# A Comparative Evaluation of a Virtual Reality Table and a HoloLens-Based Augmented Reality System for Anatomy Training

Ramiro Serrano Vergel, Pedro Morillo Tena, Sergio Casas Yrurzum, and Carolina Cruz-Neira, *Senior Member, IEEE*

***Abstract*—Anatomy training with real cadavers poses many practical problems for which new training and educational solutions have been developed making use of technologies based on real-time 3-D graphics. Although virtual reality (VR) and augmented reality (AR) have been previously used in the medical field, it is not easy to select the right 3-D technology or setup for each particular problem. For this reason, this article presents a comprehensive comparative study with 82 participants between two different 3-D interactive setups: an optical-based AR setup, implemented with a Microsoft HoloLens device, and a semi-immersive setup based on a VR Table. Both setups are tested using an anatomy training software application. Our primary hypothesis is that there would be statistically significant differences between the use of the AR application and the use of the VR Table. Our secondary hypothesis is that user preference and recommendation for the VR setup would be higher than for the HoloLens-based system. After completing two different tasks with both setups, the participants filled two questionnaires about the use of the anatomy training application. Three objective measures are also recorded (time, number of movements, and a score). The results of the experiments show that more than two-thirds of the users prefer, recommend, and find more useful the VR setup. The results also show that there are statistically significant differences in the use of both systems in favor of the VR Table.**

***Index Terms*—Anatomy, augmented reality, comparative study, Microsoft HoloLens, training, virtual reality (VR), VR table.**

## I. Introduction

RECENT advances in virtual reality (VR) and augmented reality (AR), especially regarding visualization and interaction hardware, have pushed these technologies close to the maturity point of mass adoption. Both technologies are already being used in a wide range of areas, such as education, psychology, entertainment, retail, construction, cultural heritage, tourism, etc., with many different applications such as training, skill learning, maintenance, repair, quality control, or safety awareness.

Medicine, and anatomy training in particular, is one of the many areas in which the use of VR and AR is being explored. However, these technologies still need to break out of the laboratory and become a part of day-to-day use. In addition, it is not always easy to choose the right technology or setup for each particular application given the wide spectrum of variants and flavors in which these two technologies can be instantiated. From immersive VR—with head-mounted displays (HMDs) [1] or CAVEs (CAVE Automatic Virtual Environment) [2]—and semi-immsersive VR—such as visionariums [3] or VR tables [4], to video-based AR [5], mobile AR [6], spatial AR [7], AR mirrors [8], or optical-based AR [9], a large set of options is available. Therefore, the selection of the right technology and setup for each application must be done carefully, especially in medicine, which is a highly sensitive area with many social, legal, and economic implications.

This article examines this question by providing a comprehensive comparison between two different 3-D interactive setups in the use of anatomy training: an optical-based AR setup and a semiimmersive VR setup. A Microsoft HoloLens device was chosen for the former, whereas a VR Table is utilized for the latter. The comparison is based on two tasks performed by a series of non-IT-related volunteers with a virtual cadaver on an anatomy training application that we have developed specifically for this research and that has been instantiated and customized for the aforementioned AR and VR setups. The use of a virtual cadaver is justified because of the many problems that real cadavers pose: they can be expensive to acquire and maintain, anatomy training with real cadavers is slow since many students usually need to share a single body, they raise ethical questions, they deteriorate, deflate, become black after some time, and the chemicals used to preserve them pose health risks—formaldehyde, used to preserve cadavers, is classified as carcinogenic to humans by the WHO.

Our primary hypothesis is that there would be statistically significant difference between the use of the AR application and the use of the VR Table for anatomy training. Our secondary hypothesis is that users would prefer and recommend the VR setup over the HoloLens-based setup.

Manuscript received January 7, 2019; revised August 8, 2019 and December 20, 2019; accepted February 22, 2020. This work was supported by Spanish MINECO and EU ERDF programs under Grant RTI2018-098156-B-C55 and Grant RTC-2016-5336-7. This paper was recommended by Associate Editor M. Lau. *(Corresponding author: Sergio Casas Yrurzum.)*

Ramiro Serrano Vergel is with Emerging Analytics Center, University of Arkansas at Little Rock, Little Rock, AR 72204 USA (e-mail: rxserranove@ualr.edu).

Pedro Morillo Tena and Sergio Casas Yrurzum are with Institute on Robotics and Information and Communication Technologies, University of Valencia, 46010 València, Spain (e-mail: pedro.morillo@uv.es; sergio.casas@uv.es).

Carolina Cruz-Neira is with Agere Chair in Computer Science, University of Central Florida, Orlando, FL 32816 USA (e-mail: carolina@ucf.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/THMS.2020.2984746

To the best of our knowledge, this is the first experiment performed using such a comparative analysis with these type of devices. This analysis will help researchers develop advanced VR and AR healthcare setups for instructing and educating in medical procedures using these types of technologies.

The rest of the article is organized as follows. Section II reviews related works about comparative studies with VR and AR and on the use of VR/AR for anatomy training. Section III describes the materials and methods utilized to perform the experiments. Section IV details the experimental study. In Section V, the results of these experiments are presented and discussed. Finally, Section VI outlines the future work and shows the conclusions of the article.

## II. Related Work

The technologies of VR and AR have a relatively long history with thousands of academic works and applications in which these two paradigms have been adapted to different needs. Both are based on virtual interactive 3-D elements. However, they provide two different interaction models, whereas the goal of VR is to extract the users from their actual world and make them believe they belong to a virtual synthetic world, AR aims to provide a seamless integration between real and virtual objects emphasizing the interaction of the user with the real world.

Regarding the two setups used in this research, several applications and research works have been published using either a HoloLens or a VR Table. With respect to the latter, the nomenclature in the academic literature is not consistent and many different proposals can be found based on the idea of a “virtual table” but with dissimilar setups and names, since the concept is rather broad and has been a recurrent setup in the field for decades. Examples of virtual tables can be found in [4], [10]–[12]. There are also several recent commercial VR Tables [13]–[15] for anatomy training with promising results. However, unlike our proposal, these virtual tables do not offer tracking nor stereoscopy. Therefore, their inclusion as VR devices is controversial. Regarding the HoloLens device, despite being a fairly recent product, several studies and applications of this technology can be found in the academic literature [16]–[20]. Examples in the medical field are also not uncommon [9], [21]–[25] and the use of this device is being explored for surgery and other applications.

Although the HoloLens is seen as a promising step and can be already used in its current form for different applications, some limitations related to the limited field of view (FoV) [25], the head pose estimation [26], and the hand tracking and gesture recognition accuracy [27] are reported in the academic literature. These problems need to be solved before the general public accepts this technology. Other tracking systems, such as the HTC Vive Tracker, which is also used in our experiments (in the VR-based setup), also have tracking limitations. In the case of this device, a 5 m range and a 120° FoV angle for the base to properly detect the tracker [28].

Given the large amount of possibilities that VR and AR bring, many applications of VR and AR to the medical field can be found in the literature [29]. Surgery is one of the most researched applications [30], [31] but it is not the only one. Virtual environments are also used in treatment, such as in autism spectrum disorders [32], pain-relief, such as in patients with phantom-limb pain [33], rehabilitation [34], and training [35]. Anatomy training applications based on VR and/or AR have also been proposed [36]–[45]. Although not all the proposals have been properly assessed, research suggests that VR and AR are valid alternatives for anatomy training, as they are for other medical and nonmedical uses.

On the other hand, many comparative analyses involving VR or AR have been performed and documented. Some research works perform a comparison between VR and AR versions of the same or a similar application, analyzing presence, user satisfaction, preference, or performance. Examples of this are [46]–[49]. No clear general conclusions can be drawn from this body of research, since the suitability of one or other paradigm for each specific case depends highly on the application in which they are used. Other works assess different devices for the same technology, either VR or AR, and the same application [50]–[54]. Other researchers compare traditional methods versus VR/AR alternatives [55]–[58]. The general conclusion is that VR and AR provide benefits compared to traditional methods, although these benefits should be tested individually for each particular application, since they can be significantly different.

There are also a few cases of these types of comparisons of VR/AR applications for anatomy training [59]–[62]. One recent work [63] compares VR, AR, and a tablet-based application for anatomy training. The experiment focuses on learning enhancement, although there is no control group with traditional learning. The authors do not find statistically significant differences between the three setups but show the benefits that these technologies could bring. The main conclusion that can be drawn from this research area is that AR and VR can be successfully used for anatomy training. Some studies conclude that they can be used instead of traditional approaches with textbooks, master classes, plastic models, or even real cadavers [60], whereas some authors prefer that they supplement traditional methods instead of replacing them [62]. Although it is likely that real cadavers will retain a place in medical schools, it is necessary to explore other ways for anatomy training, given the problems that they pose. For this reason, this research further explores the feasibility of using VR and AR for anatomy training, testing two particular setups. To the best of our knowledge, this is the first work comparing an optical-based AR application and a VR Table for anatomy training.

## III. Materials and Methods

Two different hardware stations were used to perform this research and evaluate the anatomy training application. The first one uses a Microsoft HoloLens as its main device; the second one is a VR Table based on a 3-D TV with stereoscopy. The hardware and software used in each of these stations is described in detail in the following sections.

