HIPS – A Surgical Virtual Reality Training System for Total Hip Arthroplasty (THA) with Realistic Force Feedback

Mario Lorenz*, Maximilian Kaluschke*, Annegret Melzer, Nina Pillen, Magdalena Sanrow, Andrea Hoffmann,
Dennis Schmidt, André Dettmann, Angelika C. Bullinger, Jérôme Perret, Gabriel Zachmann

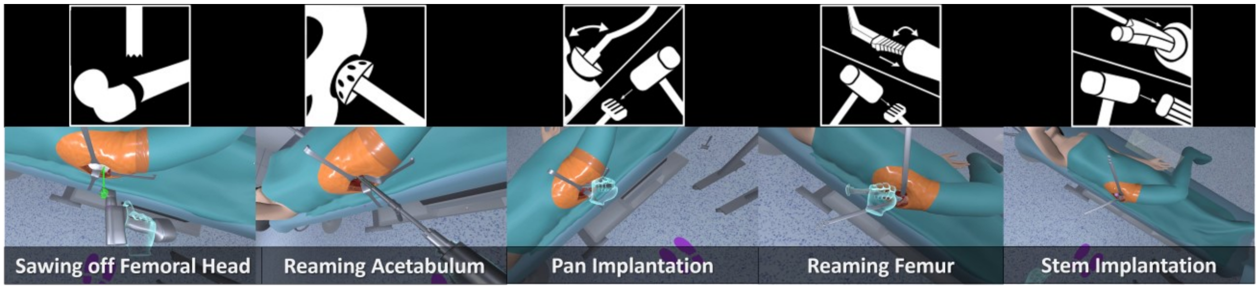


Fig. 1. The five steps of total hip arthroplasty in virtual reality: From left to right: Sawing off the femoral head, Reaming the acetabulum, Inserting the acetabulum implant, Reaming the femur, Inserting the stem implant.

Abstract— Virtual reality training simulations to acquire surgical skills are important for increasing patient safety and save valuable resources, e.g., cadavers, supervision and operating room time. However, as surgery is a craft, simulators must not only provide a high degree of visual realism, but especially a realistic haptic behavior. While such simulators exist for surgeries like laparoscopy or arthroscopy, other surgical fields, especially where large forces need to be exerted, like total hip arthroplasty (THA; implantation of a hip joint protheses), lack realistic VR training simulations. In this paper we present for the first time a novel VR training simulation for the five steps of THA (from femur head resection to stem implantation) with realis-tic haptic feedback. To achieve this, a novel haptic hammering device, an upgraded version of the Virtuose 6D haptic device from Haption, novel algorithms for collision detection, haptic rendering, and material removal are introduced. In a study with 17 surgeons of diverse experience levels, we confirmed the realism, usefulness and usability of our novel methods.

Index Terms— THA, Haptics, Force feedback, Virtual reality, Training, Hip surgery, User Experience, Usability.

◆

1 INTRODUCTION

Learning surgery requires a deep knowledge of anatomy, surgical techniques and a subtle intuition that can only be acquired through extensive practical experience. Therefore, surgical expertise is predominantly taught in a traditional master-apprentice model, where specialist surgeons closely supervise and guide residents during surgery at patients. To enhance the safety and efficacy of surgical interventions, many training methods have been devised to prepare residents for their first surgical experiences. They range from conventional cadaver training to physical anatomical models, and more recently, to virtual reality (VR) or augmented reality (AR) based simulation platforms [1–4]. Particularly, VR/AR technologies in surgical training is very promising as they inherently present fewer ethical dilemmas compared to cadaver training while also offering superior variability and reproducibility in training scenarios [5–8]. Further, body donners are usually very old so that their anatomy might not always be representative for the patients the surgery is performed on. In addition, the fact that the donners are dead and their bodies are often treated chemically for conservation alters the material properties of the tissue [9–12]. Physical models often do not adequately simulate material properties of bones [13] and livestock may have a similar but not equal anatomy of humans and comes with an ethical dilemma [14].

In the 1990s the development of surgical training simulators and have now been integrated into curricula, e.g. arthroplasty (keyhole surgery in the abdomen) [15] and laparoscopy (keyhole surgery in the abdomen) [7]. However, despite these advancements, there remain some surgical disciplines where the availability of realistic virtual training systems is notably lacking. One such discipline is total hip arthroplasty (THA; implantation of a hip joint protheses), where the primary challenge lies in creating a simulation that replicates the substantial forces encountered during the surgical procedure [16]. Given that THA is a procedure performed hundreds of thousands of

---

- *Mario Lorenz is with Chemnitz University of Technology, Professorship for Production Systems and Processes; University Hospital Leipzig, Department of Orthopaedics, Trauma and Plastic Surgery; Medical University of Graz, Division of Macroscopic and Clinical Anatomy. E-mail: mario.lorenz@mb.tu-chemnitz.de*
- *Maximilian Kaluschke and Gabriel Zachmann are with University of Bremen, Germany, Chair of Computer Graphics and Virtual Reality. E-mail: {mkalusch | zach}@uni-bremen.de.*
- *Annegret Melzer is with Chemnitz University of Technology. E-mail: annegret.melzer@mb.tu-chemnitz.de.*
- *Nina Pillen and Magdalena Sanrow are with YOUSE GmbH. E-mail: nina.pillen@youse.de | magdalena.sanrow@gmail.com*
- *Andrea Hoffmann, André Dettmann and Angelika C. Bullinger are with Chemnitz University of Technology, Chair for Ergonomics and Innovation. E-mail: {andrea.hoffmann | andre.dettmann | angelika.bullinger-hoffmann}@ mb.tu-chemnitz.de.*
- *Dennis Schmidt is with FAKT Software GmbH. E-mail: dennis.schmidt@fakt-software.de.*
- *Jérôme Perret is with Haption GmbH. E-mail: jerome.perret@haption.com.*

**Equal contributors*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.
Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx/.

times annually [17], the integration of simulators into the training regime of orthopedic surgeons becomes imperative.

In this paper, we present a number of novel methodologies necessary for a realistic virtual THA simulator, as well as results of a user study investigating its realism. Our simulator is built on top of a preliminary version [18]. This prototype uses a KUKA iiwa robot [19], incorporates tactile feedback mechanisms into the reaming tool [20], and renders the haptic feedback based on material data from biomechanical experiments [16]. Building upon the success of this initial prototype, referred to as the HIPS training simulator, we have expanded its functionalities to encompass now the five most important surgical steps for THA (see Fig. 1):

1. **Cutting off the femoral head** (ball shaped end of thigh bone) with an oscillating saw, followed by
2. **Reaming the acetabulum** (hip pan), with a half-sphere-shaped tool equipped with cutting edges similar to a cheese grate, which provides the shape for the pan implant.
3. **Inserting the acetabulum implant**, which is hammered into the prepared acetabula (press-fit fixation). Inside this implant an inlay is inserted, which poses the gliding surface.
4. **Reaming the femur** (thigh bone), with scrapping tools, which are shaped like the stem implant for preparing the stem implantation. Most often multiple tools in ascending sizes are used until the size of the stem implant is reached.
5. **Inserting the stem implant**, into the prepared femur analogous to the acetabulum implant also using a press-fit fixation, followed by mounting the ball-counterpart on the stem.

The enhanced version of HIPS now guides users through all stages of the THA procedure, with the exception of incision and closure, using novel haptic devices and haptic rendering algorithms. Only for the second step (acetabula reaming) the principles of the collision's detection and haptic rendering algorithms, the 3D-models of the operation room and the patient anatomy were reused.

In this paper, we report on a study assessing the visual and haptic realism of the simulation, its user experience, potential usefulness in medical curricula, as well as on gaining insights for further work. The research questions (RQ) this study should answer are:

RQ1: "Does the HIPS simulator provide realistic visual and haptic feedback?"

RQ2: "How is the user experience and usefulness of HIPS?"

Our systematic review of literature shows that the assessment of surgical VR/AR training simulators utilizing VR or AR often relies on quantitative methodologies [21–26]. Qualitative approaches, such as interviews, are less frequently used to evaluate their overall advantages and disadvantages [27, 28] or to evaluate specific training simulators [29]. Therefore, a mixed method design with focus on qualitative methods was chosen to assess the HIPS simulator. Overall, our key contributions are:

- First VR-simulator for training THA with realistic force feedback as a complete surgical procedure
- Presentation of novel haptic devices capable of simulating realistic forces and torques for THA
- Novel haptic rendering methods for the surgical steps of THA
- Novel insights into surgical training simulators by way of qualitative assessments

2 STATE OF THE ART

2.1 VR/AR-Training Simulators in Orthopaedic Surgery

In orthopedic surgery, the majority of VR/AR applications center around tasks such as operation planning, fracture treatment, and arthroscopy. There is a limited number of applications that specifically address arthroplasty, with an even smaller focus on THA [15, 21, 30–33]. A recent examination of orthopedic VR training simulators utilized in resident training from 2011 to 2021 revealed that only two studies out of 61 were dedicated to hip arthroplasty [21], both of which did not include haptic feedback [34, 35]. Additionally, Sun et al. [33] noted a majority of applications related to implant positioning for THA. The review by Su et al. [30] analyzing the accuracy of inclination, anteversion, and surgical duration identified five studies that stressed the advantages of using VR/AR for THA training. Previous reviews have mostly concentrate on arthroscopy, often overlooking arthroplasty [6, 36]. In a recent review by Syamlan et al. [15], only two applications for training THA [37], respectively acetabular reaming [18, 38] , were identified.

