



# Facilitating Virtual Reality Integration in Medical Education: A Case Study of Acceptability and Learning Impact in Childbirth Delivery Training

Chang Liu  
e0008909@u.nus.edu  
National University of Singapore  
Singapore, Singapore

Abhiram Kanneganti  
abhiram\_kanneganti@nuhs.edu.sg  
National University Hospital  
Singapore, Singapore

Felicia Fang-Yi Tan  
felicia.tanfy@gmail.com  
National University of Singapore  
Singapore, Singapore

Gosavi Arundhati Tushar\*  
obggat@nus.edu.sg  
National University of Singapore  
Singapore, Singapore

Shengdong Zhao\*  
shengdong.zhao@cityu.edu.hk  
City University of Hong Kong  
Hong Kong, China

Eng Tat Khoo\*  
etkhoo@nus.edu.sg  
National University of Singapore  
Singapore, Singapore



Figure 1: Medical undergraduate students trained using virtual reality childbirth delivery simulator.

## ABSTRACT

Advancements in Virtual Reality (VR) technology have opened new frontiers in medical education, igniting interest among medical educators to incorporate it into mainstream curriculum, complementing traditional training modalities such as manikin training. Despite numerous VR simulators on the market, their uptake in medical education remains limited. This paper explores the acceptability and educational effectiveness of VR in the context of vaginal childbirth delivery training, with the simulator providing

a walkthrough for the second and third stages of labour, contrasting it with established manikin-based methods. We conducted a large-scale empirical study with 117 medical students, revealing a significant 24.9% improvement in knowledge scores when using VR as compared to manikin. However, VR received significantly lower self-reported feasibility scores in Confidence, Usability, Enjoyment, Feedback and Presence, indicating low acceptance. The study provides critical insights into the relationship between technological innovation and educational impact, guiding future integration of VR into medical training curricula.

\*Corresponding authors.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs International 4.0 License.

CHI '24, May 11–16, 2024, Honolulu, HI, USA  
© 2024 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-0330-0/24/05  
<https://doi.org/10.1145/3613904.3642100>

## CCS CONCEPTS

• Empirical studies in interaction design; • Interaction design; • Human-centered computing;

## KEYWORDS

Medical Education, Virtual Reality, User Experience Design, Obstetrics & Gynaecology

**ACM Reference Format:**

Chang Liu, Felicia Fang-Yi Tan, Shengdong Zhao, Abhiram Kanneganti, Gosavi Arundhati Tushar, and Eng Tat Khoo. 2024. Facilitating Virtual Reality Integration in Medical Education: A Case Study of Acceptability and Learning Impact in Childbirth Delivery Training. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3613904.3642100>

**1 INTRODUCTION**

The progress of Virtual Reality (VR) technology has led to the development and introduction of numerous VR simulators in various fields, including medical education. Educators are eager to explore the potential of VR technology to revolutionise the existing medical training modalities and become the new standard. However, the readiness of the current medical curriculum to integrate this new technology, as well as the willingness of the students to accept it, remains uncertain.

In close collaboration with the Obstetrics and Gynaecology (O&G) department of a medical school, we developed a VR simulator for normal vaginal delivery, driven by our motivation to address challenges in standardised delivery suite experiences, enhance accessibility to hands-on learning material, and provide opportunities for distance learning during pandemics. In the scenario, users, who are undergraduate medical students, are walked through the second and third stages of labour, utilising medical instruments and perform hand manoeuvres.

The overarching exploration in this study goes beyond mere effectiveness to probe more deeply into the utility and the usability of the VR simulator: "How does the incorporation of virtual reality in medical training, particularly in training for childbirth delivery, influence both learning outcomes and the acceptance of the technology among medical students?" This larger question gives rise to 3 sub-research questions:

- **RQ1.** "How does the VR simulator affect the learning outcomes in terms of knowledge and problem-solving skills in childbirth scenarios?"
- **RQ2.** "How acceptable is the VR simulator among medical students as an educational tool for childbirth training?"
- **RQ3.** "What is the relationship between the perceived acceptability and the effectiveness of the VR and manikin simulators in enhancing learning outcomes?"

We have formulated the following hypotheses:

- **H1.** Integrating a VR simulator in childbirth delivery training will yield significantly improved learning outcomes in medical students, as evidenced by higher scores in knowledge tests, compared to traditional Manikin-based training.
- **H2.** Medical students will rate the VR simulator as significantly better in terms of the 6 feasibility domains, including learning, confidence, feedback, usability, enjoyment and presence. This affirms that the technology aligns with their educational needs and learning preferences in the context of childbirth delivery training.
- **H3.** There will be a positive correlation between the levels of acceptability and the effectiveness of the simulator; the

higher the acceptability ratings, the greater the improvement in learning outcomes.

We conducted an empirical study with 117 medical students to explore the answer to our research questions and test our hypotheses. Both knowledge gains and feasibility scores were calculated and compared between the manikin training modality and the VR training modality. Additionally, selected interviews were conducted to provide deeper insights into the nuanced relationship between technological innovation and educational impact, guiding future integration of VR into medical training curricula.

The contributions of this paper to the HCI community are:

- (1) A large-scale empirical evaluation with medical students, comparing a VR simulator with the traditional manikin method for childbirth delivery training. While participants showed significantly improved knowledge scores with the VR training, subjective ratings indicated the opposite, i.e. a low acceptance of VR training amongst medical students.
- (2) A qualitative assessment through in-depth interviews with medical instructors and students, providing insights into this discrepancy. The factors affecting VR adoption, pedagogical limitations of and manikin-based training, as well as the broader implications for integration of VR systems into medical training curricula were explored.
- (3) Details on a medical educational VR simulator's design and development process, including the design requirements, prototyping, translation of teaching materials into virtual content, and a discussion of other key features included in the simulation.

**2 BACKGROUND AND RELATED WORK**

We provide an overview of some challenges encountered in the obstetric and gynaecology training. We then discuss the conventional manikin-based simulation methods and their limitations. Finally we present the more recent VR simulation approaches in medical education.

**2.1 Challenges in securing O&G training opportunities for students**

Obstetrics and gynaecology training has long encountered challenges, hindering the active participation of medical students in the delivery suite. We highlight three primary challenges: patient acceptability, gender bias, and physician reluctance towards student involvement.

Due to the sensitive nature of childbirth delivery, only three-fifths of pregnant women are willing to have medical students participate in their intrapartum care [18]. This is contrasted with attitudes in other departments such as general medicine and ear-nose-throat centres, which welcome medical student involvement in their patient care. Even when medical students do gain entry into the delivery room, their roles are often limited to basic, routine tasks such as monitoring vital signs, rather than engaging in complex delivery procedures [33]. This limited scope hampers their learning experience and curtails opportunities for hands-on training, motivates our search for alternative educational strategies.

Furthermore, gender biases cause means that male medical students to be frequently refused entry into the delivery suite [5]. As

patient safety takes precedence and the consequences of any errors during the procedure can be significant, it becomes paramount to ensure students in the room do not pose risk. A study by Mavis et al. revealed that a quarter of surveyed physicians prefer medical students to be mere observers during these high-stakes situations, severely limiting the practical training opportunities [24]. This reluctance, although justified by patient safety concerns, leads to an unfortunate consequence: a decrease in or uneven clinical exposure for medical students. As these experiences are pivotal for the development of competence and confidence, their absence represents a significant gap in medical education.

## 2.2 Conventional manikin-based simulation training in O&G

The declining clinical opportunities in obstetrics and gynaecology stimulate the need for running simulation sessions to equip medical students with essential skills [14]. These simulations serve as a complementary to the traditional apprenticeship model, where the students directly observe and shadow the expert.

Common types of simulations used in current vaginal delivery education include [2, 9]:

- Low fidelity manikin: manikin designed to represent a particular region of the human anatomy
- Standardised patient: human actors who are trained to portray an actual patient in a consistent way
- High fidelity manikin: programmable full-body manikin with capability to give verbal cues and generate realistic physiological parameters based on user inputs

The relationship between simulation realism, cost, and accessibility presents a significant trade-off in medical training. As the level of realism in simulation-based training increases, from low-fidelity to medium-fidelity and finally to high-fidelity simulations, there is a corresponding rise in both cost and limitations on accessibility.

Several studies have shown that simulation sessions are highly beneficial for students, and that both high- and low-fidelity manikins effectively enhance competency and confidence in students, surpassing the traditional lecture and slideshow style of teaching [8, 30]. In an innovative effort to bridge the gap between realism and accessibility, a hybrid simulation model has been introduced to combine standardised patients and part-task trainers. In this setup, an actress wears a pelvic model and assumes the role of a pregnant mother, interacting authentically with medical students. The advantage of this hybrid model is that it lowers the cost of training while integrating the human element of medical practice and the hands-on skill development [10]. Previous works have shown that students leave with improved communication skills compared to those who have attended the small-group tutorial [31].

