how it will fare in actual implementation.

2.6.3 Final Thoughts

We have confirmed the use of ARTS and its current implementations, with an emphasis in the medical, rehabilitation/exercise, and industrial fields. In the future, we believe that ARTS will be utilized more in a wider variety of application fields, particularly for training that is deemed to be hazardous or complex when it is performed in conventional practices. As mentioned above, future studies on ARTS should try to evaluate the totality of the system by following the full validation process recommended by Barsom et al. As the current implementations of ARTS have heterogeneity in their study designs, it is difficult to appraise in a quantitative manner and perform statistical pooling of the data. We recommend that future ARTS implementations classify their work based on Martin et al.’s definition of training methods, while considering the utilization of AR in the training context, as described in Section 2.2. In addition, it would benefit our community to have more investigations performed on the training effects of the less frequently used methods, which include job shadowing and mentoring. When future ARTS studies follow a more structured approach, future review papers can focus more on the statistical effects of ARTS, such as conducting meta-analysis reviews.

3. Methodology

3.1 Realist Evaluation

Realist evaluation has an unique understanding of the nature of interventions and how they function, of what is involved in explaining and comprehending interventions, of the research methods required to comprehend the functioning of interventions, and of the appropriate outcomes of evaluation research. Such particulars will be discussed later, but it is important to emphasize the core objective of realist evaluation now. What are the benefits to the guideline makers? What should you anticipate if you are pursuing or using a realist evaluation? The quick answer is that this kind of assessment has an explanatory goal — intervention hypotheses are evaluated in order to refine them. Thus, the fundamental question posed, and hopefully addressed, is multifaceted. Realist evaluations do not question, “What works?” or “Does this system work?” but rather, “What works for whom under what conditions and in what ways?” [167]

A couple of key realist philosophical notions may aid comprehension of how to undertake realist research in the medical setting. These ideas consist of generative causality [166], ontological depth [23], and retroductive thinking [89]. In accordance with the concept of generative causation, the manifested universe is created by underlying processes. The iceberg metaphor of realist causality depicts the concept of ontological depth, which is the view that reality is stratified in layers, as seen in Figure 3.1 [106]. Retroductive theorizing is the process of discovering concealed action processes in these deeper levels. Realist synthesis employs the Context-Mechanism-Outcome (CMO) heuristic.

Context specifies those aspects of the circumstances under which interventions are implemented that are pertinent to the functioning of the mechanisms. Realism employs contextual thinking to answer the questions of “for whom” and “under what conditions” an intervention will be effective. The realist answer to the “one-size-fits-all” question resides in the concept of context. Certain settings will support the program theory, whilst others will not, according to realism. Realist assessment is thus tasked with separating one from the other.

Mechanisms explain what aspects of the intervention are responsible for their outcomes. Frequently, mechanisms are concealed, similar to how the inner work-

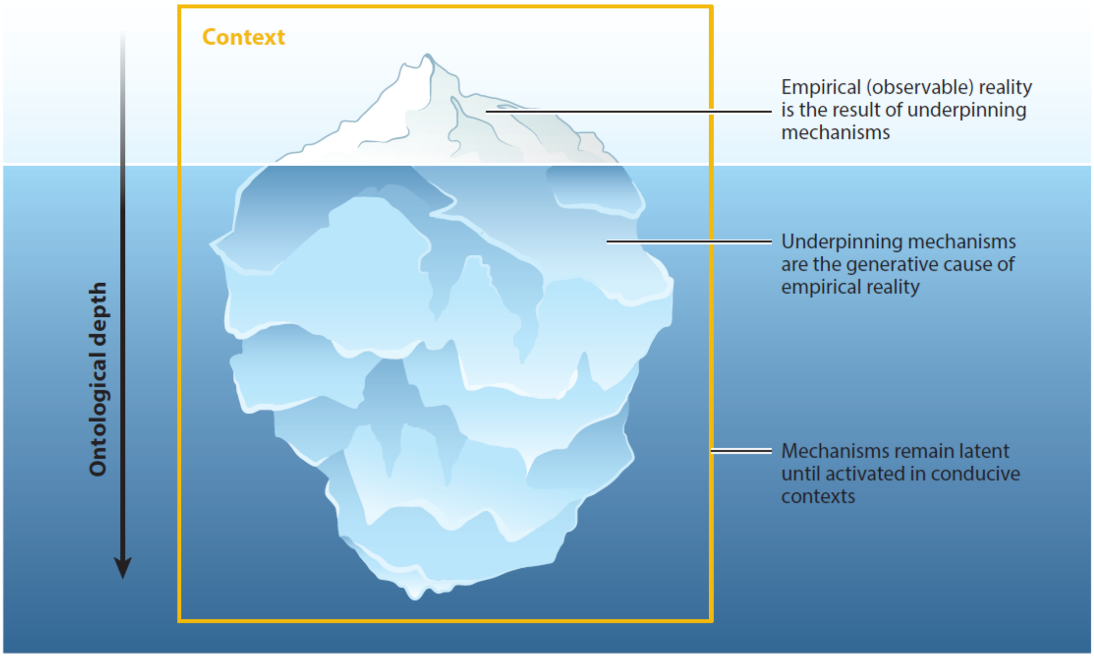


Figure 3.1: Ontological depth represented by the Iceberg metaphor to elucidate the different level of mechanisms activated by changing context, according to Jagosh. [106]

ings of a clock cannot be seen yet cause the hands' rhythmic motions. This realism idea attempts to overcome the laziness of basing assessment on the issue of whether or not "systems function." In reality, it is not the interventions themselves that are effective, but the tools they provide to allow their participants to be effective. This process of how subjects understand and respond to the intervention strategy is referred to as the program's "mechanism," and it serves as the pivot around which realist research revolves. A realist evaluation starts with the researcher hypothesizing the various mechanisms by which a program may function prior to verifying these hypotheses.

Interventions are essentially always implemented in diverse settings, in the sense that the mechanisms triggered by the interventions will change and do so in response to notably distinct situations. Due to crucial differences in context and mechanisms engaged, an intervention's outcomes are likely to be variable. To summarize, Figure 1 provides a graphic representation of such CMO configuration [106].

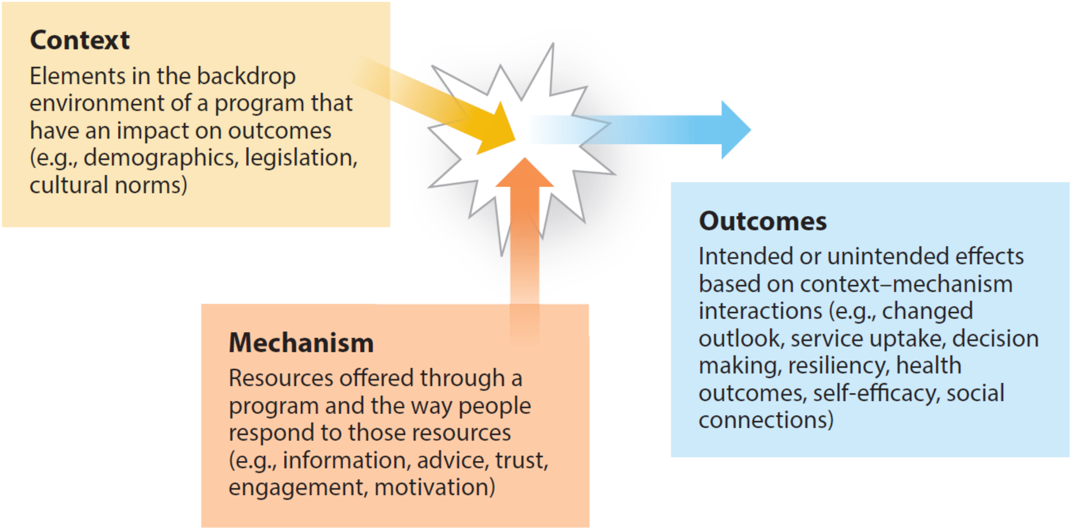


Figure 3.2: The framework of the CMO configuration for Realist evaluations, according to Jagosh. [106]

3.2 Participatory Design

Participatory design is somewhat distinct from the majority of research undertaken by scientific reporters, yet it fits well with the field of human-computer interaction. As the name suggests, the method is as much about design as it is about research, creating products, systems, work structures, and practical or tacit knowledge. In accordance with this paradigm, design is research. Participatory design utilizes a variety of research methods, such as "ethnographic observations, interviews, analysis of artifacts, and sometimes protocol analysis" [205]. However, these methods are always used to iteratively construct the emerging design, which concurrently represents and incites the scientific findings as co-interpreted by the designer-researchers and the participants who will use the design.

3.2.1 Design Stages

At Stage 1, called initial exploration of the work, designers interact with the co-designers and end-users and get acquainted with the methods in which they collaborate. This investigation encompasses not only the used technology, but also workflow, work methods, routines, collaboration, and other facets of the task.

Stage 2 is characterizes the discovery process. To comprehend and prioritize work structure and visualize the future workplace, designers and users apply a variety of methods. This phase helps designers and users to articulate the users' objectives and values and reach an agreement on the intended project result. This phase is often completed on-site or in a conference room with many users.

In Stage 3, the prototyping stage, designers and users construct artefacts iteratively to conform to the environment defined in Stage 2. If the design is a functional prototype, testing may be undertaken on-site or in a lab, with one or more users, and on-the-job.

The above three stages are iterative in nature and should provide an environment for the designers and end-users to dive into the process of co-exploration. [205]

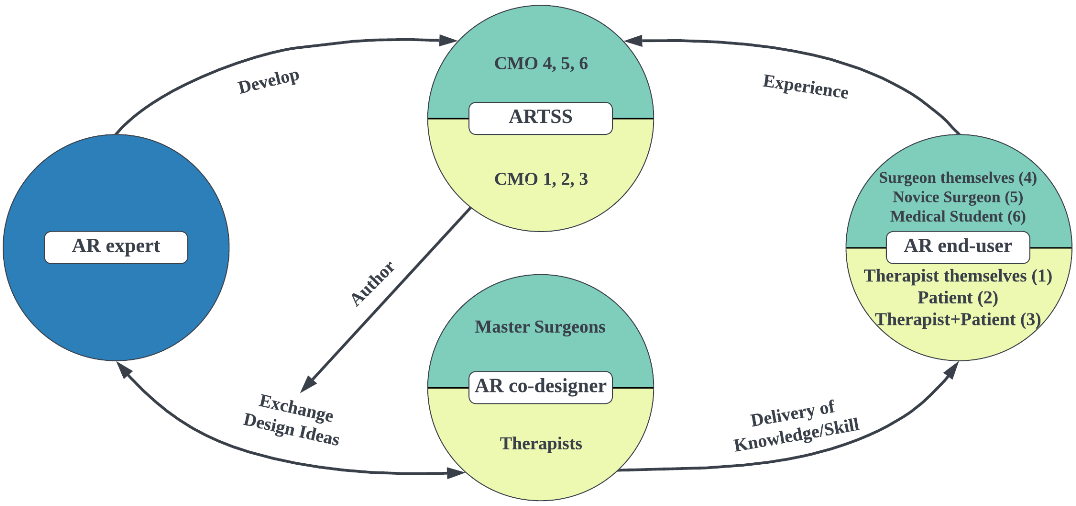


Figure 3.3: Participatory Design Overview for ARTSS (Augmented Reality Training and Support Systems) of Two Context Scenarios (Upper: Surgery Training, Lower: Occupational Therapy). A, B, C and D refer to the prototypes described in the later sections.

3.2.2 Participatory Design in AR

Especially from the standpoint of multidisciplinary collaborations, we argue that Augmented Reality (AR) designers should take more of a facilitator role rather than a dictator [205]. The reason is that, from a research point of view, there is a need to eliminate or articulate the biases present in studies. When handling multiple disciplines, however, people in their specializations hold their suppositions, making it difficult to determine these biases. By applying PD, we can generalize and conduct studies more empathetically; that is, sharing a common language between the collaborators [205].

One example of how PD bridges the gap between different educational backgrounds, cultures, and specialties is the work of ParticipArt [104]. Another instance of exciting insights gained with PD in AR is on the design for children [41, 135]. Furthermore, some studies deal with context scenarios similar to those tackled in this paper, such as surgical training [86] and therapy/rehabilitation [130, 162].

Figure 3.3 summarizes the design workflow and iterative process of the co-creation and co-design of prototype systems. The concept of our current design is inspired by the work of Santos et al.'s PD of AR Learning Objects [187]. From the book of Reimann and Cooper's About Face [178], they stand by the principle that "in the early stages of design, pretend the interface is magic." We, as AR experts, also follow through with this principle when exchanging design ideas with co-designers. PD allows us to think through the lens of our co-designers. The left side of the loop in Figure 3.3 ensures that the developed ARTSS are technically sound, while the right side of the loop ensures that it meets the end-user demands and goals. The prototype of interest are enclosed in parenthesis (i.e., co-designers of prototypes A, B, and C are surgeons, while end-user of prototype A are the surgeon themselves, end-user of prototype B are medical students, and so forth.)

For the research design, we follow the methodology of PD by Spinuzzi [205]: an initial exploration of the work, discovery process, and prototyping. In the initial exploration, we dive into the problems that are to be solved by our co-designers. We try our best to share each other's tacit knowledge, to all intents and purposes, a familiarization stage. In the discovery process, we exchanged ideas and suggested some proposals. This stage is where the vision/end goal of the co-designers is determined, and both parties agree on what the outcome should be. Finally, in the prototyping stage, abstract ideas and proposals are turned into concrete AR artifacts that prove helpful to the end-user. Iterations and descriptions of incremental redesigns are also expounded in this stage.

3.3 Design Principles when Considering the Context

3.3.1 Physical and Occupational Therapy

Under the AUSILIA project [171] in Italy, with the University of Trento and the Provincial Agency for Health Services of Trento, PD research is carried out in Villa Rosa Rehabilitation Hospital in Pergine. This project aims to restore the independence of people who have lost their autonomy after a trauma caused by a pathological status. In this regard, the rehabilitation discipline that uses assessment and treatment to develop, recover or maintain people's Activities of Daily Living (ADLs) is Occupational Therapy (OT). ADLs include personal care, such as hygiene, bathing, showering, dressing, functional mobility (walking, sitting, standing up) and self-feeding, productivity (household management, paid/unpaid work), and leisure (socialization, quiet or active recreational activities). End-users are actively involved in every step of their rehabilitation process, starting with the initial interview to identify problematic ADLs, preliminary agreement on goals and treatment plan, and self-assessment of their level of performance and satisfaction. Once problematic ADLs are identified, the OT must assess how the end-user performs each activity before implementing the treatment plan.

Initial Exploration of Work

Similar to the context scenario of the Surgical Training in the previous section, the therapists were also shown the MRTK2 Examples Hub demo [120] with HoloLens v2 as the AR device. On top of this, the AR researchers also developed an AR balancing game demo, as seen in Figure 3.4. This demo was made to stimulate the therapists' imagination, so they could draw up novel programs to give to their patients. As the developed demo application uses the HoloLens front camera for image tracking, we could not capture a first-person view of the therapist playing the actual game demonstration.

Discovery Process

An advanced measurement system combined with innovative visualization technologies such as AR is needed to assist OTs during their work. From their feedback, the ability to have more information in AR contextualized close to the patient simplifies their assessment without losing the exteroception of the scene. With these technologies, the OTs can define more reliable assessment scales based on objective parameters, increasing the effectiveness of clinical observation for more effective rehabilitation programs.

Moreover, these technologies not only increase the clinical eye of OTs and thus their final assessment of patients [210], but also the involvement of patients in daily life through gamification of some of their daily activities in AR. This gamification in AR can eliminate environmental barriers, such as social, physical, economic, or biological reasons, which are one of the causes of independence difficulties.

Finally, during all OT steps, therapist-patient interaction is kept constant or enhanced by AR because it is a crucial aspect of success in rehabilitation. As feedback collected during our demos, physical, verbal, and technical exchanges

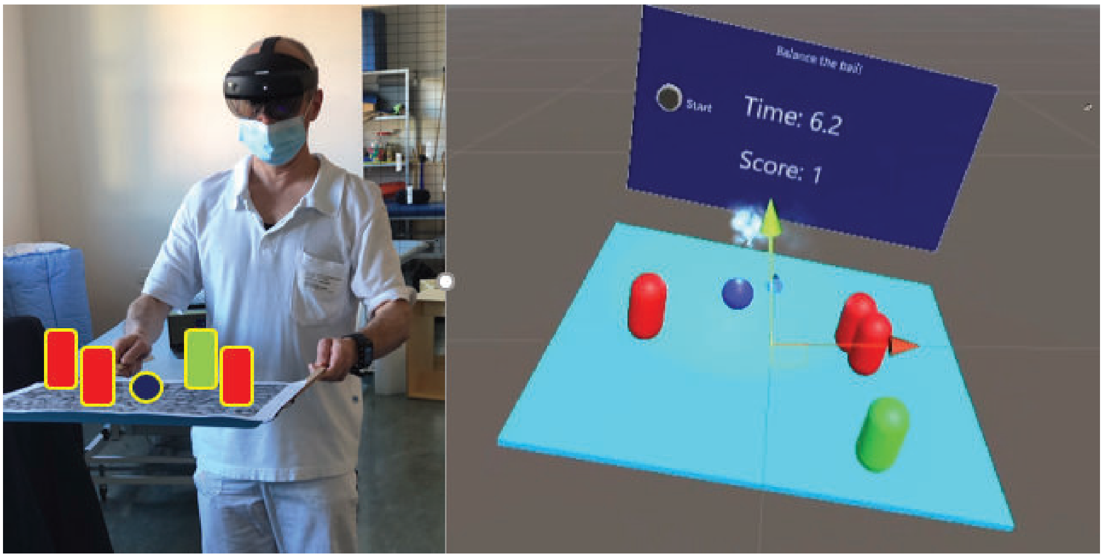


Figure 3.4: The Balance-the-ball AR game demo that was presented to the therapists.

between physical therapist and patient highly influence the outcome by promoting patient engagement and therapist involvement. Discussions led to a Shared AR demo in which not only AR cues are designed for a single end-user but instead a shared system to be viewed and managed together by the patient, therapist, or caregiver. In particular, we designed a Shared AR demo for the ADL, setting up a table where the therapist and patient can interact with visual cues superimposed on the actual environment to assess and train the patient's performance in OT.

3.3.2 Surgery Training

The first context scenario is about an ARTSS for the use of medical practitioners specializing in Periacetabular Osteotomy (i.e., hip bone surgery). The main stakeholders for this section's project are the AR developers/researchers and experienced surgeons. The sub-stakeholders to consider for this design (although participation was non-obligatory in the meetings and discussions) are the novice surgeons, medical students, and nurses who provide support during the actual surgery. There are many problems that medical practitioners want to be solved using AR, and each of these will be described in section 3.3.2. This project's ultimate goal is mainly to improve the performance of each/all of the end-users with their respective tasks.

Initial Exploration of Work

As described in subsubsection 3.2.2, PD is beneficial in sharing the tacit knowledge of expertise between the concerned parties. Besides this point, it is also vital for each member to familiarize themselves with their colleagues' explicit knowledge.

As a means of facilitating this familiarization process and bridging this gap, the medical practitioners (surgeons and medical students) were invited to the AR developer/researcher's laboratory (Interactive Media Design Lab, Nara, Japan) to experience and understand the possibilities and limitations of today's AR technology. They experienced the use of HoloLens v2 loaded with the MRTK2 Examples Hub scene [120]. Some essential points they need to understand through this experience are navigating the holograms, doing hand interactions (near and far manipulations), using their eyes to make precise selections, and so forth.

On the other hand, AR developers/researchers studied the surgical procedure of Periacetabular Osteotomy, the subject of focus for this section's design. Afterward, the self-acquired explicit knowledge gained by the developers/researchers was fact-checked by the medical practitioners through the exchange in one of the many online meetings.

Discovery Process

Constant information exchange and discussion between the AR developers/researchers and medical practitioners were upheld from the start of mid-2020, continuing up to the present. After the initial exchange of knowledge, it was agreed upon that a critical point for developing an effective training system is using actual patient computed tomography (CT) data. Another point shared was the importance of hand-eye coordination as a skill in surgery. Lastly, it is crucial to distinguish in what areas the novice and experienced surgeons differ in skill. Further explanations and solutions to these problems will be expounded and addressed in each of the following subsections of the respective prototypes below.

4. Evaluation by Empirical Evidence in Physical and Occupational Therapy

One of the main objectives of rehabilitation is to make improvements in terms of quality and quantity in the day to day activities of the independent living of the person. Three factors of motor development are early intervention, task-oriented training, and repetition intensity whereas a key therapy goal is to find ways to provide repetitive incentives for activities requiring multimodal processes (different sensory modalities like vision, haptics, proprioception, audition) and to further allow functional improvements [211]. Carr and Shepherd concentrate on motor relearning, in which relearned gestures are organized for specific tasks [112]. They suggest that the practice of special motor skills leads to the ability to perform the task and that motor tasks should be carried out in suitable environments where sensory inputs modulate their performance. Keshner and his colleagues have specifically addressed the practical importance of the specific environmental context as it applies to posture regulation [121] [122] [123]. Such scholars have shown that different postural responses vary between paradigms where discrete human control mechanisms (i.e., auditory, vestibular, somatosensory pathways) are controlled as opposed to those within a biologically relevant context where information is available from multiple paths.

In this study we will be introducing prototypes that is focused on Occupational and Physical therapy on motor rehabilitation. While the difference is slight, there is one basic difference in Occupational and Physical therapy. The main difference between occupational therapy and physical therapy is that an Occupation Therapist is focused on improving the ability of a client to perform daily living (ADL) activities, and a Physical Therapist is focused on improving the ability of a client to perform human body movement [156].

The Occupation Therapist (OT) is concerned with the person as a whole. Whether they recover from injuries, or have developmental or cognitive impairments that affect their motor skills, emotions, or behavior, OTs help people fully engage in their daily life. OT is unique in that it uses a holistic approach to investigate not only the reasons why a client's participation in operations has been compromised but also the tasks and atmosphere of the person, as mentioned in the National Board for Certification in Occupational Therapy. The method involves encouraging health, recovery, and habilitation. For example, after playing basketball you recently broke your foot and can't take part in your Wednesday night pick-up league any more. While you recover, you might meet an OT to get to the root of why you look forward to playing every week. Is it the exercise which is important? Is it the team interaction and communication with the people? The OT will help you attain your goal [156].

On the other hand, a Physical Therapist (PT) approaches the real condition of the patient from a biomechanical viewpoint. Physical therapy helps to enhance the condition by increasing mobility, aligning bones and joints, or reducing discomfort itself. The primary objective of a PT is to bring their patients back into motion with exercises, massage, and other techniques. These are focused on preventing trauma, and can help people avoid surgery or long-term drug dependence. You love playing with your kids outside however you are unable to do so anymore because of a herniated disk. Your PT will consult with you to develop a specific treatment schedule to improve your rehabilitation, including workouts and stretches you can do at home [156].

4.1 Clinical Eye Training System

4.1.1 Introduction

According to the World Health Organization (WHO), the Global Burden of Disease Study presents that 74% of the years lived with disability worldwide are related to health conditions that can be mitigated with the support of rehabilitation. Today, the number of health conditions associated with severe disability rates has increased by close to 183 million [233]. Rehabilitation needs are continuing to grow globally, especially in low and middle-income countries. Rehabilitation services demand already exceeds availability, leaving a great unmet need. To cope with these demands, education and training of therapists that had just finished schooling may find their knowledge and expertise to be lacking in the actual field [11]. This is mainly due to the differences between the scope of the theoretical knowledge in the literature about rehabilitation concepts and their application in clinical practice. Developing the "Clinical Eye" would take years and years of practicing the profession, so novice therapists would have difficulties in making difficult clinical decisions and evaluations. To offer a solution to this problem, we propose AR as a tool for these therapists to use for motor rehabilitation.

The question now arises; how can AR be used as a tool for these therapists? Simply put, AR is the technique of integrating computer graphics into the user's view of their world. By providing useful data as virtual objects, AR can improve a user's understanding of the real world and its interaction. The virtual objects display information which the user cannot detect with his own senses directly. The knowledge the virtual objects convey lets a user perform real-world tasks [14]. This is an example of what Brooks calls "Intelligence Amplification (IA): using the computer as a tool to make a task easier for a human to perform" [36], or in this case to make the rehabilitation factors and variables easier to understand.

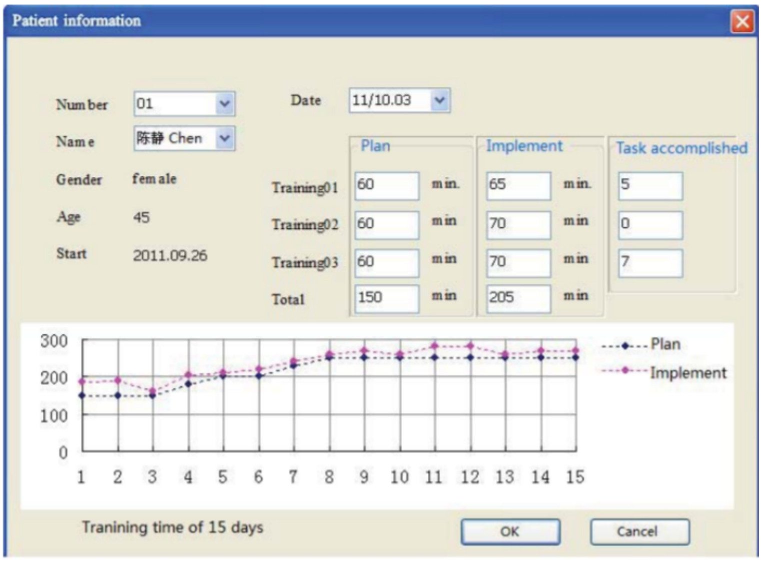
New technologies of rehabilitation can provide more flexible resources for recovery or improve the clinical process. Nevertheless, the lack of education by clinicians regarding technological advances and apprehensions related to the role of technology in the rehabilitation scene leads to wanting them sticking to the status quo. Two explanations may be given why therapists are hesitant on taking part in the development of these technologies. Firstly, the engineers who design these innovations do not recognize the value they might gain from engaging therapy practitioners to make their interfaces more functional, user-friendly and useful for certain disabilities. Second, many therapists are uneasy at using new technology and are concerned it could take the place of individualized patient experiences [73]. In this work, we will develop a prototype that is designed specifically for use of these therapists on motor rehabilitation and analyze the feasibility of this prototype as a tool by rating the prototype based on the New Technology Evaluation by Jones et. al., specifically looking at the clinical applicability, financial, marketability and safety factors [116].

4.1.2 Related Work

One example of the application of AR in motor rehabilitation is the game shown in Figure 4.1, where the patient is playing with a virtual object that is visualized as a kettle. The patient is tasked to do movements with the kettle and then keep the given pose for a given amount of time. The patient is instructed to change the position of the kettle to over the cup as he tilts the kettle simulating the action of pouring water into the cup which is empty. The computer will judge whether the kettle has entered the target location by checking the coordinates changing in real-time. The computer notifies the patient to tip the kettle at an angle and only when the requirements are met, the kettle starts to pour water into the cup. The cup will be full after waiting some time, then it will progress to the next step by generating a new virtual cup at the next position. As the patient does again this rehabilitation process, the upper limb's motor ability and joint flexibility will slowly improve over time.

Figure 4.1: AR Motor Rehabilitation example: Tea Pouring Game

To make sure that the therapist has a well-rounded understanding of the session of rehabilitation, live video, real-time data, and patient information is being delivered remotely to the system the therapist uses. Data regarding the patient includes personal and past information records history. Manpower and resources are also minimized because the therapist are able to manage multiple patient's rehabilitation training in a remote manner. Figure 4.2 demonstrates the interface for patient information within the therapist system. Personal information of the patient such as patient number, name, gender, age, and the starting date of rehabilitation training are displayed. The planned rehabilitation time for the patient, the actual training period, and the amount of activities completed in each training game was also noted. The rehabilitation training plan progress is also compared to the currently implemented plan through a graphical visualization. Using the listed data enumerated, the therapist would be able to get a general understanding of the progress of the patient. Real-time data of the patient helps the therapist track the in-depth the recovery of the patient. Figure 4.3 shows another set of information the therapist will be able to observe, which are the x, y, and z coordinate graphs of the marker. The images on the figure shows actual, not augmented images so as to avoid the limbs of the patient to be occluded by virtual objects. These coordinate data are then plotted to make it easier to understand, for example showing it as XY and XZ plots. Such graphs reflect the tempo and the stabilization of the upper limb movements of the patient. These serves to supplement the assessment of the therapist aside from the clinical eye alone, and offer an important framework for evaluating recovery of the patient [236].