Wiese et al. [39] describe a THA training simulation in VR for the surgical approach (incision to joint dislocation). Hooper et al. [34] assessed a THA VR training simulation based on the ORamaVR (Heraklion, Crete, Greece) software platform, without describing details on implementation and user feedback. Another THA VR training simulation by Younis et al. [40], allows the training from incision to relocating the joint. They generated 3D models of the operating theatre with all surgical instruments and hip anatomy. It allows to practice on either a skeleton model or a model with soft tissue. Nevertheless, the basis for anatomical modelling is not specified. The simulator gives the trainees some instructions about the procedure and also evaluates certain performance indicators. The precise configuration of the simulator, including the user feedback mechanism, is not given. Logishetty et al. [41] present a study where the THA VR Simulation v1.1 from Pixelmolkerei AG, Chur, Switzerland is assessed. It remains unspecified which stages of THA this simulator addresses or the foundation for the anatomical 3D model. However, details about the user feedback mechanism are revealed. It offers both training and assessment modes, providing guidance on surgical steps, instrument positions, labelled anatomical regions, and implant orientation in the training mode. In both modes, the trainees receive feedback on their errors in instrument selection and position, the frequency of prompts given by the simulator (activated when the surgeon fails to make incremental progress for over 30 seconds), hand path lengths, final component orientation, and the duration. In a follow up transfer study, Logishetty et al. [35] trained 32 residents over 6 weeks in a non-haptic VR scenario on the resection of the femoral head, where they improved their performance in the four measured parameter significantly. A further contribution was made by Kaluschke et al. [18] and Knopp et al. [19], presenting a training simulator for reaming the acetabula employing a KUKA iiwa robot for force feedback, representing a forerunner to the HIPS training simulator outlined in this paper. Summarizing the gaps in literature regarding XR simulators for THA: (1) there is no haptic feedback with the exception of our previous work for acetabulum reaming (1 of 5 modules presented here), (2) the realism of the simulators are not studied.

2.2 Haptic Rendering Methods for Bone Simulation

Simulating orthopedic surgeries presents the challenge that material removal needs to be calculated simultaneously with the forces to be fed back to the user. Although there exist several simulation methods that enable material removal, the majority of them are not suitable for haptic rendering, due to a low simulation frequency. Thus, in the following, we only consider methods that are applicable to haptic rendering. In fact, there are several publications on this topic in the haptic rendering literature. One of the earliest and widely utilized method for haptic rendering of arbitrary tools and environments with 6 DOF feedback is the voxmap point-shell (VPS) algorithm by McNeely et al. [42]. This approach has many advantages, such as a fast run-time and conceptual simplicity, but also many downsides, which later publications try to remedy. The approach supports material removal, by assigning each voxel a density, representing the dynamic material state. For example, an early approach by Agus et al. [43] extends VPS to consider the voxels in a closed region around the virtual tool to simulate bone dissection. A limitation of the approach is that tools can only be spherical, besides the fact that material is discretized by voxels, leading to a jagged surface appearance. Later, Morris et al. [44] further extended this approach, by introducing new features that improve the realism of the drilling simulation. Most importantly, they mitigate the stiffness-scaling problem of VPS that arises when multiple voxels overlap at the same time. They introduce a non-linear scaling of the force magnitude, depending on the number of overlapped voxels. Acosta and Liu [45] use classic VPS but tried to reduce the force discontinuities of VPS during burr hole simulation by doing additional boundary checks and separating the physics and haptics tasks into asynchronous threads. Tsai et al. [46] improved the VPS drilling behavior to be more closely related to physical laws, to reduce voxel overlap. As shown the VPS approach, suffers from multiple downsides, most importantly force discontinuities. These occur due to reliance on voxel discretization, precomputed contact normals and discrete collision detection. This issue alone makes the VPS method not suitable for high-force feedback applications, such as THA training. One reason, no high-force THA simulator existed prior to this publication.

Chan et al. [47] developed a simulator for otolaryngology surgery remedying some of the force discontinuities by integrating continuous collision detection and generating the dynamics of the tool by solving contact constraints on the acceleration level, instead of relying solely on overlap. However, another limitation of VPS remains: the discretization of the material removal, resulting in a non-continuous feed rate and visual artefacts, such as aliasing. To overcome these issues, we build upon the sphere-based approach by Kaluschke et al. [48], which does not suffer from force discontinuities and provides fully stepless material removal. Additionally, we enhance haptic rendering by integrating a rigid-body simulation to generate candidate poses, thereby improving the overall realism and effectiveness of the simulation. In summary, the key research gaps in haptic rendering for bone simulation lie in (1) the reliance on voxel-based approaches, which introduce force discontinuities and limit the stability of force feedback, and (2) the discretized material removal process, which causes visual artifacts and inconsistencies in the simulated material volume, hindering overall realism.

3 METHODS AND MATERIAL

3.1 The HIPS Training Simulation

The HIPS training simulator has a multi-user functionality and enables interactive training sessions where a trainee can virtually perform a surgery under the guidance of a supervising specialist who provides instructions through voice, pointing gestures, and visual cues within the VR environment. Moreover, multiple observers can watch the training sessions in VR. The instructions provided in HIPS occur within the simulated operating room (Fig. 2.), where the virtual patient is positioned in the center. A large interactive information display is situated behind the virtual patient to assist the trainee. Positioned to the right side of the patient is a table containing the essential tools for each module. At the beginning of each module, trainees are briefed on the surgical step and key points to observe during surgery. Pictograms are used to highlight important points with detailed information. Once the module starts, the information screen explains the necessary steps and their corresponding progress, e.g. when to activate and deactivate the saw and reamer, important angels for acetabulum reaming or the order of the femur rasps. Upon completion of a training module, the trainee is asked to evaluate their own performance, followed by the systems feedback on their objective performance. The overall training assessment is shown after completing the final module. The primary mode of operation in HIPS is the self-guided single-user mode, where the trainee is directed through the key steps of the procedure utilizing textual instructions, images, and visual aids presented at the virtual patient's location (e.g., the optimal position for placing the saw). The application provides three levels of assistance for users to select from: a beginner level offering comprehensive guidance (see Fig. 3), an intermediate level with essential support information only, and an expert level with no guidance provided. The surgical steps covered by HIPS are organized into distinct modules that can be completed individually or consecutively. Progress achieved at the end of one module does not carry over to the next, as each module begins with an ideal starting point. This approach is implemented to ensure that errors made in one step do not impact the subsequent steps, maintaining a realistic training environment. The system is designed to detect errors occurring during the procedure and provide immediate feedback (red cross popping up). Based on the severity of the errors, points were associated, which are then summed up forming a score. The anatomical representations were accurately crafted based on insights from anatomical literature, THA guidelines, and consultations with expert surgeons. The development of the HIPS received input from highly experienced surgeons at Zeisigwaldklinken Bethanien in Chemnitz, Germany, with extensive expertise in implanting thousands of hip prostheses. The workflow concept, user interface, user feedback, user interaction relevant errors, haptic behavior, validity of the anatomy and operation room were constantly discussed with these consulting surgeons throughout the development process.

Both haptic devices and the handpiece (see section 3.2) are connected to the collision detection and material removal module (see section 3.3), which handles the communication with these devices. The VR-scene is implemented in Unity3D integrating the collision detection and material removal module as an external library.

Fig. 2. VR view of the operating room with central information screen behind the patient in the center and typical equipment.

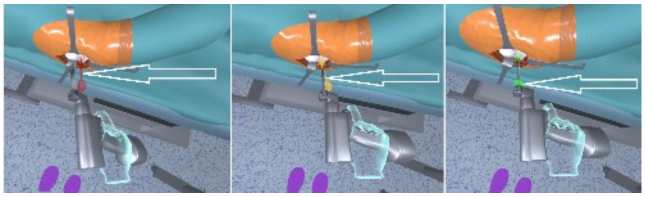


Fig. 3. White arrows pointing to visual indicators for the correct placement of the saw on the femoral neck in module 1. Left to right (red): wrong, medium deviation (yellow), correct (green).

3.2 Haptic Devices

For the haptic simulation of the sawing and reaming in modules 1 and 2 a prototype of a force-feedback device (Virtuose 6D from Haption) was used (see Fig. 8). It represents a progression from the conventional Virtuose 6D, characterized by elevated torques across all motors, resulting in a peak force of 70 N in translation (5 Nm in rotation) at the wrist within the entire workspace, which covers a fair spectrum of forces occurring during acetabula reaming [16]. These augmented motor torques are obtained through increased reduction factors, thus expanding the stability and enabling a control stiffness of up to 12 kN/m in translation (40 Nm/rad in rotation) at a refresh rate of 1 kHz. A handpiece used for the saw and the reamer was built and equipped with an Arduino microcontroller, vibration motors and trigger buttons (see Fig. 8). When the buttons are pushed, the vibration motors are turned on and a signal is sent via WIFI using TCP/IP sockets to the PC running the HIPS-simulation to activate the saw respectively the reamer. The handpiece is connected to the Virtuose 6D via a Haption standard adapter.

For the simulating the hammering steps in modules 3-5, we develop a dedicated novel device, depicted in 0 during stem implantation [49]. According to data we gathered from biomechanical experiments using fresh frozen and Thiel fixated human tissues (currently unpublished), impulses up to 20 kN are exerted during these procedures (see Fig. 7). HTC Vive trackers are used to align the hammer (a) the hand gripping the tool (b) and the hammering device in the virtual operation theater (see Fig. 4). We used Vive trackers for economic reasons and to not further increase the systems complexity by introducing another tracking system. The position of the hammering device is displayed as a ghost of the respective instrument, which needs to be aligned with the visualized instrument already correctly positioned in the situs.