## 2.3 Limitations in manikin-based simulation training in O&G

Despite the promise of cost savings, the operational demands of running simulation sessions go beyond the upfront costs of manikins [17].

To uphold the realism of training, simulation sessions are better conducted in the clinical environment [13]. This approach enables

students to directly engage with the authentic setup and become acquainted with the equipment. However, to prevent interfering with the actual patient, building a simulated clinical area becomes essential. Equipment used in these sessions demand thorough cleaning and proper storage to maximise its durability as well.

Health professionals, educators and SP trainers collaboratively prepare standardised patients and scenario cases that accurately represent clinical situations [23]. Actors involved need to be trained in accordance with the scripts provided to ensure the delivery of realistic responses and engagement in smooth communications. All of these demand a significant investment of time and effort.

Even if the standardised patient component is omitted, typical simulation sessions would still necessitate the participation and supervision of healthcare professionals, who are already occupied with their regular clinical work [1].

In obstetrics, manoeuvres, which are **psychomotor skills**, are especially in need to be demonstrated by the professionals. When the students practise on the manikins, guidance and concurrent verbal feedback from professionals are crucial to enhance acquisition and retention of skills [27].

However, not all students benefit the same from the simulation sessions, and this can be attributed to two main factors. Firstly, due to the above listed time-constraints, manpower shortage and equipment inaccessibility, not every student has the opportunity to actually perform the hands-on practice on the manikin and receive guidance from the professionals. Secondly, students exhibit varying levels of engagement with the constructed scenarios and invest differing degrees of effort in their participation [13].

## 2.4 VR simulations in medical training

Virtual reality simulation has found its place in various medical simulation training scenarios. For instance, Choi et al. conducted an assessment of manual dexterity among novices using embedded objective quantitative measures in a low-cost cataract surgery simulator [6]. Their findings suggested the potential of VR complementing the current wet-lab training in ophthalmology curriculum. Pulijala et al. conducted a randomised controlled trial comparing medical students who used a VR surgery simulator for maxillofacial anatomy with those using conventional slides and videos [29]. Their study assures positive gains in knowledge and self-confidence for the VR group. However, it's worth noting that "Enjoyment" of VR does not always correlate with improved knowledge. Xu et al. reveals that students who took the summative assessment in VR underperformed those who took it using the traditional online multiple-choice questions, despite reporting higher satisfaction and increased concentration with the VR format [34].

While web-based simulations on desktops or tablets can offer valuable case studies for learning procedures and practising decision-making [12], in contrast, VR simulations bring a different dimension to the learning experience. Their standout feature is the immersive environment they create, not just providing interactive scenarios but also familiarising users with the actual spatial orientation [26]. This allows for hands-on practice in a three-dimensional setting. The multi-sensory experience of VR, with its enhanced realism, goes beyond what web-based simulations can offer.

There are various approaches to integrate this new technology into current medical education. Pottle has outlined two case studies of universities incorporating VR simulation into their curricula [28]. The University of Northampton established a VR simulation suite where nursing students take turns participating in the VR scenarios, with their peers observing the VR view projected onto a large screen. This group setting maintains the benefits of peer learning. The University of Oxford adopts a different strategy, offering VR devices installed on trolleys as needed. Students using the devices could receive support from experienced VR users.

### 3 DESIGN PROCESS

#### 3.1 Design Requirements

The content and flow of the simulation are meticulously designed by the engineering team in collaboration with the medical team led by a senior gynaecologist with extensive experience in medical education.

**3.1.1 Team composition.** The medical team is led by the director for Undergraduate Education for the Obstetrics and Gynaecology clerkship, who also teaches medical students during their normal vaginal delivery manikin sessions. The engineering team is led by a faculty member from the College of Design and Engineering, and comprises a dedicated research engineer with Unity 3D development experience, researchers in charge of testing and user study and a group of engineering students from mechanical, electrical and biomedical background who are enthusiastic in creating educational solutions in VR. Additionally, a contract-based 3D artist, proficient in Blender and Autodesk Maya, is responsible for constructing and modifying models to meet the development needs.

**3.1.2 Learning content.** The learning objectives for the normal vaginal delivery encompass recognizing imminent delivery, preparing essential instruments and manpower, effective patient communication, ensuring proper aseptic technique, identifying the seven cardinal movements of birth, facilitating them, ensuring the safe delivery of the baby, implementing postpartum hemorrhage prophylaxis, and managing the safe delivery of the placenta. Both VR and manikin simulation sessions adhere to these learning objectives that have been formulated after a consensus taking process amongst educators in undergraduate medical education in O&G. Similar to manikin training where instructors elaborate on learning points during significant parts of their demonstration, in the VR simulator, the key learning points appear as short phrases or sentences in the form of pop-up windows at specific junctures of the procedure. For reinforcement, all learning points reappear at the conclusion of the simulation as a list.

**3.1.3 Design workflow.** We adopted a systematic and collaborative approach to design, with the overall workflow encompassing several key stages (detailed in Table 1). In the initial phase (scenario specification), an expert from the medical team formulated a detailed scenario of a 28-year-old primigravida in spontaneous labour. The engineering team then collected references from learning materials and on-site delivery ward observations, noting details of the environment and medical instruments. These elements were translated into a storyboard, visualising the delivery suite layout,

character positioning and sequence of events. This was refined and iterated on alongside the medical expert.

Subsequently, the engineering team worked closely with the 3D artist for 3D modelling and animation, where suitable models and animations were sourced and constructed. The hand manoeuvre guide involved recording an expert's actual hand movements using Unity Recorder. After expert validation, the developers began prototyping, pulling together all components of the storyboard. Finally, the validation phase involved biweekly testing and refinement by both medical and engineering teams; medical experts provided feedback on flow and accuracy of learning content which the engineering team refined on. During this phase, informal pilots were also conducted with medical students for feedback. Frame rate, instructional user interface (UI), and interactions were improved over multiple iterations.

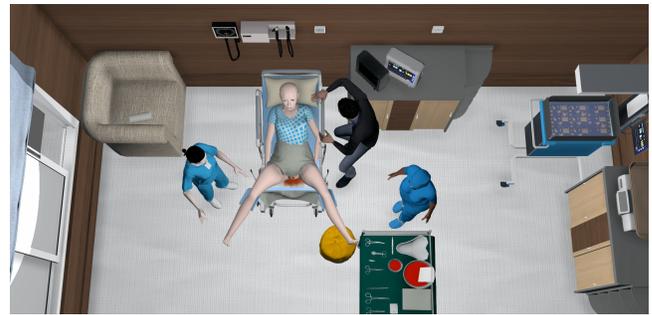


Figure 2: Layout of the virtual delivery suite.

#### 3.2 VR simulator prototyping

**3.2.1 Technology setup.** The simulation is developed using Unity Engine (Version 2020.3.13f1) and runs on Oculus Quest 2. Its hand-tracking function and interactions in the virtual environment are supported by Mixed Reality Toolkit (Version 2.7.2.0).

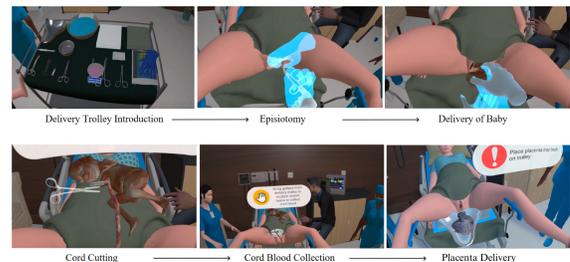


Figure 3: Six stages in the normal vaginal delivery simulation.

**3.2.2 Childbirth scenario description.** The simulation can be subdivided into six chronological stages that correspond to the different stages in a normal vaginal delivery (Figure 3).