### A. HoloLens-Based Station

This station was set up with a table, a 50″ inch Sony Bravia TV and a Microsoft HoloLens device (see Fig. 1). The TV was only used to visualize what the user was seeing during

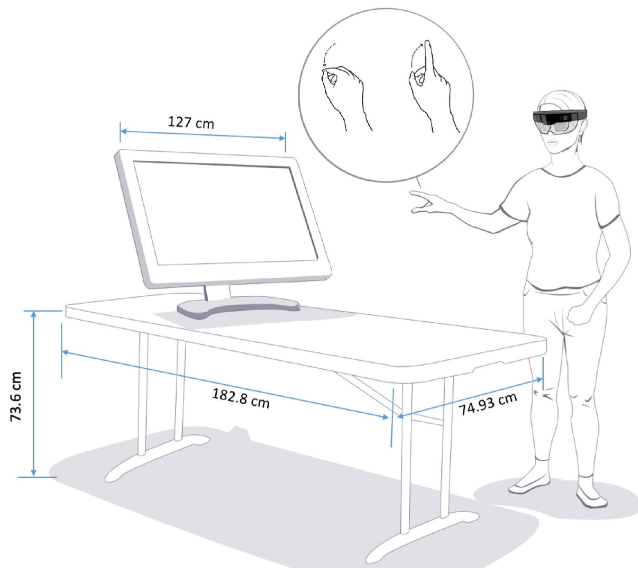


Fig. 1. Schema of the AR-based setup.

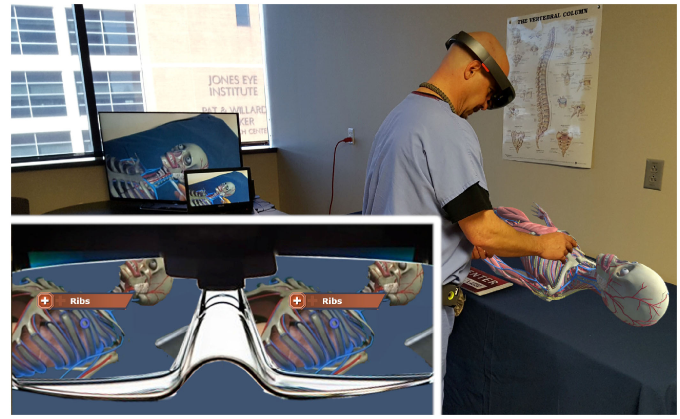


Fig. 2. User, wearing the HoloLens, using the AR-based setup.

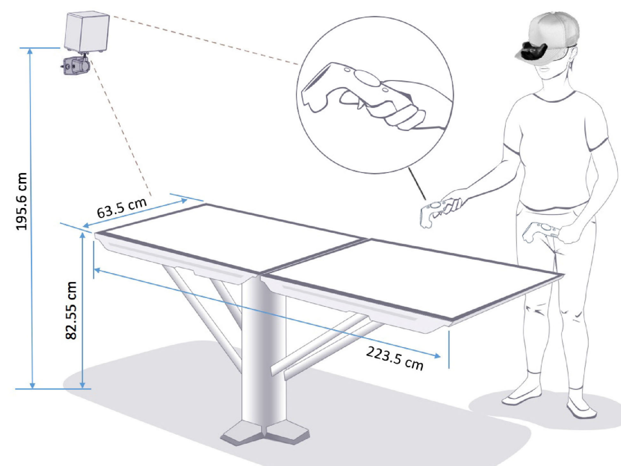


Fig. 3. Schema of the VR-based setup.

the experiment, streaming the video captured by the HoloLens onto the television. This device consists of a pair of mixed reality (MR) smart glasses developed and manufactured by Microsoft. It is a see-through HMD with a built-in PC using Windows 10. Although the device is marketed by Microsoft as a MR system, we refer to the HoloLens as an AR device, since the term AR is more specific than the broader term MR, which includes anything between the extrema of the Milgram's virtuality continuum [64].

The HoloLens features an inertial measurement unit, which includes an accelerometer, a gyroscope, and a magnetometer, four "environment understanding" sensors, a depth camera with a 120° × 120° FoV, a 2.4-megapixel camera, a four-microphone array, and an ambient light sensor.

In addition to an *Intel Cherry Trail* system on a chip, containing the CPU and the GPU, HoloLens features a custom-made *holographic processing unit*, which uses 28 custom DSPs to process and integrate data from the sensors, as well as handling tasks such as spatial mapping, tracking, gesture recognition, and voice and speech recognition. The HoloLens contains an internal rechargeable battery with an average life rated at 2–3 h of active use. It also features IEEE 802.11ac Wi-Fi and Bluetooth 4.1 low energy wireless connectivity. Fig. 2 shows a user testing the HoloLens-based station. The estimated cost of setup A is $5000.

The *a priori* advantage of the HoloLens is that it is wireless and allows avoiding the use of unnatural hardware controllers. A detailed evaluation of the accuracy of this device can be found in [26] and [65].

### B. VR Table Station

This station was set up with two Sony Bravia 50″ inch 3-D TVs and a tracking system to recognize the user's position and the interaction events (see Fig. 3). The two TVs were arranged side-to-side and placed with the display on the horizontal plane facing up. The rationale for using this setup is to provide a forensic table metaphor with a 1:1 virtual depiction of the cadaver that needs to be used instead of a real one for anatomy training. The total size of the display of the VR Table is 223.5 × 63.5 cm, enough for showing a virtual human body without scaling the model down. Stereoscopy with 3-D active shutter glasses provides depth cues since proper depth perception is essential for anatomy training. The table also provides sound cues by means of the TVs speakers.

The tracking system was based on the same system used by the HTC Vive, but instead of using an HMD, we used a tracker that was placed in a cap so we could locate the position of the user's head and estimate the user's pose. The components of the tracking and interface system were two *SteamVR Base Stations*, two *Vive Controllers,* and a *Vive Tracker*. A detailed evaluation of HTC Vive tracking accuracy and latency can be found in [28] and [66]. Fig. 4 shows a user testing the VR-based station. As it can be seen, this is not an immersive VR system, because the user is not fully immersed in the virtual world. However, it provides head tracking, a large visual field, depth cues, sound and real-time interaction with the virtual elements. Therefore, it can be classified as a semi-immersive VR system [67].

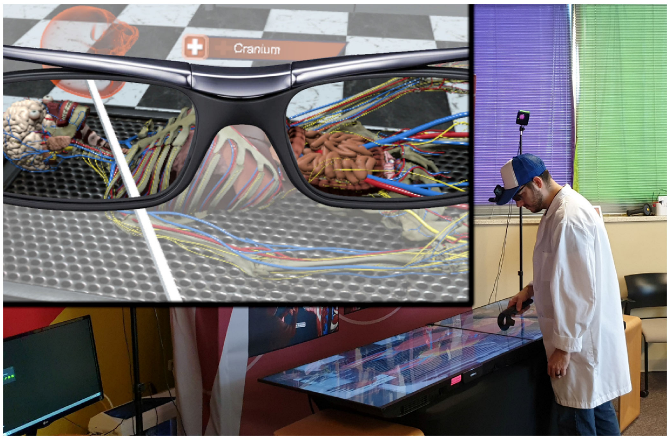


Fig. 4. User, wearing a trackable cap, using the VR-based setup.

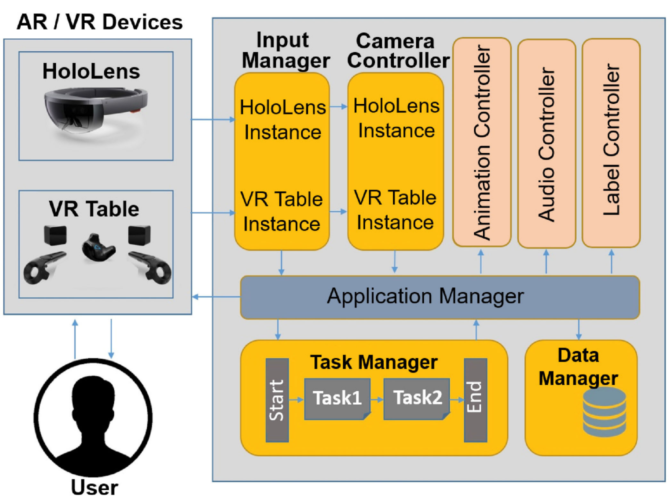


Fig. 5. Components of the anatomy training application.

The VR Table includes also a PC running Windows 10 Pro 64-bit with an Intel Core i7-4790 CPU at 3.60 GHz, 32 GB of RAM, and an Nvidia GeForce GTX 1080 graphics card. The estimated cost of setup B is $3500.

### C. Software

We used Unity3D v.2017.1.0 to build the anatomy training application. This tool was selected because it allows deploying the same software for different devices. C# was chosen to program the scripts in Unity, using Visual Studio 2010 to code and debug them. 3DS Max and Adobe Photoshop were used to create the 3-D model of the virtual cadaver. A Windows 10 operating system was used to host and run all the software. The frame rate of the application was locked at 30 frames/s, in order to provide comparable visualization systems.