The general working principle of the hammering device is an obstructed axial movement in direction of the hammering impulse (see (d) and yellow arrow in 0). The obstruction is achieved by clamping an alloy block using Shimano V-break bicycle breaks. The Bowden train of the breaks is pulled by an electrical cylinder with 100 mm stroke length allowing for a continuous control of the clamping force. The maximum movement length is 55 mm. A rotational potentiometer is used to measure the displacement using an acoustic guitar string wound around the potentiometer to transform the linear movement into a rotation. No forces were measured in the version used for this study as the clamping of the breaks was controlled by displacement alone (see section 3.3 for details). An STM32F103 BlackPill-34 microcontroller is used for controlling the hammering device and it is connected to a PC via USB. A DLL provides the control interface for the haptic rendering simulation.

Between modules 2 and 3 the user needs to change the haptic devices. Upon loading module 3 the VR scene will shift so that the situs is aligned with the hammering device. The user now just has to the situs, hence the hammering device. As both haptic devices stay close to each other (~1 m), the user can take these steps without assistance, given that no obstacles are placed on the way.

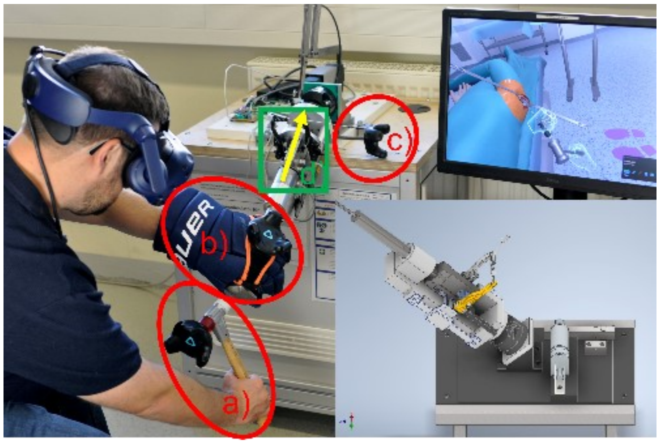


Fig. 4. Hammering device. a) Hammer with VIVE tracker b) Hand with hockey protection glove and VIVE tracker c) VIVE tracker at the hammering device d) Axial movable break component, yellow arrow indicating the movement direction.

3.3 Haptic Rendering Method

The first two surgical tasks (sawing, reaming) are rendered using traditional kinesthetic haptic feedback. To achieve this, we developed a novel rendering method that supports arbitrary tools, 6-DOF feedback, and continuous material removal at haptic rates. The latter three tasks involve hammer impacts, which we render using a simplified model specifically designed for our novel impact device.

3.3.1 Intersection-Free Bone Simulation with Contact Set

Building on Kaluschke et al. [48], we developed a rigid-body simulation (see Fig. 5). This means that before the continuous collision detection (CCD) step, an integration step is done. Thus, all candidate poses of the tool are not directly generated by the Haptic Tool (T_H), but instead by the integration of motion, based purely on physical laws that act upon the Graphic Tool (T_G). Consequently, all interactions are mediated by a user-interaction spring (UIS), similar to the virtual coupling (VC) method proposed by McNeely et al. [42]. The UIS generates external forces that align T_G with T_H. By still performing CCD, the simulation remains completely intersection-free. To detect multiple contacts, we apply a surface offset by increasing all sphere radii of one object by a constant contact margin (e.g. 2 mm). CCD integration eliminates pop-through events, to ensure that non-continuous or erroneous contact normals no longer occur, as all objects stay non-intersecting.

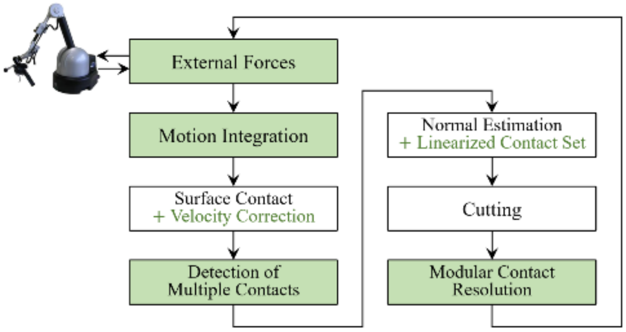


Fig. 5.The haptic rendering system, in comparison to the one presented by Kaluschke et al. [48]. The green steps are structural changes to the original system.

Contacts are resolved using a novel hybrid method. We classify contacts by their relative point velocity into colliding and resting contacts, following Mirtich and Canny [50]. For colliding contacts, we iteratively apply partial impulses, checking whether the velocities at all contacts are receding, indicating the contact will be resolved during the next integration step. In such cases, the contact resolution stops. Hence, the partial impulses are used to iteratively solve velocity constraints imposed by the contacts on the tool. For resting contacts, we found that employing a penalty-based force proportional to the intersection volume outperforms the impulse-based method, as it better handles positional drift. Overall, integrating the rigid-body simulation significantly improved the haptic rendering experience, by eliminating most of the artificial friction. This improvement is due to our faster convergence compared to the original method, which could not render lateral movement during contact without artificial friction.

We adapted the material removal process to the new method, as exact surface contact can no longer be guaranteed. Instead, the tool might slightly offset from the surface within the contact margin. Thus, during surface estimation, we compute a minimal separating distance for all contacts, which we add to the radius extension during material removal. This effectively considers the tool to be in exact contact with the environment, resulting in a consistent feed rate during material removal by removing the variable of how “close” the contact is. The material removal is instead modulated inversely by the global material density, global contact volume and directly by the local pressure that the user is applying on to the environment. Additionally, we precompute a drill direction for tool spheres, which we use to modulate the drilling contribution based on the angle between local contact normal and drill direction. This allows us to more realistically simulate the material removal of the saw, as its blade can only remove material in the direction of the saw teeth. These modulations lead to a radius extension (around 10 µm) that creates overlap between tool and environment spheres, which is resolved by directional shrinking of the spheres. Meaning we move overlapping environment spheres by half of the translational depth of penetration (dop) and reduce its radius by half of the dop. Consequently, we reduce the material volume with infinite detail, without any steps.

3.3.2 Stable Feedback without Intersection

An important part to the UIS is force saturation, which is not a part of VC. Force saturation ensures that the stiff UIS does not generate a force $\vec{\mathbf{f}}_s$ that is so large that it overpowers the contact forces $\vec{\mathbf{f}}_c$. In case

$|\vec{\mathbf{f}}_\mathrm{s}| \gg |\vec{\mathbf{f}}_\mathrm{c}|$, the sum of external forces would tend to generate already intersecting candidate poses during integration. This is especially important, as we only allow contact penetrations until the objects touch, meaning penalty forces $\vec{\mathbf{f}}_\mathrm{c}$ can only grow in a limited range (e.g. 2 mm). By saturating both forces of UIS and contacts, we make it likely that candidate poses during integration are close to $\mathrm{T_G}$'s current pose, in case the tool is in contact (see Fig. 6). Of course, if there is no contact, the UIS force will grow large enough to allow free movement without the feeling of inertia. The exact threshold for that needs to be manually tuned.

$$|\vec{\mathbf{f}}_\mathrm{c}| \leq 5\ \mathrm{N}, \qquad |\vec{\mathbf{f}}_\mathrm{s}| \leq 1.5\ \mathrm{N}, \qquad |\vec{\boldsymbol{\tau}}_\mathrm{s}| \leq 5\ \mathrm{Nm}$$

where the saturation of $\vec{\mathbf{f}}_\mathrm{c}$ also indirectly defines a saturation of

$$\vec{\mathbf{f}}_\mathrm{s} = \mathbf{T}\left({}^\mathrm{W}\mathbf{H}_{\mathrm{T_G}}\,{}^\mathrm{W}\mathbf{H}_{\mathrm{T_H}}^{-1}\right)\mathrm{k_t} - \left(\vec{\mathbf{v}}(\mathrm{T_H}) - \vec{\mathbf{v}}(\mathrm{T_G})\right)\mathrm{b_t^{rel}} - \vec{\mathbf{v}}(\mathrm{T_H})\mathrm{b_t^{abs}}$$

$$\vec{\boldsymbol{\tau}}_\mathrm{s} = \mathbf{R}\left({}^\mathrm{W}\mathbf{H}_{\mathrm{T_G}}\,{}^\mathrm{W}\mathbf{H}_{\mathrm{T_H}}^{-1}\right)\mathrm{k_r} - \left(\vec{\boldsymbol{\omega}}(\mathrm{T_H}) - \vec{\boldsymbol{\omega}}(\mathrm{T_G})\right)\mathrm{b_r^{rel}} - \vec{\boldsymbol{\omega}}(\mathrm{T_H})\mathrm{b_r^{abs}}$$

where ${}^\mathrm{W}\mathbf{H}_{\{\mathrm{T_H},\mathrm{T_G}\}}$ is a homogenous matrix that transforms the tool's centre of mass from local space to world space W, for the respective poses, and $\mathbf{T}(\mathbf{M})$, $\mathbf{R}(\mathbf{M})$ extract the translation vector and rotation axis (scaled by the rotation angle) of a transformation $\mathbf{M}$. The stiffness and damping constants need to be tuned for the individual force feedback device, and they do not necessarily need to hold the same value for the virtual simulation and the real haptic rendering.

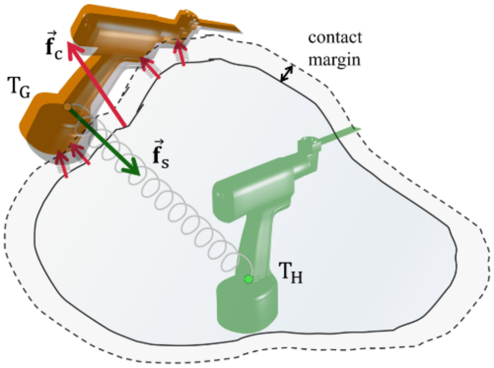


Fig. 6. The user interaction spring mediates the interaction between real and virtual world. Balancing the forces of environment and user by saturation improves stability.