Patient information

Number 01
Date 11/10.03
Name 陈静 Chen
Gender female
Age 45
Start 2011.09.26

	Plan		Implement		Task accomplished
Training01	60	min.	65	min.	5
Training02	60	min	70	min	0
Training03	60	min	70	min	7
Total	150	min	205	min	

Figure 4.2: Tea Pouring Game Therapist System: How to visualize patient information

Electromyography (EMG) can be considered one of the most significant biological signals used to track skeletal muscle performance. It can be collected from the skin and termed as surface EMG (sEMG). Aung et. al developed a system

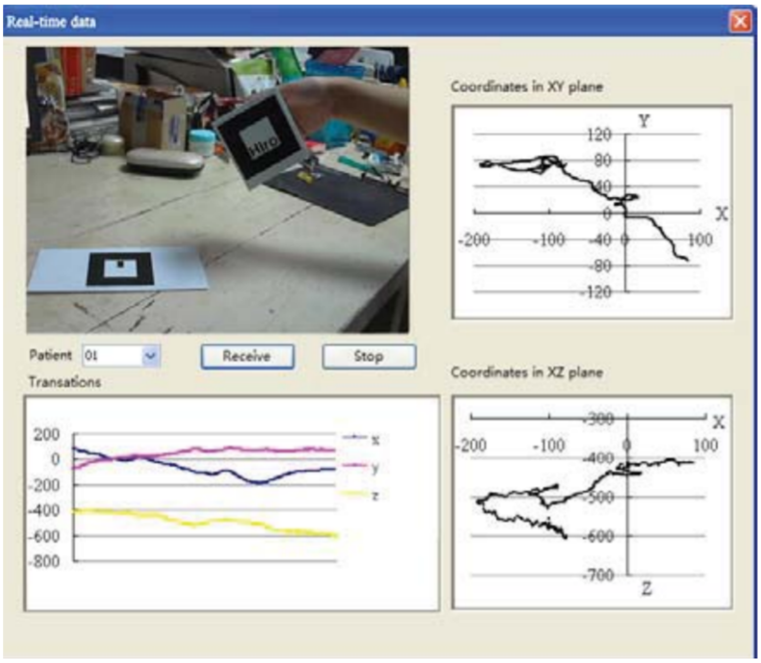
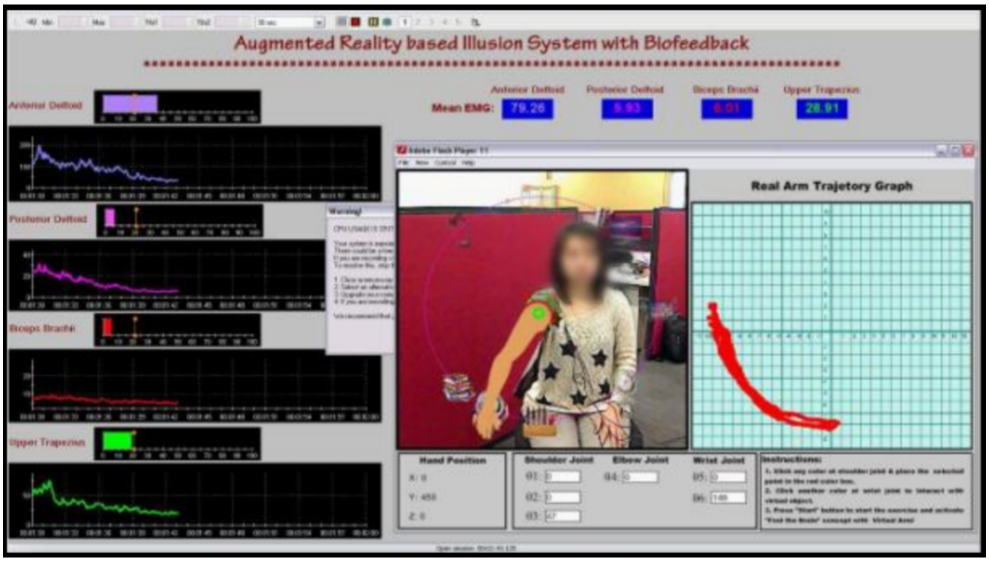


Figure 4.3: Tea Pouring Game Therapist System: Real-time data stream

which took advantage of four sEMG signals, namely anterior deltoid, posterior deltoid, bicep brachii and upper trapezius muscles. Signals that are measured are then shown on the screen in real time as shown in Figure 4.4, and also stored for future analysis. When the recorded sEMG signals go above the predefined threshold value, the simulation is triggered. At this instance, the color of the muscle will change, just like in Figure 4.5, such that the currently used muscle can be distinguished on the time of exercise. The start and stop button is also added as a form of biofeedback simulation to facilitate the patient and therapist alike to read and stop real-time view of data during muscle simulation. The biofeedback incorporates both the interests of the patient and the therapist. From a user's perspective, showing real-time added muscle activity serves as one kind of motivation strategy to imagine how the muscles work during the training. From the therapists' perspective, it helps to map and monitor the muscle strength in real-time to understand the current situation of the patient, thus providing a more informed feedback [13] [12].

The works that have been mentioned so far mention about showing the patient data and their visualizations to the patient, but at retrospect could be not



Mean EMG: Anterior Deltoid 79.26, Posterior Deltoid 9.93, Biceps Brachii 6.01, Upper Trapezius 28.91

X: 0
Y: 450
Z: 0

θ1: 0
θ2: 0
θ3: 47
θ4: 0
θ5: 0
θ6: 140

Figure 4.4: ARIS Biofeedback System: Example of EMG muscle Visualization

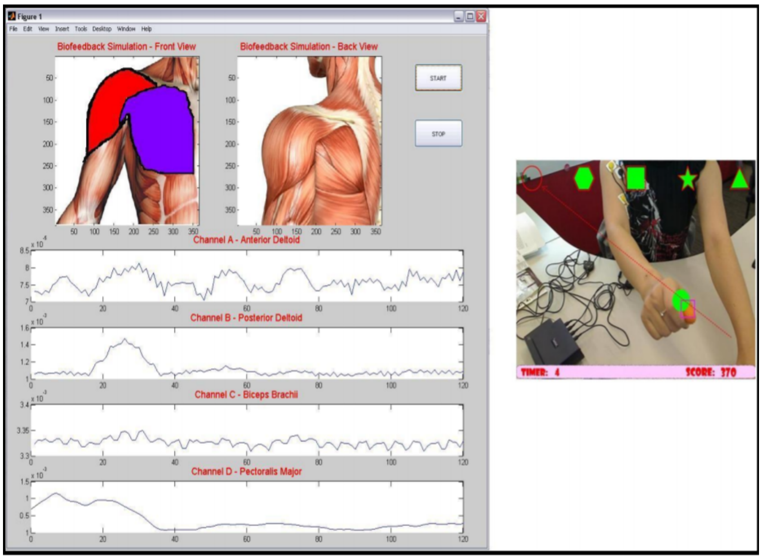


Figure 4.5: RehaBio System: Another example of EMG muscle Visualization

optimal, and have many rooms for improvement. I say this because they have implemented their system using AR and share these AR experiences with the patients. However, when they have implemented on the side of the therapist, they just show 2D visualizations in the monitor to be observed. They have missed out on the potential to also show rich visualizations using AR techniques to the therapist.

4.1.3 Prototype for Chair Training

The goal of this prototype is to create a system to assist the therapist in visualizing information of a patient with motor dysfunctions when they are performing a “chair training rehabilitation” [107]. This type of rehabilitation consists of repeated actions of the patient such as sitting, squatting, using hands as support in the action of standing up from chair, or practicing the center of gravity movement. This prototype was deployed in Kyoto University Hospital, and from there we received valuable feedback regarding our system from Occupational and Physical therapists.

In Figure 4.6, we can see the schematic diagram of the whole system. For this prototype, there are two main data that is collected and then visualized. The first type of data are force sensors attached to the chair. Each sensor node consists of a sensing module (in this case the load cell), and an Arduino. All of these data are then concatenated to the main node, which is the Raspberry Pi, and forwarded to the database which is handled by the MQTT broker (details will be described later in this chapter). The other type of data are the 3D joint positions. These data are generated by the 3D Time-of-Flight Depth camera (KinectV2) which tracks a total of 25 body joints. These data are forwarded to the Main PC where transformations of the space coordinates are processed (describe later in this chapter), and spatially correct visualizations of skeletons can be viewed in the target device, in this case the HoloLens and the tablet. The reasoning why these two devices where chosen was because we want a device that can easily provide AR experience, the device is readily available in the developer’s lab, and easily usable in the clinical setting. Finally, all AR interactions can be controlled by a smartphone which is sending MQTT messages as controlling interface, developed in Node-RED.

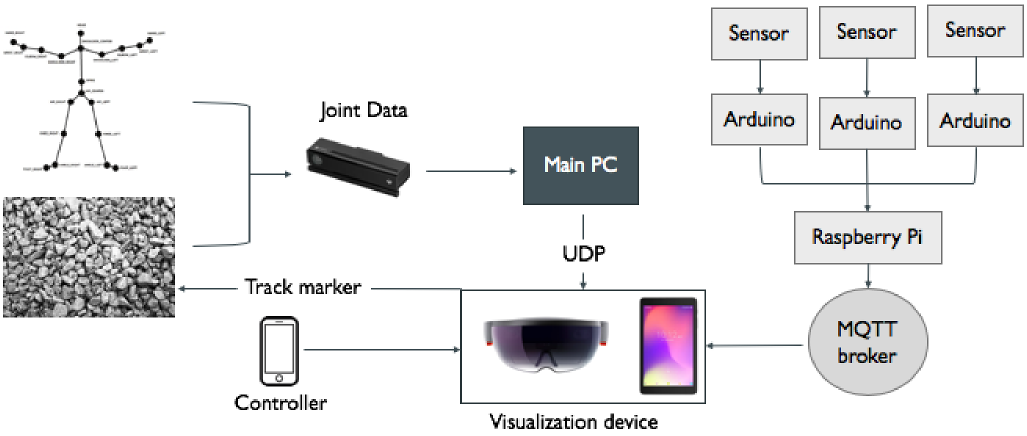


Figure 4.6: Schematic Diagram of the Whole System

Deployment

We have deployed our system at the Kyoto University Hospital. Both parties have benefited from this joint project as conducting research on new medicine and developing new medical technologies is an important mission of Kyoto University Hospital as a core clinical research hospital. The main contact we have from this hospital is Dr. Okahashi, who is working as an Occupational Therapist. The chair training visualizations consist mainly of chair force sensors and skeleton overlays. We have done the demonstrations of our chair training prototype to actual therapists as seen in Figure 4.7. The demonstration was shown to around 10 therapists, and each one was able to have the chance to try and wear the HoloLens and interact with the tablet. We can see from the figure that the HoloLens view is shared to everyone through a big monitor in the back, although there is a latency of around 5 seconds. In Figure 4.8, we have also shared our prototype of the tablet version so the therapists can compare the two AR experiences during this demonstration. The person playing the part of the patient is our Italian co-researcher, so that we can give detailed and structured explanations while he performs some actions. The duration of the demonstration took around an hour and afterwards another hour was spent on discussion with the therapists about their impressions, feedback, and suggestions regarding our prototype. This discussion was performed in a Round Table manner, and although the exchange was held in Japanese, colleagues have helped us with the note-taking and translation of the contents of the discussion. In the following subchapters, I will share some insights that I have learned after this demonstration, and also share the therapists' feedback about the prototype that we had made.

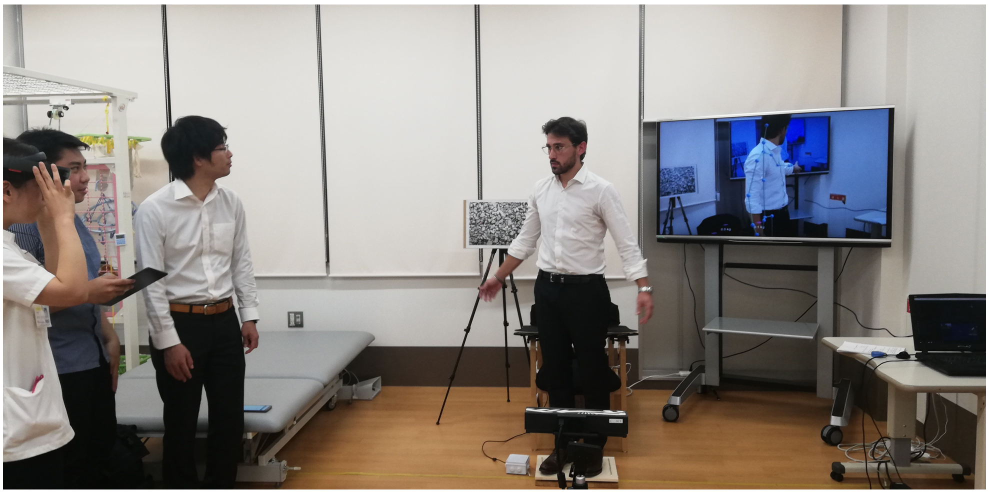


Figure 4.7: Demonstration of Prototype in Kyoto University Hospital A

Figure 4.8: Demonstration of Prototype in Kyoto University Hospital B

Data Visualization

On answering the question on how should these data be visualized, we have used augmented reality concepts to give good AR experience to the therapist.

The concept we have used is situated visualization, which is “visualizing relevant virtual data directly in the context of the physical site” [229]. This just means we have placed the angle visualizations just beside the joint of the real patient, and rendering graphs in the location of the chair area the force is exerted (therapist will be able to choose which type of visualization he wants, whether bars or arrows) as seen in Figure 4.9. We have designed the system to be as intuitive as possible, such as changing the color of the skeleton when angles reach certain thresholds, giving a warning as dangerous (also added sound feedback warning). Finally, situated visualization also suffers from the problem of data overload. To solve this problem of data overload, toggling on/off visualizations functionality has been added.

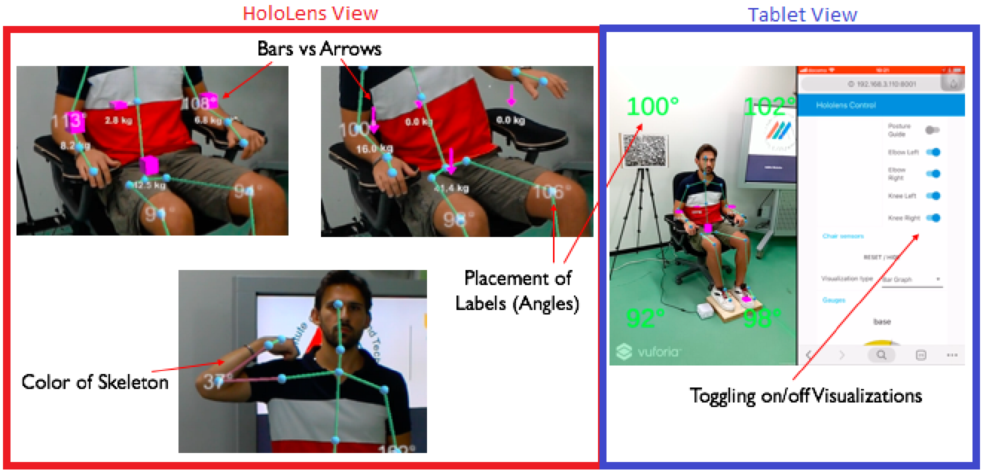


Figure 4.9: Visualizations used for the Chair Training prototype

The answer we found to what data would be helpful in the chair training are force data and 3D position of body joints data as shown in Figure 4.10 and Figure 4.11. This is because during training, the therapist wants to see how much pressure the patient is exerting on certain parts. This is important because they do not want patients to overexert muscles during rehabilitation. Another important point is tracking the joints of the patient. From this data, we are able to derive the angles in which the patient bends his joints. This is helpful for the therapist because proper posture and form is necessary when dealing with precise rehabilitation.

Figure 4.10: Chair Force Sensors developed

Figure 4.11: Kinect 3D Joints tracking used in implementation

Therapist Discussions

The main comments that we got from the therapists was that they want more kinds of information aside from the data provided by force sensors and skeleton information. Patient wearable sensors such as sensors measuring heartrate, temperature, or blood pressure is a really good addition to the system. Another suggestion they had was that instead of the total force applied to each area, they want to know the weight distribution in real-time. This can be done by having a pressure sensor matrix to gather how the distribution of force is applied. Another information they want is information about speed when the patient is moving. Another data they wish they could see is seeing which muscles is currently used by the patient. This can be done by using electromyograms to estimate the muscle activities within that area. General considerations that they suggest during future exploits is that they want a recording of data function, so that they could review the details of the rehabilitation when session is over.

4.1.4 Prototype for Walking Training

The goal of the second prototype is to create a system which is able to show useful patient data when undergoing "walking training". By walking training, the patient attempts to walk in a straight line about 5 steps while maintaining proper walking posture. Guidelines for correct walking posture that can be found on any general healthcare is listed on Figure 4.12. The goal of the walking training prototype is to give the therapists hints and indicators that the patient is properly following the correct walking posture, just like in the example of Figure 4.13, such as keeping your spine upright, or having proper spacing between your feet during walking.

The idea that implement on the walking training prototype is adding visualizations such as back posture guide, head trajectories, feet trajectories, and average speed it takes for the patient to complete one cycle of training, as seen in Figure 4.14. All of these visualizations can be made with just using the information from the 3D joint coordinates taken from the KinectV2 camera, so we just use the same prototype system that we have already discussed in the previous chapter. However, to make improvements of the system, we have also added additional sensors that therapists want to see on the duration of the train-

Figure 4.12: Common Guidelines for Correct Walking Posture

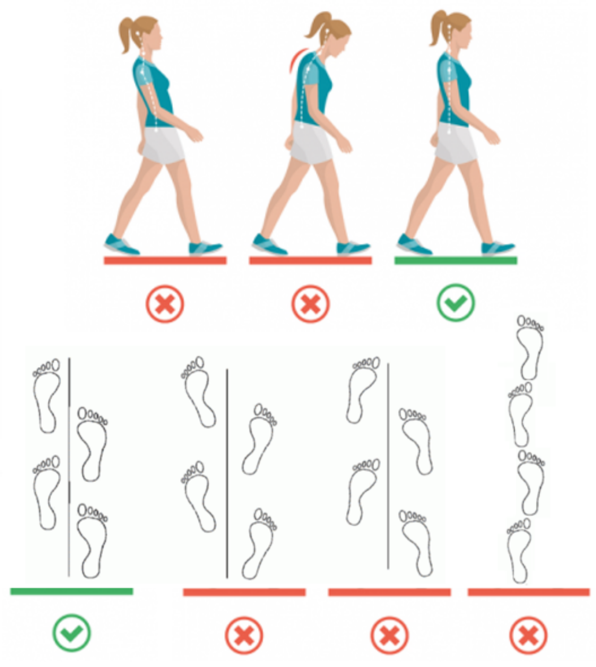


Figure 4.13: Common Indicators for Correct Walking Posture

ing. Discussed in the therapist suggestions in the previous chapter, these are heartrate and body temperature, which muscle is being used (EMG sensor), and weight distribution (Sole pressure matrix). These additional sensors are added to the system via Bluetooth, publishing all these data into the MQTT broker as seen in Figure 4.15.

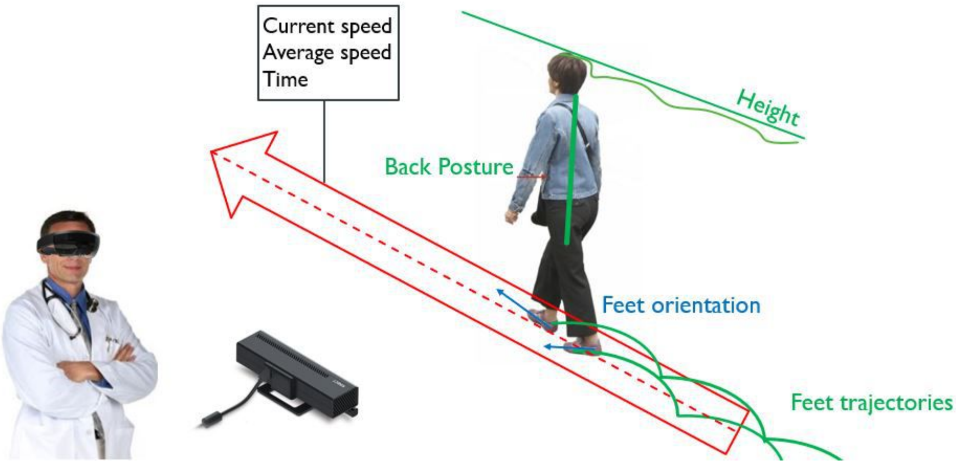


Figure 4.14: Idea for Visualizations on Walking Training

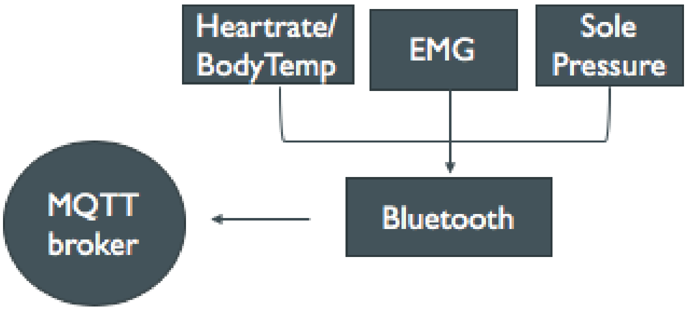


Figure 4.15: Additional Sensors Schematic

Deployment

We have deployed our system at the Takanohara Central Hospital. Unlike the previous hospital of Kyoto University Hospital which is also concerned on conducting research, the Takanohara Central Hospital is more concerned with its public service, thus for this demonstration we focus on the practicability and usability of our prototype in the clinical setting. The walking training consists mainly of weight distribution, back posture guide, head trajectories, feet trajectories, and average speed visualizations while also adding views of patient physiological parameters, such as heartrate, body temperature and muscle usage. We have done the demonstrations of our walking training prototype to actual therapists as seen in Figure 4.16, and also show quickly the previous prototype of chair training seen in Figure 4.17. The method of demonstrations was similar to that of the one held at Kyoto University Hospital. The demonstration was shown to around 15 therapists consisting of a mix between actual therapists and student therapists, and each were free to come up and try the prototype themselves. The HoloLens view is also shared through a big screen monitor, and the person playing the patient is again our Italian co-researcher. Because our demonstration was not the only demonstration in this event, therapists took turns into trying out each of the prepared demonstrations. We were then not able to get a Round Table discussion after the demonstration, however we engaged in a discussion with the therapists while they were trying out the demonstration. In the subchapters, I will share some insights that I have learned after this demonstration, and also share the therapists' feedback about the prototype that we had made.

Data Visualization

To answer the question on how to visualize these data, we have used the same concepts of situated visualization as with the earlier prototype. For the therapist to understand the whole picture of the entire training session, we have created the visualizations seen in Figure 4.18. The head and feet path trajectories allows the therapist to see the data of the patient's walking patterns, and can also allow for comparison with previous data to determine whether the patient has improved over time or not. Values for speed at which the patient is moving, heartrate, and body temperature is shown beside him for the therapist to easily monitor proper pacing of the training. A graph which changes from 0 to 100 percent shows the muscle activity of the amount of strain the patient is exerting at that located muscle. A vertical guideline is also added as a visualization to show whether the

Figure 4.16: Demonstration of Prototype in Takanohara Central Hospital A

Figure 4.17: Demonstration of Prototype in Takanohara Central Hospital B

patient is maintaining proper posture or not during the duration of the training. Lastly we show a sole visualization attached to the foot of the patient to show how the patient is distributing his weight under his feet, where green parts are areas of no force exerted while areas approaching red are areas of more force exerted.

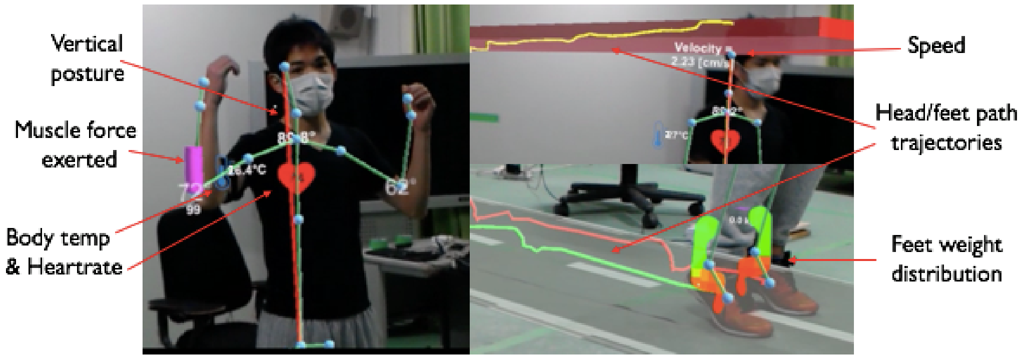


Figure 4.18: Examples of the Visualizations used in the Walking Training

The answer we found on what data would be useful for this walking training is similar to the first prototype. We need to track the 3D joint positions because we want to track whether the patient is doing the correct posture of walking or not. We also record the data of these joint positions over time, so that we can derive the path trajectory the patient has took, and also the speed at which he is walking. We have also added force sensors to determine weight distribution in the patient's sole, so that the therapist will understand more clearly the balance of the patient. Furthermore, we have added EMG sensors to measure the activity of the muscle used by the patient, again to determine if proper form is being followed. Lastly we have also included sensors for heartrate and body temperature to measure the physiological parameters of the patient whether or not the training has become a bit too intense.

Therapist Discussions

General feedback that we got from the therapists who tested our system were very good. They say that because of the variety of data, the possibility to use this prototype to other forms of training or rehabilitation is also doable. This second prototype was able to provide the data which the therapists from the first prototype wanted to see, such as the heartrate, the body temperature, the muscle activities, and the weight distribution. Their feedback was that because of the augmented information about the patient, they will be able to get a better understanding on the status of the patient. Seeing the head and feet path trajectories also proves to be very helpful for them because they are able to see past data and not rely too heavily on the memory of how the patient executed the training. The comment about having a recording function was suggested again by the therapists, however since the video recording function of the HoloLens takes too much memory when used for the duration of the whole training, we decided to adopt another solution where we do not use the video but instead we save and record the past data.

4.1.5 Feasibility in Rehabilitation

Evaluation

To determine the feasibility of the prototypes made as a support tool for motor rehabilitation, we review the clinical application and relevance, financial feasibility, and safety of the system by discussing some of the guidelines prepared by the New Technology Committee (NTC). Although marketability is part of the factors for feasibility mentioned by the NTC, we will skip this part because we are not trying to sell the system that we have developed.

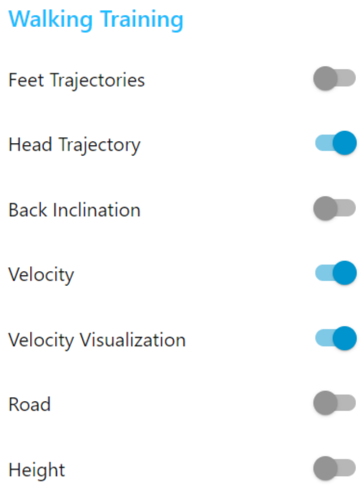
The feedback we got from the therapists when we have done these two demonstrations in the hospitals were positive, and they commented that the potential for this kind of technology to be used in the clinical setting is very exciting. We have already discussed in the Related Works section how supplementing more information about the patient to the therapist can help them understand and assess the current situation, thus making them capable of providing a more informed evaluation. We have also discussed how situated visualizations can facilitate even further their clinical eye. Providing in-situ visualization data to therapists proves to be a valuable asset when they are in the process of conducting rehabilitation and training. The 10 and 15 therapists from two different hospitals who say they want to use our system is evidence enough that this kind of technology is beneficial for medical professionals.

The cost for the development of the force sensors of the chair was really cheap because these components have been built from scratch using cheap electronic components. For the EMG, heartrate and body temperature sensors, hospitals already have these kinds of sensors so we can easily integrate them to our system since we have designed our system to be scalable, with the MQTT broker managing all communications. Sadly, we cannot put a price on the sole pressure matrix as it was a prototype system built in Italy, but this is just an upgrade to the force sensors that we have already developed. The market price for the Kinect camera is around $150, which is also relatively cheap when thinking about the capabilities that it can do. The expensive part of the system is the HoloLens, which costs $3000. The HoloLens is important in providing good AR visualizations that is properly registered into the patient's body, however another alternative we have shown is that all of these AR visualizations can also be experienced with a tablet with ARCore capabilities. Although spatial mapping and registering of virtual objects is not so accurate in the tablet version, it is still a viable option when price is of top concern.

Most of the sensors used in this system are approved for safety as these are already sold publicly in the market. The only part of our system that poses safety concerns is the part we developed ourselves from scratch, which are the force chair sensors. Safety concerns such as loose wirings from the electrical components that were soldered, or the wooden material that was used as casing for these sensors may be unstable and not very suitable for actual use of patient with serious disability, can be somewhat of a problem. In future developments we will improve our system to be accident-proof when used by disabled people, however for now our system works perfectly fine for doing demonstrations.

Insights

The first concept I want to discuss is the possibility of toggling on and off the visualizations. For example, walking training we have already introduced so many different data that can be visualized, so in Figure 4.19 the therapist has the freedom to make visible only the visualizations he wants to see and hides the rest. This is an important point because data overload can be a serious problem, where AR visualizations would instead of aiding, can hinder the understanding of the therapist. To make this process easy to use, especially to people who are new to this kind of technology, we have used the smartphone as means of interface.



Walking Training

Feet Trajectories ○

Head Trajectory ●

Back Inclination ○

Velocity ●

Velocity Visualization ●

Road ○

Height ○

Figure 4.19: Possibilities of Toggling on or off the AR Visualizations in the Smartphone

Another concept is the importance of how the data is presented, that would lead to an easier understanding. As seen in Figure 4.20, we have presented the therapist with the option to also view the stream of data in the form of gauges and charts in the smartphone, aside from viewing it as AR visualizations. The comment we got from therapists is that situated visualizations are better because it gives an overall picture, which would facilitate better understanding compared to the figures found in the smartphone.

