Design of a Mixed Reality approach to enhance understanding of Reverse Total Shoulder Arthroplasty

Alireza Sadeghi Milani[1], J. Cecil[2], Miguel Pirela-Cruz[3], Shelia Kennison[4]

[1] Center for Cyber-Physical Systems, Department of Computer Science, Oklahoma State University
[2] Center for Cyber-Physical Systems, Department of Computer Science, Oklahoma State University
j.cecil@okstate.edu
[3] Orthopedics Department, Dignity Regional Medical Center
[4] Department of Psychology, Oklahoma State University

Abstract. This paper focuses on the design of a mixed reality-based (MR) simulation environment to train health care personnel in reverse total shoulder arthroplasty (RTSA) procedure. Information-centric models involving interaction with orthopedic surgeons were created as part of a participatory design approach. These information models provided a structural foundation for the design and development of the environments. This paper concludes with a discussion of the preliminary assessment activities which includes studying the impact of such a MR approach on understanding and knowledge acquisition of the targeted surgical procedure.

Keywords: Reverse Total Shoulder Arthroplasty, Mixed Reality, Participatory Design.

1 Introduction

Extended reality tools and environments are being increasingly used to support learning and education [1]. The term Extended reality (XR) encompasses three categories: virtual reality (VR), augmented reality (AR), and mixed reality (MR). These technologies extend various levels of reality by blending real and virtual worlds to support effective immersive experiences [2]. By integrating XR into training, learners can be immersed in a multisensory environment that is more interactive, engaging, and effective. XRs can be used to support the design of Virtual Learning Environments (VLEs) which can be described as a special type of XR environment designed to support learning and training [3].

MR combines the use of AR and VR to seamlessly merge the physical and digital realms, creating an interactive environment for users. This environment can either consist of real surroundings enhanced with virtual elements (known as augmented reality) or virtual environments enriched with physical objects (known as augmented virtuality). By leveraging cutting-edge computer technology, graphics, and input systems, MR enables learners to actively engage with both physical and digital components in real-time. With the

integration of mixed reality, learners can immerse themselves in lifelike interactions with objects and individuals, blurring the boundaries between the real and virtual worlds. They are empowered to manipulate and engage with digital assets, as well as interact with real-life objects and people within the same environment. This technology opens up new possibilities for realistic and dynamic learning experiences. [3].
Glenohumeral arthritis and complicated proximal fractures can be treated surgically with reverse total shoulder arthroplasty (RTSA). The humeral and glenoid stages are the two key steps in this surgery [4]. In standard total shoulder arthroplasty (TSA) the implants (ball and cap) resemble the natural shape of the humerus and scapula however, in RTSA the implants are reversed. The ball is attached to the scapula and the cap is attached to the humerus (Fig 1).

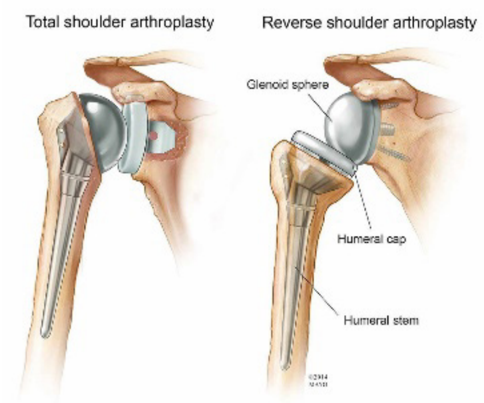

Fig. 1. TSA (on the left), RTSA (on the right) [5]

In RTSA, the proper positioning of the glenoid component helps to increase the survival rate and prevent premature loosening which is one of the procedure's key considerations [6]. Surgeons are required to have a good amount of working knowledge of these bones and surgery steps to be able to plan and perform safe RTSA. However, even after planning a complex surgery such as RTSA, its realization is challenging. Therefore, the level of experience of surgeons plays a crucial role in this process. In medical universities, the usage of XR simulators for training surgeons and residents has grown in recent years. Residents in surgical training are typically instructed to use cadavers, animals, and artificial replicas [7]. These conventional techniques have some shortcomings. Animals being used for surgical training has drawn criticism from animal rights activists. The use of cadavers increases the danger of infection. Training with synthetic mockups is expensive and impersonal when it comes to people [8]. Other methods involve residents watching a skilled surgeon perform the surgery before gradually transitioning to assist them. In this context, new technologies, such as MR, provide surgeons and residents with less expertise to catch up to those with more experience which result in reduced surgical risks and errors caused by insufficient information about anatomical variations. MR utilizes both AR and VR to blend the physical and digital worlds, with the user situated in an interactive environment that could be either real with virtual assets, or virtual with physical objects [9]. Hence, MR provides interaction with real-world information and better training as trainees do not feel completely alone because of a collaborative learning environment, and trainers do not have to be represented as virtual avatars because students are still aware of the real world [10]. Therefore, our research focuses on designing MR-based simulation environments to train residents in RTSA by creating information-centric process models. The proposed approach dealing with the MR simulator for RTSA is outlined as follows:

(a) Expert surgeons' knowledge for comprehending specific surgical procedures has not been prioritized in prior research efforts; in the strategy outlined in this study, the simulator was conceived and built in close collaboration with expert surgeons.
(b) Other earlier research initiatives have not investigated the development of advanced information-centric models of target surgical processes (as a basis to comprehend a surgical process) before developing training simulators. A few previous initiatives have investigated information-centric methodologies, but they have not gone as far as the information-centric modeling strategy used in this paper. In this work, the effect of employing such a strategy is also discussed.

Thus far, some of the key terms and concepts have been defined and described. Next, a review of the literature provided. The rest of this article is structured as follows. The research methodology is provided in Section 2. Subsequently, section 3 maps out knowledge assessment results. In section 4, a discussion of the results is provided. Further, section 4 provides a presentation of the taxonomy and a comparison of the results. In section 5, Comprehension assessment is included. Finally, the conclusion and future work is presented in section 6.