The anatomy training application is composed of several modules (see Fig. 5). The *Animation Controller* is the component that manages the animations related to the location and displacement of the parts of the virtual cadaver. For example, when a user moves a body element to a forensic tray, an animation is triggered and controlled by this component. The *Label Controller* is responsible for tagging each part of the virtual cadaver using a predefined set of labels, so each time the user selects a part of the virtual cadaver a label appears for 4 s. Labels are 20-cm camera-facing billboards, which appear next to the highlighted element. The *Camera Controller* manages the cameras of the application. A camera in the anatomy training application always represents the user's point of view. Therefore, it is important to be able to control the rotation and translation of the camera depending on the device that is being used. Since we define two different instances for the two setups (HoloLens and VR Table), the Camera Controller is a dual component, which in both cases generates an off-axis parallel stereo pair with asymmetric frusta. The *Audio Controller* is responsible for the activation of sounds according to the user's interaction in the application. For example, two different sounds are used to identify when an object has been grabbed or dropped, respectively. The *Input Manager*, like the Camera Controller, is a dual component. It has two different instances that handle the interaction of the two types of interaction systems used by the two different setups. The first instance handles the data collected by HoloLens' sensors and performs a translation of these to the events of the application. Gestures and motion are the input elements utilized in this setup. The second instance handles the inputs to the VR Table. These are the inputs generated by the Vive Controller and the Vive Tracker. The *Task Manager* is responsible for initiating and controlling the execution of the tasks defined in the application. For instance, if the user needs to grab the heart, this component is responsible for verifying the status and restrictions of the task before allowing the heart to be removed. The *Data Manager* is responsible for monitoring the time required by each user to complete the assigned tasks, as well as the registration of the events performed by the user. This information is presented on the screen for its registration and subsequent analysis. Finally, the *Application Manager* is responsible for coordinating all other components. In other words, it is responsible for communicating the status of the application to the other components and for triggering the events or elements that must be raised to respond to the actions defined in the system.

It is important to point out that, as seen in Fig. 5, both setups share most of the software modules (labels, audio, animations, etc.). Thus, the tasks and the visual (labels, color, font, size, etc.) and acoustic information about the body elements are generated in the same way for both setups.

## IV. Experimental Study

### A. Protocol Design

In order to decide which of the two setups is the most appropriate for anatomy training, we designed a set of tasks to compare these two 3-D interactive setups. We ran our experimental study in the *Jackson T. Stephens Spine and Neurosciences Institute* at the University of Arkansas for Medical Sciences (UAMS). Since the goal of this research is to study the suitability of using virtual cadavers for anatomy training, students and physicians who did not have any previous experience using AR or VR technologies (for anatomy training) were recruited for the experiment. The experiment was announced in medical schools and departments related with physiology and basic anatomy courses, performing a probability sampling to randomly choose some of the participants of these courses. As a result, 89 people volunteered and registered to perform the experiment by signing up online.

When a person arrived at the laboratory to perform the experiment, we carried out the following protocol.

*Presentation and description:* Before proceeding with the experiment, users were provided with a description about the tasks they had to complete and the maximum time available to complete them (40 min in total). Then, users were required to sign a compulsory informed consent where they declared to agree with the terms of the experiment. The experiment was divided into two similar sections, one section for each hardware station, at the end of which users had to complete a survey about the experience of using the anatomy training application with the corresponding setup (AR or VR).

*Instruction and practice:* Before the start of each section, users received a short briefing on how to use the HoloLens or the VR Table. In both cases, a free practice of 5 min was carried out on three main actions: select, drag, and drop.

*Experiment:* The experiment was designed so that the users would perform the same set of tasks in the two different setups. Keeping that in mind, users were divided into two different groups, as other similar works propose [46], [50], [54]. Group A included the users who started the experiment with the HoloLens first and then used the VR Table. Group B included the users who tried the VR Table first and then, the HoloLens. The experiment consisted in two different tasks that users needed to complete using the minimum possible time. During the experiment, user events were monitored in such a way that, at the end of each task, a score was generated based on the time required to complete each task and on the number of movements performed versus the expected number. The application is designed to count the number of times an interactive body element (those that are designed to be movable) was displaced from its previous position.

*Evaluation:* After users finished the tasks in the first station, they were prompted to complete Questionnaire 1. Table I lists the questions asked in this questionnaire. These questions were grouped in the six factors proposed by Witmer to measure presence in virtual environments [68]: sensory factors (SF), control factors (CF), distraction factors (DF), ergonomic factors (EF), realism factors (RF), and other factors (OF). This six-factor questionnaire includes questions related to the interface of the application as well as some questions concerning the actual content (especially related to the RFs) with the aim of emphasizing that there are two UX/UI levels when performing usability evaluations [69]–[71]. Once Questionnaire 1 was finished, users moved to the station employing the other setup. When the tasks in this second station were also completed, users were then asked to complete Questionnaire 2 about the second setup. The questions of Questionnaire 2 were the same questions listed in Questionnaire 1 plus six additional questions listed in Table II about user preference and recommendation regarding these two setups.

TABLE I
QUESTIONNAIRE 1

| ***Question (factor)*** |
|---|
| The information displayed on the device was adequate. (SF) |
| It was easy to handle the VR/AR application. (CF) |
| The information displayed on the device was easy to read. (SF) |
| The information displayed on the VR/AR device was clear. (SF) |
| It was easy to handle the device and its accessories. (CF) |
| I did not have to strive to recognize the instructional elements as 3D elements. (RF) |
| The 3D virtual elements looked like real. (RF) |
| The handling of the device and its accessories was simple and without complications. (CF) |
| The system responded to my actions adequately. (CF) |
| The handling of the system and its accessories was natural. (CF) |
| I did not feel delays between my actions with the device and the expected results. (CF) |
| The control mechanisms of the VR/AR (glasses, lights, surface, etc.) did not distract me. (DF) |
| I got used to the VR/AR application for medical purposes and the device. (CF) |
| The device and the application were easy to use. (EF) |
| I found very useful the information provided by the VR/AR application to complete the actions. (RF) |
| I had the impression that the aid elements appeared in 3D on the device. (RF) |
| The VR/AR application helped me to complete the require actions on the virtual human body. (RF) |
| I found the instructional elements to be useful. (RF) |
| I had the impression that the 3D labels, required for the task, were part of the scene. (RF) |
| There were moments that I thought that the elements that appeared on the device were real. (RF) |
| I did not pay attention to differences between the instructional elements and the actual device. (RF) |
| I had the impression that I could have touched the items that appear in the VR/AR. (RF) |
| I liked the visual aids to help me complete the task. (RF) |
| I liked how virtual elements correlate with the actual device. (RF) |
| I have not felt any kind of discomfort during the experience (dizziness). (EF) |
| I have felt the sensation of going in motion with this system. (OF) |
| I would like to use this technology with other uses. (OF) |
| The use of the system and its accessories was comfortable for my legs and arms. (EF) |
| I liked the experience of using a VR/AR application with a virtual human body. (OF) |
| The use of the system did not require a great effort from the legs or arms. (EF) |
| The use of the system did not require a great mental effort. (SF) |
| I have focused on the actions I had to do and not on the system or the environment. (DF) |
| My arms and legs are not tired after the experiment. (DF) |
| I felt involved during the experience. (OF) |
| At the end of the experience, I was an expert in the management of the system. (CF) |
| Rate the feeling of 3D (depth perception). |
| Rate the system as a device. |
| Rate the usefulness of the system as a system to train medical students, new surgeons or help medical professionals refresh their skills. |

## B. Task Description

As previously mentioned, the same tasks were used to compare the interaction performed on the two setups used in the experiment, which was divided into two tasks. The first task corresponds to a low level of complexity where the main objective was to locate the heart of the virtual cadaver. Then, the user

TABLE II
ADDITIONAL QUESTIONS IN QUESTIONNAIRE 2

| # | ***Question*** |
|---|---|
| Q1-Q2 | Which system did you find most useful? Why? |
| Q3-Q4 | Which system did you like the most? Why? |
| Q5-Q6 | What system would you recommend to be used as a part of an anatomy training? Why? |

had to select, grab, and place the heart on a forensic tray that was arranged next to the cadaver. In order to grab the heart, the user first needed to remove three organs (ribs, left lung, and sternum) of the virtual cadaver. Thus, this task requires a minimum of four movements. In the VR setup, the system showed the user a text note identifying the first step of the task to be performed (removal of peripheral organs). Next, a 3-D label instructed the user to remove the first organ (ribs). Since the HTC Vive Controller is used as a basic laser pointer, the user can select a body element in the virtual cadaver if it is pointed for a time longer than 1 s. Once the right element was selected, the user could grab and place the organ on the forensic tray by pressing and holding the trigger button of the HTC Vive Controller. This process is repeated for the three organs included in the first step of the task. A similar process has to be completed to grab and place the heart on a forensic tray once the peripheral organs were extracted. An analogous procedure is followed in the AR setup in order to preserve a similar visual and overall simulation quality. In this case, the pinch gesture in HoloLens allows the user to select, grab, and move the organs.

The second task involved a greater degree of complexity since the user was instructed to assemble the respiratory system from its components. This task consisted on selecting, grabbing, and moving nine parts of the respiratory system—that were previously placed in a tray—to their corresponding place in the virtual human cadaver. Thus, this task requires a minimum of nine movements. Users used the HTC Vive Controller or the pinch gesture in the VR or AR setup, respectively, to select, grab, and move the cadaver organs with an operation similar to that described for the first task.