Table 1. Impact device parameters per surgical task.

Task	*c*	i_1 (mm)	b_0	b_1
Hip implanting	0.01	5	0.35	1.0
Femur rasping	0.05	8	0.35	0.8
Femur implanting	0.2	20	0.1	0.35

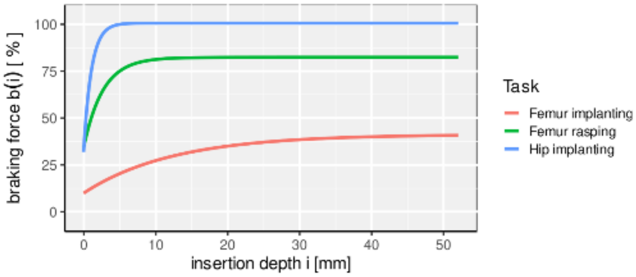


Fig. 7. The simplified impact rendering model.

3.3.3 Impact Rendering Method

We use the novel hammering device to render all surgical tasks that involve hammering, as those are impossible to render on traditional kinesthetic haptic devices. The hammering device provides the insertion depth of the sliding mechanism as output, and allows the actuation of the braking force acting on the sliding mechanism.

Through extensive biomechanical experiments, we have found the relationship between the insertion depth and hammer impact force. Based on this data, we have developed a material model that adjusts the braking force $\mathrm{b}(i)$, based on the current insertion depth i in mm:

$$\mathrm{b}(i) = \frac{\mathrm{c}^{i/i_1} - 1}{\mathrm{c} - 1}(\mathrm{b}_1 - \mathrm{b}_0) + \mathrm{b}_0$$

where c, i_1, b_0 and b_1 were tuned together with surgeons (see Tab. 1). This leads to a model that changes the insertion rate of the instrument that is struck by the hammer (see Fig. 7). The surgeon then haptically feels how far the instrument travels in relation to how hard they struck it. This ratio gives the surgeon information of how far the instrument has been inserted in relation to how far it can be inserted, without damaging any surrounding tissue.

3.4 Experimental Setup

All procedures were reviewed and approved by the local ethics board (Ethikkommission der TU Chemnitz, case-no.: #101602133). The assessment of the HIPS training simulator was carried out at the Clinic for Orthopaedics, Trauma Surgery, and Sports Traumatology at the Zeisigwaldklinik Bethaninen Hospital in Chemnitz, Germany. The majority of the Clinic's members participated in the assessment, representing a diverse range of experience levels from students via residents to seasoned surgeons. Participants were first briefed on the study's objectives and data handling procedures. Next, they provided written consent. In the initial assessment phase, participants disclosed demographic information such as age, gender, current residency year or years since completion, previous VR experience, participation in a previous non-haptic evaluation study of HIPS and their involvement in annual THA procedures. Following this, they were introduced to the VR system (HTC VIVE Pro Head Mounted Display (HMD)) and started the HIPS training simulation, completing all five modules (see Fig. 8). Participants were encouraged to comment throughout the simulation by employing the think-aloud method. After the completion of each module, the participants were asked standardized module specific questions regarding the visual, haptic and overall realism on a five-point-Likert-scale (see Tab. 3).

Fig. 8. A participant performing acetabula reaming using the Virtuose 6D.

Table 2. Questionnaire for evaluating user acceptance and intention of use of the HIPS training simulator.

Q1: The 3D presentation of the situs was realistic.
Q2: The 3D presentation of the operating room was realistic.
Q3: The HIPS simulator helps to learn how to insert an endoprosthesis.
Q4: Time passed quickly when I used the HIPS simulator.
Q5: Learning how to use the HIPS simulator is easy.
Q6: I enjoyed using the HIPS simulator.
Q7: I would recommend the HIPS simulator to medical students.
Q8: I would recommend the HIPS simulator to residents.
Q9: I would recommend the HIPS simulator to medical specialists.
Q10: The HIPS simulator should enable patient-specific training.
Likert-scale end points: 1 – strongly disagree; 5 – strongly agree

The post-assessment began immediately after the conclusion of the five training modules. Participants, still wearing the HMD, rated their level of Presence on a scale of 1 to 10 using a single-item question from Bouchard et al. [51]. We measured presence in a sense of a control variable to check, that HIPS achieves sufficiently high ratings to induce the feeling of being in an operating room. Following this, participants removed the HMD and answered a 10-item questionnaire on user acceptance and intention of use on a five-point-Likert-scale (0). This questionnaire was adapted from Fang et al. [52], who evaluated a haptic-based VR temporal bone simulator based on the Technology Acceptance Model (TAM) by Davis [53]. The reason for tailoring the questionnaire to the HIPS-simulator was to enhance participant engagement and understanding. Given the evaluation was conducted during regular hospital working days, a concise assessment was crucial to avoid prolonged participant engagement. Therefore, only essential questions were included. Finally, participants were given the opportunity to offer additional comments and summarize their overall experience. Time and accuracy were not measured as the aim of this study was to not assess any training effect or compare the surgical capabilities between residents and specialist, but rather validating the realism of HIPS. RQ1 should be answered using Q1, Q2 from Tab. 2 and Tab. 3. For answering RQ2 Q3 to Q10 from Tab. 2 were chosen. Further, the qualitative feedback should help answering both RQs.

3.5 Data Analysis

The descriptive statistics were computed with SPSS version 29.0 developed by IBM. In terms of qualitative analysis, each new input from a participant was documented by the experimenter in a spreadsheet. Subsequently, the feedback was coded, with similar responses being clustered together. The identified usability issues were quantified to assess the severity of the obstacles. Instances where a participant reiterated a previous comment from another participant were also documented. One participant could only finish module 1 and another one had to stop after completing module 2. Further, 3 residents had not performed the steps of module 3-5 at a patient. Therefore, they were only asked about the visual realism of these modules, which they know from assisting during surgery.

Table 3. Module specific questions for evaluating realism.

Module 1: Sawing off the femoral head
M1_2: Sawing off the femoral head was realistic.
M1_2: Solely visually: sawing off the femoral head looked realistic.
M1_3: Solely haptically: sawing off the femoral head felt realistic.
Module 2: Reaming the acetabulum
M2_1: Reaming the acetabulum was realistic.
M2_2: Solely visually: Reaming the acetabulum looked realistic.
M2_3: Solely haptically: Reaming the acetabulum felt realistic.
Module 3: Implanting the pan
M3_1: Implanting the pan was realistic.
M3_2: Solely visually: Implanting the pan looked realistic.
M3_3: Solely haptically: Implanting the pan felt realistic.
Module 4: Scraping out the femur
M4_1: Scraping out the femur was realistic.
M4_2: Solely visually: Scraping out the femur looked realistic.
M4_3: Solely haptically: Scraping out the femur felt realistic.
Module 5: Implanting the stem
M5_1: Implanting the stem was realistic.
M5_2: Solely visually: Implanting the stem looked realistic.
M5_3: Solely haptically: Implanting the stem felt realistic.
Likert-scale end points: 1 – strongly disagree; 5 – strongly agree

3.6 Demographics

There were 17 medical staff participating in the study with a mean age of 39.53 (SD=7.63; 29-57 [min-max]) years, and a mean height of 179.47 (SD=6.30; 171-193 [min-max]) cm. The average annual participation in THA is 86.06 (SD=106.99; 5-400 [min-max]). The 8 resident participants were on average in year 3.13 of their residency (SD=1.64; 1-6 [min-max]). One of the residents self-identified as female and 7 as male. The 9 specialist surgeons (all male) had on average 11.33 (SD=6.12; 6-24 [min-max]) years of experience after their residency. Seven participants (3 specialists; 4 residents) did not have experience with VR and did not partake in the visual-VR only evaluation of HIPS whilst 10 had previous VR experience. Eight participants partook in a previous visual-VR only evaluation of HIPS without any haptic feedback.

4 RESULTS

4.1 Realism, User Acceptance and Presence

The user acceptance results show a very overall positive reception of HIPS (see Tab. 4, Fig. 9). The 3D presentation of the situs and operating room were perceived as realistic (Q1, Q2) and, as very helpful to learn THA (Q3). Further, HIPS was fun and easy to use (Q4-Q6). From the target group of medical students and residents HIPS received high recommendation (Q7, Q8). Only for training specialist HIPS was rather not recommended (Q9). Lastly, there was a strong wish for HIPS to enable patient-specific training (Q10). Presence was rated fairly high (M = 7.20, SD = 1.90).

The module specific questions regarding the realism of HIPS are consistent with the overall good results (see Tab. 4, Fig. 10). In all modules, the visual realism was higher than the haptic realism, with the combined realism in-between or on par with the haptic realism. Despite the very different haptic tasks and devices, the ratings on haptic realism were very similar: their means ranged between 3.35 (sawing in module 1) and 3.87 (femur scraping in module 4). Similarly, although on a higher level, visual realism, was rated between 3.75 (reaming in module 2) and 4.27 (stem insertion in module 5). For the overall realism, the means, of the modules were even closer to each other: 3.56 for the sawing in module 1, and 3.87 for the femur scraping in module 4.