1.1 Related works

In recent years, the role of XR in medical training simulators has been covered in a number of studies. However, there is a lack of proper utilization of the participatory design approach during the creation of information models. Prior research efforts have not emphasized the role of expert surgeons as knowledge sources for understanding target surgical processes; in the approach discussed in this paper, the simulator was designed and developed in close interaction with expert surgeons.

A study was conducted by [11] involving a group of older adults in which they engaged in creating a VR-based ATM training simulation by using a participatory design approach. The results from the study demonstrate that VR is an effective way to directly gain valuable insights related to the design from the participants. Few studies have investigated the role of using XR-based medical training environments by conducting pre and post-test questionaries. In [12] a VR module was used for training an advanced temporal bone procedure. Their results indicate that the environment is effective in training surgical residents for the surgery. In [13] an assessment method based on cyber-humans was created for a virtual reality orthopedic surgery training simulator. Their strategy was based on the evaluation of three perspectives: knowledge, skills, and mental stability. A Mixed Reality-based surgical navigation system for orthopedic surgical navigation has been discussed by [14]. The MR-based navigation system consists of a HoloLens display, a magnetic launcher, a passenger sensor, and a processor. The MR-based system is useful in providing real-time 3D visualization. Using the MR-based system, the 3D reconstructed virtual model generated using a CT scan or MRI can be integrated with the body of the patient which can help guide the operating procedure. Such a system provides additional visual information related to the internal organ of the patient which is not visible to the naked eye. The MR-based navigation system has several advantages over the conventional image-guided surgical system such as intuitive and detailed imaging information, less time spent and mental load, and low risk of errors among others. [15] utilized VR simulation for teaching orthopedic surgery residents to see it is beneficial for their performance in cadaver Total Hip Arthroplasty (THA) surgical skills. They compared the improvement in cadaver THA performance, specific aspects of surgical skills, and knowledge and perception of surgical anatomy and indications. Participants first completed a written pretest and a single THA for establishing their baseline knowledge, then half of them randomly interacted with VR simulation. They utilized the OramaVR software platform and Oculus Rift CV1 headset for their simulation. Their results showed that VR improves surgical and technical skills but has no effect on medical knowledge. [16] presented initial outcomes of using an immersive VR-based preoperative planning tool for laparoscopic donor nephrectomy. The author stated that it was challenging to understand more than 2500 CT images hence they developed 3D models using a 3D slicer which allowed an interactive and comprehensive anatomy when viewed through an immersive headset. The CT images of seven patients were used for a study in which two surgeons assessed the preoperative understanding using CT alone and CT on the immersive headset. The results from the study indicated that immersive models enhanced the surgeons' understanding of the patient's arterial and venous anatomy. Moreover, the surgeons' overall confidence regarding the operation improved while interacting with the 3D image on the immersive headset. [17] proposed two teaching modules for their prototype 3D first Aid VR. A teaching module that described the cause and symptoms of seizure and a training module was utilized to train first aid instructions in a VR simulation. Wang and Wang [18] proposed a VR simulation for the entire chest anatomy to perform the anatomical operation with experimental steps. They divided the process into three modules: scene construction, result determination, and UI interface. [19] Created 3D-printed glenoid models with a B2 defect. Then, they used Unity to build and program 3D models for the same prints with a guide pin for installation on Microsoft HoloLens2. Their study focused on comparing the accuracy of MR holographic model–assisted glenoid guidewire placement to FreeHand (FH) and patient-specific instrumentation (PSI) methods. In all participants, the accuracy of MR was very similar to PSI and FH for the start point, and there was no statistically significant difference in version between MR/PSI and MR/FH. [20] simulated guide wire insertion procedure during dynamic hip screw surgery to depict the training potential of AR and VR. They devised a digital fluoroscopic imaging simulator that tracked colored markers tied to a guide wire using orthogonal cameras. They demonstrated that the AR overlay has a 0.4 mm precision for projecting tip apex distance. They also demonstrated that with more repetitions, the trainee's accuracy improved. [21] tracked 24 medical students who were randomly assigned to four sessions of either one-on-one instruction from a hip replacement surgeon or augmented reality instruction using the Microsoft HoloLens headgear. The accuracy, trainee perceptions, and potential training role of the AR headset were compared to hands-on, expert instruction by a surgeon. Participants underwent surgical training to position an acetabular cup on an opaque hip model in 6 distinct orientations. [22] developed a system for surgical planning and evaluated the results of unilateral, mono-segmental cervical foraminal stenosis using 3D Slicer software on a VR workstation coupled to an HTC Vive and the SteamVR tracking and controller system. The paper comprised 73 patients who used the VR system, and the surgery was planned to use traditional imaging techniques like radiography, CT, and MRI. The authors proposed that designing surgical strategies can benefit from employing VR systems to rebuild preoperative and postoperative images.

Based on the related works discussed in this section there is a need to create information-centric process models to gain better understanding of the target surgical procedures. In this research the role of creating such information-centric process models to better understand RTSA was explored. The preliminary design and implementation of our approach are based on our group's previous simulator for orthopedic surgical training [23]. The creation of information-centric models of target surgical processes is based on that work, although in a more extensive way. Also, the scope of training and assessments are more focused on MR-based simulators and by building various models under different phases to expand training sessions.

2 Research methodology

In virtual environments, virtual reality technologies are used to create a three-dimensional artificial or synthetic environment that allows users to do "what-if" analyses, comprehend target problems, and compare different solutions for a variety of fields. The overall project objective is to create and assess the impact of an MR-based simulation environment to train medical residents in orthopedic surgery. The benefits of creating such an MR simulator for training include:

- Better Preparation of medical residents for RTSA.
- Avoiding the use of human cadavers as there is the risk of infection.
- Avoiding the use of small animals such as rodents (as they are opposed from animal rights groups, which is part of the changing societal values).

- Opportunity to practice repetitively without additional costs when compared to using physical models.