- Stage 1: Scenario Introduction and Tool Familiarisation

Process Component	Role	Details
<b>Scenario specification - 2 Months</b>	Medical team (n=1)	<ul style="list-style-type: none"> <li>- Formulate the normal vaginal delivery scenario involving “a 28-year-old primigravida in spontaneous labour”</li> <li>- Develop learning points, multiple-choice questions, and explanations aligned with the Year 4 O&amp;G posting syllabus</li> <li>- Provide learning materials, including video explanations of episiotomy and hand manoeuvres</li> <li>- Arrange site-visit to delivery ward</li> <li>- Conduct regular manikin session to facilitate observation and learning of the normal vaginal delivery process for engineering team</li> <li>- Compose the script for the narration in simulation</li> </ul>
	Engineering team (n=3)	<ul style="list-style-type: none"> <li>- Capture photos of the actual delivery suite</li> <li>- Observe and document the regular manikin session</li> <li>- Record audio narration by the medical instructor, to be implemented in the simulation</li> </ul>
<b>Ideation/Storyboarding - 1 Month</b>	Engineering team (n=3)	<ul style="list-style-type: none"> <li>- Utilise Microsoft PowerPoint for storyboarding to illustrate: <ul style="list-style-type: none"> <li>- Layout of the delivery suite</li> <li>- Position of characters along with a delineation of the animations they will perform</li> <li>- Full sequence of the simulation, encompassing conditions prompting pop-up questions and animations</li> <li>- Elements for inclusion in the Instructional User Interface (UI)</li> </ul> </li> </ul>
<b>3D modelling and animation - 2 Months</b>	Engineering team (n=3)	<ul style="list-style-type: none"> <li>- Source for suitable medical-related and character models online</li> <li>- Capture model measurements and animation ranges based on medical instructor input</li> <li>- Communicate with the artist on the model and animation specifications</li> </ul>
	3D artist (n=1)	<ul style="list-style-type: none"> <li>- Construct the required models and animations utilising Blender and Autodesk Maya, referencing the provided photos and videos</li> </ul>
	Medical expert (n=1)	<ul style="list-style-type: none"> <li>- Validate the accuracy and realism of the constructed models and animations</li> </ul>
<b>Prototype development - 2 Months</b>	Engineering team (n=1)	<ul style="list-style-type: none"> <li>- Design the instructional UI to facilitate user navigation and interaction within the virtual environment</li> <li>- Construct the virtual delivery suite according to the layout of the actual room</li> <li>- Integrate different elements, including pop-up questions, audio narrations and animations together, and activating them in sequence according to the storyboard</li> <li>- Formulate the algorithm to assess the precision of the user’s hand manoeuvre</li> </ul>
<b>Validation - 2 Months</b>	Medical team (n=1)	<ul style="list-style-type: none"> <li>- Test out the preliminary prototype</li> <li>- Comment on the pace and flow of the simulation</li> <li>- Revise any incorrect presentation of learning content</li> </ul>
	Engineering team (n=3)	<ul style="list-style-type: none"> <li>- Fine tune the prototype based on preliminary tests</li> <li>- Test for prototype robustness</li> <li>- Improve usability issues</li> </ul>
	Selected medical students (n=21)	<ul style="list-style-type: none"> <li>- Test out the refined preliminary prototype</li> <li>- Provide feedback on usability issues</li> </ul>

**Table 1: Simulation design process.**

Entering the virtual delivery suite (Figure 2), the user first needs to enter the assigned identification number for progress tracking purposes. A basic question regarding the definition of “second stage labour” is asked. The scenario description is then displayed and read out by the instructor. After that, the user is guided to the

operating room trolley (delivery trolley) with various instruments that will be used later during the procedure. With the audio prompt telling the name and usage of each instrument, the user picks up each highlighted instrument respectively.

- Stage 2: Episiotomy

After draping the mother and cleaning the vaginal area, the user needs to pick up the syringe from the delivery trolley and attach it to the transparent “tool” zone to trigger the animation of injection of lignocaine for local anaesthesia. After that, the user needs to select the correct pair of episiotomy scissors to perform mediolateral episiotomy.

- Stage 3: Delivery of baby

This stage incorporates a detailed animation of how a pair of hands should support the baby on its way coming out. The user needs to overlap his/her hands with the virtual hands continuously and precisely to keep the animation playing. The baby will eventually be delivered and put onto the mother’s abdomen.

- Stage 4: Umbilical Cord Cutting

The user clips the cord using clamps, brings the artery forceps and cutting scissors to complete the cord cutting.

- Stage 5: Cord Blood Collection

The user picks up the gallipot to collect the cord blood and sends it for further evaluation.

- Stage 6: Placenta Delivery

The user needs to follow the virtual hands to pull the umbilical cord and take out the placenta. After cleaning the vulva and conducting sign-out, a list summarising all the key learning points will then appear. The user needs to click through the list one item after another to review the entire procedure. This completes the learning flow for this simulation.



**Figure 4: The user is aligning the hands with virtual hands.**

**3.2.3 Virtual hand guide.** To facilitate kinesthetic learning, a pair of animated virtual hands were implemented to depict the standard hand manoeuvres performed by doctors during the normal vaginal delivery process. This aims to guide medical students in accurately replicating these manoeuvres. The user is required to continuously align their own hands with the virtual hands, which provide support to the virtual baby, and follow the movement to guide the baby out (Figure 4).

Once the alignment of the user’s hands and the virtual hands fall below a certain predefined threshold, the animation of the baby emerging out from the vagina is paused until the desired alignment is achieved again. The alignment is examined at four specific junctures within the animation, as defined by the medical team. It is calculated using Euclidean distance between arbitrary

points located on the user’s wrists and corresponding reference points located in the same positions on virtual wrists.



**Figure 5: Target medical instrument highlighted in red.**

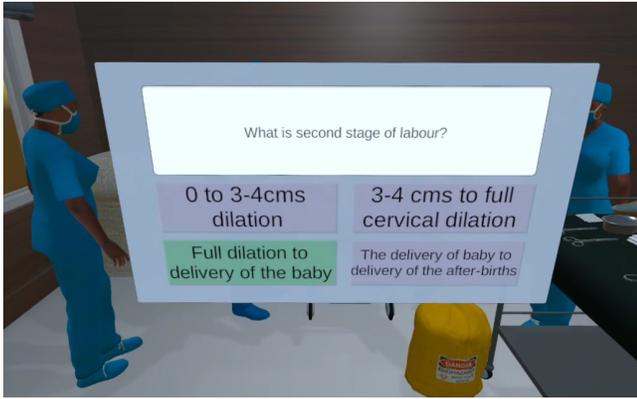
**3.2.4 Tools interaction and familiarisation.** During the childbirth delivery, multiple medical instruments are involved in various stages of the procedure. Replicating the real-world scenario, all essential instruments are grouped according to their functions and placed on a delivery trolley within the simulated delivery suite. Each instrument is introduced to the user through visual highlighting and audio description of their name and purpose (Figure 5).

For the interaction, the user needs to approach the instrument, utilise a pinch gesture by bringing their index and thumb together to pick it up, allowing for a closer observation. After that, the user needs to put the instrument back down on the designated translucent “tool zone” of the matching shape. The audio description remains paused until this action is performed and then proceeds to introduce the next instrument.

In the later stages of the simulation, UI pop-ups will appear, prompting the user to choose the appropriate instrument from the delivery trolley and place it in the translucent “tool zone” at the required location. Different from the previous instrument introduction, the required instrument is no longer highlighted. This is designed to examine the user’s comprehension subsequent to the initial introduction. The user is encouraged to inspect the shape of the “tool zone” and compare it with the instruments available on the delivery trolley. The simulation will only resume after the right instrument is attached, enhancing the learning of instruments involved.

**3.2.5 In-simulation questions.** Several multiple-choice questions crafted after the key learning concepts which are also being emphasised during the manikin session were strategically added at specific stages during the simulation. The user receives questions on the UI pop-up panels and benefits from the immediate explanation provided after each question.

The question format offers either two or four answer options, displayed as rectangular buttons. The user makes the selection by tapping the corresponding button. Upon making a correct selection, a positive auditory signal is played, and the answer button



**Figure 6: Colour indication upon user selection.**

turns green (Figure 6). Conversely, an incorrect response triggers a negative sound and the button turns red.

## 4 EVALUATION STUDY

For our evaluation study, we employed a two-part approach to assess the integration of VR in childbirth delivery training. In Part 1, a large-scale study involving 117 medical students was conducted, where participants engaged in both manikin and VR training sessions. This phase focused on quantitatively assessing improvements in knowledge scores and gathering feedback on user experience. In Part 2, we conducted in-depth interviews with a selected number of students and medical experts. This qualitative phase aimed to gain a more nuanced understanding of the factors affecting VR adoption, the pedagogical benefits and limitations of VR and manikin-based training, as well as the broader implications for medical training curricula.

All studies were approved by the Institutional Review Board (IRB). All medical students were given the opportunity to withdraw from this study and undergo legacy teaching with the manikin and were informed at the start that participation implied consent. Additionally, all medical students were informed on what data would be collected and how they would be used.

While we introduce the selected students and medical experts who participated in Part 2 (refer to 4.2), findings from the interview are integrated into the next section Discussion.