4.2 Qualitative Feedback

Here, we focus on comments about further improvements as they are most important to identify the current limitations and future work directions of HIPS. 66 aspects for improvement across all modules were documented (10 in module 1; 13 in module 2; 19 in module 3; 14 in module 4; 10 in module 5). We partitioned them into five categories, which we derived after an initial analysis of the statements: (1) haptic rendering, (2) visualization (3) auditory feedback, (4) hardware setup, and (5) teaching/user interaction concept. Most of the aspects mentioned by the participants concern category 1 (haptic rendering). In general, the participants perceived the haptic rendering as grossly realistic. However,12 participants stated that the sawing felt too hard and the resistance was too strong. 11 participants felt that their pushing on the saw had too little influence. Also, 11 participants stated that there is no (or too little) difference between cortical bone (outer, hard and dense layer) and the spongiosa (inner, sponge-like kernel) when sawing. Only the first drop in resistance, when breaking through the cortical bone into the spongiosa, could be perceived, whilst the increase in resistance from spongiosa to the cortical bone, at the back of the femoral neck, could not be felt clearly. Only 2 participants felt that the contact of bone and saw in module 1 felt too soft. For the second module, 8 participants felt no increase in resistance when reaming the acetabulum, which in reality occurs as the contact area between acetabulum and tool increases during reaming. This is an important indicator for feeling the progress of reaming. In addition, 5 participants stated that the reamer should vibrate/shake more in the beginning, and the rotational jerk was missing, which happens when the reamer makes first contact with the acetabulum. According to this demand we developed an algorithm, which simulates this behavior (see Supplement 1). For the stem implementation (module 5), 8 participants perceived the required force as too high.

The participants commented very little on the visualization (category 2). The most relevant statement here (9 participants) concerned the missing display of the reamed material that usually accumulates inside the reaming, which can be an indicator for the progress of the reaming. When inserting the rasp into the femur (module 4), 6 participants reported that they did not see a movement of the rasp, even though it was moving. Almost all participants did not care much about missing blood splattering in our simulator. Deviating from this majority, one resident requested this feature, since that would help prepare them for similar splattering in the operating room. In general, the visualization quality was praised, with only minor suggestions for further improvements.

Table 4. Results of user acceptance, intention of use, presence and module specific realism (M – mean, SD – standard deviation)

Q1_Situs_realsim	M=3.87, SD=0.83, min=2, max=5
Q2_OR_realsim	M=4.33, SD=0.62, min=3, max=5
Q3_Helps_learnimg	M=4.27, SD=0.88, min=3, max=5
Q4_Time_passed	M=4.40, SD=0.83, min=3, max=5
Q5_Easy_to_use	M=4.53, SD=0.64, min=3, max=5
Q6_Enjoyment	M=4.80, SD=0.41, min=4, max=5
Q7_Recommend_students	M=4.33, SD=0.82, min=3, max=5
Q8_Recommend_residents	M=4.00, SD=1.20, min=2, max=5
Q9_Recommend_specialists	M=2.60, SD=1.35, min=1, max=5
Q10_Train_patient_specific	M=4.20, SD=1.21, min=2, max=5
Presence	M=7.20, SD=1.90, min=4, max=10
M1_1_Saw_realism	M=3.65, SD=0.61, min=3, max=5
M1_2_Saw_visual_realism	M=4.18, SD=0.70, min=3, max=5
M1_3_Saw_haptic_realism	M=3.35, SD=0.61, min=2, max=4
M2_1_Reamer_realism	M=3.56, SD=0.73, min=3, max=5
M2_2_Reamer_visual_realism	M=3.75, SD=0.93, min=2, max=5
M2_3_Reamer_haptic_realism	M=3.56, SD=0.73, min=2, max=5
M3_1_Pan_realism	M=3.67, SD=1.07, min=2, max=5
M3_2_Pan_visual_realism	M=4.00, SD=1.00, min=1, max=5
M3_3_Pan_haptic_realism	M=3.58, SD=1.08, min=2, max=5
M4_1_Scraping_realism	M=3.67, SD=1.15, min=1, max=5
M4_2_Scraping_visual_realism	M=3.87, SD=0.74, min=3, max=5
M4_3_Scraping_haptic_realism	M=3.67, SD=1.30, min=1, max=5
M5_1_Stem_realism	M=3.58, SD=1.16, min=1, max=5
M5_2_Stem_visual_realism	M=4.27, SD=0.70, min=3, max=5
M5_3_Stem_haptic_realism	M=3.50, SD=1.38, min=1, max=5

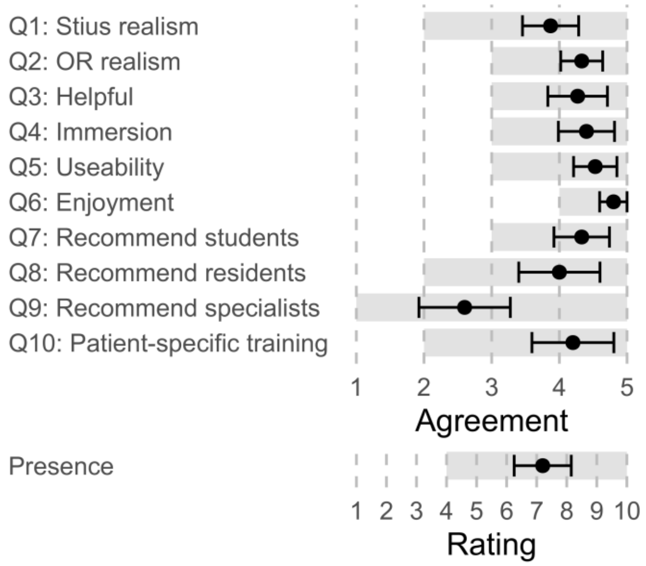


Fig. 9. Results of user acceptance, intention of use, presence (black dots – mean value, grey bar – range of the given answers, error bars – standard deviation).

In modules 1, 3, 4, and 5, the missing auditory feedback (category 3) was mentioned by 9 participants, among them 6 specialists, who stated that auditory feedback is an important reference for them to track the progress of the surgical steps.

In the category 4 (hardware setup), 7 participants reported that the positioning of the haptic devices was not ideal for applying the necessary forces. Unfortunately, the mounts for the hardware devices are not adjustable in height, hence there was no way to allow for an optimal positioning for all participants. Further, 6 participants commented that the handpiece is heavier than a real saw, even though others confirmed that the weight felt realistic. In modules 3 to 5 all participants commented on the ghost instruments alignment which jittered greatly at each hammer hit. This resulted in 5 participants not hitting as strongly as usual, as they had doubts if the virtual and real world are really in sync.

Regarding the category 5 (teaching/user interaction concept), 6 participants stated that during acetabulum reaming (module 2), only one angle was shown as correct; however, during real surgeries the reaming is performed in more than one direction.

Overall, there were many positive comments, for instance, "good visualization", "highly realistic", "I have more confidence now for my first real operation".

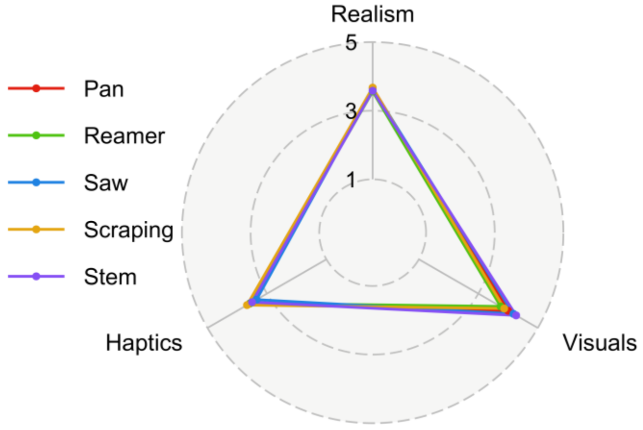


Fig. 10. Results of module specific realism.

5 DISCUSSION

5.1 Context in State of the Art

It is worth noting that our study stands out as the pioneering endeavor to comprehensively describe the haptic simulation utilized in a THA simulator, while also presenting a detailed analysis of potential enhancements in the quality of haptic simulation for THA with-in the realm of a VR-based training system. This comprehensive analysis spans from the initial step of femoral head resection to the final stage of stem implantation. Regrettably, existing THA VR training simulators such as those by Hooper et al. [34] , Logishetty et al. [35, 41, 54], and Younis et al. [40] lack detailed information regarding the capabilities of their simulators, thereby allowing only a superficial comparison with our work. Hooper et al. [34] display the patient's anatomy as a skeleton-only 3D-model and alternatively, with surrounding soft issues. Further, they also provide a realistic operation room model.

The THA VR simulator utilized by Logishetty et al. [41], offers training and assessment modes that are comparable to our simulation setup. This simulator also provides guidance on surgical instrument handling and implant positioning, while also tracking surgical errors. Logishetty et al. [41] focused primarily on evaluating the training outcomes for residents, rather than assessing the simulator itself, hence hindering a direct comparison with our study findings. In a follow-up study Logishetty et al. [35] the learning transferability for sawing off the femoral head was investigated, which corresponds to our module 1. They use a similar visual guide like us, but additionally show parameter of the cutting plane on a billboard directly above the situs. They measure four parameters defining the femoral cutting plane to assess the quality of the cut. We do not measure these parameters, as our consulting surgeons have instead defined a corridor that is considered good for our virtual patient. We decided against more detailed measures similar to Logishetty et al. [35] as the optimal cutting corridor depends on the patients anatomy, so our consulting surgeons opinion was that it is more important to identify this corridor based on the anatomy rather than focusing on training to achieve specific parameters. Additionally, Wiese et al. [39] have reported utilizing a concept similar to that of Logishetty et al. [41] and our research, focusing on the surgical approach for THA – an aspect that is not covered by our HIPS simulation. It is noteworthy that none of the aforementioned works have integrated a haptic simulation component.