Several competencies are needed to acquire anatomy knowledge [72]. One of these competencies is the location of organs and tissues, especially those that are interconnected forming an apparatus, such as the respiratory system, as in Task 2. Another important competence is the identification of body elements, and the spatial relationship between them, as in Task 1. Spatial ability is also an important factor in anatomy knowledge [73]–[75]. Therefore, given the absence of consensus about how to integrate computer-based tools in anatomy education [76], we have chosen two tasks that involve spatial ability and also location and identification.

A survey of the state-of-the-art also reveals the existence of different tasks in the assessment of computer-based anatomy tools. For instance, in [63], participants were not asked to perform a particular anatomy-related task. Instead, pre- and post-lesson knowledge tests were performed after a 10-min lesson on skull anatomy using VR, AR, and a tablet. In [45], however, participants were asked to identify and assemble canine bones with a VR application. Similarly, in [77], a 3-D virtual puzzle assembly task is used as part of an anatomy learning application. These tasks are similar to our Task 2.

### C. Participants and Groups

The first objective of this experimental study is to discover if there are differences in the use of the two systems for anatomy training. The second objective is to analyze which setup would be recommended or preferred by the users. These goals correspond to the two initial hypotheses of the research.

To achieve these objectives, we carried out a study involving 82 valid participants. We initially started with 89 participants. However, only 82 people signed the informed consent and successfully completed the experiment. On the other hand, although the experiment was completely anonymous, 14 participants refused to allow that we recorded the times and movements they required to complete the tasks.

From these 82 participants, 56 of them were women (68.29%) and 26 men (31.71%). The participants' ages ranged between 18 and 66. The mean age and standard deviation was $30.98 \pm 12.97$. We split the participants into two groups of 42 and 40 people (denoted as groups A and B), randomly assigning the participants to each group. Group A was composed of 31 women and 11 men, whereas group B included 25 women and 15 men.

Several metrics were obtained during and after the anatomy training experiment. The measurements came from the participants (through the questionnaires shown in Tables I and II) and from the application (objective measures about user performance such as time, movements, and a score calculated by the application). The questions listed in Table I are 7-scale Likert questions with 1 meaning *strongly disagree*, 2 *disagree*, 3 *somewhat disagree*, 4 *neutral*, 5 *somewhat agree*, 6 *agree,* and 7 *strongly agree*, except for the last three questions where 1 means *poor*, 2 *bad*, 3 *somewhat bad*, 4 *neutral*, 5 *positive*, 6 *good,* and 7 *excellent*. The questions shown in Table I are designed to test hypothesis 1, whereas the questions listed in Table II are asked to verify hypothesis 2.

Instead of analyzing the responses for each question separately, we grouped the first type of Likert-scale questions in the six factors described by Witmer [68], which is a well-known approach for assessing virtual environments.

As previously explained, the participants of this article were split into two groups. The reason behind this separation is to check if the order in which the two systems were used to complete the tasks of the experiment has a noticeable effect on how users perceive each system. Participants in group A first tested the AR application, filled out Questionnaire 1, then tested the VR application and finally filled out Questionnaire 2. Similarly, participants in group B first tested the VR application, filled out Questionnaire 1, then tested the AR application, and finally filled out Questionnaire 2.

## V. RESULTS AND DISCUSSION

This section presents the results of the statistical analysis of the data obtained from the experiments with the anatomy

TABLE III
STUDY OF STATISTICALLY SIGNIFICANT DIFFERENCES BETWEEN THE AR AND THE VR APPLICATION—FACTORS, 3-D PERCEPTION, SCORE, AND USEFULNESS

| | *AR (mean ± SD)* | *VR (mean ± SD)* | *t* | *p* | *Cohen's d* |
|---|---|---|---|---|---|
| *SF* | *6.214 ± 0.893* | *6.444 ± 0.530* | *-1.406* | *0.005* | *-0.323* |
| *CF* | *5.107 ± 1.419* | *5.769 ± 0.777* | *-2.600* | *0.003* | *-0.603* |
| *DF* | *5.786 ± 1.150* | *6.553 ± 0.538* | *-3.741* | *$<10^{-3}$* | *-0.886* |
| *EF* | *5.917 ± 1.222* | *6.425 ± 0.575* | *-2.391* | *$<10^{-3}$* | *-0.566* |
| *RF* | *5.704 ± 0.866* | *5.840 ± 0.678* | *0.784* | *0.111* | *-0.175* |
| *OF* | *6.030 ± 0.779* | *6.325 ± 0.633* | *-1.877* | *0.188* | *-0.418* |
| *3D* | *5.619 ± 1.324* | *6.200 ± 0.911* | *-2.303* | *0.005* | *-0.520* |
| *Score* | *5.905 ± 1.122* | *6.100 ± 0.810* | *-0.899* | *0.236* | *-0.202* |
| *Use* | *5.881 ± 1.292* | *6.000 ± 1.261* | *-0.422* | *0.871* | *-0.093* |

Groups sizes = 42 (A), 40 (B). Degrees of freedom = 80.

TABLE IV
STUDY OF STATISTICALLY SIGNIFICANT DIFFERENCES BETWEEN THE AR AND THE VR APPLICATION—TIME AND MOVEMENTS

| | *Task 1* | | *Task 2* | |
|---|---|---|---|---|
| | *Time (mean + SD)* | *Movements (mean + SD)* | *Time (mean + SD)* | *Movements (mean + SD)* |
| *AR* | *50.467 ± 30.225* | *4.639 ± 1.099* | *124.253 ± 46.027* | *10.056 ± 2.629* |
| *VR* | *19.544 ± 12.181* | *4.781 ± 2.697* | *78.056 ± 37.624* | *11.469 ± 3.742* |
| *t* | *5.407* | *-0.291* | *4.496* | *-1.818* |
| *p* | *0.004* | *0.338* | *0.013* | *0.382* |
| *Cohen's d* | *1.458* | *-0.075* | *1.105* | *-0.444* |

Groups sizes = 36 (A), 32 (B). Degrees of freedom = 66.

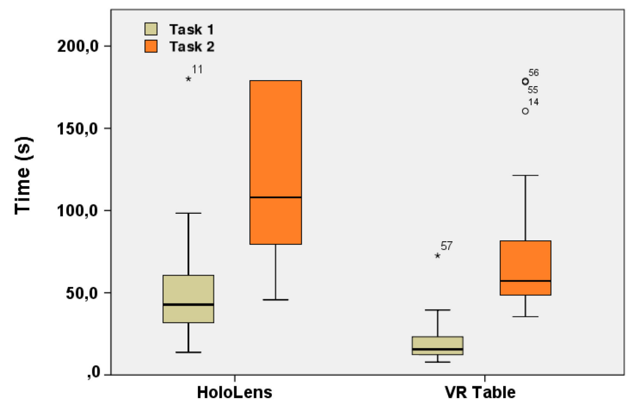


Fig. 6. Box plots for tracked times.

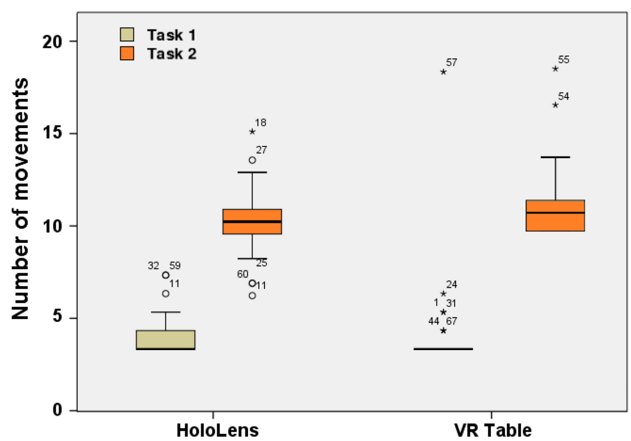


Fig. 7. Box plots for number of movements.

training application. This statistical analysis is performed with IBM SPSS 24 software. For all of the analyses detailed hereafter, significance tests were two-tailed and conducted at the 0.05 significance level.

First, we checked if the collected data follow a normal distribution. As a representative example, the Kolmogórov–Smirnov test ($D = 0.2310$ and $p$-value $= 0.4823$), the Anderson-Darling test ($A = 0.4698$ and $p$-value $= 0.1668$), and the Shapiro–Wilk test ($W = 0.5214$ and $p$-value $= 0.3188$) confirmed that the *times* dataset in Task 2 follows a normal distribution. Although for the sake of brevity we do not detail the rest of the normality tests, the same happened in the rest of datasets. Therefore, we can use parametric tests: the $t$-test and the Cohen's test for paired and unpaired data, as well as a correlation study and a multifactorial ANOVA for analyzing relationships among the parameters in the experiment.

## A. *Parametric Tests*

Next, we checked if there were statistically significant differences between the two systems for the different factors. All the participants were considered. Table III shows the results of this analysis, which clearly indicate that there are significant differences in favor of the VR Table for SF, CF, DF, EF, and for the question about depth perception. Table IV also reveals that the time needed to complete both Task 1 and Task 2 using the VR setup is significantly lower than for the AR case. The reasons that may explain these differences are the lower response time of the VR Table compared to HoloLens as well as the easy learning curve of the VR device. For instance, some users emphasized accuracy, ease of use of the VR table, as opposed to the problems that they found when manipulating the virtual organs using the HoloLens.