### 4.1 Large-scale Assessment of Knowledge Gains and User Experience

**4.1.1 Participants.** The study was undertaken with an original sample of 134 medical students. Given the dynamic nature of medical students' schedules and the constraints of fitting this study within their in-class sessions, only 117 students managed to complete both training sessions, knowledge quizzes and feedback questionnaires. This high attrition of 12.7% is not uncommon to studies conducted with medical students, providing an added layer to the study's real-world relevance [8, 10, 31].

The 117 participants (55 female, 62 male) were Year 4 medical students in the first week of their six-week clerkship with obstetrics and gynaecology, with an estimated mean age of 23.5 years.

The study employed a counterbalanced design to mitigate order effects: 61 participants started with VR training and then moved on to Manikin training, while 56 experienced the reverse sequence. All students were on the cusp of their practical rotations which included curriculum in obstetrics and gynaecology, ensuring their foundational knowledge was current and robust. They had not yet been exposed to actual childbirth delivery procedures in a professional setting, setting a baseline for the novelty of the training experience, as well as knowledge level on the topic. Additionally, the VR platform used, the Oculus Quest, was unfamiliar to all.

#### 4.1.2 Measures.

**4.1.2.1 Knowledge scores.** Pre- and post tests were administered via Google Forms before and after the first training session only. Both tests were identical, comprising 10 questions designed to evaluate participants' understanding of concepts spanning the 2nd and 3rd stages of labour in a normal vaginal childbirth delivery. Among the 10 questions, 8 took the form of multiple-choice questions presenting 4 or 5 options, while two were structured as multiple-select questions, allowing for multiple correct responses to be checked.

In addition, these questions were curated to align with the Year 4 syllabus of the medical school. For authenticity and validity, they were reviewed and approved by the director of undergraduate education for the OBGYN clerkship, who also serves as the instructor for the manikin session.

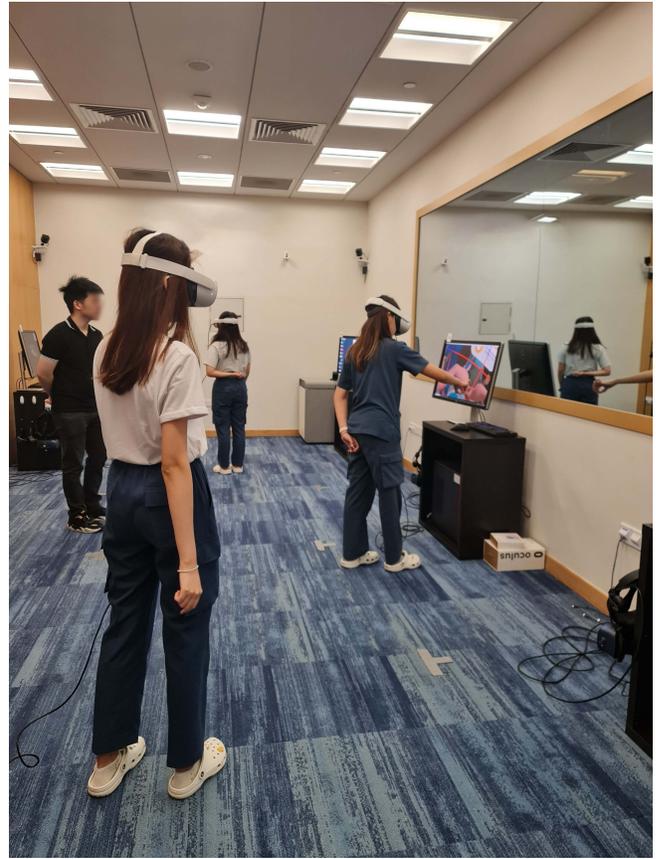
**4.1.2.2 Feasibility of training methods.** Post each training session (refer to Section 4.1.4 Design and Procedure), we administered a feasibility questionnaire via Google Forms to assess the pedagogical efficacy and user receptiveness of the two training modalities (VR and Manikin). This offered insights into the multifaceted experiences of participants, complementing the knowledge test scores and providing a comprehensive perspective on the potential advantages and limitations of each modality. There were a total of 14 questions, each classified into one of six domains [3, 15, 16, 21, 22, 25, 32]:

- **Learning** (3 Questions): Aimed to evaluate the user's understanding and internalization of the subject matter.
- **Confidence** (2 Questions): Addressed the self-belief and assurance participants felt in applying the freshly acquired knowledge.
- **Feedback** (3 Questions): Explored the clarity and direction the participants perceived during their simulation encounters.
- **Usability** (2 Questions): Assessed the user's ease of navigating and interacting with the simulation, indicating its design and operational efficiency.
- **Enjoyment** (2 Questions): Targeted at understanding the hedonic satisfaction and pleasure derived from the training sessions.
- **Presence** (2 Questions): Measured the immersive quality, realism, and sense of being "in the moment" during the sessions.

**4.1.2.3 Preferred Modality.** Post engagement with the VR and Manikin modalities, participants indicated their preferred training method: Manikin, VR, or Both. This metric captures the immediate preference and perceived efficacy of each training format.



(a) Instructor teaching in a manikin session.



(b) Students learning in a VR session.

**Figure 7: Two training modalities: manikin and VR.**

*4.1.2.4 Open Questions.* To delve deeper into the rationale behind preferences and identify potential areas for enhancement, participants responded to two open-ended questions at the end of their second feasibility questionnaire relating to:

- (1) Reason for preferred modality
- (2) Feedback on weakness or area of improvement for the training modality

#### 4.1.3 Apparatus and Training Setting.

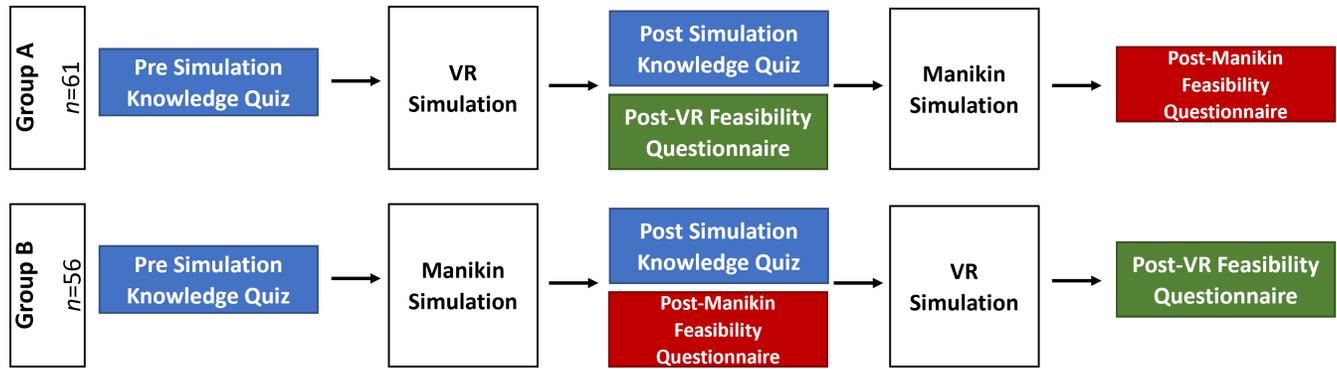
*4.1.3.1 Manikin Session.* The training was held in a simulated High Dependency Unit (HDU) room equipped with the manikin apparatus, the PROMPT Flex Birthing Simulator (Figure 7a). Each Manikin group comprised 12 to 13 students supervised by a single instructor. They received a briefing on relevant medical theories prior to the training session. While the session's structure is typical of conventional Manikin training, fewer than half of the students got the chance for hands-on engagement with the manikin. This hands-on opportunity arose in one of two ways: students volunteering when the instructor sought a demonstrator or through random selection by the instructor. Each Manikin session lasted approximately 60 minutes.

*4.1.3.2 VR Session.* The Oculus Quest 2 platform was used for the VR session (Figure 7b). To facilitate occasional real-time monitoring by the experimenters, the headset was connected to a PC using a cable. This enabled the research team to do frequent checks on participants' in-app activity and progress.

Two dedicated rooms were set up for the VR session, each accommodating three students at a time. As soon as a student concluded their session, another student was invited to occupy the newly vacant slot. Each student would spend around 15 to 20 minutes in VR.

*4.1.4 Design and Procedure.* The procedure of the user study is summarised in Figure 8.