Finally, the last concept to discuss is the capability of the therapist to choose the types of visualizations and to have some control over where to place these visualizations. In Figure 4.21, the therapist is given complete freedom to choose what type of visual representations he wants to see; for example seeing text only, graphs, or arrows. One person may understand better using one type of visualization, but another person may not. In this case, the ability to choose is of importance. Another point is choosing the location of where to visualize these virtual objects. For example, the sole visualization located beneath or beside as seen in Figure 4.22a and 4.22b is much easier to understand when the patient is doing the action of walking, as these visualizations provide directional

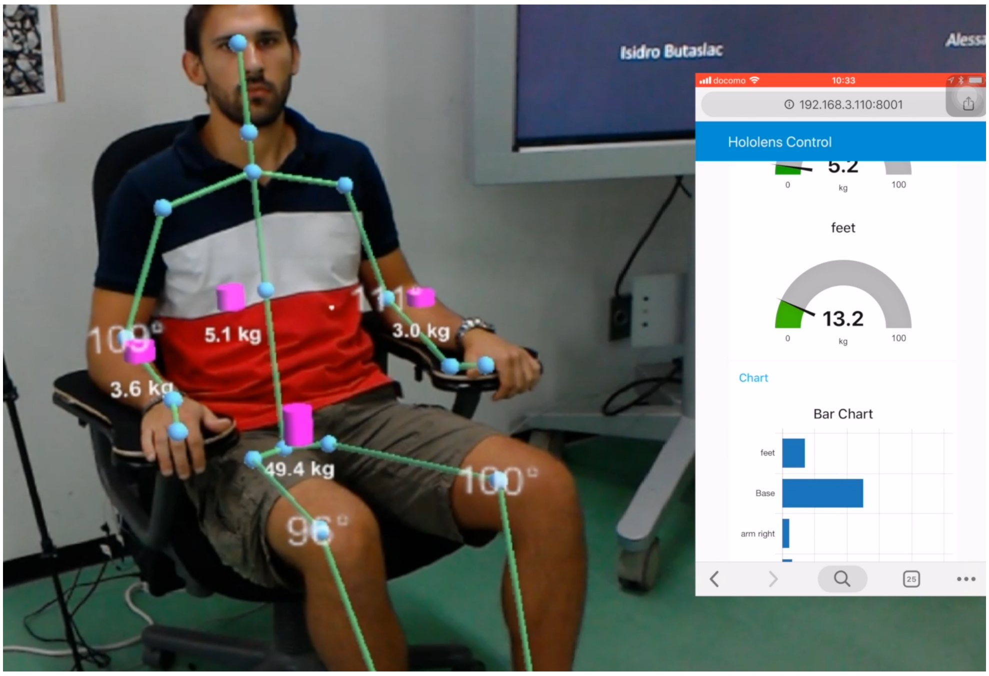


Figure 4.20: Visualizing Gauges and Charts in smartphone compared to AR

cues to where the foot is pointing. Therapists however, prefer the billboard type visualization seen in Figure 4.22c when the patient is stationary and they are the one going around and surveying the patient. This is because the billboard type visualization always faces the sole representation towards the camera/therapist view, maximizing visibility.

4.1.6 Answer to Hypothesis CMO1

This prototype for chair and walking training was made to train the physical and occupational therapist's clinical eye. This was done through a simulated clinical environment, where the therapists see real-time visualizations of invisible current conditions about the patient, such as their skeletons, the forces exerted, their heart and breath rate, and such. The strategy of this system is to assist them in the understanding of patient-related information by placing AR visualizations during their diagnosis. With this, this should help them recognize patterns easier, and increase their skill and confidence of making the correct diagnosis. We have conducted two user studies from two hospitals to use our system, and the feedback

Figure 4.21: The Ability to choose the Type of Visualizations

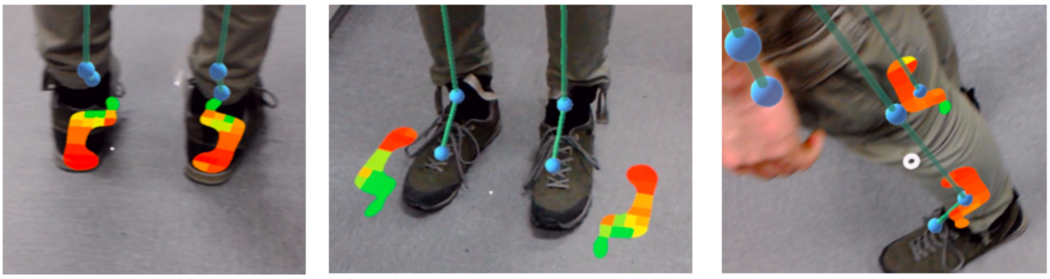

(a) (b) (c)

Figure 4.22: Example on different ways to situate sole visualization: (a) Beneath, (b) Beside, and (c) Billboard type of visualizations.

we got from the therapist's themselves were that after using the system they were able to understand and assess the current situation of the patient better. This provides subjective proof from the experts that their clinical eye has been improved.

4.2 Exproprioception Training System

4.2.1 Introduction

Extended reality applications like Virtual Reality (VR) and Augmented Reality (AR) are mostly known for their use in the fields of gaming and movie industries. However, recently the field of application of these new technologies has become so widespread that clinicians have begun exploring their potential medical applications. Examples of applications where AR is used in the context of healthcare facilities include vein visualization [102], surgical visualization [221], and education [5].

The AR advantage of ensuring that the users do not lose touch with reality since the virtual objects or holograms are displayed over the real environment, is considered one of the most distinctive features of AR compared to VR. In the aforementioned research work, AR technology was applied to Physical Therapy [110], a dynamic profession that focuses on the restoration, maintenance, and promotion of the human body' s optimal physical function. In the current trends, one of the many motivations for developing both AR and VR applications is that they provide the benefit of making the process of therapy much more engaging thanks to the gamification of the tedious and dull tasks [7]. Due to their more lightweight and user-friendly systems, patients are encouraged to do exercises at home, increasing rehabilitation effectiveness. While the development of these new technologies is continuously increasing [58], their usage in hospitals as a support for patients and therapists during standardized training is not yet a common practice. One of the few successful examples where Mixed Reality is used as support for therapists is the Italian project called AUSILIA (Assisted Unit for Simulating Independent Living Activities) [171, 210]. The AUSILIA system is an apartment-wide project fitted with a technical infrastructure that allows the individual to be tracked and supported at the premise during his stay.

In this way, the doctor can monitor and evaluate the patient through the use of Mixed Reality tools. These tools provide him with all the essential data, such as interactions with the environment or physiological parameters about the patient.

The study that was carried out mainly focused on the development of an AR system for biomechanical rehabilitation of the lower limb, specifically on stepping over obstacles; the research work was evaluated and validated from the patient' s point of view.

One possible application of the study would be the implementation of the AR system for the rehabilitation of stroke survivors or people with Parkinson' s disease. Those people have an increased risk of falls during exercise due to their locomotive disabilities, such as impaired balance, decreased stride length, decreased walking speed, compromised ability to step over objects, and decreased endurance [105, 223]. In a regular walking exercise, the user needs to see at the same time, the path to follow together with the obstacles. The user has no choice to quickly adjust his head to and from, increasing the possibility of tripping and falling. The use of virtual objects instead of real obstacles significantly reduces the possibility of tripping. This work proposes a new type of visualization: projecting paths and obstacles not on the floor but in front of the patient.

4.2.2 Related Work

In many different situations, the opportunity to see one's body's visual feedback is very helpful, depending on the specific situation, methods, and approaches that may change. For instance, a dancer uses a full-length mirror to understand the movement of one's body parts. Concerning the clinical practice, also physiotherapists may use mirrors or video-recordings to allow patients to become aware of how they move and correct any compensatory strategies. Thanks to the aforementioned full-length mirrors and offline footage, patients are able to understand better how their body parts are moving, thus enhancing motor learning in both physical and cognitive rehabilitation. However, one issue with this approach is that patients can be distracted by their self-appearance; in fact, mirrored images and clinic backgrounds increase the mental demand and cognitive load, making it hard for them to comprehend the scene.

As far as visual feedback is concerned, different novel techniques and technology are suggested, such as using projection-based AR. In the work of Sekhavat et al., the user was asked to walk on a treadmill and overcome virtual obstacles projected on the surface of a treadmill[195]. The subjects pointed out that the main limitation of the test was that they had to look down on their foot while walking on the treadmill; another limitation of the study was that the height of the obstacles was not considered since there was no information about it on the two-dimensional (2D) projected image. However, more related works have used similar approaches for walking and crossing obstacles. In Binaee et al.[26], the virtual obstacle was generated using a projector. In their implementation, the perception of the obstacle height was simulated through a perspective distortion projected on a ground plane with stereoscopic imagery that changed with the motion-tracked head position. The challenge this work faced was that both the obstacle and the observer's body cast shadows on the occluded portions of the projected image because of the limitations of using a projector.

In our research, the AR system was implemented in an HMD, solving the two problems previously pointed out. Further aspects to consider concerning the effects of using a head-mounted Virtual and Augmented Reality devices on position control and gait biomechanics.

In Chan et al. [45], the effects of VR and AR systems were explored through a simple setup using an HMD. The effect of using HMDs on hip, knee, and ankle joint kinematics, compared to normal treadmill walking, did not reveal significant differences. Indeed is pointed out how the mediolateral boundary of the center of pressure ellipse area was significantly larger in the VR condition than using AR. Instead, the subjects were more able to control their mediolateral way under the AR than under VR condition since AR provide real-time visual exteroceptive information of a person relative to the room environment, which helps in the control of locomotion.

4.2.3 Exproprioception

Over the years, there are many studies towards understanding how the human body is controlled for the purpose of everyday actions such as self-locomotion. At least three broad sources of information are required to maintain or change one' s orientation with respect to the environment. Extending from the formulation of Gibson's work about control of self-motion, Lee distinguishes three kinds of information: exteroceptive, proprioceptive and exproprioceptive [169]. The term exteroception refers to the first source of information about the layout of the environment, and the sense of understanding the position of objects, surfaces, or people around us. The term proprioception, the second source of information, refers to the information about the position and movements of the body parts relative to each other. The combination of these two is called exproprioception, which is defined as the sense of the position of the external objects relative to parts of the body, as illustrated in Figure 4.23. None of these three types of information necessarily involve movement. However, all three can and often do involve movement especially when guiding action.

According to Lee and his colleagues [33], vision typically dominates the other perceptual systems in adult motor control when all sensory inputs are available. Therefore, optical information plays a fundamental role in controlling and coordinating movement within the environment. Gibson [226] coined the term "visual proprioception" to show the purpose of vision to provide information about self-movement by the vestibular system associated with the proprioceptive cues. Using Lee' s terminology, the visual proprioception of Gibson would be referred as visual exproprioception.

In our research, all the interfaces provided to the participants were designed in a way that focuses on giving the user the needed exproprioceptive visual information. The user' s proprioception is improved with the help of a virtual Avatar that enables the user to have a better exproprioception of their own body movements in relation to the environment. Commonly, the use of Avatars is utilized mainly for gaming [216], teleconference [114] and remote collaboration [173]. Moreover, related studies show that the use of first- and third-person point-of-view Avatars generate different results in the expropriation of the user [59, 75, 145]. Although these kinds of applications are usually analyzed in VR environments, the same principles can still be applied in AR environments, with the exception of displaying the Avatar overlaid in the real world.

This design of using exproprioceptive input is supported by a lot of studies, such as the work about lower-limb trajectory during obstacle crossing [215, 165]. Patla et al. also found that during obstacle crossing, direct visual information of the lower limb and the limb’ s position in the environment (visual exproprioception) were important factors for the control of swing limb trajectory. In general, visual exteroceptive information about the environment was used in a feed-forward mode to control locomotion. On the other hand, visual exproprioceptive information about the posture and movement of the lower limb was used in an online mode to control and update the swing phase trajectory [164].

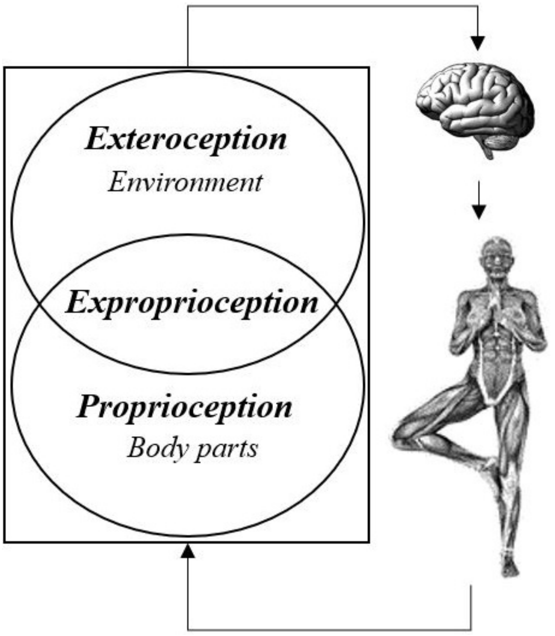
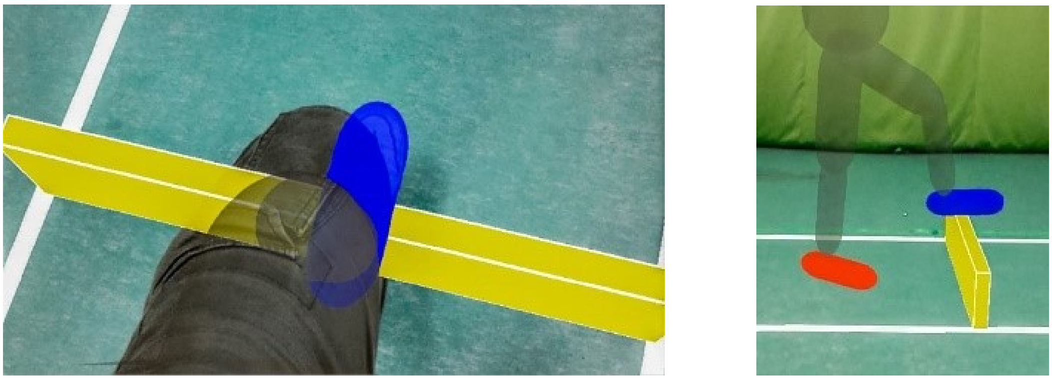


Figure 4.23: Sensory information - Posture and movement loop.

4.2.4 Development of AR Feedback System

To coordinate movement in an virtual environment without real physical references, visual information is needed for planning and ongoing control. Taking into consideration the need for exproprioceptive visual information, different interfaces were designed to increase the visual control of human locomotion and accomplish the task of overcoming a virtual obstacle. These interfaces were developed with Unity3d platform. After doing a preliminary study based on related works and current training practices in overcoming an obstacle, three different interfaces were selected and identified.

Firstly, a Baseline (V1) system which serves as a reference for the other two interfaces was developed. In this case, the virtual obstacle is simulated in the most natural way, very similar to what would be with a real obstacle. However, limitations of the HMD used made dynamic occlusion between virtual and real objects difficult to manage. As a workaround to that problem, this interface uses a virtual avatar overlaid to the real user's body that replicates the movements of the user as close as possible, seen in Figure 4.24a and Figure 4.24b. By taking advantage of a fairly accurate motion capturing method, the illusion of occluding virtual objects from the user's body was achieved.


(a)

(b)

Figure 4.24: The Baseline Interface (V1): HoloLens image from the participant' s point of view (a) and external point of view (b) for the action of walking forward.

Figure 4.24a shows how the occlusion of the leg and foot highlighted by the virtual cues allow for a better understanding of where the subject's foot is with respect to the virtual obstacle. Using this interface, during the exercise the participant have to tilt his head towards the direction of the floor to be able to put his field of view aligned with the obstacle, as seen in Figure 4.25a. In effect, this leads the user to execute a wrong walking posture.

The two other interfaces were developed considering this main point: develop a way to perceive the exproprioception information of the obstacle position while maintaining a proprioception of self through the use of virtual Avatar, all without looking down on his feet to assume correct walking posture. This was realized by simulating a dynamic screen adjusting to the body movements of the user in the AR environment, shown in Figure 4.25b. Taking this into consideration, two

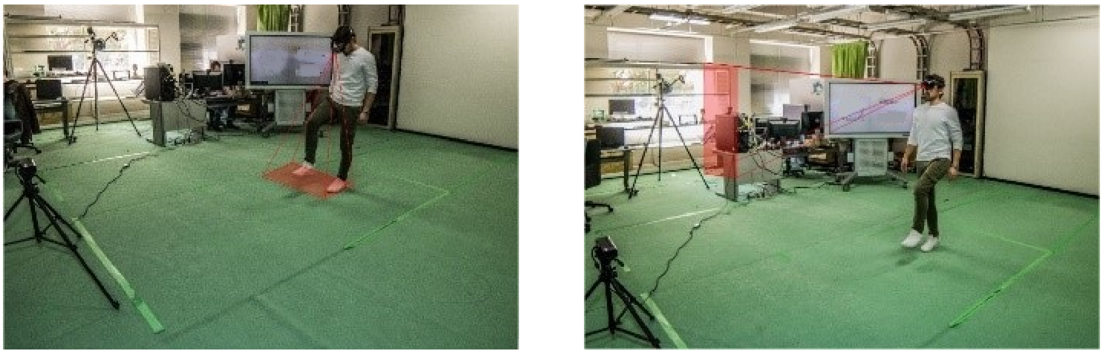

(a) (b)

Figure 4.25: HoloLens field of view in the direction of the floor (a) and in front of the participant (b).

different approaches of displaying the necessary information are determined and evaluated.

For the V2 interface, called *third-person*, the projected image shows what an ideal virtual camera would capture if placed near the user, as shown in Figure 4.27a. The isometric images are projected in real-time to the virtual screen, and the position of the virtual camera changes accordingly to the user walking direction. Thus, it is possible to avoid occlusion of the virtual avatar against the obstacle and always give the most suitable view of the scene. In Figure 4.26 is represented the position of the camera in the four walking conditions. The camera is placed in advance along the walking direction, and once the user starts to walk, the camera follows him on a fixed track along the walking direction, maintaining a fixed distance to him. This leads the user to obtain both the proprioception (i.e., his body position) and the exteroceptive (i.e., the obstacle position with respect to his body) information needed to plan the best trajectories for avoiding the obstacle. The virtual screen is always placed in front of the user body, as represented in Figure 4.27b, and it moves without rotating on ideals tracks accordingly to the walking direction, following the user's body movements and maintaining a constant distance to him. To avoid abrupt movements and makes more comfortable the visualization to the user, both the screen and the camera movements inside the augmented scene were made more smooth, using the *SmoothDamp* function of Unity. The virtual Avatar in Figure 4.27b, superimposed on the real user's body in this case is not use to increase the user's proprioception, the Avatar used to make the user's exproprioception possible is the one displayed in the virtual screen in front of the user.

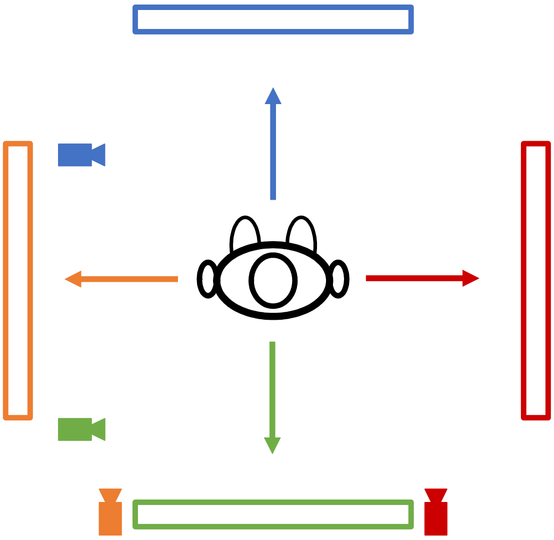
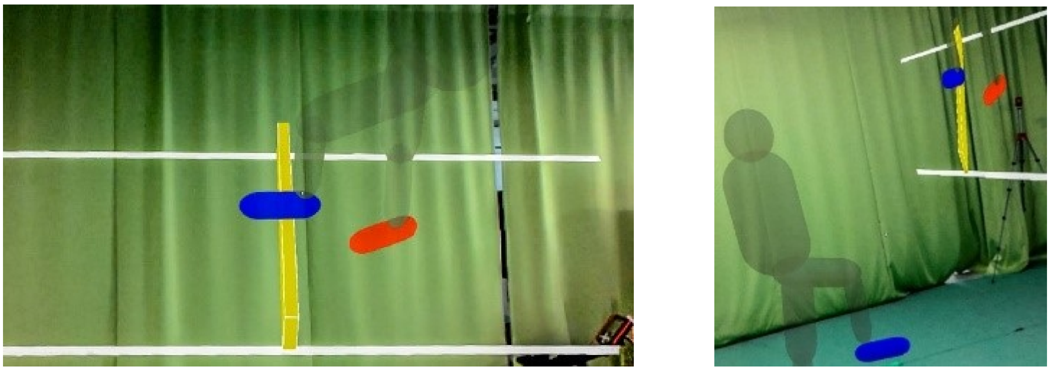


Figure 4.26: Position of the camera with respect to the user (in black) for the V2 interface. Accordingly to the walking direction the position of the camera changes: Front-Blue, Back-Green, Red-Right, Left-Orange.

In the V3 interface, namely *Top-View*, the camera is placed above the user head facing down to the ground. As in the previous case, the position of the camera is placed in advance to the user position but instead be placed in a lateral position, is aligned with the user. In this case, the projected image on the virtual screen is missing of the third dimension information, since it becomes a 2-D representation of the scene, lacking to give the user the full exteroceptive information. To avoid this condition, two sidebars (one for each foot) were placed at the side of the virtual front screen. Each bar increase its height depending on the foot's height with respect to the obstacle, and indeed the bars turn green once the foot passes the height of the obstacle Figure 4.28a. As for the V2 interface also in V3 the Avatar used for user's exproprioception is the one displayed in the virtual screen in front of him Figure 4.28b.



(a) (b)

Figure 4.27: The third-person Interface (V2): HoloLens image from the participant’ s point of view (a) and from an external point of view (b) for an example of walking forward.

(a) (b)

Figure 4.28: The TopView interface (V3): HoloLens image from the participant’ s point of view (a) and from an external point of view (b) for an example of walking forward.

4.2.5 User Study

The aim of this study is to investigate the effectiveness of a new type of interface for stepping over obstacles taking advantage of AR technology using the HMD Microsoft HoloLens.

Participants

Procedures of this study were approved by the Ethics Committee of University. Twenty-seven subjects participated with informed consent. Subjects included 3 females and 24 males from 9 different countries: 14 people with ages between 18-24 years, and 12 people with ages between 25-35 years; average leg length is 0.88 ± 0.07 m; 17 people reported to have experienced general AR, with 10 of them to have specifically used HoloLens. All participants were free from any known neurological or orthopedic disorders, or from any impediments to normal locomotion, as verified by self-report.

Method

Before starting the experiment, the experimenter explained to the participants how the system works and what is needed to accomplish the task. There was also a training session of five minutes where the subjects were free to try and familiarize with the system. This session was added as the scope of this study was not to find which interface is more intuitive for first time use, but rather which interface allows the user to perform better movements while doing repeated exercise over time.

Each session consists of 36 runs (30 seconds each). During the repetition, the subject has to walk for a few meters in a given direction and try to step over the virtual obstacle, as exemplified in Figure 4.29. The task of the subject was to step over the obstacle without touching, having the chest facing always the Kinect. There was no request for the subject to maintain a certain speed to execute the exercise. For each repetition, the obstacle height, distance and direction were randomized from predefined sets: three different heights were defined proportionally based on the length of the subject's leg; placement of obstacle from the subject was randomized between 1.5 ± 0.5 m; subjects are able to move in any of the four cardinal directions. However, the thickness of the obstacle was fixed at 5 cm. Participants have worn the Microsoft HoloLens across all the sessions.

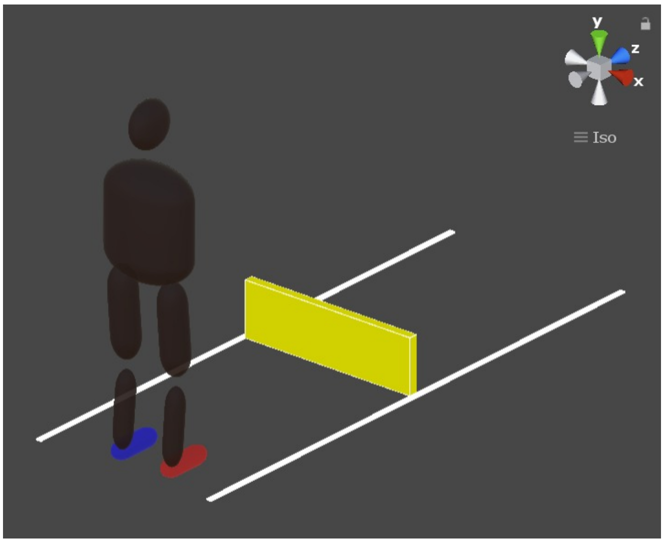


Figure 4.29: Participant Avatar with virtual obstacle.

The exercises were based on the combination of standardized exercise practice for gait rehabilitation and obstacle avoidance. The decision to limit the movable direction to the four cardinal directions were based on the series of exercises demonstrated by the famous Physical Therapist's Brad Heineck and Bob Schrupp. These exercises were chosen and decided together with the hospital's therapists. After completing the session, each participant was asked to fill out two questionnaires. The first one was about the participant's personal and physical information, together with their knowledge about AR. The second one is the NASA-TLX, which is a subjective and multidimensional assessment tool that rates the perceived workload of the given task.

Factorial Design

A full factorial experiment was designed to study the effects of each factor on the response variable and the effects of the interactions between each factor. Three factors were considered for the design of the experiment; one following between-subjects design, and two following within-subjects design as listed below. Between-subjects design factor:

- Factor **A** = Interface.

Within-subjects design factors:

- Factor **B** = Obstacle height.
- Factor **C** = Direction of walk.

Each condition is repeated thrice, totaling 36 runs for each participant (4 walking directions x 3 obstacle height x 3 repetitions = 36 runs). As factor *A* followed a between-subjects design, the 27 participants were randomly and equally divided into three groups assigned for each condition. Within-subjects design for factor *B* had three levels (small-medium-large obstacle height) while factor *C* had four levels (front-back-right-left walks). The height of the obstacles were proportional to the subject's leg length (LL) which were measured prior to the session from the floor to the location of hip joint, changing between 3 possible value: 0.15LL as small, 0.25LL as medium and 0.35LL as large (in the plots as s-m-l). To satisfy the statistical requirements of the independence of observations, the matrix for the final design was generated by randomizing the experiment order with the use of R code.

Data Processing

The evaluation of the user during each repetition is related to his body position and gait biomechanics. The body parts that mostly affect these latter parameters are associated to the user's head and feet.

Accordingly, the data collected for the evaluation of the user study are related to:

- Head position, velocity and inclinations.
- Feet positions.
- Collision between feet and obstacle.

Head movements and speed were calculated and saved from HoloLens. The feet trajectories were derived from the 3D joint positions of FootRight and FootLeft captured by the Kinect. As a means to calculate the minimum distance between the mesh surface and the obstacle, a virtual mesh capsule overlaid the tracked foot as a representation of what the system sees as foot boundaries. The foot trajectory was also used in the V3 interface as input data for the sidebars to inform the user about the height of the foot with respect to the obstacle. The collision between foot and obstacle was computed inside Unity. A flag is raised all the times that the mesh collider script detects a collision between the foot and obstacle game objects.

After each run, raw data was saved and plotted with the help of a MATLAB script. This was done in order to check the correctness of the collected data. Figure 4.30 shows an example of a test, plotted in MATLAB environment, about lateral direction walk (left) using the V1 interface. In case of collision between the subjects' foot and the obstacle, the user was informed at the end of each run.

The four graphs in Fig. 4.30 describe the data that was collected during the experiment. The upper left graph shows the y-z trajectories (side view) of the right foot, left foot, and head. The upper right graph shows the x-z trajectories (top view) of the same joints. The bottom left graph shows the instantaneous velocity of the head, while the bottom right graph shows the forward inclination of the head. Data collection was automatically started and ended once the subject entered/left within 600 mm range from the obstacle. It did not matter which foot came first in stepping over the obstacle, as the subjects were asked to perform the task in the most natural way possible.

Statistical Analysis

Unless differently specified, a three-way mixed repeated ANOVA was conducted to compare means, with an alpha set at 0.05. In the Analysis of Variance the interface type factor A was treated as between-subjects factors, while obstacle height factor B and walking direction C as within-subject factors. Significant outliers, above Q3+3x the interquartile range (IQR) or below Q1-3xIQR were removed from the cell design. Since the large number of samples sizes, the normality distribution of the data has been tested by visual inspection of the Quantile-Quantile (QQ) plot, and BoxCox method was applied to transform the outcome variable to correct for the unequal variances. A post-hoc analysis was done with Tukey' s test. For each participant, aggregated scores were computed for the questionnaires (NASA-TLX and general questions: mean values of all items). All the statistical analysis was performed in RStudio v.1.2.5033 (Boston, MA, USA), an integrated development environment for R programming language.