An information model can be described as a model that captures either the functional or process-oriented relationships between sub-activities that comprise a higher-level activity. In some situations, they can represent both the functional and process dependencies. Such an information Centric model is necessary for a formal foundation before creating VR or MR-based simulation environments. Functional relationships can enable a target process to be better understood using attributes such as information inputs controlling criteria performing agents and decision outcomes. Creation of information models to support design of software systems in engineering activities was originally proposed in [24]. In that research, a functional model of the fixture design activities was created and used as the basis to develop an automated fixture design approach, which resulted in a software tool to accomplish the same. In this research activity, the use of an engineering Enterprise modeling language (eEML) has been adopted which is used to understand the target surgical activities from the perspective of an expert surgeon[25, 26]. The importance of developing a formal modeling foundation for the participatory design approach is discussed in this study. In this study, this strategy has been used with a focus on using participatory design to build information models that help people comprehend the complexity of the surgical procedure.

2.1 Participatory design

Participatory Design is a method to involve the people who are going to be affected to have their input during the design process [27]. It is a democratic design process for the design of social and technological systems based on the argument that users should be involved in the design and all stakeholders should have input during the design process [28]. The participatory design method was first used in Scandinavia [29].

One of the key perspectives that emphasizes the role of participatory design involves the creation of information intensive function models involving human experts. The importance of the creation of such information models needs to be underscored as it provides a strong foundation for understanding the various relationships of a target function or process. Such an understanding can be captured in an information model which in turn provides a structured basis for designing and building software systems to accomplish various engineering and medical activities. One of the origins of such an approach in creating software-based systems based on participatory design involves the field of automated fixture design. In [25, 30], multiple experts served as knowledge sources to create an advanced information-rich function model of the target manual fixture design process. This information model was created based on the IDEF-0 function methodology and used as the foundational basis to design and build an automated fixture design system [25, 26, 31]. By identifying various relationships involving four key attributes (information inputs, control factors, decision outcomes and Performing mechanisms) and capturing this relationships in a function model, a better understanding of decision-making process involving the fixture design engineering approach was obtained [30]. This in turn was the basis to create an automated fixture design approach and system to support the manufacturing of prismatic parts.

A key aspect of this project involves collaborating with a leading group of medical surgeons at Dignity Regional Medical Center (Phoenix, AZ). One of the innovative aspects is the creation of an information model capturing the overall steps of the surgical procedures as well as identifying the functional and process relationships of the sub-activities and tasks within these surgical procedures. As the surgical process itself is complex, having such a representation was important for the team to understand the process and then design/build the simulator.

In the domain of medical surgery, an information-centric process model was created of a less invasive stabilization system (LISS) plating surgical procedure which is typically used to treat fractures of the femur bone [32]. This information model was used as the basis to understand the target surgical procedure (LISS plating) which in term was used to design and build an immersive VR and haptic-based simulation environment to train orthopedic surgery residents [33]. In this paper the same approach was followed; An experienced surgeon served as the knowledge source for this information model. As the 3D environments were being created, feedback was obtained from surgeons who interacted with the training simulation modules and provided changes/suggestions for the content and the manner of providing the cues /training assistance to the trainees.

2.2 Creation of Surgical training modules

The use of MR (Fig. 2) in the simulator offers a potential way to increase the precision of glenoid pin insertion while simultaneously reducing substantial practical restrictions. For this module a realistic 3d printed model of the humerus and glenoid bones used as a physical set-up, then an exact CAD model of the same bone was created by Maxon ZBrush and Autodesk Maya and imported to Unity3D for MR simulator design and to be scripted in C#.

The glenohumeral joint is a synovial joint that attaches the upper limb to the axial skeleton. It is a ball-and-socket joint, formed between the glenoid fossa of the scapula and the head of the humerus. The anatomical plane that passes vertically through the inferior angle is named the scapular body line. The Friedman line is drawn along the long axis of the scapula from the tip of the medial border to the center of the glenoid fossa. The Friedman or scapular line can be used to determine the glenoid version and glenoid bone loss. It is useful for non-navigation surgeries to detect the proper angle for positioning the k-wire into the glenoid.

In order to train residents and improve glenoid positioning, the simulation incorporated the visualization of the Friedman line, intermediate joint line, and Scapula body line. These visual cues were provided to assist residents in achieving more accurate glenoid positioning compared to freehand placement. Additionally, the simulation included the depiction of commonly observed incorrect positioning scenarios during RTSA. By highlighting these repeated incorrect positions, the simulation aimed to help learners recognize and avoid these errors, ultimately enhancing their understanding and skill in glenoid positioning during the surgical procedure. The main steps for RTSA are as below:

1. The first step is to remove soft tissue to expose the glenoid and then use the correct anatomical guide size to determine the proper spot for inserting the central guide pin.
2. Ream the glenoid fossa to remove the cartilage by Sliding the glenoid resurfacing reamer onto the central guide pin and smoothing the surface.
3. Connect the cannulated stop drill to the power tool and drill the central hole over the guide pin.
4. Assemble the internal rod of the metagene holder in the holder's main body. Insert in the final metagene implant central hole. Make the sure metagene is fully seated on the bone and screw it in. Baseplate should have good contact with the bone.
5. Attach Glenosphere.

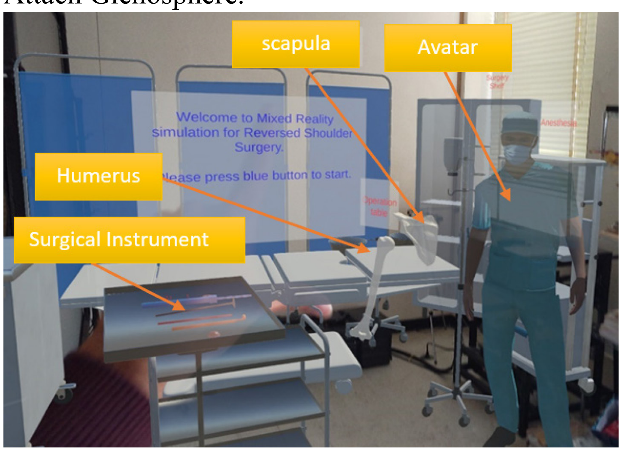

Fig. 2. A view inside the Mixed Reality environment for RTSA (Glenoid steps)

The Physical Environment: A core part of our innovative approach is to explore training using an integrated cyber-physical approach. For this, a physical surgical setup has been created. 3D-printed bones have been created and a platform has been built in which trainees can rotate the bones and position them to have a better understanding of the anatomy and surgery process. For this purpose, an inline joint linkage with steel threaded rods was used. Also, an acrylic sheet was used as a base for the setup to make it lightweight regarding the interactions.