Fig. 6 shows the box plots of the distribution of the time required for the participants to complete both tasks. Not only is the VR system more efficient in terms of required time but also there is a smaller dispersion in the results with respect to the AR system, in which, especially for Task 2, several users required a very large amount of time. Similarly, Fig. 7 shows the box plot of the distribution of the number of movements required for the participants to complete both tasks. In this case, the differences are small and not statistically significant.

Therefore, based on these results, it can be said that our primary hypothesis is correct, as there is an important number of measures (SF, CF, DF, EF, depth perception, and time) that show statistically significant differences between the two systems in the use of the anatomy training application.

This analysis is complemented by a study of the differences between the two setups, for each of the groups of the experiment. The results are listed in Table V (group A) and Table VI (group B). In this regard, users from group A show statistically significant negative (in favor of the VR setup) differences in four

TABLE V
STUDY OF STATISTICALLY SIGNIFICANT DIFFERENCES FOR GROUP A

| | *AR (mean ± SD)* | *VR (mean ± SD)* | *t* | *p* | *Cohen's d* |
|---|---|---|---|---|---|
| *SF* | *6.214 ± 0.893* | *6.554 ± 0.578* | *-2.898* | *0.006* | *-0.461* |
| *CF* | *5.107 ± 1.419* | *6.161 ± 0.727* | *-4.008* | $<10^{-3}$ | *-0.982* |
| *DF* | *5.786 ± 1.150* | *6.286 ± 0.742* | *-2.723* | *0.009* | *-0.529* |
| *EF* | *5.917 ± 1.222* | *6.458 ± 0.641* | *-3.236* | *0.002* | *-0.581* |
| *RF* | *5.704 ± 0.866* | *5.863 ± 0.789* | *-1.179* | *0.245* | *-0.192* |
| *OF* | *6.030 ± 0.779* | *6.220 ± 0.808* | *-1.309* | *0.198* | *-0.240* |
| *3D* | *5.619 ± 1.324* | *5.976 ± 1.297* | *-1.344* | *0.186* | *-0.272* |
| *Score* | *5.905 ± 1.122* | *6.190 ± 0.943* | *-1.432* | *0.160* | *-0.277* |
| *Use* | *5.881 ± 1.292* | *6.095 ± 1.246* | *-1.040* | *0.304* | *-0.169* |

Group size = 42. Degrees of freedom = 41.

TABLE VI
STUDY OF STATISTICALLY SIGNIFICANT DIFFERENCES FOR GROUP B

| | *VR (mean ± SD)* | *AR (mean ± SD)* | *t* | *p* | *Cohen's d* |
|---|---|---|---|---|---|
| *SF* | *6.444 ± 0.530* | *6.156 ± 0.695* | *2.913* | *0.006* | *0.470* |
| *CF* | *5.769 ± 0.777* | *4.919 ± 1.393* | *3.745* | *0.001* | *0.784* |
| *DF* | *6.533 ± 0.538* | *5.808 ± 1.207* | *4.351* | $<10^{-3}$ | *0.831* |
| *EF* | *6.425 ± 0.575* | *5.919 ± 1.032* | *3.296* | *0.002* | *0.630* |
| *RF* | *5.840 ± 0.678* | *5.692 ± 0.890* | *1.329* | *0.191* | *0.189* |
| *OF* | *6.325 ± 0.633* | *6.044 ± 0.928* | *2.437* | *0.019* | *0.360* |
| *3D* | *6.200 ± 0.911* | *5.800 ± 1.159* | *2.290* | *0.028* | *0.386* |
| *Score* | *6.100 ± 0.810* | *5.300 ± 1.572* | *3.663* | *0.001* | *0.672* |
| *Use* | *6.000 ± 1.261* | *5.450 ± 1.568* | *2.372* | *0.023* | *0.389* |

Group size = 40. Degrees of freedom = 39.

factors (SF, CF, DF, EF) whereas users from group B show statistically significant positive (in favor of the VR setup) differences in all the measures except from RF. The explanation for this fact may be that the users' satisfaction and friendliness of our VR Table, in terms of ergonomics, interface, information displayed, and ease of use allow a better concentration of the participants to complete the tasks, compared to the HoloLens. Thus, when users tried first the HoloLens, they provided acceptable scores and they did not penalize it later when they tried the VR Table (since the questionnaire about the AR system had already been completed). However, when they tested the VR Table first, they did not hesitate to penalize in their scores the majority of the factors of the system based on HoloLens. In any case, these results further confirm that the VR setup obtained higher values in the responses of the questions shown in Table I.

## *B. Selection, Preference, and Recommendation*

Once it had been proven that there are statistically significant differences between the VR and the AR versions of the anatomy training application, in favor of the former, we focused on the results of the questions shown in Table II. As 77 of the 82 participants answered these questions, group A was formed by 40 participants and group B by 37.

Three major questions were asked in this part of Questionnaire 2 (see Fig. 8). The first one is about the system that is considered more useful (Q1). From the 77 answers, 58 (75.32%) chose the VR Table and 19 (24.68%), the HoloLens-based system. If we study the results by group, in group A, 27 people (67.5% of the people in this group) selected the VR Table and 13 (32.5%) the AR system, whereas in group B, 31 people (83.78% of the group) selected the VR Table and only 6 (16.22%) opted for the HoloLens. As it can be seen, there is a clear majority of users who consider that the VR Table is more useful than the HoloLens for this application. This percentage is even higher for people in group B (those who tested the VR Table first). The comments of the different users and the responses to Q2 (Why?) help understand this result. In this regard, some users commented that the VR Table, compared to the HoloLens, included an easier mechanism to manipulate elements, a larger FoV, a clearer visualization of labels to identify the organs, as well as reporting less tiredness when using the system. In the same way that some participants of the experiment commented at this point the limited FoV (∼30°) of the HoloLens, as stated in [25], others indicated that they expected higher accuracy in terms of head pose estimation, hand tracking, and gesture recognition, as stated in [26] and [27]. Additionally, there were even five participants who complained about an unexpected poor performance at the head pose estimation in the HoloLens when they moved the head quickly, as stated in [26]. Precise and fast motion tracking is essential in this type of real-time applications [78]. No similar problems were reported for the VR Table and the known limitations of the HTC Vive tracking system were not meaningful in this application.

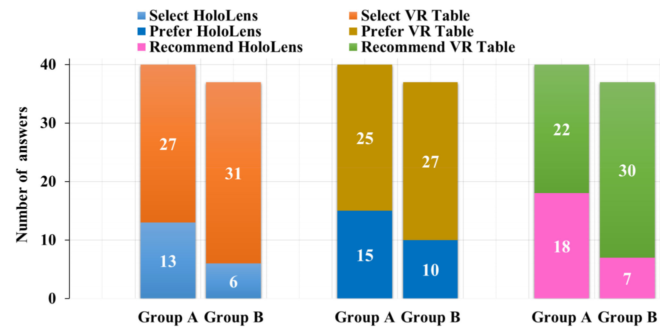


Fig. 8. Results for questions Q1, Q3, and Q5 of Table II.

The second major question is about personal preference (Q3). As in the previous question, there is a clear preference for the VR setup. From the 77 total answers, 52 (67.53%) preferred the VR application, whereas 25 people (32.47%) liked the HoloLens best. If we study the results by group, in group A, 25 people (62.5% of the people in this group) selected the VR Table and 15 (37.5%) the AR system, whereas in group B, 27 people (72.97% of the group) selected the VR Table, and 10 (27.03%) opted for the HoloLens. These results are very similar to those of Q1 and it also happens that the preference for the VR Table increases in group B, although there is also an important preference for the VR setup also in group A. When asked about their reasons for this preference (Q4), users commented the frustration of grabbing elements with the HoloLens as well as the lower system lag and the larger FoV provided by the VR Table.

Finally, the third major question is about recommendation (Q5). As expected, the VR Table is again clearly the recommended setup: 52 out of 77 participants (67.53%) recommended the VR Table and 25 people (32.47%) the AR version. If we

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



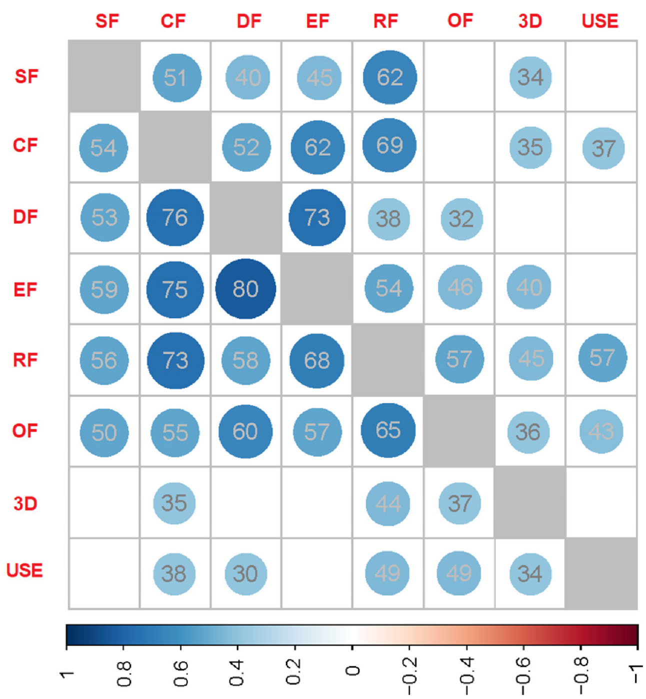

Fig. 9. Correlogram for groups A (left-down) and B (top-right).