- (1) **Briefing ( 5 minutes):** Each group of participants arrived at their designated training venue. Upon arrival and after giving their informed consent, participants underwent a structured briefing regarding the flow of the study. Specific to the VR training session, participants were given a visual and verbal demonstration by the experimenters, teaching them from how to wear the headset to how to use the VR interface. Key interaction such as pinching their fingers to



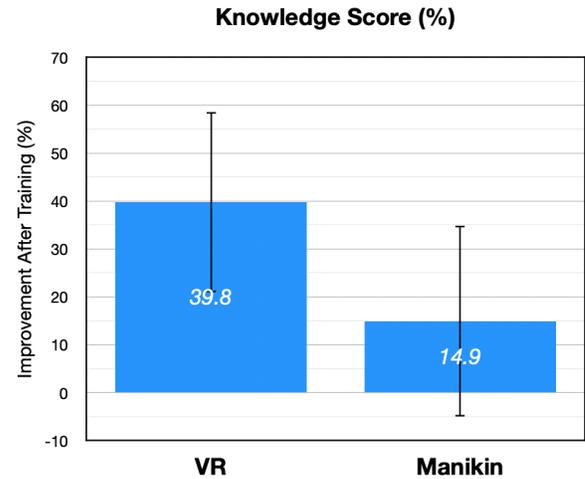
**Figure 8: Crossover user study flow.** All participants began with a pre-simulation knowledge quiz, followed by their first simulation session using either the manikin or VR training. After completing a post-simulation quiz and feasibility questionnaire, they switched to the second simulation session with the alternate training method. The study concluded with a final questionnaire and feedback session. Note that knowledge tests were administered before and after the first session, and not after the second session. Feasibility questionnaires, on the other hand, were administered after both sessions.

pick up medical instruments within the virtual environment was strongly emphasised.

- (2) **Pre-simulation Knowledge Quiz ( 3 minutes):** Following the briefing, participants filled in the pre-simulation knowledge quiz on their mobile device, which aimed to assess their baseline understanding before engaging with the training sessions.
- (3) **First Simulation Session ( 55 minutes):** For the manikin group, participants have a didactic lecture for about 15 minutes. Then, they spend 5 minutes observing the instructor’s demonstration on the hand manoeuvres and 5 minutes to orientate and familiarise themselves with the manikin. After that, the few selected or volunteered students are guided by the instructor through the process on the manikin for 25 minutes. For the VR group, participants have about 3 minutes to familiarise themselves with the headset and VR interface, and then have a 20-minute session in VR.
- (4) **Post-simulation Knowledge Quiz & Feasibility Questionnaire ( 5 minutes):** Upon completion of their first simulation, participants completed the post-simulation knowledge quiz followed by a feasibility questionnaire on their mobile device.
- (5) **Second Simulation Session ( 55 minutes):** Participants were directed to their next assigned training venue, this time experiencing the remaining modality.
- (6) **Feasibility Questionnaire & Overall Feedback ( 5 minutes):** After completing the second simulation, they filled out the feasibility questionnaire on their phone for the second modality and overall preference, including open-ended questions.

#### 4.1.5 Results.

**4.1.5.1 Knowledge scores.** We conducted a between-subjects analysis, the independent samples t-test, to compare the score improvements in VR training and Manikin training modalities, in percentage.



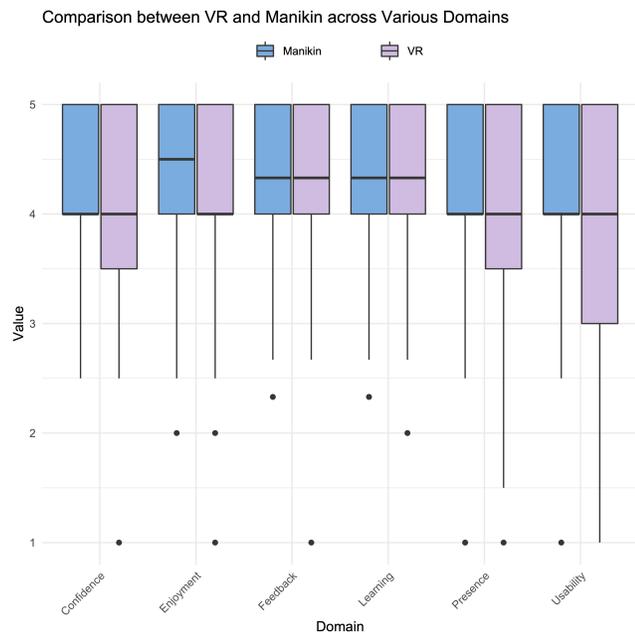
**Figure 9: Percentage improvement after training.**

Both Sphericity (Levene’s test,  $p > 0.05$ ) and Normality (Shapiro-Wilk test,  $p > 0.05$ ) assumptions were met.

There was a significant difference in score improvements between VR training ( $M = 39.8\%$ ,  $SD = 18.8\%$ ) and Manikin training ( $M = 14.9\%$ ,  $SD = 19.9\%$ );  $t(115) = 6.96$ ,  $p < 0.001$  (Figure 9).

**4.1.5.2 Feasibility Questionnaire.** We used a within-subjects analysis to compare feasibility ratings. The Shapiro-Wilk test indicated that the data for all six domains were not normally distributed ( $p < .001$ ). Thus, an Exact Wilcoxon-Pratt Signed-Rank Test was used to compare the median rankings of VR and Manikin training sessions for each of the 6 feasibility domains (Figure 10):

- **Learning:** The median of VR was 4.33 (IQR = 1) and of Manikin was 4.33 (IQR = 1). This difference was not statistically significant according to a Wilcoxon signed-rank test ( $Z = -0.067$ ,  $p > 0.05$ ).



**Figure 10: Median rankings of VR and Manikin training across 6 feasibility domains.**

- Confidence:** The median of VR was 4.0 (IQR = 1.5) and of Manikin was 4.0 (IQR = 1). This difference was statistically significant according to a Wilcoxon signed-rank test ( $Z = -2.93, p < 0.01$ ).
- Feedback:** The median of VR was 4.33 (IQR = 1) and of Manikin was 4.33 (IQR = 1). This difference was statistically significant according to a Wilcoxon signed-rank test ( $Z = -2.22, p < 0.05$ ).
- Usability:** The median of VR was 4.0 (IQR = 2) and of Manikin was 4.0 (IQR = 1). This difference was statistically significant according to a Wilcoxon signed-rank test ( $Z = -4.31, p < 0.001$ ).
- Enjoyment:** The median of VR was 4.5 (IQR = 1) and of Manikin was 4.5 (IQR = 1). This difference was statistically significant according to a Wilcoxon signed-rank test ( $Z = -3.35, p < 0.001$ ).
- Presence:** The median of VR was 4.0 (IQR = 1.5) and of Manikin was 4.0 (IQR = 1). This difference was statistically significant according to a Wilcoxon signed-rank test ( $Z = -3.43, p < 0.001$ ).

**4.1.5.3 Preferred Mode of Instruction.** After exposure to both Manikin and VR simulations, 49 participants (41.9%) preferred a combined approach, 62 participants (53.0%) favoured the Manikin alone, and 6 participants (5.1%) chose the VR Simulation. The combined and Manikin approaches were dominant, while VR was the least favoured.

**4.1.5.4 Open-ended questions.** The top two reasons why participants favoured the manikin modality are “tactility” and the presence of an “engaging tutor”. In the manikin session, they appreciated the tactile feedback from the manikin and the ability to interact

directly with the instructor, who provided immediate clarification for any doubts. This hands-on approach made them feel like they were actively participating in the delivery procedure.

On the other hand, the VR simulator excelled in “visualisation” and providing a smooth and complete “flow” of information. Participants were played in the role of a physician assisting a primigravida and learned the entire procedure of the normal vaginal delivery step by step, with important operations and hand manoeuvres being animated. Some participants even expressed a desire to have access to the recording of the VR gameplay for revision purposes.

Common suggestions for improving the manikin session included extending the duration and ensuring that everyone gets an opportunity to practise on the manikin. For the VR simulator, the most common improvement mentioned was enhancing usability such as the ease and intuitiveness of picking up and holding the virtual instruments in hand to ensure a smoother simulation experience. Other suggestions include improving the frame rate and graphics further.

## 4.2 In-depth Interviews for Qualitative Insights

In-depth interviews were conducted with medical students who participated in the previous evaluation, as well as medical experts. Consent to capture and utilise their quotes was obtained. The interviews were thematically analysed.