2.3 Validity of simulation

Simulators offer the potential for comprehensive training, gaining surgical skills, and the assessment and certification of surgical proficiency. To confirm their suitability as testing

and certification modalities, these simulators must first undergo validation investigations. The validity of a simulator must be thoroughly and objectively assessed before it can be used to measure skill. Validity assesses whether the simulator or training tool is really imparting or measuring what it is supposed to [34].
First, the realism of the simulator is typically evaluated informally by non-experts, to examine the simulator accurately weather represents what it is meant to portray, this is called Face Validity [34].Validity of simulations in this paper was examined in such a way. Next, a formal review by an expert on the subject of the training should be performed, this is referring to content validity. It is the judgment of the simulator's suitability as a teaching tool [34]. To ensure the validity of the simulation, multiple meetings were conducted with our head surgeon. During these meetings, held at Dignity Regional Medical Center in Phoenix, Arizona, our head surgeon meticulously examined the simulation step by step and provided feedback on various aspects of it. In order to enhance the realism of the simulation, more accurate 3D models were created for the bones (scapula and humerus). Corrections were made to the angles and positions of instruments for each step of the procedure, including the Friedman line angle and the angle and position for performing the k-wire procedure. Taking into account the feedback from surgeons on each main step of the surgery, we incorporated more detailed animations and instructions to further enhance the simulation's accuracy and instructional value.

3 Knowledge assessment results

To understand how and to what extent computers may aid learning, as well as how learning could be made more useful, engaging, effective, and approachable, assessments are necessary to throw light on how humans interact with virtual environments including VR/MR based simulation environments [1, 35-37].
Interactions were conducted with two distinct participant groups: first responders and nurses, as well as non-medical students. These participants were randomly assigned to either the text-based cues group or the voice cues group with a 3D medical avatar. Throughout the interaction, participants were provided with instruction on the positioning of k-wires and were given the opportunity to practice this technique on a physical setup. At the conclusion of the simulation, participants were tasked with picking up a drill and correctly indicating the direction for performing the k-wire procedure on the Glenoid. Due to the unfamiliarity of most nurses and students with the drilling process, they were not asked to perform the drilling.
Hypothesis for first responders and nurses. MR simulation of RTSA has a significant impact on the learning process of our participants. The assessment activities focusing on the understanding correctness of our hypothesis were performed at Dignity hospital, Prescott, Arizona. The expert surgeon involved in the participatory design approach also facilitated the assessment activities.

To evaluate the knowledge acquisition of first responders at Dignity Prescott and Northern Oklahoma College (NOC) following their interaction with the MR environment, a questionnaire-based pre- and post-test methodology was employed. This method involved administering a pre-test to assess participants' understanding of the RTSA concept through a series of questions. Subsequently, participants engaged in training activities within one of the MR environments featuring HL2, enabling them to learn and interact with a physical setup to determine the optimal positioning of the Glenoid. After the training session, participants completed a post-test consisting of the same set of questions from the pre-test, allowing for an assessment of knowledge gained by comparing the results of the pre- and post-tests. This knowledge assessment served as the basis for a comparative study. A total of 17 participants took part in the study, with 8 from Dignity Prescott and 9 from NOC (refer to Fig. 3 for details). Participant selection was conducted randomly to ensure unbiased representation. Welch's t-test was performed to compare the results of the pre-and post-tests (which are scored from zero to hundred points); paired-samples t-test [38] was used as it is designed for repeated measures as the pre-and post-tests are performed by same participants which is a repeated measures test.

Fig. 3. Mixed Reality RTSA Interaction at Dignity hospital – Prescott and NOC

t-test to determine critical value for t with degrees of freedom =16 and a = 0.05 was performed for participants to further analyze the results and to test if there was a significant difference in the means of answers. The absolute value of the calculated t exceeds the critical value (13.6292 > 2.12) with M = 13.7094, SD = 16.3967 for pre-tests and M=63.7024, SD = 18.2863 for post-tests, so the means are significantly different.

The experimental results comparing the means of the pre-test and post-test scores are depicted in Fig. 4 and Fig. 5. It is evident from the figures that there is a significant difference in scores between the two tests. This observation leads to the conclusion that the participants' test scores improved significantly as a result of their interaction with the simulation. Consequently, our hypothesis regarding the effectiveness of the MR simulation is confirmed.

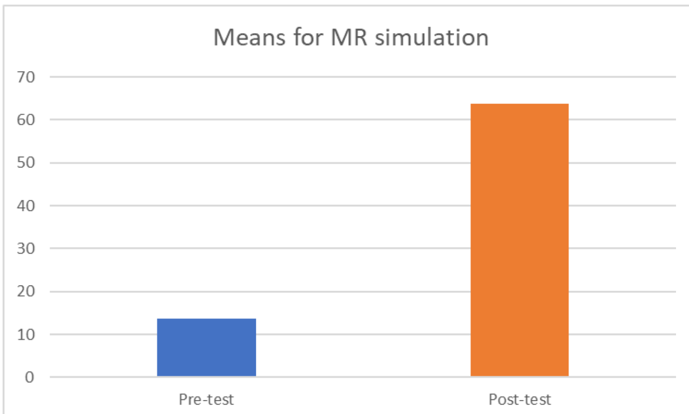

Fig. 4. Result of knowledge assessment for MR environment for first responders and nurses