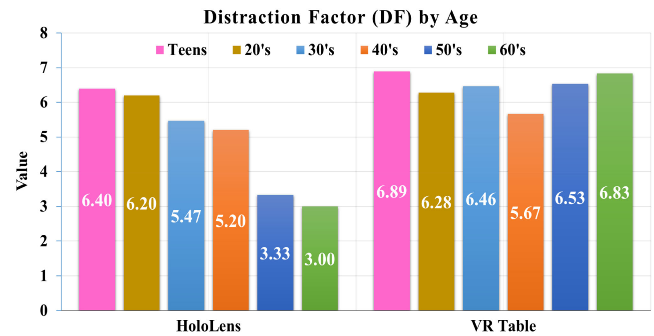


Fig. 10. Distraction factor (DF) by age.

study the results by group, in group A, 22 participants (55% of the group) recommended the VR system and 18 (45%) recommended the AR system, whereas in group B 30 people (81.08% of the group) recommended the VR version and 7 (18.92%) recommended the AR setup. These results are similar to the previous questions, although the percentage of people recommending the AR system in group A is higher than in the previous cases, and higher than in group B. This means also that the order in which the two systems are tested has an effect in the way people perceive and recommend the two alternatives. When asked about the reasons for recommending the VR system (Q6), users commented that this alternative was more intuitive to use, less distracting, and probably more reliable for a training session, where it could be used by hundreds of medical specialists in training. However, some participants concurred that the HoloLens allows observing the virtual human body from all the angles, in an easier way than with the VR Table.

In any case, given the results of these questions, it can be said that the VR setup is perceived more useful and would be the recommendation and preference of a consistent majority of users. Therefore, the second hypothesis is also validated.

### *C. Correlation Analysis*

Next, we present in Fig. 9 the correlation analysis for the responses given by the participants who tested each system first. The results of this analysis include the significance levels and the correlation factors, which are shown (only those that are statistically significant) in colored circles with numbers. These numbers and colors represent the Pearson's correlation in 0–100 units. As it can be seen, the degree of correlation between the different factors and measures is very high, especially between SF, CF, DF, EF, RF, and OF. The correlation is smaller with respect to the depth perception and usefulness, although there are statistically significant correlations in all columns for both groups. There are, in fact, some surprisingly high values such as the correlations in group A between DF and EF (0.799), CF and DF (0.757), CF and EF (0.748), CF and RF (0.731). Correlations in group B are smaller but, nevertheless, generally high, as the one between DF and EF (0.726). This means that the answers of the participants are consistent and reliable.

### *D. Multifactorial ANOVA*

Finally, a multifactorial ANOVA was performed in order to analyze if there is a significant interaction among the different features of the population and their responses to the questions shown in Table I. The following independent variables were studied: gender, age, and tested system. The dependent variables were SF, CF, DF, EF, RF, OF, depth perception, usefulness, and score. The analysis reveals that there was only one significant difference in the distraction factor with respect to age ($F[5,61] = 2.868$, $p = 0.022$, $\eta^2 = 0.190$). Indeed, as in can be seen in Fig. 10, users in their fifties and sixties assess poorly the HoloLens. In fact, there is a clear inverse correlation between age and the DF for the HoloLens-based system. This effect is not seen for the VR table. A possible explanation for this is that the interface in the HoloLens-based setup is perceived as too complex by older adults who do not feel natural to interact with holographic objects.

## VI. Conclusion

Anatomy training with real cadavers poses multiple practical problems. Therefore, it is important to find alternatives to traditional anatomy training. 3-D interactive technologies stand out as one of the most prominent alternatives for traditional training in many areas, with some previous successful attempts published in the academic literature in anatomy training that have proven the applicability of these technological paradigms. However, it is not easy to select the right 3-D technology or setup for each particular application. For this reason, this article presents a comparative study with 82 participants between two different 3-D interactive setups in the use of anatomy training: An optical-based AR setup, implemented with a Microsoft HoloLens device, and a semi-immersive setup based on a VR Table.

From the results of this article, we can conclude that our primary hypothesis ("there are statistically significant differences between the use of the AR application and the use of the VR Table") is corroborated, since significant differences can be found between the two setups for several measures, namely: SFs, CFs, DFs, EFs, depth perception, and the time required to complete the tasks proposed in the experiment. Our secondary hypothesis ("user preference and recommendation for the VR application is higher than for the HoloLens-based application") is also corroborated by the experimental data, since a sizeable amount of people prefer (67.53%), recommend (67.53%), and think (75.32%) that the setup with the VR Table is more useful than the one with the HoloLens. In addition, the VR setup is the cheapest of the two.

The reasons for these results seem to be that users' satisfaction and friendliness of the VR Table, in terms of ergonomics and interaction is higher compared to the HoloLens, which despite having a more immersive setup, is too cumbersome for some users, especially the oldest ones, who rate this setup significantly poorer than other age segments. In addition, the limited FoV featured in the HoloLens degrades the experience, which is not as immersive as it would be desired. Nevertheless, both systems obtain high scores in the different questions shown in Table I, with mean values above 5 (and in many cases above 6) in the 7-scale questions asked.

It is important to emphasize, at this point, that the aim of this research is to compare these two setups, not the technologies of AR and VR in general. The results could be different for other AR/VR setups with different devices and user interfaces. Performance, latency, and latency jitter [79] of different devices could also be different and may influence the usability and user experience [80]. More recent hardware, such as HoloLens 2 or Magic Leap, which provide a larger FoV, or Oculus Quest, which is completely wireless and provides accurate tracking based on SLAM techniques (avoiding the need to use any external hardware to provide tracking), could improve the results obtained in these experiments. In this same line, an interesting research direction is to specifically compare the usefulness—for this type of application—of natural interfaces, such as voice or gestures, with respect to other interfaces. Another important research direction is to analyze if the forensic table metaphor is essential in the AR/VR-based anatomy training or not.

In any case, the results already presented here can be helpful for future research about the successful implementation of these technologies in anatomy training. In fact, given the good results obtained by both systems, we plan to develop a transfer roadmap. First, we hope to complete more quantitative and qualitative evaluations using a different user interface for the AR setup (i.e., it would be possible to evaluate the HoloLens using the controller of the HTC Vive, instead of voice and gestures), producing further empirical evidences associated with both alternatives. Next, we plan to test an AR system based on the Magic Leap glasses and an immersive VR setup based on an Oculus Quest HMD, in order to further confirm if the VR Table, which follows the forensic table metaphor, is the preferred setup for anatomy training. Finally, we plan to introduce these technologies in anatomy academic courses to measure the effectiveness of this approach in a teaching environment.

## References


[1] J. Egger *et al.*, "HTC vive MeVisLab integration via OpenVR for medical applications," *PloS One*, vol. 12, no. 3, Mar. 2017, Art. no. e0173972.

[2] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, "The cave: Audio visual experience automatic virtual environment," *Commun. ACM*, vol. 35, no. 6, pp. 64–73, 1992.

[3] S. Casas, C. Portalés, I. García-Pereira, and M. Fernández, "On a first evaluation of ROMOT—a RObotic 3D MOvie theatre—for driving safety awareness," *Multimodal Technol. Interact.*, vol. 1, no. 2, 2017, Art. no. 6.

[4] G. Strentzsch, F. van de Camp, and R. Stiefelhagen, "Digital map table VR: Bringing an interactive system to virtual reality," in *Proc. Int. Conf. Virtual, Augmented Mixed Reality*, 2017, pp. 54–71.

[5] J. Gimeno, P. Morillo, S. Casas, and M. Fernández, "An augmented reality (AR) CAD system at construction sites," in *Augmented Reality-Some Emerging Application Areas*, Rijeka, Croatia: InTech, 2011.

[6] S. Blanco-Pons, B. Carrión-Ruiz, and J. L. Lerma, "Augmented reality application assessment for disseminating rock art," *Multimedia Tools Appl.*, vol. 78, no. 8, pp. 10265–10286, 2019.

[7] J. Lee, S. Jung, J. W. Kim, and F. Biocca, "Applying spatial augmented reality to anti-smoking message: Focusing on spatial presence, negative emotions, and threat appraisal," *Int. J. Human–Comput. Interact.*, vol. 35, no. 9, pp. 751–760, 2019.

[8] D. Kugelmann *et al.*, "An augmented reality magic mirror as additive teaching device for gross anatomy," *Annal. Anatomy-Anatomischer Anzeiger*, vol. 215, pp. 71–77, 2018.

[9] P. Pratt *et al.*, "Through the HoloLens looking glass: Augmented reality for extremity reconstruction surgery using 3D vascular models with perforating vessels," *Eur. Radiol. Exp.*, vol. 2, no. 1, 2018, Art. no. 2.

[10] C. A. Andújar Gran, M. Fairén González, and P. Brunet Crosa, "Affordable immersive projection system for 3D interaction," in *Proc. 1st Ibero-Amer. Symp. Comput. Graph.*, 2002, pp. 1–7.