The authors acknowledge the presence of OssoVR, inc. in San Francisco, USA, and FVRVS Ltd. in London, UK, who offer THA VR training similar to HIPS. Regrettably, information is only available in informal sources such as company videos and websites, making a comprehensive scientific comparison unfeasible. Thus, there is a pressing need to thoroughly document and evaluate these commercial simulators in peer-reviewed publications to allow for an unbiased evaluation and assessment.

5.2 HIPS current degree of realism

Given the divers background in knowledge and experience of the participants, we will interpret the relevance of the participants statement not solely by their frequency. Due to the master-apprentice teaching style in surgery the frequency of statements by participants at another hospital might be drastically different. We want to avoid to prematurely neglect statements as less relevant only because few participants mentioned them.

The answer to **RQ1 "Does the HIPS simulator provide realistic visual and haptic feedback?"**, is that the visual realism is high for all five modules. The haptic feedback was grossly realistic with slight differences between the modules. The results show that the situs and operating room are realistic, even though the situs was more exposed than in reality, explaining the slightly lower ratings. This is very likely due to the fact that during real surgery the soft tissue is moved and stretched to a large extend. As HIPS does not provide a soft tissue simulation allowing to morph the 3D geometry of the virtual situs, this was not simulated.

Regarding the questions directed at the haptic, visual and overall realism in all five modules, surprisingly very constant ratings were measured. No module stood out in any direction. Also, the separate realism ratings show a high consistency with visual realism always coming out top, which is not surprising as the visual simulation capabilities are superior to the haptic ones. Interestingly, the overall realism always scored in-between the visual and haptic realism or on the same level as the haptic realism. Our interpretation is, that the visual and haptic cues are roughly equally important for the participants, even though haptics might have a slightly higher weight. However, a further study is needed to explore deeper on the importance of cues for the perceived realism and the training effect.

These results also show that the haptic realism overcame an important threshold for the participants, accepting it as coarse, but still close enough to form a meaningful starting point for training. This is supported by the feedback the surgeons gave, as all improvements that were mentioned concerned details of the simulation quality. Not one participant stated the haptic feedback was totally unrealistic. Exemplarily for this are the comments regarding the sawing, where every participant felt to saw bone, even though it was perceived by some as too hard and many couldn't feel the difference between the cortical bone and spongiosa. Also, the missing increase in resistance and rattling during acetabulum reaming in module 2 or the too high forces for stem implantation in module 5 underline this point. These issues can be addressed by refining the haptic rendering. The most interesting comment regarding the haptic realism is that only 2 participants said the contact of bone and saw felt too soft. This steel/bone contact is a very hard contact that slightly exceeds the hardware capabilities of the used Virtuose 6D [55]. The fact that only 2 participants mentioned this, is highly likely because the femur and acetabula are embedded in soft tissue and move when coming in contact with the instruments. Therefore, it is perceived softer than a rigidly fixated bone and deemed as realistic by the vast majority.

Regarding the visual feedback, the most important comments concerned the collection of the cut material inside the reamer head, which is an important indicator for the surgeons. Given that the material model currently calculates exactly the chipped away volume, the most important step for visualizing is already implemented. Interestingly, there was a high number of participants who did not perceive the rasp moving inside the femur in module 4, despite it actually does. A likely explanation is that a combination of the HMD resolution, the loss of alignment that occurred at every hammer hit in combination with the concentration on the hammering, led to the participants not recognizing the rasp movement of 8 mm. An HMD with a higher resolution and more robust tracking for the hammer, hand and hammering device may solve this issue. During the concept phase of HIPS we decided against the simulation of blood splatters in consensus with the consulting specialists as it was deemed not a priority. The results from the realism questionnaires and the comments are in general supporting this decision. However, the comment of the unexperienced resident who missed blood splatter in order to be better prepared showed how simulation aspects can have different priority depending on one's experience.

The most relevant finding from our study is the importance of auditory feedback for all surgical steps except acetabula reaming, as it can be an important indicator for the surgeons to track progress. Together with the comments on the visual and haptic feedback, two important findings become clear (1) all three cues are used by the surgeons during operation and need to be simulated (2) there are individual differences on how heavy the surgeons lean on these three cues for tracking the progress of the surgical steps. Simulating realistic audio feedback is a very challenging feat, as it requires material property dependent computation of the auditory feedback. However, the currently implemented material model for haptic rendering provides a good starting point. By extending the haptic properties with sound properties, a parallel simulation pipeline for auditory feedback can be implemented, based on audio samples from real surgeries.

RQ2: "How is the user experience and usefulness of HIPS?", can be answered, that a highly engaging user experience was achieved as well as high approval rates for the targeted user groups of medical students and residents. To no surprise, the current implementation was deemed as currently less usable for specialist training. This is backed up by the numerous comments of the participants pointing out many details for improving the realism, whilst considering it as sufficiently good to gather first meaningful hands-on experiences. In this context, we can also report a surprisingly high demand for patient specific training capabilities. For the hardware setup the most pressing need for improvement concerns height adjustable mounts for the haptic devices, allowing to operate in comfortable positions. Due to the high forces exerted, the mounts for the haptic devices needs to be sturdy and specially developed, which we decided to put on a lower priority in order to concentrate on the haptic feedback. Another interesting finding are the comments on the seemingly too high weight of the saw. There are two kinds of saws available in reality, electrical and pneumatical. Both types of saws differ significantly in their weight. Depending on what type the surgeons are used to, they perceived the weight of the handpiece used in HIPS as either too heavy or realistic.

As for the teaching/user interaction concept the alignment of the visual models of hammer, hand and instruments with their haptic counterparts yielded the most pressing need for improvement. The hammer hits cause the VIVE trackers to vibrate leading to a temporarily loss of tracking for 2-3 seconds. As the time between hammer hits is much faster than the cool down time of the VIVE trackers, implementing compensation approaches is not feasible. Instead, a more robust tracking method must be implemented. An optical infrared tracking could be a solution. However, it would significantly increase costs and technical complexity of HIPS, and is sensitive to occlusion. RGB cameras and AI-based object recognition and tracking is another approach for improving the alignment, but also

IMU, accelerometers, depth cameras or VR HMD with integrated hand tracking are possibly solutions. A robust alignment will consequently help to improve trust in HIPS and reduce the participants' fear of hitting themselves. Another point for improving the acetabulum reaming in module 2, is to clearly point out, that in the current version only the reaming with the last reamer tool is simulated. This would make it clear that no other angle than the indicated must be reamed. In a next step, the capabilities of HIPS should be expanded to allow for reamer tool changes. Aside from this ambiguity the teaching/user interaction concept was perceived as very good, which emerged from the questionnaire results as well as from the participants comments. Lastly, the presence rating suggest that the participants felt as being inside an operation room.

Following this initial assessment of HIPS realism, the actually transferability of the skills learned in the simulation to real surgery has to be investigated. This means to determine if the current haptic simulation with all its current impairments is still enough to detect a learning effect or if further improvements in the simulation realism have to be achieved first.

The results of our study provide deep insights for researchers and practitioners alike that are valuable also for building training simulators for other surgical procedures. The introduced novel haptic hammering device is relevant for haptic simulations of other surgical procedures, e.g. knee and shoulder prostheses.

5.3 Study Limitations

The limitations and measures for improvements of the HIPS simulator itself have been discussed previously, so that here only limitations concerning the user study are discussed. The participants all were from the same hospital and therefore biased by the way THA is performed there. It is strongly recommended to verify our results with a larger number of participants from other hospitals and countries. Only 1 of 17 participants were female, which is however representative for the field of orthopedic surgeons where country dependent 3 % (England) to 11.2 % (Canada) are female [56].

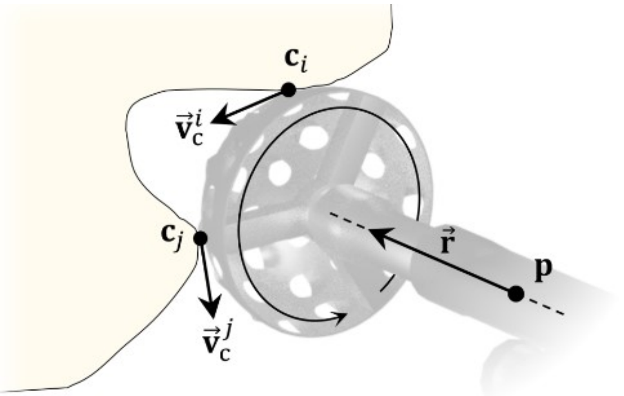
6 CONCLUSION

In this work, we present a novel VR training simulator, called HIPS, for learning the implantation of hip protheses (THA), including realistic haptic feedback. It is the first time that realistic feedback for the training of this surgery is presented, which required the novel development of a haptic device for the hammering surgical steps in conjunction with novel algorithms for haptic rendering and material removal simulations. HIPS enables the training of five surgical steps ranging from cutting off the femoral head to implanting the stem. In a user study with 17 surgeons from all experience levels, diverse and excessive feedback regarding the simulation realism, usability and further improvement were gathered. HIPS was very well received by the participants and they confirmed that it is feasible to gather first valuable practical experiences. Most importantly, many points for increasing the realism of the simulation quality even further were gathered and will be introduced in future developments. In a next step, HIPS should be investigated in an orthopedic resident curricular to determine skill transferability. Further, improving the haptic devices to allow the simulation of higher forces (Virtuose 6D) and more degrees of freedom (hammering device) along enhanced haptic rendering and material removal algorithms are important future work. Lastly, the simulation of the incision would represent a giant leap in simulation quality demanding novel algorithms for visual and haptic rendering.