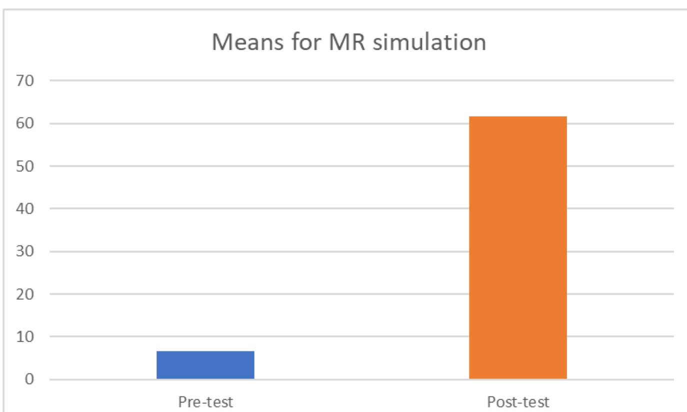

Fig. 5. Result of knowledge assessment for MR environment for non-medical students

Hypothesis for non-medical students. MR simulation of RTSA has a significant effect on the learning process of the student participants. Additionally, it is expected that they would perform the simulation either equally as well or slightly less well compared to first responders and nurses. The assessment activities to evaluate the validity of this hypothesis were carried out at Oklahoma State University, located in Stillwater, Oklahoma. The

assessment process was facilitated by an expert surgeon who was involved in the participatory design approach.
To assess the knowledge acquisition of non-medical students at Oklahoma State University following their interaction with the MR environment, the same questionnaire-based pre and post-test methodology was employed. A total of 10 participants engaged with the MR environment as part of this study (refer to Fig. 6 for details). The selection of participants was conducted randomly to ensure unbiased representation. The pre-test and post-test questionnaires were administered to assess the participants' knowledge, and the results were scored on a scale from zero to one hundred points. To analyze the data and determine the significance of any observed differences, Welch's t-test was conducted. The paired-samples t-test, which is specifically designed for repeated measures, was utilized as the pre-test and post-test were performed by the same group of participants, constituting a repeated measures design.

Fig. 6. Mixed Reality RTSA Interaction at Oklahoma State University.

t-test to determine the critical value for t with degrees of freedom =18 and a = 0.05 was performed for participants to further analyze the results and to test if there was a significant difference in the means of answers. The absolute value of the calculated t exceeds the critical value (8.6614 > 2.101) with M = 6.652, SD = 11.0402 for pre-tests and M=61.63, SD = 16.7634 for post-tests, so the means are significantly different.
The experimental results comparing the means of the pre-test and post-test scores for the non-medical student participants are presented in Figure 6. The figures indicate a significant difference in scores between the two tests. From these findings, it can be concluded that the participants' test scores improved significantly as a result of their interaction with the MR simulation. This supports the hypothesis that the MR simulation has a positive impact on the learning process of non-medical students.

4 Results and discussion

Through collaborations with surgeons, residents, and medical students at Dignity Hospital in Arizona, and non-medical students at Oklahoma State University the effectiveness and impact of employing MR simulation for RTSA as a teaching and training resource was investigated. The learning exercises were conducted with the following two objectives in mind:

(1) The first focus was on making sure the simulator environments were accurate; this was done by senior surgeons interacting with the different training environments and giving thorough feedback on changes to the content of the simulation environments.

(2) The second focus was on researching the learning impact on medical residents and students, which will be done in stages. With the assistance of a skilled surgeon, the simulation's material was continuously revised and improved. 8 medical participants and 10 non-medical participants took part in the activity for phase 1 once these adjustments had been made and validated. Following the pre-test, the participants used the MR simulator to practice glenoid steps and learn about the surgery procedures. The individuals were then assessed using a post-test. The pre-test and post-test each took 20 minutes to complete for the participants. The participants were only required to complete the training once during the 20 minutes of allowed interaction with the MR simulation. Every activity was carried out on the same day, and a 5-min break was given after the pre-test and after all interactions with the MR simulation for glenoid processes were finished. Scores for each participant are shown on the Y-axis in Fig. 13 and Fig.14 on a scale from 0 to 100. (Participants are on the X-axis). If participants improved by at least 40 points, the lead surgeon involved in the learning designated the improvement as "substantial”. "Moderate" improvement was defined as a change of 10 to 40 points. The outcomes of the evaluation activity are as follows:

- A majority of medical participants (13 of 17 for medical and 8 of 10 for non-medical) demonstrated an improved understanding of the glenoid process.
- 13 medical participants and 8 non-medical participants showed significant improvement, and 4 medical and 2 non-medical participants performed at a level termed moderate improvement.
- 9 medical and 7 non-medical participants received a score of 0 in the pre-test as they had no prior understanding of the surgical process (Fig. 7 and Fig. 8). After interacting with the VSE, 8 participants showed an improvement of 50 points, respectively (as shown in Fig. 7 and Fig 8). Zero scores for non-medical participants are relatively high regarding medical participants which seems logical as non-medical participants have no prior knowledge of surgery procedures. The reason that a few non-medical participants answered the pre-test question is either they have prior experience with this surgery, or they know someone that experienced it, and they have some basic knowledge about the procedure.

- Non-medical and medical participants were randomly assigned to two various groups. Group 1 with text cues for instructions and group 2 with voice cues and a 3d medical avatar for training instructions. Based on the results participants in group 2 gained slightly better knowledge during simulation. For the next phase, the plan is to have more participants from each group to examine if voice cues and Avatar roles could have a significant impact on the learning process for simulation.

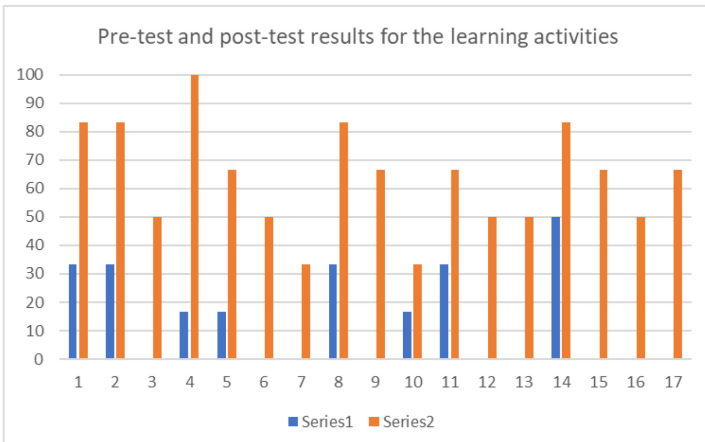

Fig. 7. Pre-test and post-test results for the learning activities for first responders and nurses