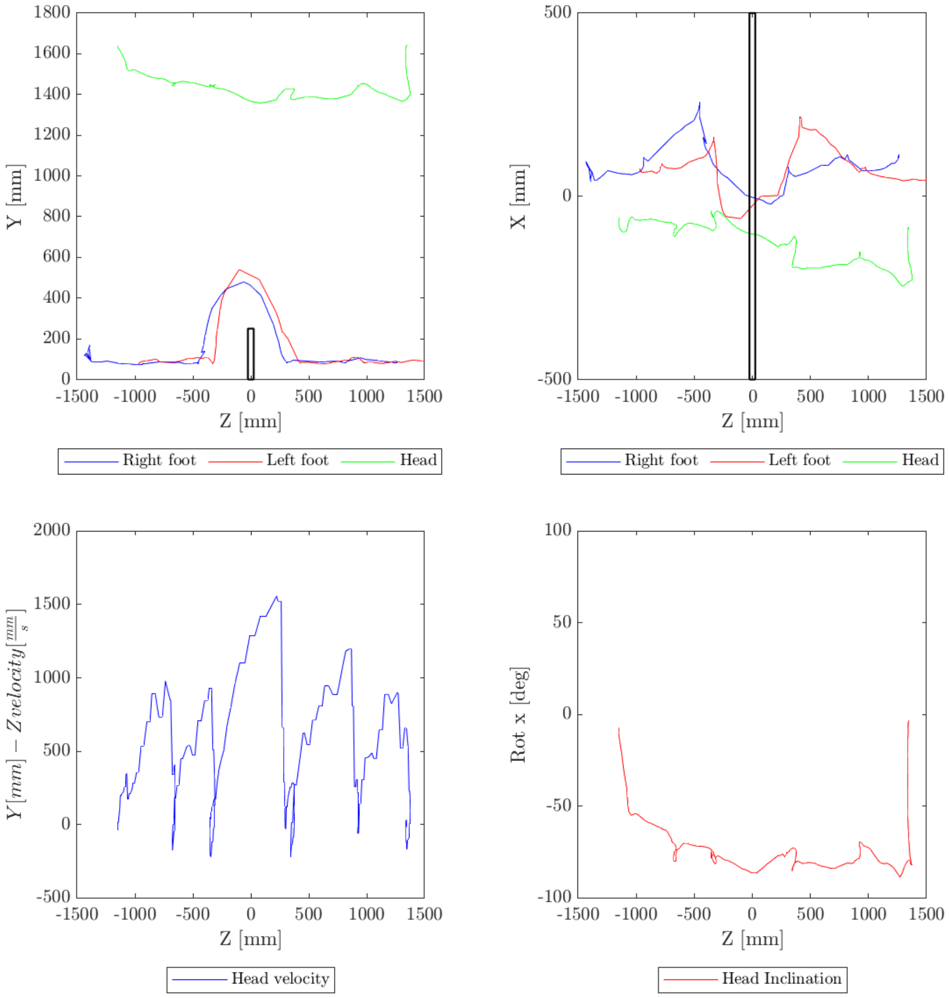


Figure 4.30: Acquired parameters from Baseline test left direction.

Table 4.1: Statistical tests for head kinematics parameters.

	Head inclination		**Head Y positioning**		**Head speed**	
	F value	p value	F value	p value	F value	p value
A	F(2,954) = 20225.621	<2.20*e*-16	F(2,955) = 304.059	<2.20*e*-16	F(2,955) = 348.593	<2.20*e*-16
B	F(2,954) = 6.904	1.00*e*-03	F(2,955) = 16.853	6.43*e*-08	F(2,955) = 15.011	3.81*e*-07
C	F(3,954) = 3.091	2.60*e*-02	F(3,955) = 17.526	4.44*e*-11	F(3,955) = 20.153	1.14*e*-12
A:B	F(4,954) = 8.831	5.30*e*-07	— —	—	— —	—
A:C	F(6,954) = 4.422	2.02*e*-04	F(6,955) = 3.81	9.28*e*-04	— —	—
B:C	— —	—	— —	—	— —	—

Results

The statistical analysis results are reported in Table 4.1 for the Head kinematics, and in Table 4.2 for the feet one.

The mean of the *head inclination* around the x-axis in the *V1* condition is much higher (mean ± standard deviation for all the values: 71.29±7.46°) compared to *V2* (-2.82± 2.16°) and *V3* (-1.88±1.92°), resulting in a strong skewness of the data (Figure 4.31). A square root transformation was applied to the full linear model ($\sqrt{|Y|} \sim A * B * C$) in order to normalize the residual distribution and then the no significant parameters were removed. There is a strong main effect of *A* factor and a slighter effect of both *B* and *C*. Significant interaction is present between *A:B* and between *A:C*.
Post-hoc test results in a strong group' s means difference for factor *A* among all the three possible levels combinations: *V2-V1* ($p =< 2.2e-16$), *V3-V1* ($p =< 2.2e-16$), *V3-V2* ($p = 5.37e-11$); for *B*: *s-l* ($p = 6.27e-04$); for *C*: *Dir3-Dir2* ($p = 0.0356$).

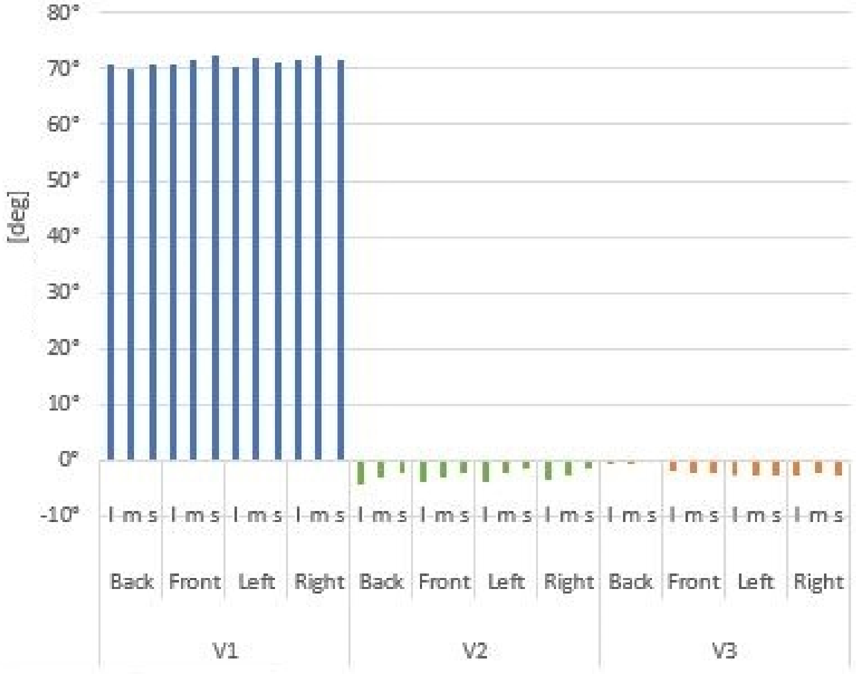


Figure 4.31: Mean head rotation around x-axis with s-m-l obstacle's heights, back-front-left-right walking directions and V1-V2-V3 interfaces.

Table 4.2: Statistical tests for first and second foot.

	First Foot			**Second Foot**		
				Collisions		
	F value		p value	F value		p value
A	—	—	—	F(2,960) =	8.913	1.46*e*-04
B	—	—	—		—	—
C	—	—	—	F(3,960) =	5.955	5.00*e*-04
A:B	—	—	—	—	—	—
A:C	F(6,960) =	8.913	9.00*e*-03	F(6,960) =	3.018	6.00*e*-03
B:C	—	—	—	—	—	—
				Clearance		
	F value		p value	F value		p value
A	F(2,952) =	4.012	1.80*e*-02	F(2,943) =	6.02	3.00*e*-03
B	F(2,952) =	25.099	2.38*e*-11	F(2,943) =	4.05	1.80*e*-02
C	—	—	—	F(3,943) =	62.617	2.20*e*-16
A:B	F(4,952) =	8.045	2.22*e*-06	F(4,943) =	10.337	3.39*e*-08
A:C	F(9,952) =	4.819	2.59*e*-06	F(6,943) =	8.47	5.64*e*-09
B:C	—	—	—	F(6,943) =	2.725	1.20*e*-02
				Peak position Z		
	F value		p value	F value		p value
A	F(2,949) =	51.364	<2.20*e*-16	F(2,949) =	24.37	6.56*e*-11
B	—	—	—	F(2,949) =	15.082	3.56*e*-07
C	F(3,949) =	111.59	<2.20*e*-16	F(3,949) =	19.622	<2.20*e*-16
A:B	—	—	—	—	—	—
A:C	F(6,949) =	5.725	7.31*e*-06	F(6,949) =	6.458	1.10*e*-06
B:C	—	—	—	F(6,949) =	3.862	8.15*e*-04

Head rotation is reflected in the *lowering of the head* during the exercise. Compared to the maximum height reached by the subject, body tilt led to a general decrease of the head height. Much higher for the subject with *V1* interface (-48.6±18.0mm) with respect to *V2* (-31.1±14mm)) and *V3* (-25.8±11.2mm). The logarithmic transformation is used to stabilize the variance ($\log Y \sim A * B * C$). Each factor presents a main effect, stronger for *A* and *C*, slightly lower for *B*, and a strong evidence of interaction between *A:C* were highlighted. A comparison of residual vs run order revealed a trend of the values even among the different interfaces as shown in Figure 4.32, with a positive slope of the *V1*

($y = 1.10x + 314.88$) and *V2* ($y = 1.70x + 265.66$) trend line, and with a negative slope for *V3* ($y = -0.39x + 212.51$).

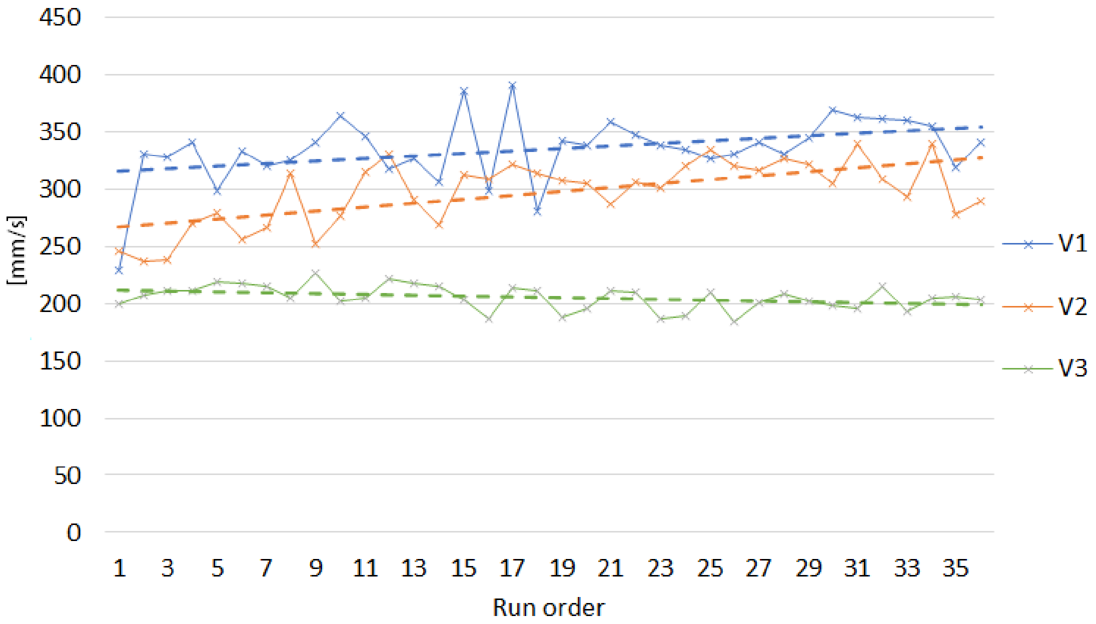


Figure 4.32: Trend of mean velocity during each session, for each interfaces.

Another analysed output parameter was the participant's *average speed* during their repetitions. The average values of the head inclinations for interface type in decreasing order are *V1* (335.25±92.88mm/s), *V2* (297.13±101.26mm/s) and *V3* (205.32±66.35mm/s), Figure 4.33. The model is described with a logarithmic transformation. The main effect is given by single factors. It is greater for A than C and B. Between levels of factor A the averages that are significantly different are: *V2*-*V1* ($p = 2.56e-13$), *V3*-*V1* ($p = 2.12e-13$), *V3*-*V2* ($p = 2.12e-13$); for B: *s*-*l* ($p = 1.91\text{e}-08$), *s*-*m* ($p = 4.99\text{e}-4$) and for C: *Back*-*Front* ($p = 1.12e-12$), *Right*-*Front* ($p = 5.18e-06$), *Left*-*Front* ($p = 4.73e-06$).

Feet analysis is kept separate between the first and second foot that overcomes the obstacle (Table 4.2). For the *number of collisions*, Figure 4.34, a linear model was used to describe the relations between factors. For the first foot collisions, the means among the individual factors are not relevant, but instead, a weak relation is given by the interaction *A*:*C*. For the second foot instead, there is a significant effect of singles factors A and C. It is also significant their interaction

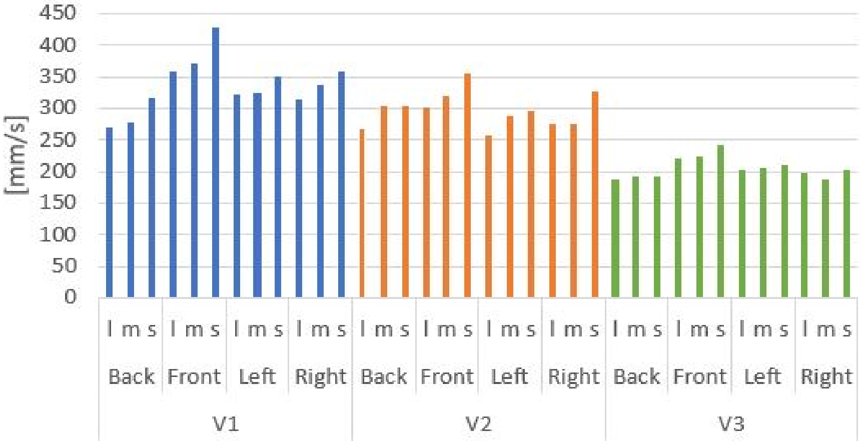


Figure 4.33: Mean speed.

A:C. Between levels of *A* there is a significant difference in *V3-V1* ($p = 0.003$) and *V3-V2* ($p = 0.00026$). Instead for *C* in *Right-Front* ($p = 0.0056$) and *Left-Front* ($p = 0.0005$).

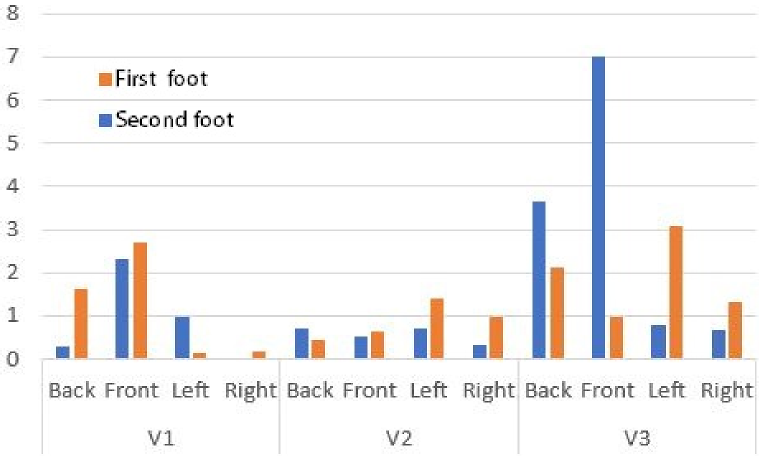


Figure 4.34: Feet collisions with obstacle.

The *clearance* between the position of the foot that overcome the obstacle and its height has been evaluated. Due to the small thickness of the obstacle, 5 cm, the average height of the foot was used by computing the mean between the maximum and minimum values above the obstacle. The mean among the different conditions of factor *A* are almost equal: *V1* (236.713± 87.761mm), *V2* (238.400±

91.275mm), *V3* (223.963± 91.913mm). The same for the second foot with *V1* (198.713± 86.410mm), *V2* (190.646± 129.102mm), *V3* (219.558± 115.502mm), Figure 4.35. For the first foot, a linear model was applied. The main single factors' effect is given by B and a low influence from A. A significant interaction there is also between $A{:}B$ and between $A{:}C$. The levels of A significantly different are *V3*-*V2* ($p = 0.0250$). For the second foot, a power transformation of λ=0.67 was applied to the linear model. The main significant effect is C, less effect A and B. About the interactions, the greater effects are $A{:}B$ and $A{:}C$ than $B{:}C$. For each factor the levels significantly different are *V3*-*V2* ($p = 0.002$) for A,*Back-Front*, *Left-Front*, *Right-Front* with ($p = 0$) for C and *m-l* ($p = 0.22$), *s-l* ($p = 0.013$), *s-m* ($p = 0.47$) for B.

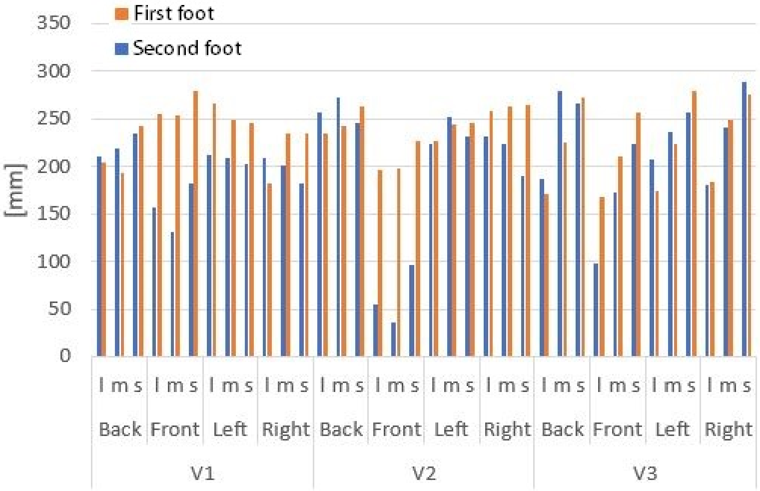


Figure 4.35: Clearance between the feet and obstacle on the step over the obstacle.

The last significative parameter is *the horizontal distance of the foot peak respect to the obstacle*, Figure 4.36. For both feet models, it was applied a linear model. The main factors' effects for the first foot are given by C, less by A, and by their interaction $A{:}C$. For the second foot, the significant effect is C followed by A, B, and their significant interaction $A{:}C$ and $B{:}C$. The averages of these distances for each foot respect to A have standard deviation values very big because the peak position changes mainly with C and not with A.

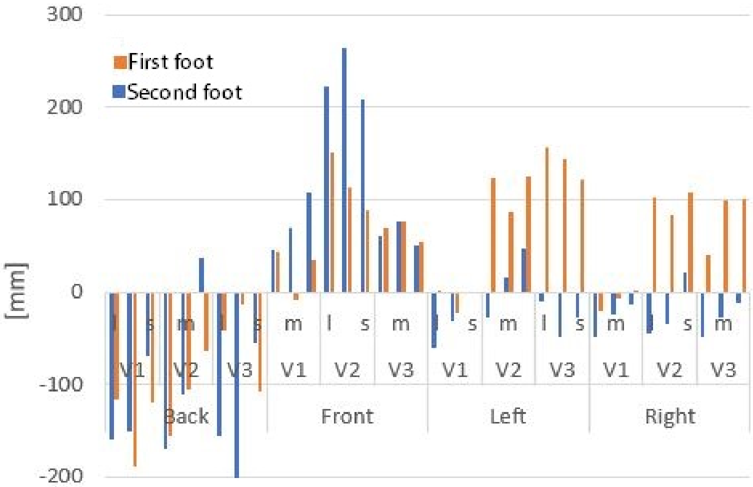


Figure 4.36: Peaks horizontal distances feet-obstacle.

4.2.6 Discussion

The purpose of this study is to evaluate the different approaches of AR interfaces and to propose a new way of providing information to the user when performing rehabilitation exercises, specifically for stepping over obstacles. However, the problem of putting into practice this kind of rehabilitation is that the movements and rotations of the patients' head are limited. When looking at objects placed on the ground, the risk of injuries due to head-related movements becomes very high. These type of movements are further amplified in the AR scenario as HMDs have a really small field of views (e.g. HoloLens having 34" diagonal in 16:9). Our data reveal that the V1 interface (the interface similar to the current approach) shows a larger mean head rotation around the x-axis as compared to V2 and V3 interface, as indicated in Figure 4.31. As proof of posture discomfort, seven participants reported neck pain in the subjective questionnaire when using the V1 interface. The pain, as they write, is greater when they have to walk backwards because besides tilting the head they have to rotate the neck to see the obstacle keeping the frontal torso.

The average speed at which the task is executed is proportional to the confidence of the user when performing the task. Results show that the factors are independent with each other, as seen in Figure 4.37. As expected, people overcame small obstacles faster and walking in a frontward direction provided the best results.

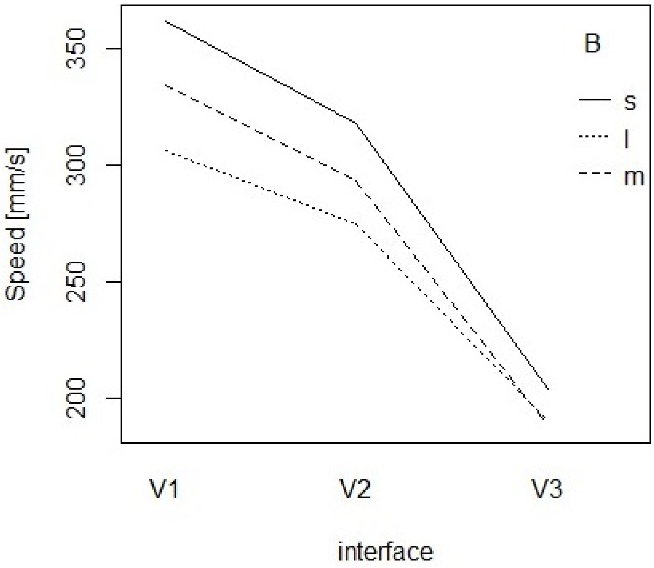


Figure 4.37: Interface vs. Obstacle height

The variable that most influenced the speed is the interface type. This is evident in the graph of Figure 4.37 as V3 interface shows a sudden drop in speed as compared to V1 and V2 interface. The reason for this is that the V3 interface splits the dimensional information into two, obstacle position and height. As the task is split into two subtasks, execution would inevitably be slower. This is also apparent in Figure 4.38 as participants using the V3 interface executed the motion in an unnatural manner; firstly, they approached near the obstacle, stopped for a moment, rose the first foot in an upwards direction, stepped over the obstacle, and then finally repeated the same actions with the second foot. With the V1 and V2 interface, however, the participants overcame the obstacle in a more natural manner following a parabolic trajectory.

By analyzing the trends in run order for each interface, V3 follows a negative slope meanwhile V1 and V2 follow a positive slope as seen in Figure 4.32. Users of the V3 interface indicate no improvements over time. Moreover, V2 shows a steeper slope as compared to V1. This implies that training has stronger effects on the V2 interface. Users of V2 interface may find it difficult to understand at first, but as they get used to, it becomes more proficient. Another important statistic to look at is the number of times the feet collide with the obstacle, seen in Figure 4.34. Even though there is no significant difference when looking at the first foot, there are significantly less collisions for the V2 interface when looking at the second foot.

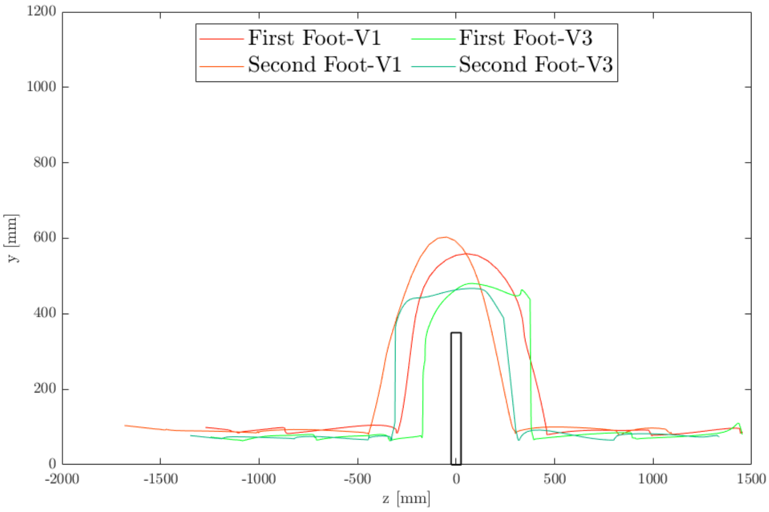


Figure 4.38: Feet trajectories with the V1 and V3 interface. For the V1 condition the trajectory is smoother and the lead foot start to rise before the one with the V3.

When considering the vertical distance between the foot and the obstacle, obstacle height and walking direction affect the first and second foot but not the interface type. Moreover, there is no significant difference among the interfaces implying similar perception in the obstacle difficulty, as well as among each factor. However, the V2 interface has less vertical distance when looking at the second foot going in the frontward direction seen in Figure 4.35. This suggests that the user for the V2 interface has a better understanding of the obstacle height.

The walking direction affects the horizontal positions of the maximum peaks as compared to the interface type. Figure 4.36 may show the location at which the peak occurs, but it does not describe which interface type is better. Peaks occur before the obstacle for both feet in the backward direction, meanwhile, they occur after the obstacle in the frontwards direction. For both cases regarding the lateral directions (with exception of the V1 interface), peaks occur after the obstacle for the first foot and near the obstacle for the second foot. The descriptions above about the feet trajectories characterize the general behavior of how the participants step over the obstacle. However, this only portraits the body gait biomechanics of stepping over obstacles rather than describing a concise evaluation of the different interfaces.

Aside from an objective evaluation, a subjective one was also done using the NASA-TLX questionnaire. Users evaluated the Temporal Demand and Expected Performance as similar across each interface type. The Mental Demand and Effort to complete the task is rated high for the V3 interface. This happens because with V3 users have to combine two different types of information, namely horizontal and vertical positions of the foot with respect to the obstacle, consequently bringing about a clear segmentation of the movement. This division of the trajectory leads to a slower speed in accomplishing the task. The Frustration and Physical Demand were rated higher for V1 as compared to V2 and V3 due to neck pain caused by the high head inclination.

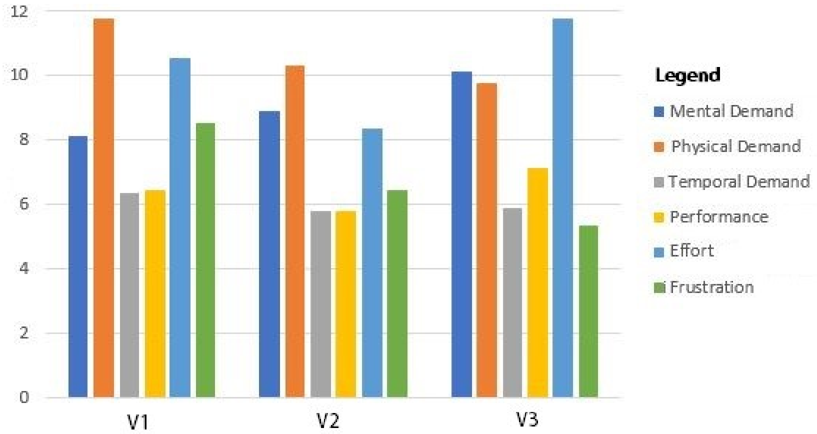


Figure 4.39: NASA-TLX parapeters avarage results.

4.2.7 Answer to Hypothesis CMO2

This prototype was designed to train the patient to train their motor skills and their sense of exproprioception. The patients were provided real-time feedback of their motion, and different ways of visualizations were presented to assist them in their walking training. Overall, the third-person point of view interface V2 proved to be the best interface. It not only enabled the user to walk in a fast and natural manner similar to the Baseline interface, but it also enabled the user to be much faster and precise as compared to the TopView interface. People using the third-person point of view interface were more capable of understanding obstacle heights and planning their next steps. Positioning the visualization in front of the user puts him in a situation to look forward while performing the task, resulting to reduced head movements and better posture when walking as compared to the Baseline interface. Finally, results from the subjective questionnaires indicated no signs of difficulties for the users when using the HMD AR interface, even for users without prior AR knowledge and experience. With both these quantitative and qualitative data, it shows that the system was able to train the locomotion skill of an "alternative" patient as they would perform in the actual exercise.

4.3 Shared AR Activities of Daily Living System

4.3.1 Introduction

Nowadays, with the development of innovative technologies that are increasingly high-performing and affordable, the possibility of using them in many applications is growing day by day. Among these technologies, the field of Augmented Reality (AR) supports a wide range of innovative use cases, especially for real-time tasks [140, 39, 139].

This paper analyzed an AR framework for the metrological assessment of activities of daily living (ADLs)[118] for the occupational therapist/caregiver and patient/fragile end-user. The field of interest of our use case is rehabilitation but can be extended to an immersive AR scenario to support fragile end-users at home by their caregivers.

Acquired brain injuries such as stroke, traumatic brain injury, or brain tumor impact people's functions from mild to very severe limitation and thus their independence during ADLs. Limitations can include motor/sensory deficits, cognitive/perceptual deficits, behavioral deficits, or visual deficits. The assessment of patients' skills is based on a non-standardized approach that quantifies patients' performance using the following qualifiers: safety, efficiency, effort, and independence [69]. However, occupational therapists administer standardized assessments with manuals such as the Assessment of Motor and Process Skills (AMPS) [46] which makes them more reliable and objective but never perfect. The AMPS assessment also requires an ongoing education courses to learn how to administer it. Therefore, an assessment using these modalities is influenced by the experience of the clinician and is therefore error-prone for less experienced therapist.