ACKNOWLEDGEMENTS

The project on which this paper is based was funded by the German Federal Ministry of Education and Research under the funding code 16SV8356. The responsibility for the content of this publication lies with the authors.

SUPPLEMENT S1: TURBULENT DRILLING FRICTION



Supplement Figure 1: Possible contact situation for the hip reamer showing the point velocity augmentation, which leads to emergent behavior that represents a key haptic sensation to learn.

The mentioned partial impulses in section 3.3.1 include a frictional component, which allows us to render turbulent torques, as those described by Lorenz et al. [57]. The basic idea is to augment the point velocity computation by the theoretical angular velocity of spinning cutting blade's angular velocity. Given the cutting blade's rotation axis $\vec{\mathbf{r}}$, passing through a point $\mathbf{p}$ and a contact point $\mathbf{c}_\mathrm{I}$ with contact normal $\vec{\mathbf{n}}_\mathrm{I}$, we define the point velocity $\vec{\mathbf{v}}_\mathrm{p}^i(\mathrm{T})$ of tool T at contact i (see Supplement Fig. 1) during contact classification and tangential impulse in the following manner:

$$\vec{\mathbf{v}}_\mathrm{p}^i(\mathrm{T}) = \vec{\mathbf{v}}(\mathrm{T}) + \vec{\boldsymbol{\omega}}(\mathrm{T}) \times \left(\mathbf{c}_i - \mathbf{o}(\mathrm{T})\right) + \vec{\mathbf{v}}_\mathrm{c}^i$$

$$\vec{\mathbf{v}}_\mathrm{c}^i = \left(\mathbf{c}_i - \mathbf{p} - \hat{\mathbf{r}}((\mathbf{c}_i - \mathbf{p})^\mathrm{T}\hat{\mathbf{r}})\right) \times \hat{\boldsymbol{r}}\, \omega_\mathrm{d}\, \mathrm{clamp}\left(\frac{1}{V_d}, 0,1\right)$$

where $\mathbf{o}(\mathrm{T})$ is the tool's centre of mass, ω_d is the angular velocity of the cutting blade (for a hip reamer, we use 28.27 rad/s [58]) and V_d is the contact volume of the cutting blade with the bone (in m^3). We factor in the inverse of the global contact volume, as large friction in a localized area tends to exacerbate the turbulent torque. The additional friction is only displayed during material removal and leads to emergent behavior from the simulation that mimics the real phenomenon. It should be noted, that this adjusted point velocity is used to classify colliding and resting contacts, but not to evaluate the velocity constraints, as this allows us to still maintain high stability, despite the volitional turbulence.

REFERENCES

[1] F. LeBlanc *et al.,* "A comparison of human cadaver and augmented reality simulator models for straight laparoscopic colorectal skills acquisition training," *Journal of the American College of Surgeons*, vol. 211, no. 2, pp. 250–255, 2010, doi: 10.1016/j.jamcollsurg.2010.04.002.

[2] G. Huri, M. R. Gülşen, E. B. Karmış, and D. Karagüven, "Cadaver versus simulator based arthroscopic training in shoulder surgery," *Turkish journal of medical sciences*, vol. 51, no. 3, pp. 1179–1190, 2021, doi: 10.3906/sag-2011-71.

[3] J. Gilbody, A. W. Prasthofer, K. Ho, and M. L. Costa, "The use and effectiveness of cadaveric workshops in higher surgical training: a systematic review," *Annals of the Royal College of Surgeons of England*, vol. 93, no. 5, pp. 347–352, 2011, doi: 10.1308/147870811X582954.

[4] C. Melcher *et al.,* "Use of a High-Fidelity Training Simulator for Minimally Invasive Lumbar Decompression Increases Working Knowledge and Technical Skills Among Orthopedic and Neurosurgical Trainees," *Global spine journal*, vol. 13, no. 8, pp. 2182–2192, 2023, doi: 10.1177/21925682221076044.

[5] D. Escobar-Castillejos, J. Noguez, L. Neri, A. Magana, and B. Benes, "A Review of Simulators with Haptic Devices for Medical Training," *Journal of medical systems*, vol. 40, no. 4, p. 104, 2016, doi: 10.1007/s10916-016-0459-8.

[6] F. Aïm, G. Lonjon, D. Hannouche, and R. Nizard, "Effectiveness of Virtual Reality Training in Orthopaedic Surgery," *Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association*, vol. 32, no. 1, pp. 224–232, 2016, doi: 10.1016/j.arthro.2015.07.023.

[7] S. R. Dawe, J. A. Windsor, J. A. J. L. Broeders, P. C. Cregan, P. J. Hewett, and G. J. Maddern, "A systematic review of surgical skills transfer after simulation-based training: Laparoscopic cholecystectomy and endoscopy," *Annals of Surgery*, vol. 259, no. 2, pp. 236–248, 2014, doi: 10.1097/SLA.0000000000000245.

[8] K. E. Darras *et al.,* "Is there a superior simulator for human anatomy education? How virtual dissection can overcome the anatomic and pedagogic limitations of cadaveric dissection," *Medical teacher*, vol. 40, no. 7, pp. 752–753, 2018, doi: 10.1080/0142159X.2018.1451629.

[9] N. Hammer, S. Löffler, I. Bechmann, H. Steinke, C. Hädrich, and C. Feja, "Comparison of modified Thiel embalming and ethanol-glycerin fixation in an anatomy environment: Potentials and limitations of two complementary techniques," *Anatomical sciences education*, vol. 8, no. 1, pp. 74–85, 2015, doi: 10.1002/ase.1450.

[10] N. Hammer *et al.,* "Ethanol-glycerin fixation with thymol conservation: a potential alternative to formaldehyde and phenol embalming," *Anatomical sciences education*, vol. 5, no. 4, pp. 225–233, 2012, doi: 10.1002/ase.1270.

[11] N. Hammer, C. Schröder, and S. Schleifenbaum, "On the suitability of Thiel-fixed samples for biomechanical purposes: Critical considerations on the articles of Liao et al. "elastic properties of Thiel-embalmed human ankle tendon and ligament" and Verstraete et al. "impact of drying and thiel embalming," *Clin. Anat.*, vol. 29, no. 4, pp. 424–425, 2016, doi: 10.1002/ca.22679.

[12] N. Hammer *et al.,* "Ethanol and formaldehyde fixation irreversibly alter bones' organic matrix," *Journal of the mechanical behavior of biomedical materials*, vol. 29, pp. 252–258, 2014, doi: 10.1016/j.jmbbm.2013.09.008.

[13] R. Zdero, P. Brzozowski, and E. H. Schemitsch, "Biomechanical properties of artificial bones made by Sawbones: A review," *Medical engineering & physics*, vol. 118, p. 104017, 2023, doi: 10.1016/j.medengphy.2023.104017.

[14] D. Stefanidis, T. C. Yonce, J. M. Green, and A. P. Coker, "Cadavers versus pigs: which are better for procedural training of surgery residents outside the OR?," *Surgery*, vol. 154, no. 1, pp. 34–37, 2013, doi: 10.1016/j.surg.2013.05.001.

[15] A. Syamlan, Fathurachman, K. Denis, E. Vander Poorten, B. Pramujati, and T. Tjahjowidodo, "Haptic/virtual reality orthopedic surgical simulators: a literature review," *Virtual Reality*, 2022, doi: 10.1007/s10055-022-00666-y.

[16] L. Pelliccia *et al.,* "A cadaver-based biomechanical model of acetabulum reaming for surgical virtual reality training simulators," *Scientific reports*, vol. 10, no. 1, p. 14545, 2020, doi: 10.1038/s41598-020-71499-5.

[17] R. J. Ferguson, A. Palmer, JR, A. Taylor, M. L. Porter, H. Malchau, and S. Glyn-Jones, "Hip replacement," *Lancet (London, England)*, vol. 392, no. 10158, pp. 1662–1671, 2018, doi: 10.1016/S0140-6736(18)31777-X.

[18] M. Kaluschke *et al.,* "A Virtual Hip Replacement Surgery Simulator with Realistic Haptic Feedback," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Tuebingen/Reutlingen, Germany, 2018, pp. 759–760.

[19] S. Knopp, M. Lorenz, L. Pelliccia, and P. Klimant, "Using Industrial Robots as Haptic Devices for VR-Training," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Tuebingen/Reutlingen, Germany, 2018, pp. 607–608.

[20] D. Weik *et al.,* "Integrating Tactile Feedback in an Acetabular Reamer for Surgical VR-Training," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Osaka, Japan, 2019, pp. 1227–1228.

[21] G. Cate *et al.,* "Current status of virtual reality simulation education for orthopedic residents: the need for a change in focus," *Global surgical education : journal of the Association for Surgical Education*, vol. 2, no. 1, p. 46, 2023, doi: 10.1007/s44186-023-00120-w.

[22] R. Q. Mao *et al.,* "Immersive Virtual Reality for Surgical Training: A Systematic Review," *The Journal of surgical research*, vol. 268, pp. 40–58, 2021, doi: 10.1016/j.jss.2021.06.045.

[23] M. Pfandler, M. Lazarovici, P. Stefan, P. Wucherer, and M. Weigl, "Virtual reality-based simulators for spine surgery: a systematic review," *The spine journal : official journal of the North American Spine Society*, vol. 17, no. 9, pp. 1352–1363, 2017, doi: 10.1016/j.spinee.2017.05.016.

[24] A. J. Lungu, W. Swinkels, L. Claesen, P. Tu, J. Egger, and X. Chen, "A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery," *Expert review of medical devices*, vol. 18, no. 1, pp. 47–62, 2021, doi: 10.1080/17434440.2021.1860750.