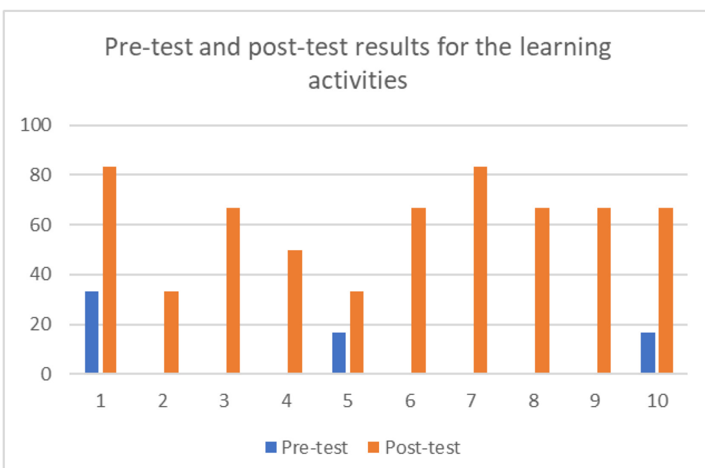

Fig. 8. Pre-test and post-test results for non-medical students

5 Comprehension assessment

Comprehension measures whether a participant can understand the intended meaning of a step (label or voice) and can draw the correct conclusions from the instructions [39]. Participants in the study were also assessed on their ability to perform the intended actions following the learning process through the simulation. Two different scenarios were conducted for the comprehension test. Non-medical students were randomly divided into two groups.

In Group 1, students underwent training using the first simulation under normal circumstances in a relatively quiet room. Following the training session, they were asked to complete pre- and post-tests to assess their understanding. Subsequently, they engaged in a second simulation, referred to as the challenge test, where they were observed to determine if they could identify any incorrect steps during the simulation. In Group 2, additional auditory distractors were introduced during the training simulation. These distractors included ambient sounds from a surgery room, random intensive care unit sounds, and random voices related to various surgical processes. Similar to Group 1, participants in Group 2 completed pre- and post-tests after the training session. They were then asked to perform a second simulation and observe if they could recognize any incorrect steps during the simulation, despite the presence of distractors. These allowed for a comparison of performance between the two groups, assessing the impact of the added auditory distractors on the participants' ability to recognize and execute the correct steps during the simulation.

For Phase 1 of the study involved 10 students in each group. To evaluate the results, the Mann-Whitney test [40] was utilized. Unlike the t-test that compares the means, the Mann-Whitney U test compares a randomly selected value from group 1 to a randomly selected value from group 2. A maximum of 100 scores for pre/post-tests and a maximum of 100 scores for challenge tests were assigned. For the second phase 10 new students randomly assigned in these two groups. In order to perform a more in-depth study of distractors and stressors, also assessment of training simulation this time there were more challenge tests.

For Phase 2, a new set of 10 students was randomly assigned to the two groups. This phase aimed to conduct a more comprehensive examination of distractors, stressors, and the assessment of the training simulation. Additionally, more challenging tests were introduced for both new groups to further evaluate their performance. In this phase, a maximum score of 100 was considered as the criterion for evaluation.

Phase 1 hypothesis: H0: The distractors don't have any effect on participants' learning process in group 2 versus H1: The distractors affect the learning process of group 2.
Since p-value < α, H0 is rejected. The difference between the randomly selected value of Group 1 and the Group 2 populations is big enough to be statistically significant. The p-

value equals 0.04387, (p (x ≤ Z) = 0.9781). It means that the chance of a type I error (rejecting a correct H0) is small: 0.04387 (4.39%). The test statistic Z equals 2.0154, which is not in the 95% region of acceptance: [-1.96: 1.96]. U=77, is not in the 95% region of acceptance: [24.2285: 75.7715]. The observed results suggest that there is a medium-level difference between the values obtained from Group 1 and Group 2. This indicates a significant distinction in the performance between the two groups. The histograms for both groups are presented in Figure 9, providing a visual representation of the distribution of scores or values within each group. These histograms offer further insights into the performance characteristics of the participants in each group.

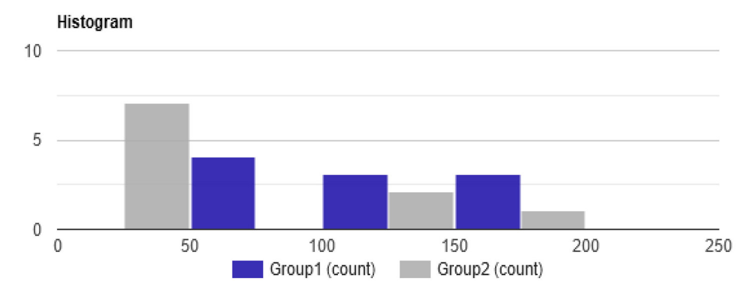

Fig. 9. Histogram for two groups in phase 1

Phase 2 hypothesis: H0: The distractors don't have any effect on participants' learning process in group 2 versus H1: The distractors affect the learning process of group 2.
Since p-value < α, H0 is rejected. The difference between the randomly selected value of Group1 and the Group2 populations is big enough to be statistically significant. The p-value equals 0.04344, (p(x≤ Z) = 0.9783). It means that the chance of type I error (rejecting a correct H0) is small: 0.04344 (4.34%). The test statistic Z equals 2.0195, which is not in the 95% region of acceptance: [-1.96: 1.96]. U=22, is not in the 95% region of acceptance: [3.7653: 21.2347]. The observed standardized effect size, Z/√(n1+n2), is large (0.64). The observed results suggest that there is a medium-level difference between the values obtained from Group 1 and Group 2. This indicates a significant distinction in the performance between the two groups. The histograms for both groups are presented in Figure 10, providing a visual representation of the distribution of scores or values within each group. These histograms offer further insights into the performance characteristics of the participants in each group.