**4.2.1.1 Medical experts.** Three experts from the department of obstetrics and gynaecology were invited to be interviewed online. One of them is the director of the Year 4 O&G posting who collaborated with us. The other two are associate consultants who are keen in exploring VR technology in the childbirth delivery education. The interview durations were 55 minutes, 15 minutes, and 10 minutes, respectively. We interviewed them to understand:

- (1) Student behaviour during the manikin session
- (2) Importance of simulation
- (3) Opinion on traditional manikin training and VR training
- (4) Potential use of VR simulations in O&G training

**4.2.1.2 Selected medical students.** We conducted in-depth interviews with 5 randomly selected participants, 2 of whom were in Year 4 and 3 in Year 5 of the OG settle-in programme. Interviews were conducted either online or face-to-face. The interview durations were about 10 to 15 minutes each. All interviewees have attended four VR anatomy sessions in year 1 and one VR surgical safety session in year 3. Each of their VR sessions lasted less than half an hour, including briefing on theories and debrief to facilitate content reflection. All sessions used HTC Vive as the head-mounted display and its controllers as the input method. We interviewed them to understand:

- (1) Prior experience with VR simulations in their education
- (2) Importance of simulation
- (3) Opinion on training manikin training and VR training
- (4) Potential use of VR simulations in the O&G training

## 5 RESULTS AND DISCUSSION

Our hypotheses, H1 and H2, proposed that medical students would both demonstrate improved learning outcomes and show a high level of acceptance of our VR simulator for childbirth training.

While the data confirms H1, providing positive feedback in response to RQ1 ("How does the VR simulator affect the learning outcomes in terms of knowledge and problem-solving skills in childbirth scenarios?"), H2 was not empirically supported by the results. The significantly lower ratings for VR training compared to Manikin training in 5 out of 6 feasibility domains: Confidence, Usability, Enjoyment, Feedback and Presence answers RQ2 ("How acceptable is the VR simulator among medical students as an educational tool for childbirth training?"). This discrepancy between learning outcomes and user acceptance (RQ3 and H3) highlights a pivotal challenge in the integration of VR technology into traditional medical curriculum. Below, we explore relevant themes that emerged from our in-depth interviews with medical experts and students. Insights from our studies, reflections and suggestions provide valuable guidance for addressing this challenge and enhancing the effectiveness and acceptance of VR in medical education.

### 5.1 The Irreplaceable Role of Instructor Engagement

Despite the advancements in VR technology, thematic analysis of open-ended responses from our feasibility questionnaire (Section 4.1.5.2) reveal a preference for face-to-face interaction with instructors. For instance P8963 noted that a "VR system is not as interactive as a tutorial with an instructor," highlighting the essential role of human instructors in the learning process. This is potentially tied to the lower "Feedback" ratings in the VR simulation, and a critical reason why 94.9% of the participants were not willing to dispense with in-person sessions, i.e. participants preferred Manikin-only or both Manikin and VR modalities as part of their childbirth delivery training.

Expert 1 concurred to the vital role of instructor engagement during their interview, expressing that in-person experiences simply cannot be replaced, much less by "videos on YouTube". This medical expert emphasised the importance of the apprenticeship model, which offers a blend of observation, experience, and inspiration, that is effective not just for knowledge transfer, but for contextual experience and decision-making capabilities that are difficult to replicate in simulated environments. This may explain why H3 ("There will be a positive correlation between the levels of acceptability and the effectiveness of the VR simulator; the higher the acceptability ratings, the greater the improvement in learning outcomes.") was not supported by the data.

To address this limitation while still leveraging VR's advantages, we suggest future explorations of a hybrid training model combining modern technology's benefits with the wisdom and adaptability of live instruction. A next direction worth exploring includes remote VR training incorporating real-time instructor feedback. The use of cloud-based technology could alleviate the need for dedicated physical spaces, and a multi-player login feature could allow instructors to interact with multiple students in a simulated environment. This approach ensures the instructor remains an integral part of the training process.

### 5.2 Adaptability and Learning Curve in VR

While the VR simulation led to improved knowledge scores (thereby supporting H1) in our large-scale empirical study, a steep learning

curve might have played a role in the discrepant results between knowledge and experiential outcome. Specifically, given students' lack of experience with VR systems, they were less confident, affecting the system's perceived usability. Feedback provided by students substantiates this – P2107 for instance, acknowledged VR for its immersive qualities but also cautioned that it "takes some time to get used to [the experience]". Multiple students echoed P8884's report of feeling "dizzy" initially, a common issue when adapting to VR simulations. Past research indeed alludes to the importance of designing VR experiences to help alleviate these early-stage challenges [19], suggesting that future implementations of the VR simulator can have in-built design adaptations to improve the usability and acceptability of the VR simulator as part of the core training modality [7].

In addition, two out of three medical expert interviewees proactively recommended longitudinal studies to monitor how users' comfort and proficiency evolve with sustained exposure to VR environments. Such studies could serve as a feedback mechanism, leading to design iterations that better accommodate students' learning. By understanding these factors more comprehensively, we can address the initial barriers that impede VR's effectiveness.

### 5.3 Integration of Different Realism Aspects

Based on student feedback in the feasibility questionnaire's open-ended responses and interviews, there is a general consensus that VR simulations provide a degree of visual realism that benefits learning. This is evidenced by comments such as "[VR makes it] more visualisable" (P8903). However, VR simulations fall short in aspects of tactile realism, a salient aspect in the training of procedures like surgery or childbirth, whereby the precision and subtlety of hand movements are crucial [20].

For instance, P2108 highlighted the lack of tactile engagement with medical equipment. Specifically, this participant referred to the pinching gesture used in the simulation as non-naturalistic: "Performing the steps and usage of instruments is not very realistic", failing to simulate the actual experience of using medical instruments. To this end, the VR simulator's lack of tactile realism is likely to have adversely affected the 'Presence' and 'Enjoyment' ratings in the feasibility questionnaire.

To address the gap between interactions in the digital and physical world, the potential of the Mixed Reality (MR) platform becomes evident. The MR platform may be suitable for incorporating tactile elements from the real world into the simulation, thereby offering an enriched sensory experience that better represents real-world tools and procedures. This has the potential to improve both learning outcome and user acceptance, allowing students to effectively learn about the tasks and tool usage that medical professionals routinely perform.

### 5.4 Chaos Simulation

In keeping with the need for integrated realism, all of our medical expert interviewees also highlighted the importance of integrating real-world scenarios into the VR design as a means of improving the level of realism. Expert 1 emphasised the importance of replicating the "natural chaos" found in labor wards during her interview; this is a key need for effective training in realistic high-stakes, high-stress

medical environments. This includes simulations of medical complexities, unpredictable interpersonal and environmental variables such as interruptions from colleagues, input from family members, and unforeseen emergencies. Students tend to lack exposure to these naturalistic elements in classroom teaching, and it increases the challenge of navigating real-life medical situations.

To meet this educational need, we foresee that future iterations of our VR simulations will leverage AI-driven scenarios and randomized elements. This approach aims to simulate the rapidly evolving, multi-variable environment that undergraduates need to become proficient in before entering their residency. Additionally, integrating biometric feedback systems could offer real-time monitoring of stress and performance, providing a nuanced understanding of each trainee's adaptability to simulated 'chaos'.

One of our medical expert interviewees also highlighted the importance of students' psychological well-being. High-fidelity simulations must stimulate sufficient levels of stress, preparing trainees for the kinds of high-pressure situations they might encounter in the future. However, the simulations should not be so stressful as to be traumatize or to jeopardise their mental health, a balance supported by previous research [11]. To ensure this balance, real-time feedback mechanisms can be implemented, allowing either students or instructors to adjust the simulation's complexity during training, based on student performance and stress levels. We recommend safe and iterative user testing as part of the implementation process.

## 5.5 Resource Constraints

The VR simulation was initially conceptualized as a potential solution to resource constraints, given the possibilities of scalability and standardisation. While the data does support the VR simulator's efficacy in improving learning outcomes, the challenges related to Usability suggest that resource constraints remain an issue, impacting the simulation's overall acceptance too.

Interviews with medical experts revealed a trend in the landscape of medical training; there has been an increasing emphasis on breadth, with "more postings added to curriculum" resulting in shorter periods of exposure to specialized areas such as obstetrics and gynaecology, from "12 weeks to 6 weeks". As a result, there is the challenge of standardisation, i.e. medical students' exposure to real-world situations differ widely due to the unpredictable nature of childbirth events, such as "how many patients come in [those 6 weeks]" or "how many patients actually give birth on the day they visit the ward".

These findings from the interviews highlight the need for more standardised, accessible, and scalable training solutions, and VR simulations can function as an invaluable supplement to traditional training. With the development of scenario-based learning exercises, VR simulation platforms can offer controlled environments where students can repeatedly practice, enhance competency on both knowledge and psychomotor skills, and be assessed without taxing already constrained resources like physical space, instructor time, and training equipment. Training using such simulations has the potential to reduce the theory-practice gap that Brown [4] suggests in his study with graduate nurses through the use of high-fidelity simulations. Whilst not a replacement for real patient

scenarios, technology can certainly supplement traditional education modalities, such as the Manikin.