[25] W. Zhang, X. Liu, and B. Zheng, "Virtual reality simulation in training endoscopic skills: A systematic review," *Laparoscopic, Endoscopic and Robotic Surgery*, vol. 4, no. 4, pp. 97–104, 2021, doi: 10.1016/j.lers.2021.09.002.

[26] M. Iskander, T. Ogunsola, R. Ramachandran, R. McGowan, and L. A. Al-Aswad, "Virtual Reality and Augmented Reality in Ophthalmology: A Contemporary Prospective," *Asia-Pacific journal of ophthalmology (Philadelphia, Pa.)*, vol. 10, no. 3, pp. 244–252, 2021, doi: 10.1097/APO.0000000000000409.

[27] M. Färber, F. Hummel, C. Gerloff, and H. Handels, "Virtual reality simulator for the training of lumbar punctures," *Methods of information in medicine*, vol. 48, no. 5, pp. 493–501, 2009, doi: 10.3414/ME0566.

[28] A. W. Kuhn, J. K. Yu, K. M. Gerull, R. M. Silverman, and A. W. Aleem, "Virtual Reality and Surgical Simulation Training for Orthopaedic Surgery Residents: A Qualitative Assessment of Trainee Perspectives," *JB & JS open access*, vol. 9, no. 1, 2024, doi: 10.2106/JBJS.OA.23.00142.

[29] A. Koch *et al.,* "Say, What Is on Your Mind? Surgeons' Evaluations of Realism and Usability of a Virtual Reality Vertebroplasty Simulator," *Surgical innovation*, vol. 26, no. 2, pp. 234–243, 2019, doi: 10.1177/1553350618822869.

[30] S. Su, R. Wang, R. Zhou, Z. Chen, and F. Zhou, "The effectiveness of virtual reality, augmented reality, and mixed reality training in total hip arthroplasty: a systematic review and meta-analysis," *Journal of orthopaedic surgery and research*, vol. 18, no. 1, p. 121, 2023, doi: 10.1186/s13018-023-03604-z.

[31] E. Clarke, "Virtual reality simulation-the future of orthopaedic training? A systematic review and narrative analysis," *Advances in simulation (London, England)*, vol. 6, no. 1, p. 2, 2021, doi: 10.1186/s41077-020-00153-x.

[32] J. T. Verhey, J. M. Haglin, E. M. Verhey, and D. E. Hartigan, "Virtual, augmented, and mixed reality applications in orthopedic surgery," *The international journal of medical robotics + computer assisted surgery : MRCAS*, vol. 16, no. 2, e2067, 2020, doi: 10.1002/rcs.2067.

[33] P. Sun *et al.,* "Application of Virtual and Augmented Reality Technology in Hip Surgery: Systematic Review," *Journal of medical Internet research*, vol. 25, e37599, 2023, doi: 10.2196/37599.

[34] 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, doi: 10.1016/j.arth.2019.04.002.

[35] K. Logishetty, W. T. Gofton, B. Rudran, P. E. Beaulé, and J. P. Cobb, "Fully Immersive Virtual Reality for Total Hip Arthroplasty: Objective Measurement of Skills and Transfer of Visuospatial Performance After a Competency-Based Simulation Curriculum," *The Journal of bone and joint surgery. American volume*, vol. 102, no. 6, e27, 2020, doi: 10.2106/JBJS.19.00629.

[36] M. Morgan, A. Aydin, A. Salih, S. Robati, and K. Ahmed, "Current Status of Simulation-based Training Tools in Orthopedic Surgery: A Systematic Review," *Journal of surgical education*, vol. 74, no. 4, pp. 698–716, 2017, doi: 10.1016/j.jsurg.2017.01.005.

[37] G. Turini *et al.,* "A Microsoft HoloLens Mixed Reality Surgical Simulator for Patient-Specific Hip Arthroplasty Training," in *Lecture Notes in Computer Science, Augmented Reality, Virtual Reality, and Computer Graphics*, L. T. de Paolis and P. Bourdot, Eds., Cham: Springer International Publishing, 2018, pp. 201–210.

[38] D. Panariello *et al.,* "Using the kuka lbr iiwa robot as haptic device for virtual reality training of hip replacement surgery," in *2019 Third IEEE International Conference on Robotic Computing (IRC)*, 2019, pp. 449–450.

[39] A. Wiese *et al.,* "Virtual Medical Instruments for Orthopedic Surgery Training: A Hip Arthroplasty Application," in *2021 IEEE Sensors Applications Symposium (SAS)*, Sundsvall, Sweden, 2021, pp. 1–6.

[40] S. B. Younis and E. H. Al-Hemiary, "Immersive Virtual Reality Application For Total Hip Replacement Surgical Training," in *2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, Paris / Créteil, France, 2021, pp. 1–5.

[41] K. Logishetty, B. Rudran, and J. P. Cobb, "Virtual reality training improves trainee performance in total hip arthroplasty: a randomized controlled trial," *The bone & joint journal*, 101-B, no. 12, pp. 1585–1592, 2019, doi: 10.1302/0301-620x.101b12.bjj-2019-0643.r1.

[42] W. A. McNeely, K. D. Puterbaugh, and J. J. Troy, "Six degree-of-freedom haptic rendering using voxel sampling," in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99*, Not Known, 1999, pp. 401–408.

[43] M. Agus, A. Giachetti, E. Gobbetti, G. Zanetti, and A. Zorcolo, "Real-Time Haptic and Visual Simulation of Bone Dissection," *Presence: Teleoperators and Virtual Environments*, vol. 12, no. 1, pp. 110–122, 2003, doi: 10.1162/105474603763835378.

[44] D. Morris, C. Sewell, F. Barbagli, K. Salisbury, N. H. Blevins, and S. Girod, "Visuohaptic simulation of bone surgery for training and evaluation," *IEEE Comput. Grap. Appl.*, vol. 26, no. 6, pp. 48–57, 2006, doi: 10.1109/MCG.2006.140.

[45] E. Acosta and A. Liu, "Real-time Volumetric Haptic and Visual Burrhole Simulation," in *2007 IEEE Virtual Reality Conference*, Charlotte, NC, USA, 2007, pp. 247–250.

[46] M.-D. Tsai and M.-S. Hsieh, "Accurate visual and haptic burring surgery simulation based on a volumetric model," *Journal of X-ray science and technology*, vol. 18, no. 1, pp. 69–85, 2010, doi: 10.3233/XST-2010-0242.

[47] S. Chan, F. Conti, N. H. Blevins, and K. Salisbury, "Constraint-based six degree-of-freedom haptic rendering of volume-embedded isosurfaces," in *2011 IEEE World Haptics Conference*, Istanbul, 2011, pp. 89–94.

[48] M. Kaluschke, R. Weller, N. Hammer, L. Pelliccia, M. Lorenz, and G. Zachmann, "Realistic Haptic Feedback for Material Removal in Medical Simulations," in *2020 IEEE Haptics Symposium (HAPTICS)*, Crystal City, VA, USA, 2020, pp. 920–926.

[49] M. Lorenz, S. Knopp, G. Matschewsky, and S. Murali, "Einschlagapparat zur haptischen Simulation von hämmernden Arbeitsschritten," DE 10 2022 128 868, Deutschland 10 2022 128 868.4, May 2, 2024.

[50] B. Mirtich and J. Canny, "Impulse-based simulation of rigid bodies," in *Proceedings of the 1995 symposium on Interactive 3D graphics - SI3D '95*, Monterey, California, United States, 1995, 181-ff.

[51] S. Bouchard, G. Robillard, J. St-Jacques, S. Dumoulin, M. J. Patry, and P. Renaud, "Reliability and validity of a single-item measure of presence in VR," in *Proceedings. Second International Conference on Creating, Connecting and Collaborating through Computing*, Ottawa, Ont., Canada, 2004, pp. 59–61.

[52] T.-Y. Fang, P.-C. Wang, C.-H. Liu, M.-C. Su, and S.-C. Yeh, "Evaluation of a haptics-based virtual reality temporal bone simulator for anatomy and surgery training," *Computer methods and programs in biomedicine*, vol. 113, no. 2, pp. 674–681, 2014, doi: 10.1016/j.cmpb.2013.11.005.

[53] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, vol. 13, no. 3, p. 319, 1989, doi: 10.2307/249008.

[54] K. Logishetty, B. Rudran, W. Gofton, P. Beaule, and J. P. Cobb, "Total hip arthroplasty training in virtual reality transfers to the real world: a randomized blinded study," in *Orthopaedic Proceedings*, 2019, p. 27.

[55] M. Lorenz *et al.,* "Perceived realism of haptic rendering methods for bimanual high force tasks: original and replication study," *Scientific reports*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-38201-x.

[56] P. M. Lutz, J. Lenz, A. Achtnich, and S. Geyer, "Ärztinnen in der Orthopädie und Unfallchirurgie in Deutschland: ein aktueller Status quo," (in ger), *Der Orthopäde*, vol. 50, no. 9, pp. 713–721, 2021, doi: 10.1007/s00132-020-04048-7.

[57] M. Lorenz *et al.,* "Wrist at risk? – Considerations derived from a novel experimental setup to assess torques during hip reaming with potential implications on the orthopedic surgeons' health," *Journal of the mechanical behavior of biomedical materials*, p. 104160, 2021, doi: 10.1016/j.jmbbm.2020.104160.

[58] W. F. Sherman *et al.,* "Variability of Cutting and Thermal Dynamics Between New and Used Acetabular Reamers During Total Hip Arthroplasty," *Arthroplasty today*, vol. 7, pp. 91–97, 2021, doi: 10.1016/j.artd.2020.12.002.