[11] Y. Fei, D. Kryze, and A. Melle, "Tavola: Holographic user experience," in *Proc. ACM SIGGRAPH Emerg. Technol.*, 2012, Art. no. 21.

[12] O. Bimber, L. M. Encarnação, and P. Branco, "The extended virtual table: An optical extension for table-like projection systems," *Presence, Teleoperators Virtual Environ.*, vol. 10, no. 6, pp. 613–631, 2001.

[13] M. Rosenzweig, M. MacEachern, and C. Masters, "The anatomage table: An innovative approach to anatomy education," in *Proc. Med. Library Assoc. 117th Annu. Meeting Exhib.*, Seattle, WA, USA, 2017.

[14] S. Fyfe, G. Fyfe, D. Dye, and H. Radley-Crabb, "The anatomage table: Differences in student ratings between initial implementation and established use," *Focus Health Professional Educ., Multi-Disciplinary J.*, vol. 19, no. 2, 2018, Art. no. 41.

[15] D. Shi *et al.*, "An exploratory study of sectra table visualization improves the effectiveness of emergency bedside echocardiography training," *J. Ultrasound Med.*, vol. 38, pp. 363–370, 2018.

[16] A. Coppens, "Merging real and virtual worlds: An analysis of the state of the art and practical evaluation of Microsoft HoloLens," Software Engineering Lab, Department of Computer Science, University of Mons, Mons, Belgium, 2017.

[17] M. Hoover, "An evaluation of the Microsoft HoloLens for a manufacturing-guided assembly task," Department of Mechanical Engineering Iowa State University, Ames, IA, USA, 2018.

[18] M. Kalantari and P. Rauschnabel, "Exploring the early adopters of augmented reality smart glasses: The case of Microsoft HoloLens," in *Augmented Reality and Virtual Reality*. New York, NY, USA: Springer, 2018, pp. 229–245.

[19] H. Xue, P. Sharma, and F. Wild, "User satisfaction in augmented reality-based training using Microsoft HoloLens," *Computers*, vol. 8, no. 1, 2019, Art. no. 9.

[20] U. Riedlinger, L. Oppermann, and W. Prinz, "Tango vs. HoloLens: A comparison of collaborative indoor AR visualisations using hand-held and hands-free devices," *Multimodal Technol. Interact.*, vol. 3, no. 2, 2019, Art. no. 23.

[21] E. Rae, A. Lasso, M. S. Holden, E. Morin, R. Levy, and G. Fichtinger, "Neurosurgical burr hole placement using the Microsoft HoloLens," *Med. Imag., Image-Guided Procedures, Robotic Interventions, Model.*, 2018, vol. 10576, Art. no. 105760T.

[22] H. F. Al Janabi *et al.*, "Effectiveness of the HoloLens mixed-reality headset in minimally invasive surgery: A simulation-based feasibility study," *Surgical Endoscopy*, vol. 34, pp. 1143–1149, 2020.

[23] R. Affolter, S. Eggert, T. Sieberth, M. Thali, and L. C. Ebert, "Applying augmented reality during a forensic autopsy—Microsoft HoloLens as a DICOM viewer," *J. Forensic Radiol. Imag.*, vol. 16, pp. 5–8, 2019.

[24] M. G. Hanna, I. Ahmed, J. Nine, S. Prajapati, and L. Pantanowitz, "Augmented reality technology using Microsoft HoloLens in anatomic pathology," *Archives Pathol. Laboratory Med.*, vol. 142, no. 5, pp. 638–644, 2018.

[25] I. Kuhlemann, M. Kleemann, P. Jauer, A. Schweikard, and F. Ernst, "Towards X-ray free endovascular interventions—Using HoloLens for online holographic visualisation," *Healthcare Technol. Lett.*, vol. 4, no. 5, pp. 184–187, 2017.

[26] Y. Liu, H. Dong, L. Zhang, and A. El Saddik, "Technical evaluation of HoloLensfor multimedia: A first look," *IEEE MultiMedia*, vol. 25, no. 4, pp. 8–18, Oct.–Dec. 2018.

[27] G. Evans, J. Miller, M. I. Pena, A. MacAllister, and E. Winer, "Evaluating the Microsoft HoloLens through an augmented reality assembly application," in *Proc. Degraded Environ., Sens., Process., Display*, 2017, vol. 10197, Art. no. 101970V.

[28] D. C. Niehorster, L. Li, and M. Lappe, "The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research," *i-Perception*, vol. 8, no. 3, pp. 1–23, May 2017.

[29] D.-N. Le, C. Van Le, J. G. Tromp, and G. N. Nguyen, *Emerging Technologies for Health and Medicine: Virtual Reality, Augmented Reality, Artificial Intelligence, Internet Things, Robotics, Industry 4.0*. Hoboken, NJ, USA: Wiley, 2018.

[30] W. S. Khor, B. Baker, K. Amin, A. Chan, K. Patel, and J. Wong, "Augmented and virtual reality in surgery—The digital surgical environment: Applications, limitations and legal pitfalls," *Ann. Transl. Med.*, vol. 4, no. 23, 2016, Art. no. 454.

[31] B. Fida, F. Cutolo, G. di Franco, M. Ferrari, and V. Ferrari, "Augmented reality in open surgery," *Updates Surgery*, vol. 70, no. 3, pp. 389–400, 2018.

[32] G. Herrera, L. Vera, J. Sevilla, C. Portalés, and S. Casas, "On the development of VR and AR learning contents for children on the autism spectrum: From real requirements to virtual scenarios," in *Proc. Augmented Reality Enhanced Learn. Environ.*, 2018, pp. 106–141.

[33] J. Dunn, E. Yeo, P. Moghaddampour, B. Chau, and S. Humbert, "Virtual and augmented reality in the treatment of phantom limb pain: A literature review," *NeuroRehabilitation*, vol. 40, no. 4, pp. 595–601, 2017.

[34] M. C. Howard, "A meta-analysis and systematic literature review of virtual reality rehabilitation programs," *Comput. Human Behav.*, vol. 70, pp. 317–327, 2017.

[35] E. Z. Barsom, M. Graafland, and M. P. Schijven, "Systematic review on the effectiveness of augmented reality applications in medical training," *Surgical Endoscopy*, vol. 30, no. 10, pp. 4174–4183, 2016.

[36] S. Yeom, "Augmented reality for learning anatomy," in *Proc. Ascilite Hobart*, 2011, pp. 1377–1383.

[37] M. Hackett and M. Proctor, "Three-dimensional display technologies for anatomical education: A literature review," *J. Sci. Educ. Technol.*, vol. 25, no. 4, pp. 641–654, 2016.

[38] M. Aebersold *et al.*, "Interactive anatomy-augmented virtual simulation training," *Clin. Simul. Nursing*, vol. 15, pp. 34–41, 2018.

[39] T. Blum, V. Kleeberger, C. Bichlmeier, and N. Navab, "Mirracle: An augmented reality magic mirror system for anatomy education," in *Proc. Virtual Reality Short Papers Posters*, 2012, pp. 115–116.

[40] P. Boonbrahm, C. Kaewrat, P. Pengkaew, S. Boonbrahm, and V. Meni, "Study of the hand anatomy using real hand and augmented reality," *Int. J. Interactive Mobile Technol.*, vol. 12, no. 7, pp. 181–190, 2018.

[41] C.-H. Chien, C.-H. Chen, and T.-S. Jeng, "An interactive augmented reality system for learning anatomy structure," in *Proc. Int. Multiconf. Eng. Comput. Scientists*, 2010, vol. 1, pp. 17–19.

[42] R. Codd-Downey, R. Shewaga, A. Uribe-Quevedo, B. Kapralos, K. Kanev, and M. Jenkin, "A novel tabletop and tablet-based display system to support learner-centric ophthalmic anatomy education," in *Proc. Int. Conf. Augmented Reality, Virtual Reality Comput. Graph.*, 2016, pp. 3–12.

[43] J. Falah, V. Charissis, S. Khan, W. Chan, S. F. Alfalah, and D. K. Harrison, "Development and evaluation of virtual reality medical training system for anatomy education," in *Intelligent Systems in Science and Information 2014*, New York, NY, USA: Springer, 2015, pp. 369–383.

[44] J. Ferrer-Torregrosa, J. Torralba, M. A. Jimenez, S. García, and J. M. Barcia, "Arbook: Development and assessment of a tool based on augmented reality for anatomy," *J. Sci. Educ. Technol.*, vol. 24, no. 1, pp. 119–124, 2015.

[45] J. H. Seo *et al.*, "Anatomy builder VR: Applying a constructive learning method in the virtual reality canine skeletal system," in *Proc. IEEE Virtual Reality*, 2017, pp. 399–400.

[46] J.-J. Arino, M.-C. Juan, J.-A. Gil-Gómez, and R. Mollá, "A comparative study using an autostereoscopic display with augmented and virtual reality," *Behav. Inf. Technol.*, vol. 33, no. 6, pp. 646–655, 2014.

[47] N. Gavish *et al.*, "Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks," *Interactive Learn. Environ.*, vol. 23, no. 6, pp. 778–798, 2015.

[48] B. Blissing, F. Bruzelius, and O. Eriksson, "Driver behavior in mixed and virtual reality—A comparative study," *Transp. Res. Part F, Traffic Psychol. Behav.*, vol. 61, pp. 229–237, 2017.