To overcome this limitation, it is necessary to use innovative technologies and advanced measurement systems to obtain an objective assessment of the efficiency of treatment and training strategies. The use of AR as assistive technology in clinical settings is widely discussed in literature also for ADLs support [181, 230, 154]. In most cases, the choice to use AR technologies instead of Augmented Virtuality technologies is due to the fact that the performance of subjects in AR compared to Virtual Reality (VR) is higher due to the extraneous hand-eye coordination that exists in VR while it is eliminated in AR [124]. Moreover, in our prototype, the patient, while viewing virtual information in AR, is asked to handle real objects in order to guarantee not only the perception of the entire environment but also their weight.

The main innovation of the proposed framework lies in the increment of the clinical eye [44] in a shared real/virtual environment that enables the evaluation/support in AR contexts for future ADLs scenarios by increasing empathy between actors [174]. The developed prototype enhances the therapist with all the patient's multidimensional data in AR and the patient himself by giving him the possibility to interact with both virtual cues and real objects during ADLs. From the proposed framework we developed a specific application for occupational therapists where we combined a system that leverages AR technologies with a robust and accurate measurement system based on computer vision algorithms to ensure the high metrological quality of the assessment.

4.3.2 ADL Framework

The prototype has been developed in the laboratory of Measurement, Instrumentation and Robotics of the University of Trento and set up inside the home automation apartment AUSILIA (Assisted Unit for the Simulation of Independent Activities) [85] at the rehabilitation hospital Villa Rosa in Pergine Valsugana

(Italy). Fig. 4.40 shows the framework tools used during the ADL assessment.

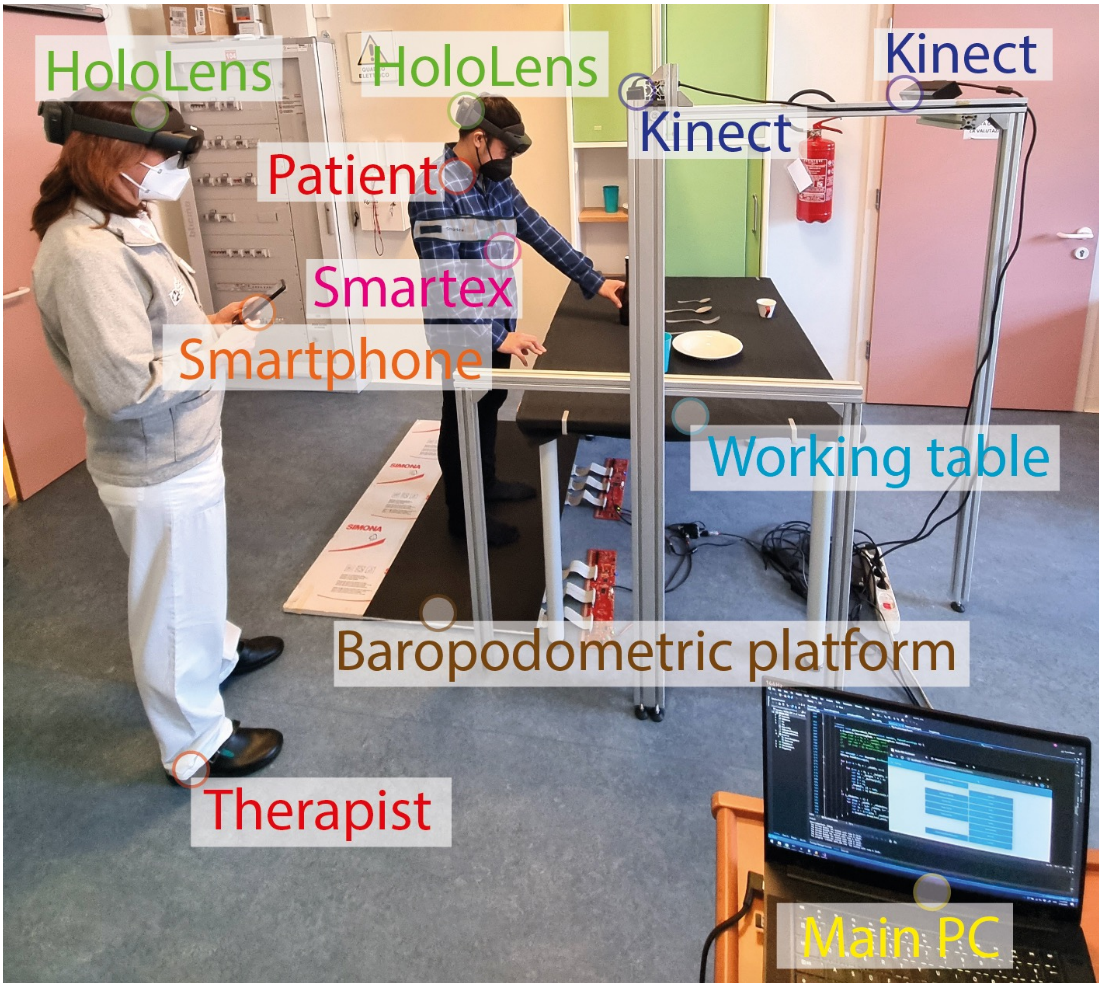


Figure 4.40: Framework setup in the AUSILIA apartment.

Visualization devices

The devices used to display the information are:

- Two Microsoft HoloLens 2 head-mounted displays to show AR cues to both therapist and patient.
- A handheld device such as a smartphone facilitates the management and the interactivity with the information to be displayed by the therapist.

Distributed Measurement system

The measurement system consists of:

- Two time-of-flight depth cameras such as the Kinect v2, one above the table used only to capture an RGB image (1920 × 1080) for the computer vision based-algorithm, and one in front of the patient used to obtain the position of the body joints in 3D space [157].
- A wearable band system designed by Smartex company to continuously monitor a cluster of physiological parameters. In particular, the system is able to simultaneously acquire the electrocardiographic signal (ECG) and the respiratory signal of the patient.
- The baropodometric platform used for static and dynamic non-invasive pressure measurements and body stability analysis is a custom model from the FreeMed family manufactured by Sensor Medica. The platform of size 56x120 cm is composed of two units whose sum leads to 6000 resistive sensors coated in 24k gold with frequency acquisition up to 400 Hz.
- The main PC, where all raw sensor data are processed, stored, and sent.

Software development and communication protocols

The control interface for handheld devices such as smartphones was developed with the programming tool Node-RED.

All the devices such as HoloLens, a smartphone, and the main computer are connected over the same LAN. Communication of data that deals with logic control (i.e. interface buttons, switches, etc.) use the MQTT (MQ Telemetry Transport) protocol running over TCP/IP because of its reliability and lightweight. On the other hand, data that deal with a large and continuous stream of data (i.e. platform data, Kinect data) are broadcasted over standard UDP (User Datagram Protocol).

The raw ECG and respiratory signal acquired by the Smartex band were filtered and analyzed via Bluetooth Low Energy (BLE) in the main PC and then sent their average values in a 2-second time window to HoloLens.

The Kinect is connected to the main PC via USB. Only the joints of the upper half of the body are considered. These data are then broadcasted via UDP to

Hololens at a rate of 30 Hz. The FreeMed platform is also connected to the main PC via USB and broadcasted via UDP to Hololens. Fig. 4.41 shows the data processing pipeline.

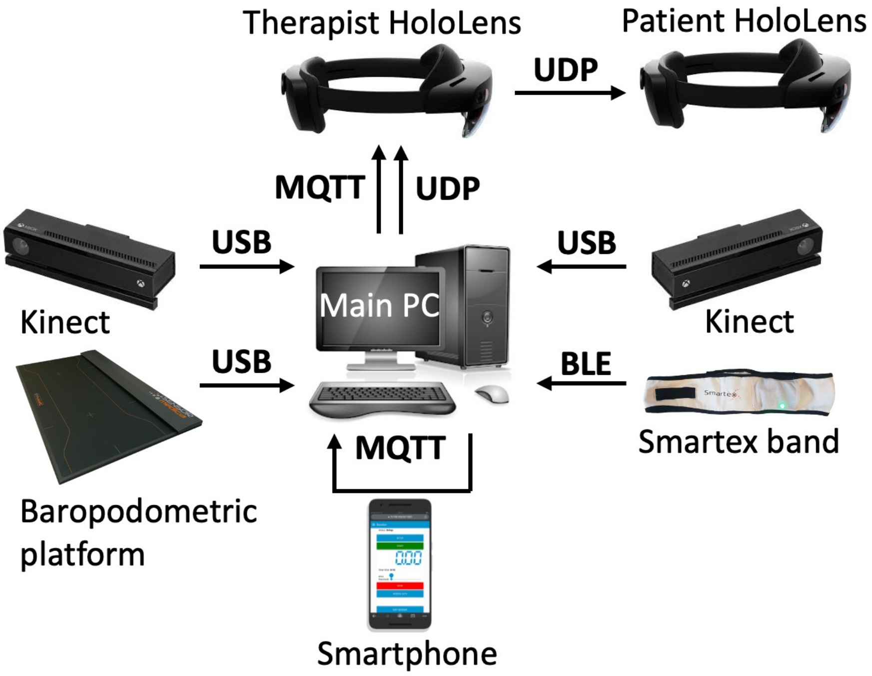


Figure 4.41: Data processing pipeline

Extrinsic parameters calibration

An initial calibration (Fig. 4.42) is required for the Kinect used to track patient kinematics and the therapist's HoloLens to operate in the same reference system. To do this, a marker with enough detectable feature points was used, during the set-up phase, to derive a transformation matrix from Kinect camera coordinates to marker coordinates. This calibration process is repeated until an acceptable reprojection error is reached. For the HoloLens part, Vuforia SDK handles all the image target tracking. A spatial anchor is saved in the HoloLens using the same marker used for Kinect calibration. In this way both the Kinect and HoloLens are able to operate in the same reference system. Once calibrated, the marker can be removed anytime. Additional spatial anchors are saved in the therapist's HoloLens to define the reference systems of the working table and baropodometric platform. On the patient's HoloLens, however, only the spatial anchor related to the working table reference system is saved to operate in the same reference system as the therapist's HoloLens.

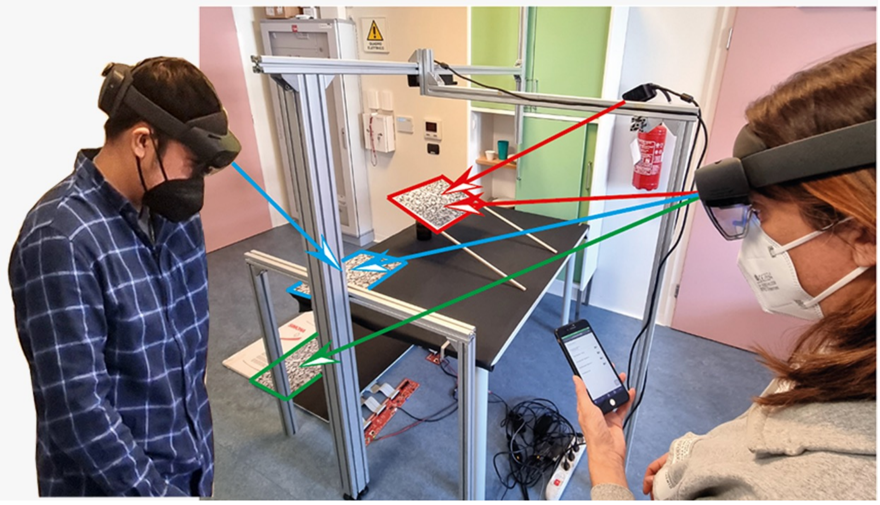


Figure 4.42: Spatial Anchors setting: the red image target is used by the therapist's HoloLens and the Kinect to operate in the same reference system; the blue target is used by the therapist's and the patient's Hololens to have the same reference system of the working plane; the green target is used only by the therapist's HoloLens to localize the baropodometric platform in space.

4.3.3 Specific ADL in the Kitchen Environment

The therapist assessing the patient during the instrumental ADL of setting a table is aided by a shared scenario in AR that can enhance his clinical assessment in an immersive and engaging way for the patient. The evaluation process involves the following steps:

1. The therapist, wearing a head-mounted Microsoft HoloLens 2, sets the table with virtual objects. The number and type of objects can be selected from a graphical interface designed for handheld devices. This allows therapists to flexibly increase or decrease the complexity of the setup based on the type of patient they are assessing.

2. Once finished, the patient wearing another HoloLens 2 can view the virtual environment previously set up by the therapist and is asked to try to match the virtual objects with the real ones.

3. After that, the therapist by pressing a button on the smartphone starts the process of estimating the position and angle error of the real objects compared to the virtual ones that will appear in AR next to each object with numbers following the therapist's gaze in Fig. 4.43. The color of the numbers displayed (green-yellow-red) depends on the tolerance and thus the threshold of error acceptability that the therapist sets on the smartphone.

4. Another panel in AR summarizes the average angles and the average distances between the barycenters of the virtual and real models. In addition, an estimation of the total time to perform the task is shown Fig. 4.43.

5. The therapist can decide with the smartphone whether to display additional information about the patient in AR during the exercise session, such as the reconstruction of the patient's kinematics and angles between the limbs, the load distribution of the legs, and his physiological parameters.

6. At the end of each session, the therapist can decide to save all captured data to a text file.

4.3.4 Algorithm for Object Detection

To identify and locate real objects of interest placed on a table by a user, an algorithm was developed in a MATLAB environment. This algorithm, following the processing of an RGB image, can detect and identify such objects.

Items to identify include polished stainless steel cutlery. To overcome the problem of reflections on their surface that could affect the result of the algorithm, they were treated with a sandblasting process that made them opaque and unaffected by the direction of light.

The algorithm can be divided into the following steps:

Figure 4.43: Errors visualization in AR via therapist's HoloLens 2.

Segmentation and Localization

First, the RGB image captured by the Kinect fixed on top of the table was captured and processed. In particular, the image was processed in the following order:

1. The images acquired with the Kinect of the empty table and the one of the same table covered with the real objects were converted to grayscale.
2. Images were cropped to take into account only the table region of interest (ROI).
3. A background subtraction was then applied by subtracting each pixel of the previous two images and then converting the result to a binary image by selecting an appropriate threshold.
4. The resulting mask was applied to the initial RGB image of the table set, and a color-based threshold was applied to remove object shadows from the image.

5. A flood-fill operation was then performed on the hole pixels of the closed regions [203].
6. A boundary label was applied to the filtered image [92].
7. A threshold on the minimum number of pixels over the area of each labeled object was applied to remove noisy regions.
8. The outer boundaries of each object were then traced [82], Fig. 4.44.
9. Each object was localized in position by taking the mean of its boundary coordinates and in rotation from the Singular Value Decomposition (SVD).
10. At the end, a mask with each object-centered and aligned was stored.

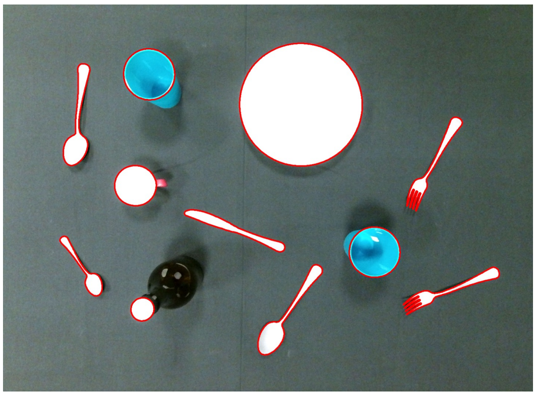


Figure 4.44: Initial RGB cropped image with red outer boundaries of all found elements.

Identification

The result of the previous image processing is a binary image of each segmented, center-aligned object in the initial image. To identify the object, a cost function was developed that compares the object under consideration with a previously created database. The same segmentation and realignment method seen in the previous subsection was used to create the database, and the final labeling of the object was given manually in this step. In addition to the background image, only one image for each object was needed to create the database that will be referenced during the matching (REF_{IM}). The image of the object under analysis as input (IN_{IM}) was compared to all objects in the database to find the best match and finally identified it or not based on a set threshold. Initially, it is checked that the areas between IN_{IM} and REF_{IM} are similar within 30%. If this first step is overcome, the cost function (CF) between them is calculated as follows:

$$CF = \frac{(1 - SC) + (1 - SA) + (1 - SSIM)}{3} \tag{1}$$

where SC is the score of similarity related to the object contours. In particular, the contour of IN_{IM} is smoothed with a 2D Gaussian smoothing kernel with a standard deviation whose value changes according to the size of the object. Then, the resulting image is converted to a binary image and multiplied by the contour of REF_{IM} to check how many points of the two contours are in common. SA is the score of similarity related to the object areas. It consists to the product of the two binary images of IN_{IM} and REF_{IM} to check how many points of the two areas are in common. $SSIM$ calculates the score related to the structural similarity between the IN_{IM} and REF_{IM}. This score is a multiplicative combination of the three terms, namely the luminance term, the contrast term, and the structural term [228]. For this value to fall within an acceptable order of magnitude, both source images were cropped before this comparison with dimensions 50 % larger than the largest dimension between IN_{IM} and REF_{IM}. All the scores in Equation 1 are normalized and a 1 value was subtracted from all the terms because we are looking for the minimum value of the CF.

4.3.5 Metric Calibration of the Working Table

To evaluate the performance of the implemented algorithm for identifying real objects and estimating their position and orientation, metric analyses were performed after an initial camera calibration process.

Camera Calibration

The purpose of camera calibration is to find a correlation between the coordinates of each pixel in the CCD image sensor with their real-world measurements, taking into account lens distortions which are the most common monochromatic optical aberrations. For the calibration, the Kinect camera, attached perpendicular to the table and in its center at a fixed height of 80 cm, captures an image of a planar pattern placed on top of the table. The planar pattern consists of 55 Aruco markers located at the vertices of a grid with known positions. After the detection of the Aruco markers [76] their geometrical centers and identifiers were saved and compared with their positions in the environment.

An additional planar Aruco model (Fig. 4.45) was used to evaluate the calibration process and thus the accuracy of a random position on the table plane of dimensions 750x1020 mm. Once the set of random Aruco markers in the four corners was taken, the second time the set of randomly placed Aruco markers in the center was taken and the corresponding two-dimensional covariance matrices were computed. The plot of the covariance matrices is shown in Fig. 4.46. As expected, the uncertainty ellipse around the corners is larger due to the higher camera distortion.

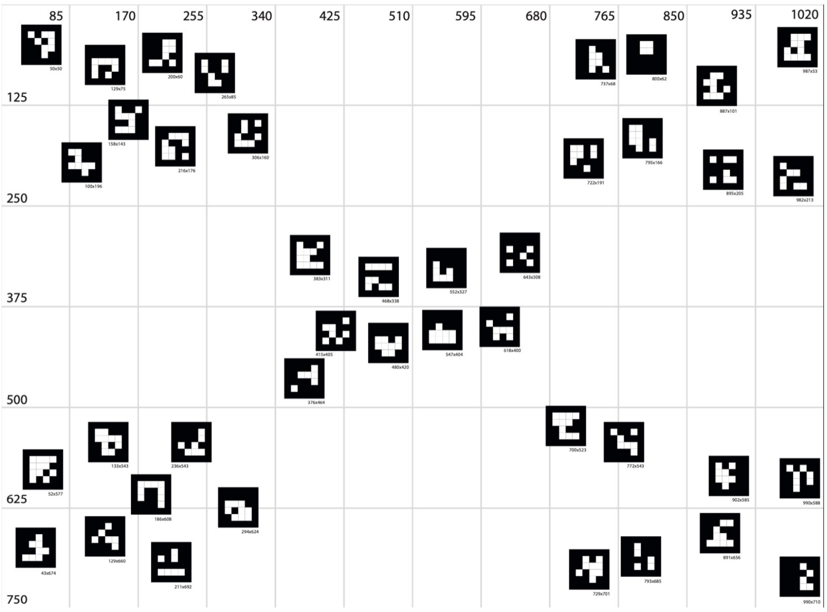


Figure 4.45: Aruco markers plane for accuracy checking.

Algorithm Accuracy

To evaluate the performance of the developed computer vision-based algorithm, rotation tests were performed using a knife. In particular, Fig. 4.47 shows a

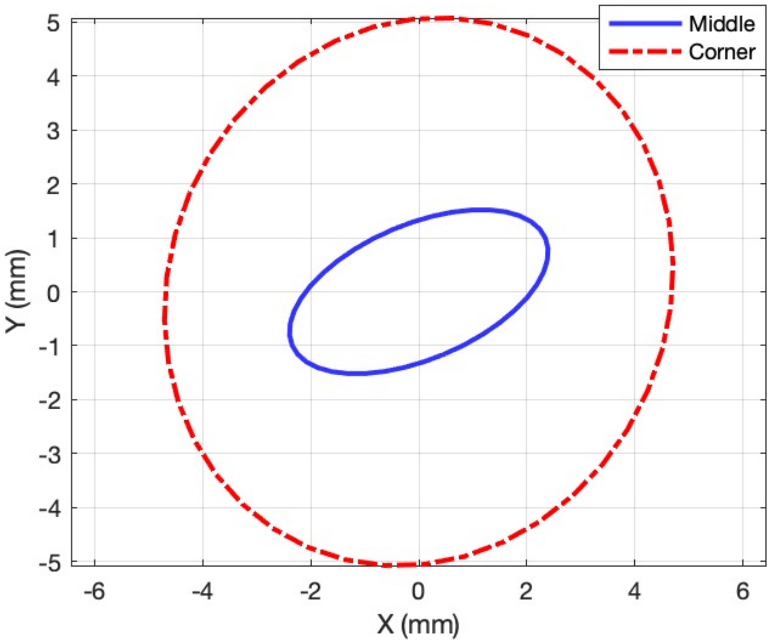


Figure 4.46: Ellipses of uncertainty in position (95 % confidence level with k= 2.4478 [202]).

cropped RGB acquisition image of tests conducted using a manual rotation motion platform (Standa 126865) covered with black to facilitate the background subtraction and filtering process. Fig. 4.47 shows both setups: one for tests conducted in the center of the table and one for tests near a corner of the table.

For the rotation tests, 180 acquisitions were performed for each of the two setups from 0 to 360° with a step size of 2°. The decision to do these tests on both the center and sides of the acquired images was to get a better estimation of the algorithm's performance over the entire table surface. The differences between the obtained rotations from the SVD algorithm and the one from the rotation motion platform taken as ground truth are shown as histograms of residuals for the two different setups in Fig. 4.48. The histogram spread for the setup at the center of the table is smaller than that for the setup near the corner due to the higher camera distortion. However, the residual in the estimation of rotations for the localization algorithm in general over the entire table surface is less than 1°.

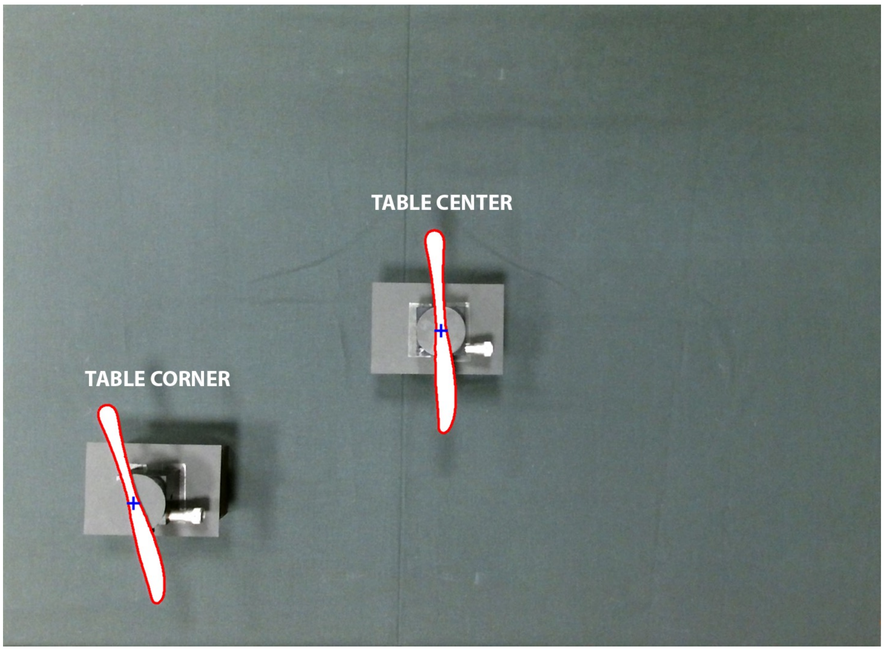


Figure 4.47: Cropped RGB image acquired for rotation tests.

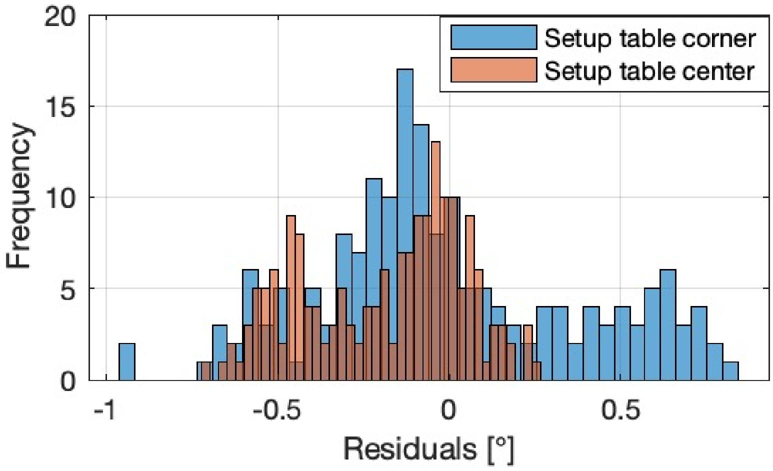


Figure 4.48: Histograms of residuals in the two setups.

4.3.6 Answer to Hypothesis CMO3

This prototype was designed for both the use of the therapist and the patient. The purpose of this is to train the ADL of the patient, while at the same time improving the clinical diagnosis and guidance of the therapist. This was done by having a behavioral role modeling of what is the ideal execution of the ADL the patient should follow. This was shown by the therapist through a mentoring system where the therapist place the virtual objects as targets while being with them physically and verbally throughout the guidance. The patient were able to improve their motor skills through this demonstration and live feedback from the therapist strategy. The therapist were also able to make better diagnosis about the patient condition through the evaluation system that we have created, were we validated the accuracy and reliability of the evaluation. By adding our system to the usual rehabilitation workflow between the therapist and the patient, expert feedback from the therapist indicated that they were again able to improve their clinical eye. They also indicated better communication with the patient as it was easier to explain what they expect from the patient visually rather than verbally. As an attitude outcome of this prototype, the patient and therapist was able to have a technological acceptance of AR systems within the traditional implementation.

5. Evaluation by Logical Induction in Surgery Training

5.1 Pre-Operative Rehearsal System

5.1.1 Background

Through the consolidation of patient-specific data and the experience of novel interactions in 3-D space, AR and VR can enable the “ contextual and detailed preplanning of medical procedures, better identification of incidentals, collaborative postoperative review, and advanced remote interaction [60].” Generally, surgeons use a variety of 2-D digital information. This 2-D representation of 3-D patient data necessitates that surgeons develop a 3-D mental picture of the surgery and its possible problems.

As depicted in Figure 5.1, numerous research investigations are done about the use of AR in the setting of highly complex exploration of medical imaging data, such as computed tomography (CT) scans, magnetic resonance angiography (MRA), and magnetic resonance imaging (MRI) data. The objective is to assist the diagnostic and treatment planning with the capacity to rebuild and display complicated information in 3-D, as well as simulate treatments. This is particularly significant in the context of surgical planning, where a complicated patient-specific operation may be rehearsed and alternative possibilities, such as tumor resection margins and distance evaluation, can be envisioned, as examined by Hansen et al. in liver surgery utilizing an AR technique [91].

Knowledge is acquired by storing information in a person’s memory for subsequent retrieval. It comes in two forms, either as an explicit memory which is “a visually recallable memory in time” or as an implicit skill recall where the “body responds to the tasks at hand without recourse to visual memory” [218]. Medical doctors may need to access memory related with the recollection of particular case histories and have a strong implicit awareness of muscle memory in relation to the handling of tools and the patient being treated. In the same way that a space agency simulates a spacecraft’s course before pressing the launch button, performers practice a play before taking the stage. UCLA and Surgical Theatre, for instance, employ Oculus Rift headsets to simulate complex and delicate

Figure 5.1: Example of medical data shown in AR, from Miyake et al.'s VeinViewer prototype. [149]

neurosurgeries. [168].

5.1.2 Model Conceptualization

The main goal of this system is for the medical practitioners to have a deeper grasp of the real patient's particularities before proceeding with the actual surgical operation. This goal is achieved by illustrating a virtual model of the actual patient CT data as seen in Figure 5.2. Inside this mode, the medical practitioners can freely move, rotate and scale the virtual model, pick up and manipulate each of the muscles, and reorient the cut hip socket. Because the model represents data about the actual patient, it creates an environment where the surgeons can familiarize themselves with the patient's uniqueness and gain confidence through the practice of "image training". Image training involves immersing oneself in their imagination and rehearsing a specific skill. It is the daydreaming practice inside one's head, and skill acquisition becomes more prevalent the closer the parameters are to actuality (i.e., following the PETTLEP model of imagery) [225]. By mimicking the seven components of the PETTLEP model as close as possible corresponding to the surgical procedure itself, we can maximize the effects of skill acquisition gained through image training.