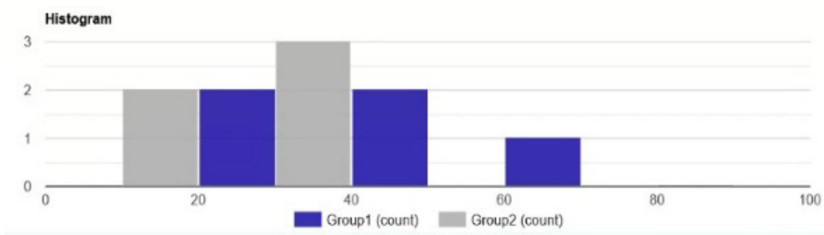

Fig. 10. Histogram for two groups in phase 2

While the sample size in this study may be limited, the results obtained provide initial evidence supporting the conclusion that the modeled distractors, based on real-life stressors, have an impact on the performance of both non-medical and medical students. These findings suggest that when designing training applications for surgical procedures, it is crucial to consider the inclusion of realistic stressors. However, further research with larger sample sizes would be necessary to draw more comprehensive and robust conclusions. These initial results highlight the importance of incorporating realistic stressors in training simulations to enhance the preparation and performance of students in real-life surgical scenarios.

6 Conclusion and future work

In this paper the design process for a distinct module that serves as a training simulator for RTSA has been explained. Using the engineering Enterprise Modeling Language (eEML), an information-centric modeling approach was suggested for developing this simulator. Detailed information-centric models of the surgical process were also built using eEML, providing a structural foundation for designing and developing this simulator. Validity of MR simulation has been performed and through interactions with medical and non-medical students, the usefulness of this simulator-based training technique was investigated. The majority of participants demonstrated improvement in their comprehension of the RTSA process after interaction activities. Also, it indicated how various types of affordances affect the users' comprehension and skills and knowledge acquisition. Assessment of distractors and impact of MR simulation for learning of complex processes studied in the simulation. Future plan is to create 3D surgical skills challenge tests within the VR as well as MR environments to study the role of various factors in affecting the acquisition of surgical skills.

Acknowledgment

The authors would like to acknowledge funding for the research activities discussed in this paper through grants from the National Science Foundation (2050960 and 2106901), and the Oklahoma Center for Advanced Science and Technology (OCAST). We also wanted to thank the collaborators at Dignity Regional Medical Center (Phoenix, AZ), Dignity Health, Yavapai Regional (Prescott Valley, AZ) and Northern Oklahoma College (NOC) (Enid and Tonkawa campuses). We also wanted to express our thanks to Vern McKinney (head ER nurse at Dignity Health, Yavapai Regional) and Dr. Nikole Hicks (Nursing Division Chair at Northern Oklahoma College) for interacting with us as part of these research activities.

References

[1] J. Cecil, S. Kauffman, A. Gupta, V. McKinney, and M. M. Pirela-Cruz, "Design of a human centered computing (HCC) based virtual reality simulator to train first responders involved in the Covid-19 pandemic," in *2021 IEEE International Systems Conference (SysCon)*, 2021: IEEE, pp. 1-7.

[2] J. Cecil, S. Albuhamood, P. Ramanathan, and A. Gupta, "An Internet-of-Things (IoT) based cyber manufacturing framework for the assembly of microdevices," *International Journal of Computer Integrated Manufacturing,* vol. 32, no. 4-5, pp. 430-440, 2019.

[3] A. Sadeghi Milani, A. Cecil-Xavier, A. Gupta, J. Cecil, and S. Kennison, "A Systematic Review of Human–Computer Interaction (HCI) Research in Medical and Other Engineering Fields," *International Journal of Human–Computer Interaction,* pp. 1-22, 2022.

[4] M. Walker, J. Brooks, M. Willis, and M. Frankle, "How reverse shoulder arthroplasty works," *Clinical Orthopaedics and Related Research®,* vol. 469, pp. 2440-2451, 2011.

[5] drgordongroh.https://www.drgordongroh.com/orthopaedic-injuries treatment/shoulder/standard-total-shoulder-replacement/ (accessed.

[6] K. Singhal and R. Rammohan, "Going forward with reverse shoulder arthroplasty," (in eng), *J Clin Orthop Trauma,* vol. 9, no. 1, pp. 87-93, Jan-Mar 2018, doi: 10.1016/j.jcot.2017.10.002.

[7] K. R. Gertsch *et al.*, "Description and validation of a structured simulation curriculum for strabismus surgery," *Journal of American Association for Pediatric Ophthalmology and Strabismus,* vol. 19, no. 1, pp. 3-5. e3, 2015.

[8] K. Kunkler, "The role of medical simulation: an overview," *The International Journal of Medical Robotics and Computer Assisted Surgery,* vol. 2, no. 3, pp. 203-210, 2006.

[9] A. Gupta, J. Cecil, M. Pirela-Cruz, and P. Ramanathan, "A virtual reality enhanced cyber-human framework for orthopedic surgical training," *IEEE Systems Journal,* vol. 13, no. 3, pp. 3501-3512, 2019.

[10] C. E. Hughes, C. B. Stapleton, D. E. Hughes, and E. M. Smith, "Mixed reality in education, entertainment, and training," *IEEE computer graphics and applications,* vol. 25, no. 6, pp. 24-30, 2005.

[11] W. Kopeć *et al.*, "VR with older adults: participatory design of a virtual ATM training simulation," *IFAC-PapersOnLine,* vol. 52, no. 19, pp. 277-281, 2019.

[12] S. Wijewickrema *et al.*, "Design and evaluation of a virtual reality simulation module for training advanced temporal bone surgery," in *2017 IEEE 30th international symposium on computer-based medical systems (CBMS)*, 2017: IEEE, pp. 7-12.

[13] A. Gupta, J. Cecil, and M. Pirela-Cruz, "A cyber-human based integrated assessment approach for orthopedic surgical training," in *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*, 2020: IEEE, pp. 1-8.

[14] X. Wu *et al.*, "Mixed Reality Technology–Assisted Orthopedics Surgery Navigation," *Surgical Innovation,* vol. 25, no. 3, pp. 304-305, 2018.