[49] M. C. Juan and D. Pérez, "Using augmented and virtual reality for the development of acrophobic scenarios: Comparison of the levels of presence and anxiety," *Comput. Graphics*, vol. 34, no. 6, pp. 756–766, 2010.

[50] M. Juan, I. García-García, R. Mollá, and R. López, "Users' perceptions using low-end and high-end mobile-rendered HMDs: A comparative study," *Comput.*, vol. 7, no. 1, 2018, Art. no. 15.

[51] R. Jose, "A comparative study of using augmented reality interfaces for vehicle navigation," Human Interface Technology Lab, College of Engineering, University of Canterbury, Christchurch, New Zealand, 2015.

[52] L. Figueiredo, E. Rodrigues, J. Teixeira, and V. Techrieb, "A comparative evaluation of direct hand and wand interactions on consumer devices," *Comput. Graph.*, vol. 77, pp. 108–121, 2018.

[53] L. Qian *et al.*, "Comparison of optical see-through head-mounted displays for surgical interventions with object-anchored 2D-display," *Int. J. Comput. Assisted Radiol. Surgery*, vol. 12, no. 6, pp. 901–910, 2017.

[54] D. Rodríguez-Andrés, M.-C. Juan, M. Méndez-López, E. Pérez-Hernández, and J. Lluch, "Mnemocity task: Assessment of childrens spatial memory using stereoscopy and virtual environments," *PloS One*, vol. 11, no. 8, 2016, Art. no. e0161858.

[55] E. Jiménez, G. Mariscal, M. Heredia, and G. Castilla, "Virtual reality versus master class: A comparative study," in *Proc. 6th Int. Conf. Technol. Ecosystems Enhancing Multiculturality*, 2018, pp. 568–573.

[56] D. H.-L. Goh, C. S. Lee, and K. Razikin, "Comparative evaluation of interfaces for presenting location-based information on mobile devices," in *Proc. Int. Conf. Asian Digital Libraries*, 2011, pp. 237–246.

[57] P. Safadel and D. White, "A comparative analysis of augmented reality and two-dimensional using molecular modeling on student learning," in *Proc. Soc. Inf. Technol. Teacher Educ. Int. Conf.*, 2017, pp. 1774–1776.

[58] A. Miloff *et al.*, "Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial," *Behav. Res. Therapy*, vol. 118, pp. 130–140, 2019.

[59] S. F. Alfalah, J. F. Falah, T. Alfalah, M. Elfalah, N. Muhaidat, and O. Falah, "A comparative study between a virtual reality heart anatomy system and traditional medical teaching modalities," *Virtual Reality*, vol. 23, pp. 229–234, 2018.

[60] Chung Van Le, J. G. Tromp, and V. Puri, "Using 3D simulation in medical education: A comparative test of teaching anatomy using virtual reality," *Emerg. Technol. Health Med., Virtual Reality, Augmented Reality, Artif. Intell., Internet Things, Robotics, Industry 4.0*, vol. 12, 2018, Art. no. 21.

[61] J. Ferrer-Torregrosa, M. Á. Jiménez-Rodríguez, J. Torralba-Estelles, F. Garzón-Farinós, M. Pérez-Bermejo, and N. Fernández-Ehrling, "Distance learning ECTs and flipped classroom in the anatomy learning: Comparative study of the use of augmented reality, video and notes," *BMC Med. Educ.*, vol. 16, no. 1, 2016, Art. no. 230.

[62] A. M. Codd and B. Choudhury, "Virtual reality anatomy: Is it comparable with traditional methods in the teaching of human forearm musculoskeletal anatomy?" *Anatomical Sci. Educ.*, vol. 4, no. 3, pp. 119–125, 2011.

[63] C. Moro, Z. Štromberga, A. Raikos, and A. Stirling, "The effectiveness of virtual and augmented reality in health sciences and medical anatomy," *Anatomical Sci. Educ.*, vol. 10, no. 6, pp. 549–559, 2017.

[64] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Trans. Inf. Syst.*, vol. 77, no. 12, pp. 1321–1329, 1994.

[65] N. Chaconas and T. Höllerer, "An evaluation of bimanual gestures on the Microsoft HoloLens," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces*, 2018, pp. 1–8.

[66] M. Borges, A. Symington, B. Coltin, T. Smith, and R. Ventura, "HTC vive: Analysis and accuracy improvement," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2018, pp. 2610–2615.

[67] T. Rose, C. S. Nam, and K. B. Chen, "Immersion of virtual reality for rehabilitation-review," *Appl. Ergonom.*, vol. 69, pp. 153–161, 2018.

[68] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

[69] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. London, U.K.: Pearson, 2016.

[70] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice*. Reading, MA, USA: Addison-Wesley, 2017.

[71] M. E. Latoschik and W. Stuerzlinger, "On the art of the evaluation and presentation of RIS-engineering," in *Proc. IEEE 7th Workshop Softw. Eng. Archit. Realtime Interactive Syst.*, 2014, pp. 9–17.

[72] J. F. Morris, "Competencies for teaching anatomy effectively and efficiently," in *Teaching Anatomy*. New York, NY, USA: Springer, 2015, pp. 39–44.

[73] M. A. Vorstenbosch, T. P. Klaassen, A. R. T. Donders, J. G. Kooloos, S. M. Bolhuis, and R. F. Laan, "Learning anatomy enhances spatial ability," *Anatomical Sci. Educ.*, vol. 6, no. 4, pp. 257–262, 2013.

[74] A. Guillot, S. Champely, C. Batier, P. Thiriet, and C. Collet, "Relationship between spatial abilities, mental rotation and functional anatomy learning," *Adv. Health Sci. Educ.*, vol. 12, no. 4, pp. 491–507, 2007.

[75] A. X. Garg, G. Norman, and L. Sperotable, "How medical students learn spatial anatomy," *Lancet*, vol. 357, no. 9253, pp. 363–364, 2001.

[76] M. Estai and S. Bunt, "Best teaching practices in anatomy education: A critical review," *Annal. Anatomy-Anatomischer Anzeiger*, vol. 208, pp. 151–157, 2016.

[77] E. Messier, J. Wilcox, A. Dawson-Elli, G. Diaz, and C. A. Linte, "An interactive 3D virtual anatomy puzzle for learning and simulation-initial demonstration and evaluation," *Studies Health Technol. Inf.*, vol. 220, 2016, Art. no. 233.

[78] J.-L. Lugrin, D. Wiebusch, M. E. Latoschik, and A. Strehler, "Usability benchmarks for motion tracking systems," in *Proc. 19th ACM Symp. Virtual Reality Softw. Technol.*, 2013, pp. 49–58.

[79] J.-P. Stauffert, F. Niebling, and M. E. Latoschik, "Towards comparable evaluation methods and measures for timing behavior of virtual reality systems," in *Proc. 22nd ACM Conf. Virtual Reality Softw. Technol.*, 2016, pp. 47–50.

[80] J.-P. Stauffert, F. Niebling, and M. E. Latoschik, "Effects of latency jitter on simulator sickness in a search task," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces*, 2018, pp. 121–127.


**Ramiro Serrano Vergel** was born in Bucaramanga, Colombia, in 1975. He received the M.S. degree in computer science from the Andes University, Bogotá, Colombia, in 2013. He is currently working toward the Ph.D. degree with the University of Arkansas at Little Rock, Little Rock, AR, USA.

Before moving to UALR, he worked as a Developer for video games in the private sector and as a Researcher with Andes University and CIDLIS-Colombia. His research focuses on finding new ways to display data in real time on augmented reality and virtual reality systems for training and medical purposes.

**Pedro Morillo Tena** was born in Valencia, Spain, in 1975. He received the M.S. and Ph.D. degrees in computer engineering from the University of Valencia (UV), Valencia, in 1998 and 2004, respectively.

He is currently an Associate Professor with the Computer Science Department, UV, Spain. He is also member of the Institute of Robotics and Information and Communication Technologies, UV, Spain. His research interests include distributed virtual environment systems, cluster computing, virtual and augmented reality.

**Sergio Casas Yrurzum** was born in Valencia, Spain, in 1978. He received the B.S. degree in telecommunications engineering, the M.S. degree in computer engineering, and the Ph.D. degree in computational mathematics from the University of Valencia (UV), Valencia, in 2001, 2006, and 2014, respectively.

He is currently an Associate Professor with the Computer Science Department, UV. He is also member of the Institute of Robotics and Information and Communication Technologies, UV. His expertise is in virtual reality and augmented reality.

**Carolina Cruz-Neira** (Senior Member, IEEE) was born in Alicante, Spain. She received the M.S. and Ph.D. degrees in electrical engineering and computer science from the University of Illinois at Chicago, Chicago, IL, USA, in 1991 and 1995, respectively.

She is currently the Agere Chair of the Computer Science Department with the University of Central Florida, Orlando, FL, USA. She is also the former Executive Director of the Emerging Analytics Center with the University of Arkansas at Little Rock (UALR), Little Rock, AR, USA. She is pioneer in the areas of virtual reality and interactive visualization, having created and deployed a variety of technologies that have become standard tools in industry, government, and academia. She is known worldwide for being the creator of the CAVE virtual reality system.

Dr. Cruz-Neira is a member of the National Academy of Engineering and has received the IEEE Virtual Reality Technical Achievement Award, among many other recognitions.