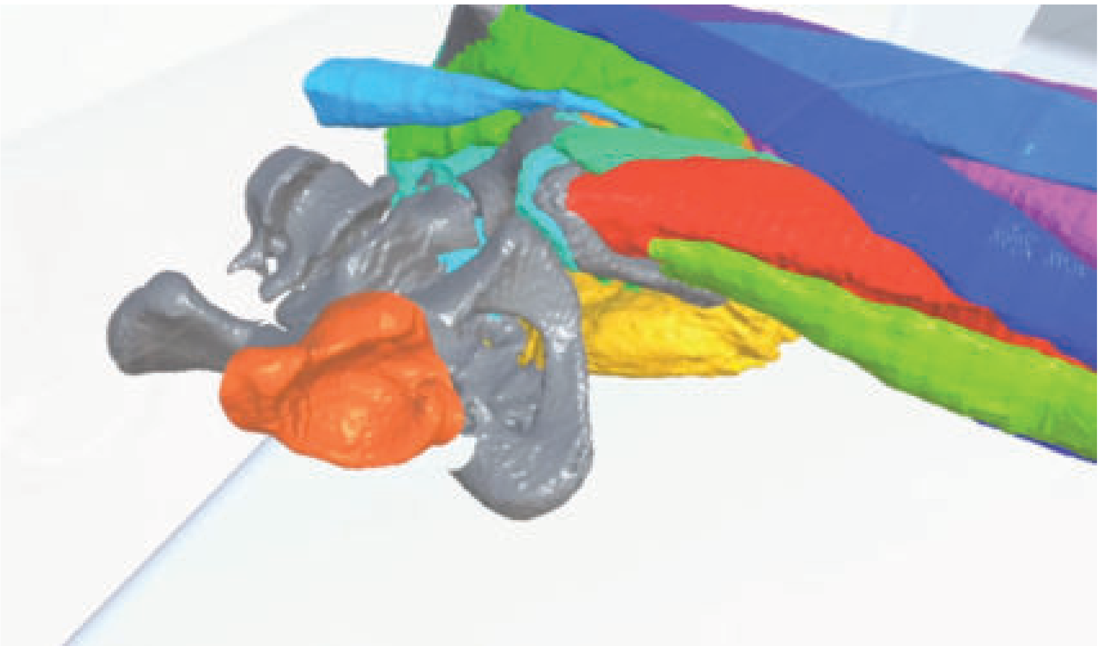


Figure 5.2: This prototype uses a Virtual model representing the actual patient's CT data. Gray color for the pelvic bone; orange for the reoriented hip socket; rainbow color for the muscles.

Future iterations of this prototype would be to add the functionality of placing long screws and wires. Aside from bone cutting, the metal plate and screw placement are also essential to the surgical procedure. Another idea proposed by the medical practitioners was to mix soft tissues with 3D printed hard bones instead of all entirely virtual. However, the cost for this implementation is high; thus, more discussion is needed before fully committing to this approach.

5.1.3 Answer to Hypothesis CMO4

This prototype was created with the intention of providing the master surgeon with a safe rehearsal environment in which to perform surgical procedures. The training was implemented with the simulation method, with the goal of stimulating the PETTLEP imagery training. The training strategy taken was by facilitating the transfer of skill through the imitation of patient's real CT data. As the master surgeon already has enough skill to perform the surgical operation, there is no such need to increase their knowledge and skills. Rather we focus on their attitudes, specifically on them gaining their confidence during the rehearsal so they can perform at optimal and with full confidence during the actual surgical operation. As the PETTLEP imagery of training dictates that as the variables in the practice are tied as close to actuality, we theorize that not only will they gain more insights on the knowledge about the patient's body uniqueness, but also gain the comfort of this knowledge and skill performance through their boosted confidence during the repeated rehearsal.

5.2 Eye-gaze Sharing and Review System

5.2.1 Background

Gaze is a typical communication tool for strengthening coordination methods in collaborative endeavors. Eye fixation may function as a rapid and accurate pointer, to confirm and clarify the item of focus, or as a reference that simplifies linguistically complicated things with deictic allusions (e.g. "this" or "here"). Shared Gaze Visualizations (SGVs) aid in grounding and are assisted by joint visual attention (JVA), resulting in the effective establishment of common ground in fostering cooperation [57].

In prior studies, visualisation approaches such as a cursor, spotlight, scanpath, trail, or heatmap were employed to depict real-time gaze cues in conventional 2D displays [237, 56]. Due to the scale of the display and the quantity of information shown, these visualisations may also serve as a mental distraction [132]. Researchers have begun using immersive technologies to show gaze visualisations for placed items in physical settings, although the most prevalent way is to utilize a basic cursor (e.g., a circle or a crosshair) to depict the current gaze point recorded in real-time [87]. Visualizing a virtual gaze ray or merging a view frustum with a gaze ray in the shared AR and VR work area is another common method [172].

Gaze is one of the most often utilized communication cues in the real world and has considerable promise for Computer Supported Collaborative Work (CSCW) study. Due to the implicit nature of gaze signals in both the physical and virtual worlds, it is crucial to investigate how gaze might be expressed more visibly and naturally during cooperation. Jing et al. [113] investigated about how users intuitively see gaze signals in order to obtain a better knowledge of one another as seen in Figure 5.3. Providing a good way of sharing gaze visualization coud result in improved co-presence, coordination awareness, task engagement, and overall collaborative experience.

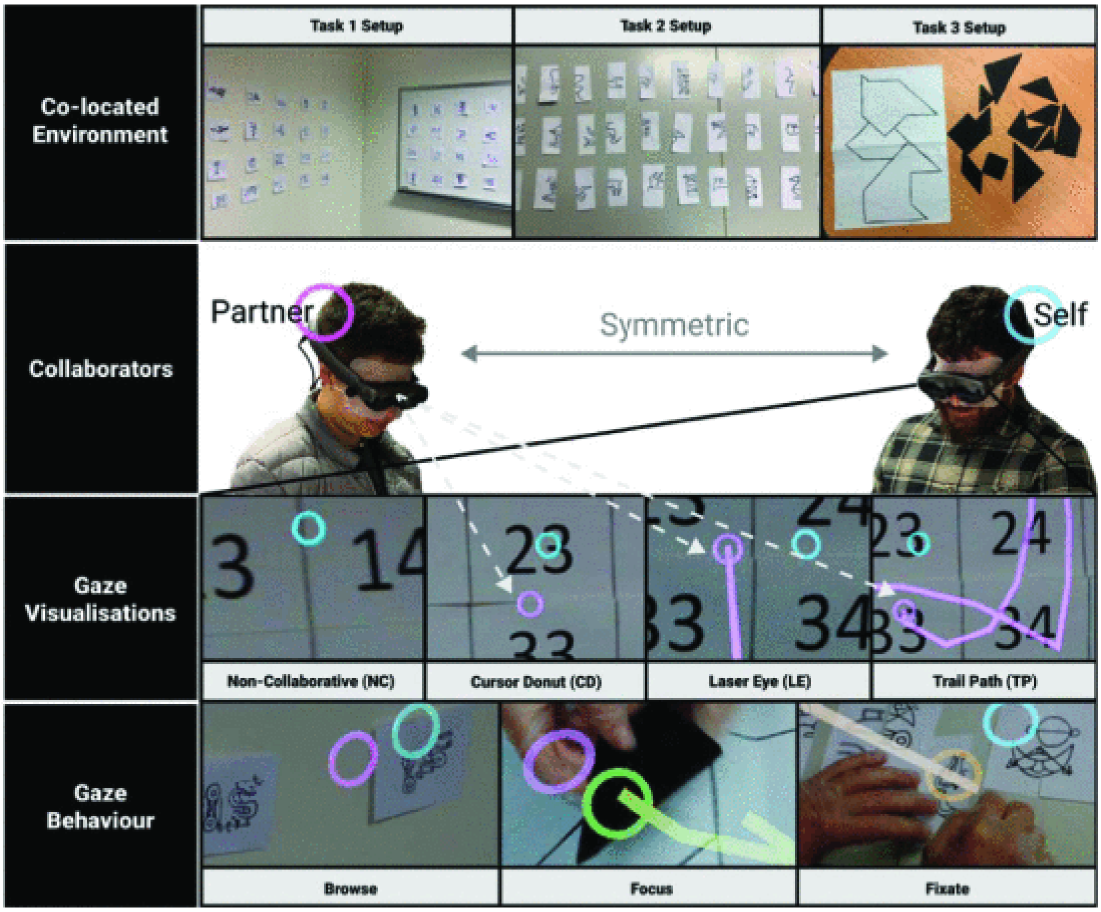


Figure 5.3: Prototype of Jing et al. to study the gaze visualizations and behaviors in a shared AR environment [113].

5.2.2 Model Conceptualization

The system developed with the medical practitioners was a system that can highlight the difference between the actions of an experienced surgeon and a novice surgeon. The idea of this AR artifact is to share the eye-gaze information of the experienced surgeon to the novice surgeon observing from a third-person perspective, shown in Figure 5.4. The system can record this eye-gaze information, and the novice can playback certain parts of the recording as a review. The sharing of eye-gaze information provides an excellent medium for learning how people with higher skill think and also facilitate the reflection of the causes of this skill difference.

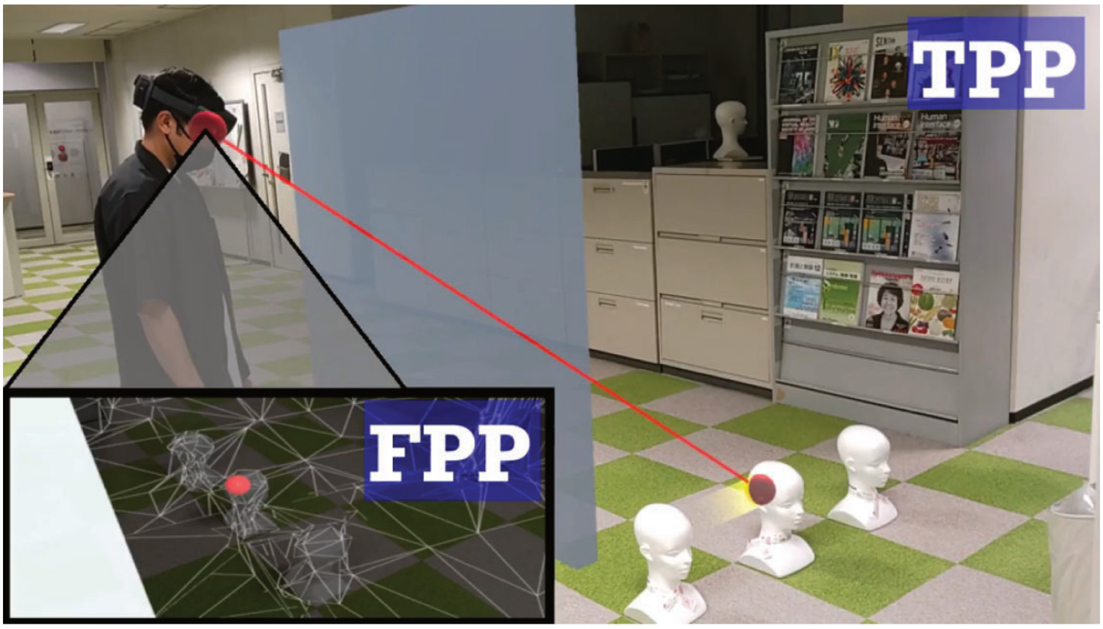


Figure 5.4: This prototype is a 2-HoloLens system for sharing the eye-gaze perspective of the experienced surgeon in the first-person perspective (FPP) to the novice surgeon in the third-person perspective (TPP).

Figure 5.4 represents the previous iteration of this design. The current prototype iteration provides a cross-hair visualization to recognize easier which is the head-point and which is the eye-gaze hit point. Also, instead of relying on the spatial map provided by MRTK, 3D scans of the mannequins were used for better accuracy. Other ideas for the next iteration of this design are also to record the point cloud data captured by a depth camera and combine the replay data of the point cloud with the eye-gaze implementation, providing a “person-like figure” during the playback of the eye-gaze recordings.

5.2.3 Answer to Hypothesis CMO5

This prototype was designed for the use of surgeons who are already considered professionals and have licenses, but are still relatively less experienced as compared to their peers in this profession (around 1-3 years of experience). This system was made for these less experienced surgeons to have a way to observe the TPP gaze of the eye of those who are considered much higher skilled than them in the hopes that they would get realizations and insights by themselves from observation. The training method that was appropriate for this was the use of the job shadowing system. This is because the strategy to share the skill from master to less experienced surgeon was through demonstration of the actual surgical operation itself, taking advantage of the behavioral role modeling. This approach was taken because there is a lot of tacit knowledge the master surgeons cannot explain that they have gained only through the years of doing this profession. We theorize that through the observation of the gaze visualizations and gaze behaviors of the master surgeon, less experienced surgeons would be able to understand these tacit knowledge that are hidden which are only unlocked through their years of experience.

5.3 Hand-Eye Coordination Training System

5.3.1 Background

Attention often refers to the capacity to focus mental powers on an object, such as cautious observation or attentive listening, or the mental capacity to concentrate. It has been known for at least half a century that the human attention span is limited [35]. This implies that we can only focus on a limited number of stimuli or pieces of information at any one moment.

Figure 5.5 demonstrates that the master surgeon uses less attentional capacity than the beginner for fundamental psychomotor, spatial, and decision-making activities. The gap between the top of the master surgeon column and the attentional capacity threshold symbolizes an attentional resource buffer zone that the master surgeon employs to handle problems and maintain track of extra data, such as instrument readouts or patient physiologic monitor information. When a novice is learning new abilities, such as those necessary for surgery, he or she

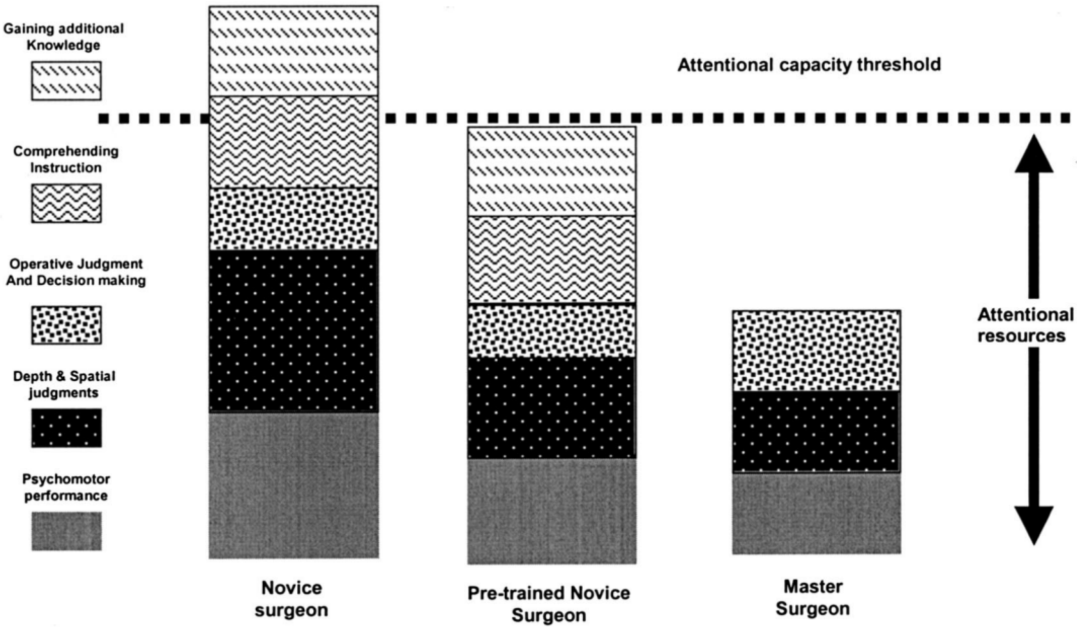


Figure 5.5: The hypothetical attention resource model according to Gallagher et al. [74]

must utilize these attentional resources to actively monitor what their hands are doing in addition to making spatial judgements and operating decisions, although making fewer decisions than an attending master surgeon. This leads in a modest increase in the novice's attentional ability. Given that most novice surgeons are expected to acquire judgment and decision-making from the master surgeon, the novice's attentional barrier is soon exceeded by these extra attentional resource demands. As shown in Figure 5.5, simulation skills training enables the formation of the "pretrained novice." This individual has been trained with simulation to the point where many psychomotor skills and spatial judgments have been automated, requiring significantly less attentional resources. This allows the novice to focus more on learning the steps of the operation and how to handle complications, rather than wasting valuable operating room time on the initial refinement of technical skills [74].

5.3.2 Model Conceptualization

Another system developed with the suggestions of the medical practitioners was an AR artifact that can nurture the hand-eye coordination skill of medical students learning surgery. Acquiring this hand-eye coordination skill is a must before becoming a full-fledged surgeon. The problem with acquiring this skill is that the jump in difficulty from textbook explanation to actual hands-on operation is too vast of a gap. As a method of softening the gap in skill level between these two extremities, this system visualizes a 4-panel AR head-locked view seen in Figure 5.6. This 4-panel head-locked view represents what is otherwise projected on a large monitor display in the surgical room. An outside-in tracking was implemented using a 3D printed tool with attached reflective markers and an Optitrack multi-camera motion capture system.

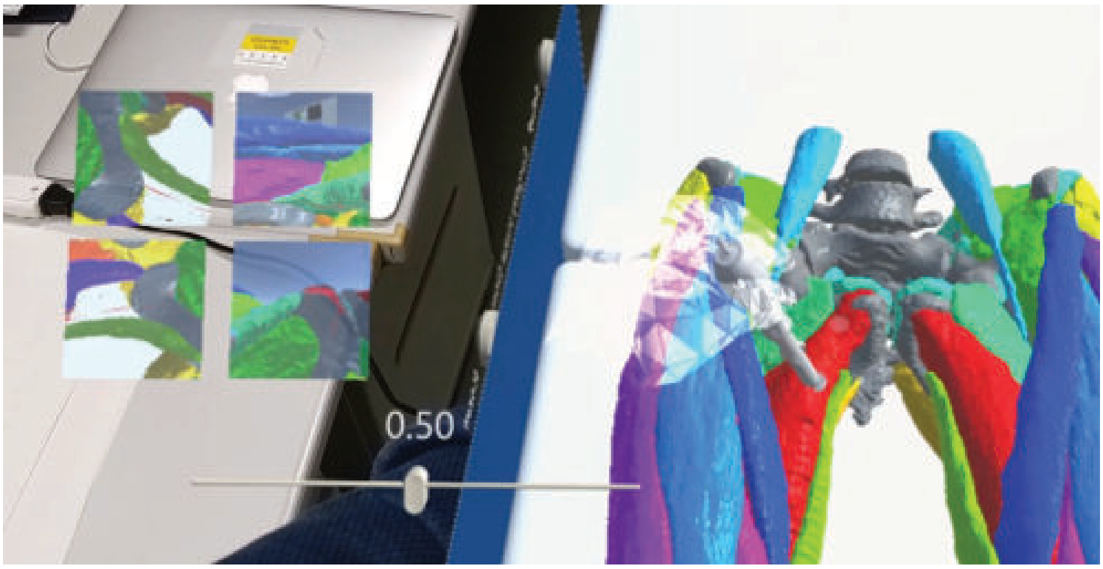


Figure 5.6: This prototype is a Navigation tool that visualizes a head-locked 4-panel view (coronal, sagittal, axial, and oblique).

Further iterations of this prototype include the operation of a virtual C-arm machine, which can be adjusted akin to that of a real C-arm machine. This virtual C-arm machine can be used to project a digitally reconstructed radiograph (DRR) image of the virtual patient model instead of this 4-panel view. However, because the manipulation and adjustment of this virtual C-arm takes a huge learning curve for those new to AR, we have also decided to add a smartphone interface that acts like a remote controller to rotate the virtual C-arm machine.

5.3.3 Answer to Hypothesis CMO6

This prototype was designed for the medical student who still lack the skill of hand-eye coordination, which is one of the most important skill to master before becoming a professional surgeon. Based on earlier hypothetical attention resource model in Figure 5.5, for the student to gain additional knowledge we need to lower the allocation of the other resources, in this case the student's depth and spatial judgements during the tool operation, as well as their psyhomotor performance. This is lowered by using the strength of AR, which is repositioning data visualizations in 3D. The method used in this training is to have a games-based approach, which is to increase progressively the difficulty of where these data visualizations are positioned, with real-time feedback for the motion of the manipulated tools. This progressive difficulty should promote the deliberate practice, which is practicing things you cannot currently do, with effort. With this balance of deliberate practice, and softening of the attentional resources, we theorize that students would gain the skill and knowledge without losing their confidence through the proper balance and matching of the extraneous cognitive load with their current skill level.

6. Conclusion

As additional technical obstacles are overcome over years of study, computer-mediated training employing AR is gradually being added to and utilized in conjunction with conventional training techniques. Although the use of AR for training has a great deal of promise, its efficacy is not as clear-cut due to a number of controversial and poorly understood ideas that need additional exploration. In this thesis, we have explored the application of ARTS by investigating how and under what conditions ARTS are effective for educating people involved in the medical ecosystem. Through the development of a program theory, this work utilized a realism approach. This is accomplished by the theoretical elicitation of CMO combinations, which are subsequently tested by confirming each of their separate assumptions. The training effects of different CMO setups were evaluated using both empirical data and inductive reasoning. Below, we show again the hypotheses we explored in an attempt to understand what are effective ARTS, and re-summarize them in Table 6.1.

6.1 Reviewing Answers to Hypotheses

Again, the hypotheses we tested in our realist approach are as follows:

CMO-1 The presentation of real-time visualizations of invisible current conditions about the patient can train the clinical eye of the therapist.

CMO-2 Providing real-time feedback that supports the understanding of exproprioceptive information can train the locomotion skill as performed in actual exercise.

CMO-3 Incorporating a valid and reliable shared AR experience into the usual rehabilitation workflow improves therapist's diagnosis and communication with the patient.

CMO-4 The use of patient's real CT data can trigger perceptions of deep immersion, which will result in improved learning, knowledge, and comfort with knowledge and skill performance.

CMO-5 Observing different perspectives of higher skilled operative performance can facilitate the transfer of tacit knowledge and skill to learners.

CMO-6 Progressive difficulty of AR presentation can promote deliberate practice, which will make the transfer of knowledge and skills from practice to actual performance smoother.

Correspondingly, Table 6.1 are the CMO configurations and summary of the answers we learned from testing the above hypotheses. The mechanisms part are covered and explained in Chapter 2, the context on Chapter 3, the first three CMO's in Chapter 4, and the last three CMO's in Chapter 5.

Table 6.1: The CMO configurations we explored in this thesis. This table summarizes how ARTS work from the realist perspective.

Hypothesis	**Context**		**Mechanisms**			**Outcome**
	User	Purpose	Methods (Martin et al. [142])	Principles (Salas et al. [184])	Strategies (Salas et al. [184])	KSA
CMO-1	Therapist	Clinical Eye training	Simulation	Transfer appropriate processing	- Assistance in task execution through the presentation of task-related information - Real-time visualization of invisible current conditions	Skill of making correct diagnosis
CMO-2	Patient	Motor and exproprioception training	Games-based	Transfer appropriate processing	- Assistance in task execution through the presentation of task-related information - Real-time feedback for motion compensation	Motor skill and exproprioception
CMO-3	Therapist + Patient	ADL training of patient	Mentoring	Behavioral role modeling	- Demonstration - Extend the flexibility of training content by adding information to the physical object - Feedback on the performance, problems, and ways to improve	Patient: motor skill Therapist: diagnosis Both: Tech acceptance in workflow
CMO-4	Master Surgeon	Image training	Simulation	Self-regulation	- Facilitation of skill transfer by imitating real objects and their superimposed display	Gain confidence in rehearsal
CMO-5	Less experienced Surgeon	Observational training	Job shadowing	Behavioral role modeling	- Demonstration (Observe TPP eyegaze of higher skill)	Gain tacit knowledge of expert
CMO-6	Medical Student	Hand-eye coordination training	Games-based	Transfer appropriate processing	- Facilitation of skill transfer by imitating real objects and their superimposed display - Real-time feedback for motion compensation	Gain skill without losing confidence

6.2 Generalizations for Effective ARTS

The CMO configurations are explicable by the application of learning theories mentioned in the studied literature. Constructivism implies that learning is an active process that builds on prior skills, knowledge, and contact with the physical and social environment [170]. In order to adapt and learn, trainees engage with the environment via active building and experiential learning. Similarly, AR may be used by health professionals with prior clinical field experience, acquired knowledge, and/or acquired abilities. By immersing health professionals in simulated real-world contexts, AR technologies may facilitate active learning for competency, that is increasing their knowledge, skills, and attitudes. This is represented in the method of deep immersion for learners. Constructivism also explains the mechanics of repeated practice, skill enhancement, and interactive experiences since learners may engage with AR environments to practice their abilities [77].

The cognitive load theory may also be used to describe the mechanics of training humans in AR. According to Gallagher's attention resource model [74], this theory posits that humans have a limited quantity of accessible working memory. We have a limitless capacity for long-term memory, which stores cognitive schemas (experiential knowledge). The process of learning is subsequently the construction and automatization of these schemas so that they may be kept in long-term memory. Intrinsic load (task-specific cognitive effort) is distinguished from extraneous load (irrelevant cognitive effort) and germane load (residual working memory capacity) [213]. By giving real-time signals and feedback, AR may be able to minimize extraneous load, which consists of non-learning activities. However, it is also feasible that AR learning aids may raise task-specific or unnecessary cognitive burden accidentally since they may complicate learning processes [77]. This is due to the fact that learners unfamiliar with VR or AR technologies may need adaptation.

To summarize and generalize everything, what it takes for ARTS to be effective in training humans is valid environment, repeated attempts with feedback, and deliberate practice. Valid environment means that there is a possibility to recognize patterns, as this is how the human brains learn. As a counter-example, a casino is not a valid environment because in gambling there is no point in trying to predict the unpredictable. We have done the confirmation of the validity of the training systems through the use of empirical evidences such as doing user studies and feedback from experts for CMO 1-3, and doing theory constructions and model conceptualizations for CMO 4-6. The other takeaway we have learned is that a many repetitions with timely and appropriate feedback is needed. Humans learn from mistakes and failure, that is why experience is very valuable in the learning process. As all the training systems operated in a simulated environment, the users could practice as many times as they want. Extending this, timely and appropriate feedback works together with this because this feedback points to where the mistake and failure stem from, allowing for the learner to do corrective measures. For example, in CMO 3, the patient gets real-time feedback from the therapist about the exercise's current performance, problems, and ways to improve. Furthermore, the therapist get feedback from the evaluation system regarding the patient's performance during the exercise. The last insight we have is on deliberate practice. Simply put, the training has to be challenging and the learner should not get too comfortable. There should be a delicate balance between the difficulty of the training and the current skill of the learner that is ever changing and adapting. This is best exemplified by CMO 6 such that we are constantly adjusting the visualization of the AR contents with respect to the ability of the learner. With the slow and deliberate conscious practice, repeated enough, leads any human of any skill level to expert performance and growth.

In this work, we have explored a number of CMO configurations, for example on training people with a wide variety of skills/experience, or upskilling different kinds of end-users in a training program (i.e., not only increase the skill of the trainee, but also the trainer, thus increasing overall quality of training). There are however still quite a lot of unexplored areas that need further investigations. One such is on the context of the quality of the AR device used. Depending on the quality of the device, it will affect the level of fidelity and immersion experienced by the user, thus affecting the outcomes of training. Although it largely depends on the task at hand, a HoloLens will have different training effects as compared to a smartphone-based OST AR device such as the Google Cardboard. From Burdea et. al. [38], they stated that the learners in their study were not satisfied with the simulator because it was not perceived as realistic. In addition, the lack of perceived realism might be why their simulator group performed worse than the control group (using a rubber simulator) in diagnosing prostate cancer (33% vs 92%, respectively). It was expected that a more realistic simulator would have improved performance and learner satisfaction. We have also not considered about the starting intrinsic motivation of the user, as we have not measured the willingness to learn for each individual. This is important because this determines the skills-challenge balance theory which is used for deliberate practice, which eventually leads to the outcomes of training.

In the CMO prototypes we have developed, we have also not taken advantage of the training mechanism of programmed instruction of Salas [184]. This just so happens because during the participatory process with the co-designers, this mechanism was not needed to meet the purpose of the training (e.g., clinical eye training, etc...). The same can also be said for the error training, which is incorporating errors to the training program intentionally so that we can train the users how to respond to solve the errors when they are placed in difficult situations out of the blue. The outcomes of training that we have measured are also only for specific skill sets. We have not measured the KSA outcomes in general terms, such as skills related to teamwork and interaction with the other learners. We have assumed that improvements in these specific skill sets will contribute to the overall general improvement in competency and performance, however we have not yet considered about the bigger picture outcome of training. For example, we have not conducted any evaluation and observations regarding the long term use of our training systems, as the user studies lasted only for short periods of time, mostly within just a month.

6.3 Limitations and Challenges

As we have just discussed that there are still a lot of unexplored area of AR with regards to training, this discussion should be extended while keeping in mind the limitations and the challenges encountered during training. In this section, limitations and challenges about technological, user-related, and evaluation will be discussed.