The financial requirements for implementing our childbirth delivery training module can be categorised into development and deployment expenses. The development aspect encompasses funds allocated for recruiting a dedicated developer and a contract-based artist. Additionally, it includes contributions from engineering students who are integrating this project into their design and research curriculum. On the deployment front, the costs involve purchasing standard VR headsets, each estimated at around \$299, with no extra expenditure on software owing to in-house development. The human resources needed for facilitating the VR sessions are relatively modest, possibly just two engineering students, with each supervising one of the two designated training environments.

However, several administrative and logistical challenges still arise that need careful consideration, in order to effectively blend VR technology into current curriculum. Venue limitations, particularly concerning the capacity to accommodate large student cohorts that come through each year, pose a barrier to the technology's scalability. While VR equipment has become more portable and user-friendly, the necessity for individualised learning environments remains a point of contention. The medical experts emphasised the value of such personalised settings, arguing that they allow for greater immersion in the VR scenario. To integrate rather than replace opportunities for face-to-face peer learning, we recommend exploring into blended learning approach. In the blended model, VR could be used for individualised, high-fidelity simulations to allow deep immersive learning while traditional educational methods could provide opportunities for peer-to-peer interaction and team-based training.

Medical experts also alluded to the issue of differing paces at which students complete VR simulations during the experiment. The variability in pace can be attributed to both unfamiliarity with the technology and individual learning styles, such as how thoroughly students read through explanations or take time to contemplate questions. To address this in future training implementations, instructors could offer tiered scenarios or implement adaptive algorithms that modify the simulation's difficulty and guidance in real-time, thereby accommodating individual learning needs even in a group-based setting.

## 5.6 Limitations

There are inherent limitations in this user study that warrant attention.

Firstly, due to constraints related to time and manpower, the pre- and post-tests were designed to only examine the basic knowledge using multiple-choice questions. More advanced psychomotor skills, such as hand manoeuvres, were not evaluated after the sessions. We could not conclude the effectiveness of the design of teaching the hand manoeuvre in VR.

Secondly, our methodology did not include an assessment of long-term retention rates for learning through VR or manikins.

Thirdly, recruiting volunteers for interviews proved challenging due to the busy schedules of medical students, and the limited number of selected interviewees may not be fully representative for the entire cohort. A potential solution could involve interviewing

participants who already completed their simulation and are waiting in the briefing room. This approach can significantly increase in the number of interviewees to enhance the generalisability of interview results.

## 5.7 Future Work

As the semesters progress, more batches of students will be using our simulator in their O&G posting. We have several plans for our future work.

Firstly, we aim to improve the user experience, especially for novice users in VR. Our goal is to explore a more natural interaction so that users can quickly engage with the content without being distracted by complex gestures and mechanisms. More time can be spent on learning the content instead of system familiarisation.

Secondly, we are currently developing a new VR simulation focused on shoulder dystocia complication, which contains more complicated situations than the normal vaginal delivery. We plan to assess the cognitive load experienced by the users during various sudden events within this scenario and study how these events impact their learning.

Lastly, we are also investigating enhancing the transfer of psychomotor skills in VR, particularly through an episiotomy simulation. This involves recording and analysing detailed hand movement trajectories to improve skill acquisition.

These initiatives reflect our commitment to addressing the needs of simulation in O&G training and attempts to ensure a smooth integration of the technology.

## 6 CONCLUSION

In close collaboration with Obstetrics and Gynaecology (O&G) department of a medical school, we developed a VR simulator for normal vaginal delivery. This paper explores the acceptability and educational effectiveness of the VR simulator through a large-scale empirical study with 117 medical students and a follow up interview with 3 O&G medical professional and 5 medical students. While our results revealed a significant 24.9% improvement in knowledge scores when using VR as compared to manikin training, VR training received significantly lower self-reported feasibility scores. This discrepancy between learning outcomes and user acceptance highlights a pivotal challenge in the integration of VR technology into traditional medical curriculum, the themes of which we explored further in our in-depth interviews with medical experts and students. We propose the following design considerations to overcome the challenges in integrating VR in medical education:

- (1) blending the role of instructor in simulation
- (2) improve realism by replicating the visual fidelity, medical complexity and interpersonal interaction
- (3) incorporating tactile feedback from reality into the simulation using mixed reality approach

## REFERENCES

- [1] Robert D. Acton, Jeffrey G. Chipman, Michelle Lunden, and Connie C. Schmitz. 2015. Unanticipated Teaching Demands Rise with Simulation Training: Strategies for Managing Faculty Workload. *Journal of Surgical Education* 72, 3 (may 2015), 522–529. <https://doi.org/10.1016/j.jsurg.2014.10.013>
- [2] H S Barrows. 1993. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Academic Medicine* 68, 6 (jun 1993), 443–51. <https://doi.org/10.1097/00001888-199306000-00002>
- [3] Maik Bieleke, Katarzyna Gogol, Thomas Goetz, Lia Daniels, and Reinhard Pekrun. 2021. The AEQ-S: A short version of the Achievement Emotions Questionnaire. *Contemporary Educational Psychology* 65 (apr 2021), 101940. <https://doi.org/10.1016/j.cedpsych.2020.101940>
- [4] Joset E. Brown. 2019. Graduate Nurses' Perception of the Effect of Simulation on Reducing the Theory-Practice Gap. *SAGE Open Nursing* 5 (jan 2019), 237796081989696. <https://doi.org/10.1177/2377960819896963>
- [5] Judy C. Chang, Michele R. Odrobina, and Kathleen McIntyre-Seltman. 2010. The Effect of Student Gender on the Obstetrics and Gynecology Clerkship Experience. *Journal of Women's Health* 19, 1 (jan 2010), 87–92. <https://doi.org/10.1089/jwh.2009.1357>
- [6] Kup-Sze Choi, Sophia Soo, and Fu-Lai Chung. 2009. A virtual training simulator for learning cataract surgery with phacoemulsification. *Computers in Biology and Medicine* 39, 11 (nov 2009), 1020–1031. <https://doi.org/10.1016/j.combiomed.2009.08.003>
- [7] Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13 (sep 1989), 319–340. <https://doi.org/10.2307/249008>
- [8] Ashlesha K. Dayal, Nelli Fisher, Diane Magrane, Dena Goffman, Peter S. Bernstein, and Nadine T. Katz. 2009. Simulation Training Improves Medical Students' Learning Experiences When Performing Real Vaginal Deliveries. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare* 4, 3 (2009), 155–159. <https://doi.org/10.1097/sih.0b013e3181b3e4ab>
- [9] Sharon Decker, Susan Sportsman, Linda Puetz, and Lynda Billings. 2008. The Evolution of Simulation and Its Contribution to Competency. *The Journal of Continuing Education in Nursing* 39, 2 (feb 2008), 74–80. <https://doi.org/10.3928/00220124-20080201-06>
- [10] Christopher C. DeStephano, Betty Chou, Silka Patel, Rebecca Slattery, and Nancy Hueppchen. 2015. A randomized controlled trial of birth simulation for medical students. *American Journal of Obstetrics and Gynecology* 213, 1 (jul 2015), 91.e1–91.e7. <https://doi.org/10.1016/j.ajog.2015.03.024>
- [11] James E. Driskell and Joan H. Johnston. 1998. Stress exposure training. In *Making decisions under stress: Implications for individual and team training*. American Psychological Association, Washington, DC, 191–217. <https://doi.org/10.1037/10278-007>
- [12] Anna Eleftheriou, Aikaterini Rokou, Christos Argyriou, Nikolaos Papanas, and George S. Georgiadis. 2022. Web-based Medical Education During COVID-19 Lockdown: A Step Back or a Leap to the Future? *Int J Low Extrem Wounds* 3 (sep 2022), 272–274. <https://doi.org/10.1177/15347346211011848>
- [13] Christopher S. Ennen and Andrew J. Satin. 2010. Training and assessment in obstetrics: the role of simulation. *Best Practice & Research. Clinical Obstetrics & Gynaecology* 24, 6 (dec 2010), 747–758. <https://doi.org/10.1016/j.bpobgyn.2010.03.003>
- [14] Elise N. Everett, David A. Forstein, Susan Bliss, Samantha D. Buery-Joyner, LaTasha B. Craig, Scott C. Graziano, Brittany S. Hampton, Laura Hopkins, Margaret L. McKenzie, Helen Morgan, Archana Pradhan, and Sarah M. Page-Ramsey. 2019. To the Point: The expanding role of simulation in obstetrics and gynecology medical student education. *American Journal of Obstetrics and Gynecology* 220, 2 (feb 2019), 129–141. <https://doi.org/10.1016/j.ajog.2018.10.029>
- [15] Fong-Ling Fu, Rong-Chang Su, and Sheng-Chin Yu. 2009. EGameFlow: A scale to measure learners' enjoyment of e-learning games. *Computers & Education* 52, 1 (jan 2009), 101–112. <https://doi.org/10.1016/j.compedu.2008.07.004>
- [16] Maribeth Gandy, Richard Catrambone, Blair MacIntyre, Chris Alvarez, Elsa Eiriksdottir, Matthew Hilimire, Brian Davidson, and Anne Collins McLaughlin. 2010. Experiences with an AR evaluation test bed: Presence, performance, and physiological measurement. In *2010 IEEE International Symposium on Mixed and Augmented Reality*. IEEE, Seoul. <https://doi.org/10.1109/ismar.2010.5643560>
- [17] M L Good. 2003. Patient simulation for training basic and advanced clinical skills. *Medical Education* 37 (oct 2003), 14–21. <https://doi.org/10.1046/j.1365-2923.37.s1.6.x>
- [18] Devlka Grasby and Julie A Quinlivan. 2001. Attitudes of patients towards the involvement of medical students in their intrapartum obstetric care. *The Australian and New Zealand Journal of Obstetrics and Gynaecology* 41, 1 (feb 2001), 91–96. <https://doi.org/10.1111/j.1479-828x.2001.tb01302.x>
- [19] Ilanit Hochmitz and Nirit Yuviler-Gavish. 2011. Physical Fidelity Versus Cognitive Fidelity Training in Procedural Skills Acquisition. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53, 5 (jul 2011), 489–501. <https://doi.org/10.1177/0018720811412777>
- [20] Omar M. Young Katherine H. Bligard, Kim L. Lipsey. 2019. Simulation Training for Operative Vaginal Delivery Among Obstetrics and Gynecology Residents: A Systematic Review. *Obstet Gynecol* 134 Suppl 1 (oct 2019), 16S–21S. <https://doi.org/10.1097/AOG.00000000000003431>
- [21] Robin Kay. 2011. Evaluating learning, design, and engagement in web-based learning tools (WBLTs): The WBLT Evaluation Scale. *Computers in Human Behavior* 27, 5 (sep 2011), 1849–1856. <https://doi.org/10.1016/j.chb.2011.04.007>
- [22] James R. Lewis. 2018. Measuring Perceived Usability: The CSUQ, SUS, and UMUX. *International Journal of Human-Computer Interaction* 34, 12 (jan 2018), 1148–1156. <https://doi.org/10.1080/10447318.2017.1418805>