[15] J. Hooper *et al.*, "Virtual reality simulation facilitates resident training in total hip arthroplasty: a randomized controlled trial," *The Journal of arthroplasty,* vol. 34, no. 10, pp. 2278-2283, 2019.

[16] F. Jefferson* *et al.*, "MP35-08 IMMERSIVE VIRTUAL REALITY (IVR) RENAL MODELS AS AN EDUCATIONAL AND PREOPERATIVE PLANNING TOOL FOR LAPAROSCOPIC DONOR NEPHRECTOMY: INITIAL EXPERIENCE," *The Journal of Urology,* vol. 201, no. Supplement 4, pp. e508-e508, 2019.

[17] N. N. Al-Hiyari and S. S. Jusoh, "Healthcare Training Application: 3D First Aid Virtual Reality," in *International Conference on Data Science, E-learning and Information Systems 2021*, 2021, pp. 107-116.

[18] X. Wang and X. Wang, "Virtual Reality training system for surgical anatomy," in *Proceedings of the 2018 International Conference on Artificial Intelligence and Virtual Reality*, 2018, pp. 30-34.

[19] J. Erickson *et al.*, "Mixed-reality holographic-assisted placement of glenoid guidewire in shoulder arthroplasty: preliminary comparison to patient-specific instrumentation in B2 glenoid model," in *Seminars in Arthroplasty: JSES*, 2022, vol. 32, no. 4: Elsevier, pp. 688-696.

[20] B. Van Duren, K. Sugand, R. Wescott, R. Carrington, and A. Hart, "Augmented reality fluoroscopy simulation of the guide-wire insertion in DHS surgery: A proof of concept study," *Medical engineering & physics,* vol. 55, pp. 52-59, 2018.

[21] K. Logishetty, L. Western, R. Morgan, F. Iranpour, J. P. Cobb, and E. Auvinet, "Can an augmented reality headset improve accuracy of acetabular cup orientation in simulated THA? A randomized trial," *Clinical Orthopaedics and Related Research,* vol. 477, no. 5, p. 1190, 2019.

[22] S. Z. Alsofy *et al.*, "Virtual reality-based evaluation of surgical planning and outcome of monosegmental, unilateral cervical foraminal stenosis," *World Neurosurgery,* vol. 129, pp. e857-e865, 2019.

[23] J. Cecil, A. Gupta, and M. Pirela-Cruz, "An advanced simulator for orthopedic surgical training," *International journal of computer assisted radiology and surgery,* vol. 13, no. 2, pp. 305-319, 2018.

[24] J. Cecil, "Computer aided fixture design: using information intensive function models in the development of automated fixture design systems," *Journal of manufacturing systems,* vol. 21, no. 1, pp. 58-71, 2002.

[25] J. A. Cecil, *Fixture design in a computer-integrated manufacturing (CIM) environment.* Texas A&M University, 1995.

[26] J. Cecil, "A clamping design approach for automated fixture design," *The International Journal of Advanced Manufacturing Technology,* vol. 18, no. 11, pp. 784-789, 2001.

[27] P. Ehn *et al.*, "The Envisionment Workshop-from visions to practice," in *Proceedings of the Participatory Design conference*, 1996: MIT Boston, pp. 141-152.

[28] M. J. Muller and S. Kuhn, "Participatory design," *Communications of the ACM,* vol. 36, no. 6, pp. 24-28, 1993.

[29] P. L. Jensen, "The Scandinavian Approach in Participatory Ergonomics," in *13th Triennial Congress of the International Ergonomics Association*, 1997: Finnish Institute of Occupational Health, pp. 13-15.

[30] J. Cecil, "A Functional Model of Fixture Design to aid in the design and development of automated Fixture Design systems’," *Journal of Manufacturing Systems,* vol. 21, no. 1, pp. 58-72, 2002.

[31] J. Cecil, "TAMIL: an integrated fixture design system for prismatic parts," *International Journal of Computer Integrated Manufacturing,* vol. 17, no. 5, pp. 421-434, 2004.

[32] J. Cecil, A. Gupta, and M. Pirela-Cruz, "An advanced simulator for orthopedic surgical training," *International journal of computer assisted radiology and surgery,* vol. 13, pp. 305-319, 2018.

[33] A. Gupta, "Investigation of a holistic human-computer interaction (HCI) framework to support the design of extended reality (XR) based training simulators," PhD, Computer Science, Oklahoma State university, 2021.

[34] E. M. McDougall, "Validation of surgical simulators," *Journal of endourology,* vol. 21, no. 3, pp. 244-247, 2007.

[35] J. Cecil, S. Kauffman, A. Cecil-Xavier, A. Gupta, V. McKinney, and M. Sweet-Darter, "Exploring Human-Computer Interaction (HCI) criteria in the design and assessment of Next Generation VR based education and training environments," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2021: IEEE, pp. 524-525.

[36] A. Gupta, J. Cecil, and M. 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)*, 2022: IEEE, pp. 652-653.

[37] A. Gupta, J. Cecil, M. Pirela-Cruz, R. Shamsuddin, S. Kennison, and C. Crick, "An Investigation on the Role of Affordance in the Design of Extended Reality based Environments for Surgical Training," in *2022 IEEE International Systems Conference (SysCon)*, 2022: IEEE, pp. 1-7.

[38] A. I. Tergas, S. B. Sheth, I. C. Green, R. L. Giuntoli, A. D. Winder, and A. N. Fader, "A pilot study of surgical training using a virtual robotic surgery simulator," *JSLS: Journal of the Society of Laparoendoscopic Surgeons,* vol. 17, no. 2, p. 219, 2013.

[39] M. A. Gernsbacher, K. R. Varner, and M. E. Faust, "Investigating differences in general comprehension skill," *Journal of Experimental Psychology: Learning, Memory, and Cognition,* vol. 16, no. 3, p. 430, 1990.

[40] N. Nachar, "The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution," *Tutorials in quantitative Methods for Psychology,* vol. 4, no. 1, pp. 13-20, 2008.