One of the limitations when using AR in training people is that most of the developed AR technology focus mostly on the visual perception; while only a decent amount on auditory and tactile, and only a countable amount of research work on olfactory and gustatory. However, this is expected because humans are considered to be highly visual creatures [16], that is, we rely heavily on our visual cues for our behaviour formation, which includes training. To get the most efficiency out of training however, it makes sense that we take advantage of all five of our senses of our body, to stimulate the key components of the PETTLEP model[225]. It would also be to our advantage if we also make use of the four other senses during training. For example, we could use hearing aids in AR as a support tool in rehabilitation, or we could use haptic gloves that mimic the sense of touching organs and blood vessels for surgical training. As for the sense of smell and taste however, these still remain a relatively harder challenge today as it is still yet to be discovered how to seamlessly link these chemically-based experiences connected to our brain receptors with digital input [95].

To give an example, a training scenario in AR that does not involve the visual sense is the context of hearing rehabilitation. Auditory support such as hearing aids can be thought of as a kind of AR, as we are digitizing the sound data accordingly from the dynamic changes in the environment and amplify the parts of interest to a comfortable level so that the user with hearing impairment may be able to hear much clearer [129]. The work by Mehra et al. explored the potential of the auditory aspect of AR by trying to tackle the Cocktail-Party problem. This problem is described as the effect when multiple sources of sound in a noisy environment mix in the air just before reaching the person's ear, just as in a Cocktail Party [146]. The brain of that person must then be able to distinguish between the distinct sound sources and be able to focus, understand, and lock on to one sound source. Mehra et al. tried to solve this by creating an AR platform that has a machine-learning backbone with input such as multimodal sensors and motion tracking systems to determine the "intent detection, speaker separation, and noise separation" [146]. In short, this auditory AR system is able to isolate the wanted audio cues while filtering out the unwanted ones.

Extending to the example above, the described auditory AR system can be useful for training in hearing rehabilitation. It is possible to apply the progressive difficulty concept to the design of the auditory AR training [40]. For example, we could follow Erber's proposal on the level of a person's listening skills, from order of low to high skill: “sound awareness, sound discrimination, sound identification, and sound comprehension [64].” Sound awareness, at the most fundamental level, is the ability to judge the existence or absence of a sound. Sound discrimination involves the capacity to determine sounds to be identical or distinct, despite the ability to link interpretation to these sounds. Sound identification refers to the capacity to appropriately identify a sound, or a set of consecutive sounds (forming a phrase). Finally, sound comprehension pertains to one’s capacity to interpret identifiable sounds, from the level before, but now being able to comprehend the meaning when these sounds are spoken in order [40]. Adjusting these four levels of difficulty can be a good auditory AR training example for hearing rehabilitation.

There are also user-related limitations that should be considered when designing AR systems used for training. For example, in CMO 3, we have found that there is a big effect on the fatigue on the users when they are wearing the HoloLens for an extended period of time [53]. This was measured from the high levels of frustration and physical demand scores of the NASA-TLX. This may be attributed to the form factor of the HoloLens, which can be quite heavy and bulky, leading to the uncomfortable feeling when used long. Another user-related case to consider is on the problem of the vergence-accommodation conflict (VAC), which is common in the context of AR. VAC happens when the brain receives contradicting visual cues about the distance of a 3D virtual object (vergence) and the required focusing distance (accommodation) for the eyes to concentrate on the aformentioned object [240]. There are a lot of research trying to solve this problem, but it is a fact that a lot of AR devices still hold this issue, which leads to cybersickness.

Lastly, there are also the evaluation-related limitations when doing training on AR. For example, one big problem on medicine when testing early prototypes of AR training systems to critical-care patients is the ethical issue [37]. Questions about whether we can guarantee the safety of the patient during the use of prototypes that are still at the early stages of development is still hard to answer. A human-to-human interaction is still promoted favorably as compared to a human-to-machine interaction. This means that when creating AR systems for training, a user-centric design should be instilled, as discussed in subsection 3.2. In this kind of research, it is also very difficult to get the quality of evaluation testing that is needed before we can say for certainty it is ready for mass production and social implementation. This is because it it very troublesome to setup randomized-control trials, blinding studies, and placebo studies with actual users [186]. This stems from the limited number of users that are available for a certain study, for example critical-care patients for rehabilitation, and also the earlier issue about the ethics of using early prototype systems to highly sensitive users.

6.4 Future Work

In this work, we took the approach of authoring ARTS with participatory design with a realist evaluation through the CMO configurations. We have learned and explored the different examples of combinations of the context plus mechanisms which lead to different outcomes, that were confirmed either through empirical evidence or logical induction. However, this work only covered some of the mechanisms in training, and when they are activated or not. For example, this work has not explored the effects of the level of fidelity of the simulations, or the level of immersion the trainees experience, like when comparing AR to varied level of VR immersion. There are still a lot of areas that need to be explored and understood in the ARTS domain, however this should only be a secondary priority as compared to the needs of the end-user. When developing and implementing effective ARTS, our philosophy should still be that it should fit in the current and traditional workflow of social structures, and at the least should be "useful" to the end-users in a consistent basis.

Acknowledgements

First and foremost, I would like to extend my deepest gratitude to Professor Hirokazu Kato for giving me the opportunity to study under his laboratory and for guiding and supporting me through the pursuits of my research work. This is even more true to Assistant Professor Yuichiro Fujimoto who was always there, especially during the brainstorming and writing of my journal paper. I am also very grateful for Professor Kiyokawa, Associate Professor Kanbara, and Assistant Professor Sawabe for the advice they have given me for this research and for the time they took to read this relatively long thesis. I would also like to thank our collaborators from the medical side, Takao-sensei, Uemura-sensei, Okahashi-sensei, plus their team, and also from the Imaging-based Computational Biomedicine lab, Sato-sensei and Otake-sensei, for their invaluable feedback and ideas in this continuous exchange of participatory design.

I would like to also thank the members of Interactive Media Laboratory, who has been with me through the tough times, and the fun times. I am really grateful to be a part of the Tsundere Project, which is one of my memorable experiences as a doctoral student. This is why I cannot be grateful enough to Keishi Tainaka and other members of the Tsundere Project, including the people from maid cafe CCOCha. Furthermore, I would like to thank Jayzon Flores and Nicko Caluya, my two Filipino senpai from this lab that I always look up to. I would also like to thank my travel buddies in the lab for all the fun exploration around Japan, and for the late night ramen expeditions. I also thank our collaboration partner, MIRo lab in Italy, for the opportunity to do research work together. I especially like to give my deepest appreciation to Alessandro Luchetti, for not only being a colleague in research, but also as a dear friend who is there when the times get tough. I also want to mention Manuel Rosi and friends, for making my internship stay in Italy a very unforgettable experience. I also would like to return my love to my family in the Philippines, my mom and my dad, who supported me in my pursuits of studying in Japan. I especially thank my brother, Ian Butaslac, who was always there to listen to my problems, especially when I have had mental and emotional breakdowns. Finally, I would like to thank all the friends I have made here in Japan, as I consider my stay here a very memorable part of my life.

References

[1] Kamyar Abhari et al. Training for planning tumour resection: Augmented reality and human factors. *IEEE Trans. Biomed. Eng.*, 62(6):1466–1477, June 2015.

[2] Louise Ada, Simone Dorsch, and Colleen G Canning. Strengthening interventions increase strength and improve activity after stroke: A systematic review. *J. Physiother.*, 52(4):241–248, February 2006.

[3] Michelle Aebersold et al. Interactive anatomy-augmented virtual simulation training. *Clin. Simul. Nurs.*, 15(1):34–41, February 2018.

[4] Ali Alaraj et al. Role of cranial and spinal virtual and augmented reality simulation using immersive touch modules in neurosurgical training. *Neurosurgery*, 72(S1):115–123, January 2013.

[5] Urs-Vito Albrecht, Kristian Folta-Schoofs, Marianne Behrends, and Ute Von Jan. Effects of mobile augmented reality learning compared to textbook learning on medical students: randomized controlled pilot study. *Journal of medical Internet research*, 15(8):e182, 2013.

[6] Raed S. Alsawaier. The effect of gamification on motivation and engagement. *Int. J. Inf. Learn. Technol.*, 35(1):56–79, January 2018.

[7] Chang Man An and Dae Hyun Kim. Clinical application of ar system in early rehabilitation program after stroke: 2 case study. *The Journal of Korean Physical Therapy*, 31(3):141–146, 2019.

[8] Daniel Andersen et al. Medical telementoring using an augmented reality transparent display. *Surgery*, 159(6):1646–1653, January 2016.

[9] Jose Manuel Andujar, Andrés Mejias, and Marco Antonio Marquez. Augmented reality for the improvement of remote laboratories: An augmented remote laboratory. *IEEE Trans. Educ.*, 54(3):492–500, August 2011.

[10] Brad Astbury and Frans Leeuw. Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation - AM J EVAL*, 31:363–381, September 2010.

[11] Osnat Atun-Einy and Michal Kafri. Implementation of motor learning principles in physical therapy practice: Survey of physical therapists' perceptions and reported implementation. *Physiotherapy Theory and Practice*, 35:1–12, March 2018.

[12] Yee Aung and Adel Al-Jumaily. Augmented reality based illusion system with biofeedback. *Middle East Conference on Biomedical Engineering, MECBME*, pages 265–268, February 2014.

[13] Yee Aung and Adel Al-Jumaily. Augmented reality-based rehabio system for shoulder rehabilitation. *Int. J. of Mechatronics and Automation*, 4:52–62, January 2014.

[14] Ronald T. Azuma. A survey of augmented reality. *Presence (Camb)*, 6(4):355–385, August 1997.

[15] Bettina Babbitt, William Pieper, Clarence Semple, and Dianne Swanson. Effectiveness and efficiency of computer-based training: Recent research findings. In *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, volume 29, pages 198–201, October 1985.

[16] Pooja Balaram and Jon Kaas. Current research on the organization and function of the visual system in primates. *Eye and Brain*, 6:1–4, September 2014.

[17] Timothy T. Baldwin. Effects of alternative modeling strategies on outcomes of interpersonal-skills training. *Journal of Applied Psychology*, 77:147–154, 1992.

[18] Steve Balian, Shaun K. McGovern, Benjamin S. Abella, Audrey L. Blewer, and Marion Leary. Feasibility of an augmented reality cardiopulmonary resuscitation training system for health care providers. *Heliyon*, 5(8), August 2019. Art. no. e02205.

[19] Albert Bandura. *Social learning theory.* Englewood Cliffs, N.J. Prentice-Hall, 1977.

[20] E. Z. Barsom, M. Graafland, and M. P. Schijven. Systematic review on the effectiveness of augmented reality applications in medical training. *Surg. Endosc.*, 30(10):4174–4183, February 2016.

[21] Anil Ufuk Batmaz, Aunnoy K Mutasim, Morteza Malekmakan, Elham Sadr, and Wolfgang Stuerzlinger. Touch the wall: Comparison of virtual and augmented reality with conventional 2D screen eye-hand coordination training systems. In *2020 IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*, pages 184–193, March 2020.

[22] Kirsten Berthold, Matthias Nückles, and Alexander Renkl. Do learning protocols support learning strategies and outcomes? the role of cognitive and metacognitive prompts. *Learning and Instruction*, 17:564–577, October 2007.

[23] Roy Bhaskar. *Realist Theory of Science.* UK: Harvester, 1978.

[24] Isabela Bianchi, Alexandre Lazaretti Zanatta, and Rafael Rieder. Augmented reality in medical teaching-learning process content: A systematic review. In *2020 22nd Symp. Virtual and Augmented Reality (SVR)*, pages 129–133, November 2020.

[25] Paolo Bifulco, Fabio Narducci, Raffaele Vertucci, Pasquale Ambruosi, Mario Cesarelli, et al. Telemedicine supported by augmented reality: An interactive guide for untrained people in performing an ECG test. *Biomed. Eng. Online*, 13(1), November 2014. Art. no. 153.

[26] Kamran Binaee and Gabriel J Diaz. Assessment of an augmented reality apparatus for the study of visually guided walking and obstacle crossing. *Behavior research methods*, 51(2):523–531, 2019.

[27] James Menzies Black. *Developing Competent Subordinates.* American Management Association, New York, 1961.

[28] Sven Blomqvist, Stefan Seipel, and Maria Engström. Using augmented reality technology for balance training in the older adults: A feasibility pilot study. *BMC Geriatr.*, 21(1), February 2021. Art. no. 144.

[29] Sanne M.B.I. Botden, Sonja N. Buzink, Marlies P. Schijven, and Jack J. Jakimowicz. Augmented versus virtual reality laparoscopic simulation: What is the difference? *World J. Surg.*, 31(4):764–772, March 2007.

[30] Klaudia Bovermann and Theo J. Bastiaens. Towards a motivational design? Connecting gamification user types and online learning activities. *Res. Pract. Technol. Enhanc. Learn.*, 15, January 2020. Art. no. 1.

[31] Mehmet Gazi Boyaci et al. Augmented reality supported cervical transpedicular fixation on 3D-printed vertebrae model: An experimental education study. *Turk. Neurosurg.*, 30(6):937–943, August 2020.

[32] Thiago Braga Rodrigues, Ciarán Ó Catháin, Noel E. O'Connor, and Niall Murray. A quality of experience assessment of haptic and augmented reality feedback modalities in a gait analysis system. *PLoS One*, 15(3), March 2020. Art. no. e0230570.

[33] J Gavin Bremner and Alan Slater. *Theories of infant development.* John Wiley & Sons, 2008.

[34] Yassine Brij and Hicham Belhadaoui. Virtual and augmented reality in school context: A literature review. In *2021 3rd Int. Conf. Transportation and Smart Technologies (TST)*, pages 16–23, May 2021.

[35] D. Broadbent. Selective and control processes. *Cognition*, 10:53–58, 1981.

[36] F. P. Brooks. The computer scientist as toolsmith ii. *Communications ofthe ACM*, 39:61–68, 1996.

[37] Raphael Bruno, Georg Wolff, Bernhard Wernly, Maryna Masyuk, Kerstin Piayda, et al. Virtual and augmented reality in critical care medicine: the patient's, clinician's, and researcher's perspective. *Critical Care*, 26, October 2022.

[38] G. Burdea, G. Patounakis, V. Popescu, and R.E. Weiss. Virtual reality-based training for the diagnosis of prostate cancer. *IEEE Transactions on Biomedical Engineering*, 46:1253–1260, 1999.

[39] Isidro III Butaslac, Alessandro Luchetti, Edoardo Parolin, Yuichiro Fujimoto, Masayuki Kanbara, et al. The feasibility of augmented reality as a support tool for motor rehabilitation. In Lucio Tommaso De Paolis and Patrick Bourdot, editors, *Augmented Reality, Virtual Reality, and Computer Graphics*, pages 165–173, Cham, August 2020. Springer International Publishing.

[40] Rachel Caissie. Chapter 5: Auditory training, 2019. Accessed 26 February 2020.

[41] Brendan Cassidy, Gavin Sim, Matthew Horton, and Daniel Fitton. Participatory design of wearable augmented reality display elements for children at play. In *2015 7th Computer Science and Electronic Engineering Conference (CEEC)*, pages 53–58, September 2015.

[42] T. P. Caudell and D. W. Mizell. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *Proc. 25th Hawaii Int. Conf. System Sciences*, volume 2, pages 659–669, January 1992.

[43] Mariolino Cecco, Alberto Fornaser, Paolo Tomasin, Matteo Zanetti, G. Guandalini, et al. Augmented reality to enhance the clinician's observation during assessment of daily living activities. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pages 3–21, June 2017.

[44] Mariolino Cecco, Alberto Fornaser, Paolo Tomasin, Matteo Zanetti, G. Guandalini, et al. Augmented reality to enhance the clinician's observation during assessment of daily living activities. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pages 3–21, June 2017.

[45] Zoe YS Chan, Aislinn JC MacPhail, Ivan PH Au, Janet H Zhang, Ben MF Lam, et al. Walking with head-mounted virtual and augmented reality devices: Effects on position control and gait biomechanics. *PLoS one*, 14(12), 2019.

[46] Gill Chard. An investigation into the use of the assessment of motor and process skills (amps) in clinical practice. *British Journal of Occupational Therapy*, 63(10):481–488, 2000.

[47] Meiling Chen, Qingfeng Tang, Shoujiang Xu, Pengfei Leng, and Zhigeng Pan. Design and evaluation of an augmented reality-based exergame system to reduce fall risk in the elderly. *Int. J. Environ. Res. Public Health*, 17(19), October 2020. Art. no. 7208.

[48] Po-Jung Chen, I-Wen Penn, Shun-Hwa Wei, Long-Ren Chuang, and Wen-Hsu Sung. Augmented reality-assisted training with selected Tai-Chi movements improves balance control and increases lower limb muscle strength in older adults: A prospective randomized trial. *J. Exerc. Sci. Fit.*, 18(3):142–147, May 2020.

[49] I A Chiang, Rajiv S Jhangiani, and Paul C Price. *Research Methods in Psychology*. BCcampus, 2nd canadian ed. edition, October 2015.

[50] Winyu Chinthammit et al. Ghostman: Augmented reality application for telerehabilitation and remote instruction of a novel motor skill. *BioMed Res. Int.*, 2014, April 2014. Art. no. 646347.

[51] Mathieu Chollet, Stacy Marsella, and Stefan Scherer. Training public speaking with virtual social interactions: Effectiveness of real-time feedback and delayed feedback. *J. Multimodal User Interfaces*, 16:17–29, March 2021.

[52] Ashirwad Chowriappa et al. Augmented-reality-based skills training for robot-assisted urethrovesical anastomosis: A multi-institutional randomised controlled trial. *BJU Int.*, 115(2):336–345, August 2014.

[53] Carole Cometti, Christos Païzis, Audrey Casteleira, Guillaume Pons, and Nicolas Babault. Effects of mixed reality head-mounted glasses during 90 minutes of mental and manual tasks on cognitive and physiological functions. *PeerJ*, 6:e5847, November 2018.

[54] Sara Condino et al. How to build a patient-specific hybrid simulator for orthopaedic open surgery: Benefits and limits of mixed-reality using the

Microsoft HoloLens. *J. Healthc. Eng.*, 2018, November 2018. Art. no. 5435097.

[55] Richard Culatta and Greg Kearsley. Sequencing of instruction. Accessed 28 June 2021.

[56] Sarah D'Angelo and Darren Gergle. An eye for design: Gaze visualizations for remote collaborative work. In *2018 CHI Conference*, pages 1–12, April 2018.

[57] Sarah D'Angelo and Bertrand Schneider. Shared gaze visualizations in collaborative interactions: Past, present and future. *Interacting with Computers*, 33, May 2021.

[58] Mariolino De Cecco, Alberto Fornaser, Paolo Tomasin, Matteo Zanetti, G Guandalini, et al. Augmented reality to enhance the clinician's observation during assessment of daily living activities. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pages 3–21. Springer, 2017.

[59] Henrique G Debarba, Eray Molla, Bruno Herbelin, and Ronan Boulic. Characterizing embodied interaction in first and third person perspective viewpoints. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 67–72. IEEE, 2015.

[60] Mathilde R. Desselle, Ross A. Brown, Allan R. James, Mark J. Midwinter, Sean K. Powell, et al. Augmented and virtual reality in surgery. *Comput. Sci. Eng.*, 22(3):18 – 26, February 2020.

[61] Arindam Dey, Mark Billinghurst, Robert W. Lindeman, and J. Edward Swan. A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. *Front. Robot. AI*, 5, April 2018. Art. no. 37.

[62] Margaret Duff et al. An adaptive mixed reality training system for stroke rehabilitation. *IEEE Trans. Neural Syst. Rehabilitation Eng.*, 18(5):531–541, October 2010.

[63] Pelle Ehn. *Work-Oriented Design of Computer Artifacts.* Almqvist and Wiksell International for Arbetslivcentrum, Stockholm, May 1988.

[64] Norman Erber. *Auditory training.* Alexander Graham Bell Association For The Deaf, 1982.

[65] Alberto J. Espay et al. At-home training with closed-loop augmented-reality cueing device for improving gait in patients with Parkinson disease. *J. Rehabil. Res. Dev.*, 47(6):573–582, August 2010.

[66] Steven Feiner, Blair Macintyre, and Dorée Seligmann. Knowledge-based augmented reality. *Commun. ACM*, 36:53–62, July 1993.

[67] Vincenzo Ferrari et al. Augmented reality visualization of deformable tubular structures for surgical simulation. *Int. J. Med. Robot.*, 12(2):231–240, July 2015.

[68] Javier Ferrer-Torregrosa, Miguel Ángel Jiménez-Rodríguez, Javier Torralba-Estelles, Fernanda Garzón-Farinós, Marcelo Pérez-Bermejo, et al. Distance learning ECTS and flipped classroom in the anatomy learning: Comparative study of the use of augmented reality, video and notes. *BMC Med. Educ.*, 16(1), September 2016. Art. no. 230.

[69] A. G. Fisher and A. Marterella. *Powerful practice: A model for authentic occupational therapy.* CIOTS-Center for Innovative OT Solutions, 2019.

[70] Paul M. Fitts. Perceptual-motor skill learning. In Arthur W. Melton, editor, *Categories of Human Learning*, pages 243–285. Academic Press, 1964.

[71] Nizan Friedman et al. Retraining and assessing hand movement after stroke using the MusicGlove: Comparison with conventional hand therapy and isometric grip training. *J. Neuroeng. Rehabilitation*, 11(1), April 2014. Art. no. 76.

[72] Henry Fuchs, Young Cha, Rohan Chabra, Federico Menozzi, Andrei State, et al. Immersive learning experiences for surgical procedures. *Studies in health technology and informatics*, 220:55–62, April 2016.

[73] Andrea Gaggioli, Emily Keshner, Patrice Weiss, and Giuseppe Riva. *Advanced Technologies in Rehabilitation.* IOS Press - Studies in Health Technology and Informatics, January 2009.

[74] Anthony G. Gallagher et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. *Ann. Surg.*, 241(2):364–372, February 2005.

[75] Henrique Galvan Debarba, Sidney Bovet, Roy Salomon, Olaf Blanke, Bruno Herbelin, et al. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PloS one*, 12(12):e0190109, 2017.

[76] Sergio Garrido-Jurado, Rafael Muñoz Salinas, Francisco Madrid-Cuevas, and Manuel Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47:2280 – 2292, June 2014.

[77] Norina Gasteiger, Sabine N van der Veer, Paul Wilson, and Dawn Dowding. A realist synthesis on how, for whom and in which contexts or conditions augmented and virtual reality training 'works' in upskilling healthcare workers. *JMIR Serious Games*, 10:e31644, June 2021.

[78] Michele Gattullo, Alessandro Evangelista, Antonio Emanuele Uva, Michele Fiorentino, and Joe Gabbard. What, how, and why are visual assets used in industrial augmented reality? A systematic review and classification in maintenance, assembly, and training (from 1997 to 2019). *IEEE Trans. Vis. Comput. Graph.*, 28(2):1443–1456, August 2020.

[79] Hamza Ghandorh, Justin Mackenzie, Roy Eagleson, and Sandrine de Ribaupierre. Development of augmented reality training simulator systems for neurosurgery using model-driven software engineering. In *2017 IEEE 30th Canadian Conf. Electrical and Computer Engineering (CCECE)*, pages 1–6, April 2017.

[80] Shawn M. Glynn and Francis J. di Vesta. Outline and hierarchical organization as aids for study and retrieval. *J. Educ. Psychol.*, 69(2):89–95, 1977.

[81] I. L. Goldstein and J. K. Ford. *Training in Organizations: Needs Assessment, Development, and Evaluation.* Wadsworth, Belmont, 4 edition, 2002.

[82] R. E. Woods Gonzalez R. C. and S. L. Eddins. *Digital Image Processing Using MATLAB.* Pearson Prentice Hall, 2004.

[83] Michihiko Goto, Yuko Uematsu, Hideo Saito, Shuji Senda, and Akihiko Iketani. Task support system by displaying instructional video onto AR workspace. In *2010 IEEE Int. Symp. Mixed and Augmented Reality (ISMAR)*, pages 83–90, October 2010.

[84] Thomas Gregory, Jules Gregory, John Sledge, Romain Allard, and Olivier Mir. Surgery guided by mixed reality: presentation of a proof of concept. *Acta orthopaedica*, 89:480–483, October 2018.

[85] Andrea Grisenti, Giandomenico Nollo, Michela Dalprá, Francesco Natale, Mariolino Cecco, et al. *Technological Infrastructure Supports New Paradigm of Care for Healthy Aging: The Living Lab Ausilia*, pages 85–99. Springer, Cham, February 2021.

[86] Avinash Gupta, Joe Cecil, and Miguel Pirela-Cruz. Role of dynamic affordance and cognitive load in the design of extended reality based simulation environments for surgical contexts. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 652–653, March 2022.

[87] Kunal Gupta, Gun Lee, and Mark Billinghurst. Do you see what i see? the effect of gaze tracking. *IEEE Transactions on Visualization and Computer Graphics*, 22:1–1, July 2016.

[88] Sayera Hafsa, Mazlina Abdul Majid, and Ragad M Tawafak. Learnability factors of AR usage performance: Validating through survey. In *2021 Int.*

Conf. Software Engineering and Computer Systems and 4th Int. Conf. Computational Science and Information Management (ICSECS-ICOCSIM), pages 371–376, August 2021.

[89] Brian Haig and Colin Evers. *Realist inquiry in social science.* London: Sage, January 2016.

[90] Tansel Halic, Sinan Kockara, Coskun Bayrak, and Richard Rowe. Mixed reality simulation of rasping procedure in artificial cervical disc replacement (ACDR) surgery. *BMC Bioinform.*, 11(S6), October 2010. Art. no. S11.

[91] Christian Hansen, Jan Wieferich, Felix Ritter, Christian Rieder, and Heinz-Otto Peitgen. Illustrative visualization of 3d planning models for augmented reality in liver surgery. *International journal of computer assisted radiology and surgery*, 5:133–41, March 2010.

[92] Robert M. Haralick and Linda G. Apiro. *Computer and Robot Vision*, volume 1. Addison-Wesley, 1992.

[93] Jeremia Philipp Oskar Held et al. Augmented reality – based rehabilitation of gait impairments: Case report. *JMIR mhealth uhealth*, 8(5), May 2020. Art. no. e17804.

[94] Thomas Hermann, Andy Hunt, and John G Neuhoff. *The Sonification Handbook.* Logos Verlag, Berlin, 2011.

[95] Alex Hertel. Augmented reality isn't just visual, January 2019. Accessed 26 February 2023.

[96] J. E. Hirsch. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. U.S.A.*, 102(46):16569–16572, November 2005.

[97] Petr Horejsi, Konstantin Novikov, and Michal Simon. A smart factory in a smart city: Virtual and augmented reality in a smart assembly line. *IEEE Access*, 8:94330–94340, May 2020.

[98] C. E. Hughes, C. B. Stapleton, D. E. Hughes, and E. M. Smith. Mixed reality in education, entertainment, and training. *IEEE Comput. Graph. Appl.*, 25(6):24–30, November 2005.

[99] Thomas Hulin, Volker Schmirgel, Eldad Yechiam, Uwe E. Zimmermann, Carsten Preusche, et al. Evaluating exemplary training accelerators for programming-by-demonstration. In *19th Int. Symp. Robot and Human Interactive Communication (RO-MAN)*, pages 440–445, September 2010.

[100] W.E. Ichinose, David Reinkensmeyer, D. Aoyagi, J.T. Lin, K. Ngai, et al. Proceedings of the 25th annual international conference of the ieee engineering in medicine and biology society. In *25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, volume 2, pages 1690–1693 Vol.2, October 2003.

[101] Dal Jae Im et al. Utility of a three-dimensional interactive augmented reality program for balance and mobility rehabilitation in the elderly: A feasibility study. *Ann. Rehabil. Med.*, 39(3):462–472, June 2015.

[102] AccuVein Inc. Accuvein® vein visualization. `https://www.accuvein.com`. Accessed 5 March 2020.

[103] Pier Luigi Ingrassia et al. Augmented reality learning environment for basic life support and defibrillation training: Usability study. *J. Med. Internet Res.*, 22(5), May 2020. Art. no. e14910.

[104] Giulio Jacucci et al. ParticipArt: Exploring participation in interactive art installations. In *2010 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH)*, pages 3–10, November 2010.

[105] David L Jaffe, David A Brown, Cheryl D Pierson-Carey, Ellie L Buckley, and Henry L Lew. Stepping over obstacles to improve walking in individuals with poststroke hemiplegia. *Journal of Rehabilitation Research & Development*, 41, 2004.

[106] Justin Jagosh. Realist synthesis for public health: Building an ontologically deep understanding of how programs work, for whom, and in which contexts. *Annual Review of Public Health*, 40:361–372, April 2019.

[107] Wim GM Janssen, Hans BJ Bussmann, and Henk J Stam. Determinants of the Sit-to-Stand Movement: A Review. *Physical Therapy*, 82(9):866–879, September 2002.

[108] Anthony M. Jarc, Andrew A. Stanley, Thomas Clifford, Inderbir S. Gill, and Andrew J. Hung. Proctors exploit three-dimensional ghost tools during clinical-like training scenarios: A preliminary study. *World J. Urol.*, 35(6):957–965, September 2016.

[109] Allan Javaux et al. A mixed-reality surgical trainer with comprehensive sensing for fetal laser minimally invasive surgery. *Int. J. Comput. Assist. Radiol. Surg.*, 13(12):1949–1957, July 2018.

[110] Gail Jensen, Jan Gwyer, Laurita Hack, and Katherine Shepard. *Expertise in physical therapy practice.* Elsevier Inc., 2007.