- [23] C Macedonia. 2003. Simulation laboratories for training in obstetrics and gynecology. *Obstetrics & Gynecology* 102, 2 (aug 2003), 388–392. [https://doi.org/10.1016/s0029-7844\(03\)00483-6](https://doi.org/10.1016/s0029-7844(03)00483-6)
- [24] Brian Mavis, Peter Vasilenko, Rae Schnuth, Joseph Marshall, and Madeline Colavito Jeffs. 2006. Medical Students' Involvement in Outpatient Clinical Encounters: A Survey of Patients and Their Obstetricians-Gynecologists. *Academic medicine : journal of the Association of American Medical Colleges* 81, 3 (apr 2006), 290–296. <https://doi.org/10.1097/00001888-200603000-00023>
- [25] Jiri Motejlek and Esat Alpaly. 2023. The retention of information in virtual reality based engineering simulations. *European Journal of Engineering Education* 48, 5 (jan 2023), 929–948. <https://doi.org/10.1080/03043797.2022.2160968>
- [26] S. Pastel, D. Bürger, C. H. Chen, K. Petri, and K. Witte. 2022. Comparison of spatial orientation skill between real and virtual environment. *Virtual Reality* 26 (mar 2022), 91–104. <https://doi.org/10.1007/s10055-021-00539-w>
- [27] Mark C. Porte, George Xeroulis, Richard K. Reznick, and Adam Dubrowski. 2007. Verbal feedback from an expert is more effective than self-accessed feedback about motion efficiency in learning new surgical skills. *The American Journal of Surgery* 193, 1 (jan 2007), 105–110. <https://doi.org/10.1016/j.amjsurg.2006.03.016>
- [28] Jack Pottle. 2019. Virtual reality and the transformation of medical education. *Future Healthcare Journal* 6, 3 (oct 2019), 181–185. <https://doi.org/10.7861/fhj.2019-0036>
- [29] Yeshwanth Pulijala, Minhua Ma, Matthew Pears, David Peebles, and Ashraf Ayoub. 2018. Effectiveness of Immersive Virtual Reality in Surgical Training - A Randomized Control Trial. *Journal of Oral and Maxillofacial Surgery* 76, 5 (may 2018), 1065–1072. <https://doi.org/10.1016/j.joms.2017.10.002>
- [30] Jeanelle N. Sabourin, Rhonda Van Thournout, Venu Jain, Nestor Demianczuk, and Cathy Flood. 2014. Confidence in Performing Normal Vaginal Delivery in the Obstetrics Clerkship: A Randomized Trial of Two Simulators. *Journal of Obstetrics and Gynaecology Canada* 36, 7 (jul 2014), 620–627. [https://doi.org/10.1016/s1701-2163\(15\)30542-9](https://doi.org/10.1016/s1701-2163(15)30542-9)
- [31] Dimitrios Siassakos, Timothy Draycott, Katherine O'Brien, Charlotte Kenyon, Christine Bartlett, and Robert Fox. 2010. Exploratory Randomized Controlled Trial of Hybrid Obstetric Simulation Training for Undergraduate Students. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare* 5, 4 (aug 2010), 193–198. <https://doi.org/10.1097/sih.0b013e3181d3ee0b>
- [32] Penelope Sweetser and Peta Wyeth. 2005. GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment* 3, 3 (jul 2005), 3–3. <https://doi.org/10.1145/1077246.1077253>
- [33] John L. Vaughn, Lanette R. Rickborn, and John A. Davis. 2015. Patients' Attitudes Toward Medical Student Participation Across Specialties: A Systematic Review. *Teaching and Learning in Medicine* 27, 3 (jul 2015), 245–253. <https://doi.org/10.1080/10401334.2015.1044750>
- [34] Xuanhui Xu, Xingyu Pan, David Kilroy, Arun Kumar, Eleni Mangina, and Abraham G. Campbell. 2022. Using HMD-based Hand Tracking Virtual Reality in Canine Anatomy Summative Assessment: a User Study. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 287–296. <https://doi.org/10.1109/ISMAR55827.2022.00044>
- (2) Simulation learning helps me to remember the information and procedures in the third stage of labour.
  - (3) The (voice) instructions in the simulation are clear.
  - (4) I am comfortable with the pace of the simulation.
  - (5) This training modality has helped me acquire knowledge about normal labour.
  - (6) This training modality has helped me understand the steps involved in performing a normal vaginal delivery.
  - (7) The equipment worked smoothly and was easy to use.
  - (8) The instructions were clear and the pace of the simulation was appropriate.
  - (9) I was able to complete the simulation independently.
  - (10) The training experience was realistic and facilitated my orientation of actual delivery suite room.
  - (11) I felt completely absorbed (forgot about time passing) when performing the simulation.
  - (12) This teaching method was enjoyable.
  - (13) I feel confident to go to a real delivery setting after this training session.
  - (14) The feedback provided during the simulation was immediate, consistent at each attempt and promoted my learning outcomes.

## ACKNOWLEDGMENTS

We like to extend our heartfelt appreciation to Associate Professor Mary Rauff for her contribution to the development of the initial prototype. Additionally, we express our gratitude to both Associate Professor Mahesh Choolani and Associate Professor Mary Rauff for their unwavering support and guidance in medical education throughout our journey. We would also like to thank the research engineer and students from the College of Design and Engineering for their invaluable contributions to our project. Specifically, we acknowledge Teo Tian Song for his contribution in software development, and Lee Chin Hao, Nicole Lim Minghui, Varshini Saravanan, Yan Tianai, Catherine Koh Sin Hoon, Sheryl Ha Wen Qi, and Sahana Srinivasan, for their collaborative efforts in facilitating the design of the simulator and user studies.

## A FEASIBILITY QUESTIONNAIRE

Rate from 1 to 5: 5 – Strongly Agree; 4 – Agree; 3 – Neutral; 2 – Disagree; 1 – Strongly Disagree

- (1) The simulation helps me understand the steps involved in the third stage of labour.