[111] Seokhee Jeon, Seungmoon Choi, and Matthias Harders. Rendering virtual tumors in real tissue mock-ups using haptic augmented reality. *IEEE Trans. Haptics*, 5(1):77–84, January 2012.

[112] Carr J.H. and Shepherd R.B. *A motor relearning programme for stroke (2nd edition).* Oxford: Butterworth Heinemann, 1987.

[113] Allison Jing. The impact of gaze cues in mixed reality collaborations. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 473–475, 2021.

[114] Dongsik Jo, Ki-Hong Kim, and Gerard Jounghyun Kim. Spacetime: adaptive control of the teleported avatar for improved ar tele-conference experience. *Computer Animation and Virtual Worlds*, 26(3-4):259–269, 2015.

[115] Kyle Johnsen et al. Mixed reality virtual pets to reduce childhood obesity. *IEEE Trans. Vis. Comput. Graph.*, 20(4):523–530, April 2014.

[116] Mike Jones, James Mueller, and John Morris. Advanced technologies in stroke rehabilitation and recovery. *Topics in stroke rehabilitation*, 17:323–7, September 2010.

[117] Peng-Yuan Kao, Ping-Hsuan Han, Yao-Fu Jan, Zhi-Wei Yang, Chun-Hsien Li, et al. On learning weight distribution of Tai Chi Chuan using pressure sensing insoles and MR-HMD. In *2019 IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*, pages 1457–1464, March 2019.

[118] SIDNEY KATZ. Assessing self-maintenance: Activities of daily living, mobility, and instrumental activities of daily living. *Journal of the American Geriatrics Society*, 31(12):721–727, 1983.

[119] Joseph R. Keebler, Travis J. Wiltshire, Dustin C. Smith, Stephen M. Fiore, and Jeffrey S. Bedwell. Shifting the paradigm of music instruction: Implications of embodiment stemming from an augmented reality guitar learning system. *Front. Psychol.*, 5, May 2014. Art. no. 471.

[120] Sean Kerawala, Vinnie Tieto, and Qian Wen. Try out the MRTK2 examples hub, June 2022.

[121] Emily Keshner and Robert Kenyon. Using immersive technology for postural research and rehabilitation. *Assistive technology : the official journal of RESNA*, 16:54–62, February 2004.

[122] Emily Keshner, Robert Kenyon, and Y Dhaher. Postural research and rehabilitation in an immersive virtual environment. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 7, pages 4862–4865, February 2004.

[123] Emily Keshner, Robert Kenyon, and Jessica Langston. Postural responses exhibit multisensory dependencies with discordant visual and support surface motion. *Journal of vestibular research : equilibrium and orientation*, 14:307–319, February 2004.

[124] Maryam Khademi, Hossein Mousavi Hondori, Lucy Dodakian, Steve Cramer, and Cristina Lopes. Comparing 'pick and place' task in spatial augmented reality versus non-immersive virtual reality for rehabilitation setting. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 2013:4613–4616, July 2013.

[125] Gary Kielhofner. *Conceptual Foundations of Occupational Therapy Practice.* F.A. Davis Co., Philadelphia, 4 edition, 2009.

[126] Seongki Kim, JinHo Ryu, Youngchyul Choi, YooSeok Kang, Hongle Li, et al. Eye-contact game using mixed reality for the treatment of children with attention deficit hyperactivity disorder. *IEEE Access*, 8:45996–46006, March 2020.

[127] Youngjun Kim, Hannah Kim, and Yong Oock Kim. Virtual reality and augmented reality in plastic surgery: A review. *Arch. Plast. Surg.*, 44(3):179–187, May 2017.

[128] Lena von Koch, Annica Wohlin Wottrich, and Lotta Widén Holmqvist. Rehabilitation in the home versus the hospital: The importance of context. *Disabil. Rehabil.*, 20(10):367–372, October 1998.

[129] Birger Kollmeier and Jürgen Kiessling. Functionality of hearing aids: state-of-the-art and future model-based solutions. *International Journal of Audiology*, 57:S3–S28, December 2016.

[130] Varsha Koushik and Shaun Kane. Towards augmented reality coaching for daily routines: Participatory design with individuals with cognitive disabilities and their caregivers. *International Journal of Human-Computer Studies*, 165(2), May 2022. Art. no. 102862.

[131] Bireswar Laha, Doug A. Bowman, and James D. Schiffbauer. Validation of the MR simulation approach for evaluating the effects of immersion on visual analysis of volume data. *IEEE Trans. Vis. Comput. Graph.*, 19(4):529–538, April 2013.

[132] Gun Lee, Seungwon Kim, Youngho Lee, Arindam Dey, Thammathip Piumsomboon, et al. Improving collaboration in augmented video conference using mutually shared gaze. In *2017 International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, November 2017.

[133] Gun A. Lee, Seungjun Ahn, William Hoff, and Mark Billinghurst. Enhancing first-person view task instruction videos with augmented reality cues. In *2020 IEEE Int. Symp. Mixed and Augmented Reality (ISMAR)*, pages 498–508, November 2020.

[134] Nicole Lehrer, Yinpeng Chen, Margaret Duff, Steven L Wolf, and Thanassis Rikakis. Exploring the bases for a mixed reality stroke rehabilitation system, part II: Design of interactive feedback for upper limb rehabilitation. *J. Neuroeng. Rehabilitation*, 8(1), September 2011. Art. no. 54.

[135] Jingya Li, Erik Spek, Jun Hu, and Loe Feijs. Extracting design guidelines for augmented reality serious games for children. *IEEE Access*, 10:66660–66671, January 2022.

[136] Bibeg Hang Limbu, Halszka Jarodzka, Roland Klemke, and Marcus Specht. Using sensors and augmented reality to train apprentices using recorded expert performance: A systematic literature review. *Educ. Res. Rev.*, 25:1–22, November 2018.

[137] Chengyuan Lin et al. How about the mentor? Effective workspace visualization in AR telementoring. In *2020 IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*, pages 212–220, March 2020.

[138] Robert F. Lorch and Elizabeth P. Lorch. Topic structure representation and text recall. *J. Educ. Psychol.*, 77(2):137–148, April 1985.

[139] Alessandro Luchetti, Edoardo Parolin, Isidro Butaslac, Yuichiro Fujimoto, Masayuki Kanbara, et al. Stepping over obstacles with augmented reality based on visual exproprioception. In *2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 96–101, 2020.

[140] Alessandro Luchetti, Paolo Tomasin, Alberto Fornaser, Paolo Tallarico, Paolo Bosetti, et al. The human being at the center of smart factories thanks to augmented reality. In *2019 IEEE 5th International forum on Research and Technology for Society and Industry (RTSI)*, pages 51–56, 2019.

[141] D. Magee, Y. Zhu, R. Ratnalingam, P. Gardner, and D. Kessel. An augmented reality simulator for ultrasound guided needle placement training. *Med. Biol. Eng. Comput.*, 45(10):957–967, July 2007.

[142] Barbara Ostrowski Martin, Klodiana Kolomitro, and Tony C. M. Lam. Training methods: A review and analysis. *Hum. Resour. Dev. Rev.*, 13(1):11–35, October 2013.

[143] Richard E Mayer. *Multimedia Learning.* Cambridge University Press, Cambridge, 2 edition, 2009.

[144] Andreas Mayr, Markus Kofler, Quirbach Ellen, Heinz Matzak, Katrin Fröhlich, et al. Prospective, blinded, randomized crossover study of gait rehabilitation in stroke patients using the lokomat gait orthosis. *Neurorehabilitation and neural repair*, 21:307–14, May 2007.

[145] Daniel Medeiros, Rafael K dos Anjos, Daniel Mendes, João Madeiras Pereira, Alberto Raposo, et al. Keep my head on my shoulders! why third-person is bad for navigation in vr. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pages 1–10, 2018.

[146] Ravish Mehra, Owen Brimijoin, Philip Robinson, and Thomas Lunner. Potential of augmented reality platforms to improve individual hearing aids and to support more ecologically valid research. *Ear & Hearing*, 41:140S–146S, October 2020.

[147] Jonathan Mercier-Ganady, Fabien Lotte, Emilie Loup-Escande, Maud Marchal, and Anatole Lecuyer. The Mind-Mirror: See your brain in action in your head using EEG and augmented reality. In *2014 IEEE Virtual Reality (VR)*, pages 33–38, March 2014.

[148] M. D. Merrill. Component display theory. In C. Reigeluth, editor, *Instructional Design Theories and Models.* Erlbaum Associates, Hillsdale, New Jersey, 1983.

[149] Roberto Miyake, Herbert Zeman, Henrique Duarte, Rodrigo Kikuchi, Eduardo Ramacciotti, et al. Vein imaging: A new method of near infrared

imaging, where a processed image is projected onto the skin for the enhancement of vein treatment. *Dermatologic surgery : official publication for American Society for Dermatologic Surgery [et al.]*, 32:1031–8, September 2006.

[150] Michael Molenda. The programmed instruction era: When effectiveness mattered. *TechTrends*, 52(2):52–58, May 2008.

[151] Theerapat Muangpoon, Reza Haghighi Osgouei, David Escobar-Castillejos, Christos Kontovounisios, and Fernando Bello. Augmented reality system for digital rectal examination training and assessment: System validation. *J. Med. Internet Res.*, 22(8), August 2020. Art. no. e18637.

[152] Francisco Munoz-Montoya, M. Juan, Magdalena Mendez-Lopez, and Camino Fidalgo. Augmented reality based on SLAM to assess spatial short-term memory. *IEEE Access*, 7:2453–2466, December 2018.

[153] Som Naidu, John Barrett, and Peter Olsen. Improving instructional effectiveness with computer-mediated communication. *Res. Learn. Technol.*, 3(2):63–75, 1995.

[154] Yuki Nakamura, Shinya Yamamoto, Morihiko Tamai, and Keiichi Yasumoto. Supporting daily living activities using behavior logs and augmented reality. In *2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 658–663, 2013.

[155] Arthur A. Nelson. Research design: Measurement, reliability, and validity. *Am. J. Health Syst. Pharm.*, 37(6):851–857, June 1980.

[156] St. Catherine University Online OTA. OT vs PT: The difference between fields of occupational therapy and physical therapy. Accessed 5 March 2020.

[157] Karen Otte, Bastian Kayser, Sebastian Mansow-Model, Julius Verrel, Friedemann Paul, Alexander U. Brandt, and Tanja Schmitz-Hübsch. Accuracy and reliability of the kinect version 2 for clinical measurement of motor function. *PLOS ONE*, 11(11):1–17, November 2016.

[158] Matthew J Page et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Brit. Med. J.*, 372, March 2021. Art. no. n71.

[159] J. Paillard. Vectorial versus configural encoding of body space. In V. Knockaert and H. D. Preester, editors, *Body Image and Body Schema*, pages 89–109. John Benjamins Publishing Co., Amsterdam, 2005.

[160] Riccardo Palmarini, John Ahmet Erkoyuncu, Rajkumar Roy, and Hosein Torabmostaedi. A systematic review of augmented reality applications in maintenance. *Robot. Comput. Integr. Manuf.*, 49:215–228, February 2018.

[161] Christos Papakostas, Christos Troussas, Akrivi Krouska, and Cleo Sgouropoulou. Exploration of augmented reality in spatial abilities training: A systematic literature review for the last decade. *Inform. Educ.*, 20(1):107–130, March 2021.

[162] Helder Paraense, Bernardo Marques, Paula Amorim, Paulo Dias, and Beatriz Santos. Whac-A-Mole: Exploring virtual reality (VR) for upper-limb post-stroke physical rehabilitation based on participatory design and serious games. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 716–717, March 2022.

[163] Eunhee Park et al. Effects of a mixed reality-based cognitive training system compared to a conventional computer-assisted cognitive training system on mild cognitive impairment: A pilot study. *Cogn. Behav. Neurol.*, 32(3):172–178, September 2019.

[164] Aftab E Patla. How is human gait controlled by vision. *Ecological psychology*, 10(3-4):287–302, 1998.

[165] Aftab E Patla, Shirley Rietdyk, Christine Martin, and Sandra Prentice. Locomotor patterns of the leading and the trailing limbs as solid and fragile obstacles are stepped over: some insights into the role of vision during locomotion. *Journal of motor behavior*, 28(1):35–47, 1996.

[166] Ray Pawson. *The Science of Evaluation: A Realist Manifesto.* London: Sage, 2013.

[167] Ray Pawson and Nick Tilley. *Realistic evaluation.* Sage, 1997.

[168] Panayiotis Pelargos, Daniel Nagasawa, Carlito Lagman, Stephen Tenn, Joanna Demos, et al. Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. *Journal of Clinical Neuroscience*, 35:1–4, October 2016.

[169] Gert-Jan Pepping and Madeleine A Grealy. *Closing the gap: The scientific writings of David N. Lee.* Psychology Press, 2007.

[170] D. C. Phillips. The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher*, 24:5–12, October 1995.

[171] Tommaso Pisoni et al. AUSILIA: Assisted unit for simulating independent living activities. In *2016 IEEE Int. Smart Cities Conf. (ISC2)*, pages 1–4, September 2016.

[172] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. [poster] covar: Mixed-platform remote collaborative augmented and virtual realities system with shared collaboration cues. In *2017 IEEE International Symposium on Mixed and Augmented Reality*, pages 218–219, October 2017.

[173] Thammathip Piumsomboon, Gun A Lee, Jonathon D Hart, Barrett Ens, Robert W Lindeman, et al. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–13, 2018.

[174] Thammathip Piumsomboon, Youngho Lee, Gun Lee, Arindam Dey, and Mark Billinghurst. Empathic mixed reality: Sharing what you feel and interacting with what you see. In *2017 International Symposium on Ubiquitous Virtual Reality*, pages 38–41, June 2017.

[175] Philip Pratt, Matthew Ives, Graham Lawton, Jonathan Simmons, Nasko Radev, et al. Through the hololens looking glass: augmented reality for extremity reconstruction surgery using 3d vascular models with perforating vessels. *European Radiology Experimental*, 2:2, December 2018.

[176] Zhibao Qin et al. Towards virtual VATS, face, and construct evaluation for peg transfer training of box, VR, AR, and MR trainer. *J. Healthc. Eng.*, 2019, January 2019. Art. no. 6813719.

[177] John Quarles, Samsun Lampotang, Ira Fischler, Paul Fishwick, and Benjamin Lok. Collocated AAR: Augmenting after action review with mixed reality. In *2008 7th IEEE/ACM Int. Symp. Mixed and Augmented Reality (ISMAR)*, pages 107–116, September 2008.

[178] Robert Reimann, Alan Cooper, and Dave Cronin. *About Face: The Essentials of Interaction Design.* John Wiley & Sons, 4 edition, August 2014.

[179] Jun Rekimoto and Katashi Nagao. The world through the computer: Computer augmented interaction with real world environments. *UIST (User Interface Software and Technology): Proceedings of the ACM Symposium*, November 1995.

[180] Jonas Roessel, Moritz Knoell, Jannic Hofmann, and Ricardo Buettner. A systematic literature review of practical virtual and augmented reality solutions in surgery. In *2020 IEEE 44th Annu. Computers, Software, and Applications Conf. (COMPSAC)*, pages 489–498, July 2020.

[181] Nina Rohrbach, Philipp Gulde, Alan Armstrong, Linda Hartig, Anas Abdelrazeq, et al. An augmented reality approach for adl support in alzheimer's disease: a crossover trial. *Journal of NeuroEngineering and Rehabilitation*, 16:11, June 2019.

[182] Edgar Rojas-Muñoz et al. Evaluation of an augmented reality platform for austere surgical telementoring: A randomized controlled crossover study in cricothyroidotomies. *NPJ Digit. Med.*, 3(1), May 2020. Art. no. 75.

[183] Michael Sailer, Jan Ulrich Hense, Sarah Katharina Mayr, and Heinz Mandl. How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Comput. Hum. Behav.*, 69:371–380, April 2017.

[184] Eduardo Salas, Scott I. Tannenbaum, Kurt Kraiger, and Kimberly A. Smith-Jentsch. The science of training and development in organizations. *Psychol. Sci. Public Interest*, 13(2):74–101, June 2012.

[185] Naveen Kumar Sankaran et al. Efficacy study on interactive mixed reality (IMR) software with sepsis prevention medical education. In *2019 IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*, pages 664–670, March 2019.

[186] Danielle Santarelli. Clinical trials - medical device trials, September 2018.

[187] Marc Ericson Santos, Goshiro Yamamoto, Takafumi Taketomi, Jun Miyazaki, and Hirokazu Kato. Authoring augmented reality learning experiences as learning objects. In *2013 IEEE 13th International Conference on Advanced Learning Technologies (ICALT)*, pages 506–507, July 2013.

[188] Marc Ericson C. Santos, Angie Chen, Takafumi Taketomi, Goshiro Yamamoto, Jun Miyazaki, et al. Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Trans. Learn. Technol.*, 7(1):38–56, December 2013.

[189] Joshua Warren Sappenfield et al. Visualization improves supraclavicular access to the subclavian vein in a mixed reality simulator. *Anesth. Analg.*, 127(1):83–89, July 2018.

[190] Ratna Candra Sari, Mahfud Sholihin, Nurhening Yuniarti, Ida Ayu Purnama, and Hardika Dwi Hermawan. Does behavior simulation based on augmented reality improve moral imagination? *Educ. Inf. Technol. (Dordr)*, 26(1):441–463, July 2020.

[191] Dieter Schmalstieg and Tobias Höllerer. *Augmented Reality: Principles and Practice.* Addison-Wesley, Boston, June 2016.

[192] Richard Schmidt and Robert Bjork. New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science - PSYCHOL SCI*, 3:207–217, July 1992.

[193] Richard A. Schmidt. A schema theory of discrete motor skill learning. *Psychol. Rev.*, 82(4):225–260, July 1975.

[194] D. S. Schoeb et al. Mixed reality for teaching catheter placement to medical students: A randomized single-blinded, prospective trial. *BMC Med. Educ.*, 20(1), December 2020. Art. no. 510.

[195] Yoones A Sekhavat and Mohammad S Namani. Projection-based ar: Effective visual feedback in gait rehabilitation. *IEEE Transactions on Human-Machine Systems*, 48(6):626–636, 2018.

[196] Ramiro Serrano Vergel, Pedro Morillo Tena, Sergio Casas Yrurzum, and Carolina Cruz-Neira. A comparative evaluation of a virtual reality table and a HoloLens-based augmented reality system for anatomy training. *IEEE Trans. Hum. Mach. Syst.*, 50(4):337–348, June 2020.

[197] T. Sielhorst, T. Blum, and N. Navab. Synchronizing 3D movements for quantitative comparison and simultaneous visualization of actions. In *4th IEEE/ACM Int. Symp. Mixed and Augmented Reality (ISMAR)*, pages 38–47, October 2005.

[198] Roland Sigrist, Georg Rauter, Laura Marchal-Crespo, Robert Riener, and Peter Wolf. Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. *Exp. Brain Res.*, 233(3):909–925, December 2014.

[199] Bipasha Sinha. Programmed instruction as a teaching strategy. *Int. Educ. Sci. Res. J.*, 5(6):34–38, June 2019.

[200] Traci Sitzmann, Bradford Bell, Kurt Kraiger, and Adam Kanar. A multilevel analysis of the effect of prompting self-regulation in technology-delivered instruction. *CAHRS Working Paper Series*, 62, December 2009.

[201] Traci Sitzmann and Katherine Ely. Sometimes you need a reminder: The effects of prompting self-regulation on regulatory processes, learning, and attrition. *The Journal of applied psychology*, 95:132–44, January 2010.

[202] Randall C. Smith and Peter Cheeseman. On the representation and estimation of spatial uncertainty. *The International Journal of Robotics Research*, 5(4):56–68, 1986.

[203] P. Soille. *Morphological Image Analysis: Principles and Applications.* Springer-Verlag, 1999.

[204] Usha Valli Somasundaram and Toby Marshall Egan. Training and development: An examination of definitions and dependent variables. In *Academy of Human Resource Development Int. Conf. (AHRD)*, pages 850–857, March 2004.

[205] Clay Spinuzzi. The methodology of participatory design. *Technical Communication*, 52(2):163–174, May 2005.

[206] Daniele Sportillo, Alexis Paljic, and Luciano Ojeda. On-road evaluation of autonomous driving training. In *2019 14th ACM/IEEE Int. Conf. Human-Robot Interaction (HRI)*, pages 182–190, March 2019.

[207] Andrei State, David Chen, Chris Tector, Andrew Brandt, Hong Chen, et al. Case study: Observing a volume rendered fetus within a pregnant patient. In *Proceedings Visualization '94*, volume 18, pages 364–368, January 1994.

[208] Andrei State, Mark Livingston, William Garrett, Gentaro Hirota, Mary Whitton, et al. Technologies for augmented reality systems: Realizing ultrasound-guided needle biopsies. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 439–446, August 1996.

[209] James Stewart and Mark Billinghurst. A wearable navigation display can improve attentiveness to the surgical field. *International Journal of Computer Assisted Radiology and Surgery*, 11:1193 —1200, March 2016.

[210] Michele Stocco et al. Augmented reality to enhance the clinical eye: The improvement of ADL evaluation by mean of a sensors based observation. In P. Bourdot, V. Interrante, L. Nedel, N. Magnenat-Thalmann, and G. Zachmann, editors, *Virtual Reality and Augmented Reality*, volume 11883, pages 291–296, Cham, October 2019. Springer International Publishing.

[211] Katherine Sullivan and Pamela Duncan. New perspectives for locomotor training after stroke. *Neurology Report*, 24:55–59, January 2000.

[212] Colin Sutherland, Keyvan Hashtrudi-Zaad, Rick Sellens, Purang Abolmaesumi, and Parvin Mousavi. An augmented reality haptic training simulator for spinal needle procedures. *IEEE. Trans. Biomed. Eng.*, 60(11):3009–3018, December 2012.

[213] John Sweller. Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22:123–138, April 2010.

[214] Paul Taylor, Darlene Russ-Eft, and Wai Lok Chan. A meta-analytic review of behavior modeling training. *The Journal of applied psychology*, 90:692–709, August 2005.

[215] Matthew A Timmis and John G Buckley. Obstacle crossing during locomotion: visual exproprioceptive information is used in an online mode to update foot placement before the obstacle but not swing trajectory over it. *Gait & posture*, 36(1):160–162, 2012.

[216] Daniel Makoto Tokunaga, Ricardo Nakamura, and Romero Tori. Nonphotorealistic 3d video-avatar. In *SIGGRAPH'09: Posters*, pages 1–1, 2009.

[217] Jörg Trojan et al. An augmented reality home-training system based on the mirror training and imagery approach. *Behav. Res. Methods*, 46(3):634–640, December 2013.

[218] Phil Turner. *HCI Redux: The Promise of Post-Cognitive Interaction.* Springer Publishing Company, Incorporated, 1st edition, 2016.

[219] James Van Patten, Chun Chao, and Charles M. Reigeluth. A review of strategies for sequencing and synthesizing instruction. *Rev. Educ. Res.*, 56(4):437–471, December 1986.

[220] Andres Vargas Gonzalez et al. A comparison of desktop and augmented reality scenario based training authoring tools. In *2019 IEEE Int. Symp. Mixed and Augmented Reality (ISMAR)*, pages 339–350, October 2019.

[221] Petr Vávra, J Roman, Pavel Zonča, Peter Ihnát, Martin Němec, et al. Recent development of augmented reality in surgery: a review. *Journal of healthcare engineering*, 2017.

[222] Petr Vávra, Jan Roman, P. Zonča, Peter Ihnát, Martin Nemec, et al. Recent development of augmented reality in surgery: A review. *Journal of Healthcare Engineering*, 2017:1–9, August 2017.

[223] R Vitório, E Lirani-Silva, F Pieruccini-Faria, Renato de Moraes, LTB Gobbi, et al. Visual cues and gait improvement in parkinson's disease: which piece of information is really important? *Neuroscience*, 277:273–280, 2014.

[224] Juan Wachs. Gaze, posture and gesture recognition to minimize focus shifts for intelligent operating rooms in a collaborative support system. *Int. J. of Computers*, V:106–124, March 2010.

[225] Caroline Wakefield and Dave Smith. Perfecting practice: Applying the PETTLEP model of motor imagery. *Journal of Sport Psychology in Action*, 3(1):1–11, March 2012.

[226] Richard D Walk and Herbert L Pick. *Intersensory perception and sensory integration*, volume 2. Springer Science & Business Media, 2013.

[227] Shiyao Wang et al. Augmented reality as a telemedicine platform for remote procedural training. *Sensors*, 17(10), October 2017. Art. no. 2294.

[228] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[229] Sean White and Steven Feiner. Sitelens: situated visualization techniques for urban site visits. *Conference on Human Factors in Computing Systems - Proceedings*, pages 1117–1120, April 2009.

[230] Dennis Wolf, Daniel Besserer, Karolina Sejunaite, Matthias Riepe, and Enrico Rukzio. care: An augmented reality support system for dementia

patients. In *31st User Interface Software and Technology Symposium*, pages 42–44, October 2018.

[231] Geoff Wong, Trish Greenhalgh, Gill Westhorp, Jeanette Buckingham, and Ray Pawson. Rameses publication standards: realist syntheses. *BMC Medicine*, 11, January 2013.

[232] Geoff Wong, Gill Westhorp, Ana Manzano, Joanne Greenhalgh, Justin Jagosh, et al. Rameses ii reporting standards for realist evaluations. *BMC Medicine*, 14:1–18, June 2016.

[233] World Health Organization. The need to scale up rehabilitation. Accessed 10 January 2020.

[234] Trinette Wright, Sandrine Ribaupierre, and Roy Eagleson. Design and evaluation of an augmented reality simulator using Leap Motion. *Healthc. Technol. Lett.*, 4(5):210–215, September 2017.

[235] C T Yeo, T Ungi, P U-Thainual, A Lasso, R C McGraw, et al. The effect of augmented reality training on percutaneous needle placement in spinal facet joint injections. *IEEE. Trans. Biomed. Eng.*, 58(7):2031–2037, July 2011.

[236] Wang Ying and Wang Aimin. Augmented reality based upper limb rehabilitation system. In *13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, pages 426–430, October 2017.

[237] Yanxia Zhang, Ken Pfeuffer, Ming Chong, Jason Alexander, Andreas Bulling, et al. Look together: using gaze for assisting co-located collaborative search. *Personal and Ubiquitous Computing*, 21, February 2017.

[238] Shang Zhao et al. An intelligent augmented reality training framework for neonatal endotracheal intubation. In *2020 IEEE Int. Symp. Mixed and Augmented Reality (ISMAR)*, pages 672–681, November 2020.

[239] Chao Zhou, Hong Li, and Yulong Bian. Identifying the optimal 3D display technology for hands-on virtual experiential learning: A comparison study. *IEEE Access*, 8:73791–73803, April 2020.

[240] Yao Zhou, Jufan Zhang, and Fengzhou Fang. Vergence-accommodation conflict in optical see-through display: review and prospect. *Results in Optics*, 5:100160, December 2021.

[241] Mark Zolotas, Joshua Elsdon, and Yiannis Demiris. Head-mounted augmented reality for explainable robotic wheelchair assistance. In *2018 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, pages 1823–1829, October 2018.

Publication List

Peer Review Journal Paper

1. **Isidro Butaslac**, Yuichiro Fujimoto, Taishi Sawabe, Masayuki Kanbara, Hirokazu Kato, "Systematic Review of Augmented Reality Training Systems", *IEEE Transactions on Visualization and Computer Graphics*, Early Access, August 2022. (Chapter 2)

Peer Review International Conference

1. **Isidro Butaslac**, Alessandro Luchetti, Junya Ino, Yuichiro Fujimoto, Taishi Sawabe, Masayuki Kanbara, Hirokazu Kato, Keisuke Uemura, Yoshito Otake, Yoshinobu Sato, Masaki Takao, Nobuhiko Sugano, "Application of Participatory Design Methodology in AR: Developing Prototypes for Two Context Scenarios", In *Proceedings of the 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct*, Early Access, Singapore, October 2022. (Chapter 3)
2. Alessandro Luchetti, **Isidro Butaslac**, Manuel Rosi, Damiano Fruet, Giandomenico Nollo, Patrizia Ianes, Francesco Pilla, Barbara Gasperini, Giovanni Guandalini, Jacopo Bonavita, Hirokazu Kato, Mariolino De Cecco, "Multidimensional assessment of daily living activities in a shared Augmented Reality environment", In *Proceedings of the 2022 IEEE International Workshop on Metrology for Living Environment*, pp. 60–65, Cosenza, Italy, May 2022. (Chapter 4)
3. Alessandro Luchetti, Edoardo Parolin, **Isidro Butaslac**, Yuichiro Fujimoto, Masayuki Kanbara, Paolo Bosetti, Mariolino De Cecco, Hirokazu Kato, "Stepping over Obstacles with Augmented Reality based on Visual Exproprioception", In *Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct*, pp. 96–101, Recife, Brazil, November 2020. (Chapter 4)
4. **Isidro Butaslac**, Alessandro Luchetti, Edoardo Parolin, Yuichiro Fujimoto, Masayuki Kanbara, Mariolino De Cecco, Hirokazu Kato, "The Feasibility of Augmented Reality as a Support Tool for Motor Rehabilitation", In *Augmented Reality, Virtual Reality, and Computer Graphics*, P. Bourdot, L. De Paolis, Eds. Springer Cham, vol 12243, pp. 165–173, August 2020. (Chapter 4)