IJCCS (Indonesian Journal of Computing and Cybernetics Systems)
Vol. x, No. x, July xxxx, pp. 1–5
ISSN (print): 1978-1520, ISSN (online): 2460-7258
DOI: 10.22146/ijccs.xxxx

Development and Validation of a Virtual Reality Circumcision Training Simulator: Simulator Sickness, User Experience, and Clinical Performance in Bali, Indonesia

I Gede Partha Sindu*[1] , Ni Ketut Kertiasih[2] , I Gede Surya Dinata[3], Putu Sugiartawan[4]
[1,2]Department of Informatics, FTK, Universitas Pendidikan Ganesha, Singaraja, Indonesia
[3]Faculty of Medicine, Universitas Pendidikan Ganesha, Singaraja, Indonesia
[3]Postgraduate Program, Master of Informatics, INSTIKI, Denpasar, Indonesia
e-mail: ***[1]partha.sindu@undiksha.ac.id**, [2] ketut.kertiasih@undiksha.ac.id, [3] surya.dinata@undiksha.ac.id, [4]putu.sugiartawan@instiki.ac.id

Abstrak

Virtual Reality (VR) semakin banyak digunakan dalam pendidikan kedokteran, namun implementasinya di Indonesia masih terbatas. Penelitian ini mengembangkan dan memvalidasi simulator sirkumsisi berbasis VR untuk menilai gejala simulator sickness, pengalaman pengguna, dan kinerja klinis. Studi mixed-methods dengan desain repeated measures melibatkan 74 peserta (25 Novice, 24 Intermediate, 25 Expert). Peserta menjalani tiga mode simulasi (Autonomous, Guided, Haptic). Instrumen mencakup SSQ, FMS, VRNQ, UEQ-S, Checklist, dan OSATS. Analisis menggunakan ANOVA berulang, uji nonparametrik, serta korelasi Spearman. Gejala simulator sickness lebih tinggi pada Autonomous Mode. Skor UX meningkat seiring pengalaman, dengan korelasi positif terhadap performa dan negatif terhadap sickness. Expert menunjukkan skor kinerja tertinggi, dan peningkatan performa bertahan hingga satu bulan. Simulator VR sirkumsisi memiliki validitas konstruk dan dampak edukatif signifikan. Mode instruksional terbukti menurunkan sickness, sedangkan integrasi haptic meningkatkan orientasi spasial. Studi lanjutan perlu mengeksplorasi pengukuran fisiologis serta uji transfer keterampilan ke praktik klinis.

Kata kunci*— Virtual Reality, simulator sirkumsisi, pengalaman pengguna, simulator sickness, performa klinis*

Abstract

Virtual Reality (VR) is increasingly integrated into medical education, yet its application in Indonesia remains limited. This study developed and validated a VR-based circumcision simulator to evaluate simulator sickness, user experience, and clinical performance. A mixed-methods, repeated-measures design was conducted with 74 participants (25 Novices, 24 Intermediates, 25 Experts). Participants engaged in three simulation modes (Autonomous, Guided, Haptic). Instruments included SSQ, FMS, VRNQ, UEQ-S, Checklist, and OSATS. Analyses employed repeated-measures ANOVA, nonparametric tests, and Spearman correlations. Simulator sickness was highest in Autonomous Mode. User experience scores improved with expertise, showing positive correlations with performance and negative correlations with sickness. Experts consistently outperformed other groups, and skill improvements were retained for up to one month. The VR circumcision simulator demonstrated strong construct validity and educational impact. Instructional modes effectively reduced sickness, while haptic integration enhanced spatial orientation. Future studies should incorporate physiological measures and assess real-world skill transfer.

Keywords— *Virtual Reality, circumcision simulator, user experience, simulator sickness, clinical performance*

Received June 1, 2012; Revised June 25, 2012; Accepted July 10, 2012

1. INTRODUCTION

The development of digital technology over the past decade has driven significant transformation in medical education. Technology-based simulations are now seen as one of the main strategies for overcoming the limitations of hands-on practice with patients, which is often hampered by ethical, safety, and access considerations to real cases[1], [2]. Virtual Reality (VR), in particular, enables the creation of immersive environments that resemble actual clinical situations so that students can perform repeated exercises without posing clinical risks [3]. Recent studies show that VR contributes positively to the improvement of both technical and non-technical skills in various procedures, including laparoscopy, endoscopy, and emergency skills[4], [5].

Although its use has become widespread in developed countries, the adoption of VR in medical education in developing countries, including Indonesia, is still limited [6], [7]. The main obstacles include the cost of procuring hardware, the complexity of software development, and the lack of local validation evidence supporting the effectiveness of VR as a learning medium [8], [9]. In the Indonesian context, circumcision procedures are highly clinically relevant as they are one of the most frequently performed minor surgical procedures in primary health care facilities [10]. Medical students and general practitioners are required to master this skill early on, but opportunities for hands-on learning are often limited due to the limited number of patients, cultural norms, and variations in the techniques used [5], [11].

To address this need, the research team developed a VR-based circumcision simulator using Oculus Quest 2 as the main hardware. This headset was chosen because it is standalone, portable, and relatively more affordable than high-performance PC-based devices [12]. The application was developed using Unity 2021 LTS, with 3D anatomical designs and surgical instruments created using Blender 3.6, resulting in realistic procedural representations [13]. The development process was carried out in early 2024 at the Faculty of Engineering and Vocational Studies, while trials took place between December 2024 and January 2025 at the Faculty of Medicine, Ganesha University of Education (Undiksha), Bali.

At the time of the study, the laboratory facilities at Undiksha were not equipped with physiological measurement instruments such as heart rate variability (HRV), electrodermal activity (EDA), or eye-tracking. Therefore, this study focused its evaluation on three main domains, namely simulator sickness symptoms, user experience, and clinical performance. The instruments used included the Simulator Sickness Questionnaire (SSQ), Fast Motion Sickness Scale (FMS), Virtual Reality Neuroscience Questionnaire (VRNQ), User Experience Questionnaire–Short (UEQ-S), and clinical skill instruments in the form of a procedural checklist and Objective Structured Assessment of Technical Skills (OSATS) [14] [15].

The strength of this study lies in its relatively large sample size (74 participants) with a balanced distribution between males and females, as well as the stratification of experience into three groups: Novice, Intermediate, and Expert. This approach allows for a more robust construct validity analysis, in line with literature recommendations that generally involve 20–30 participants per group [16], [17]. By evaluating the VR-based circumcision simulator through various parameters, this study aims to assess the validity, feasibility, and educational impact of this local innovation. The results of this study are expected to not only enrich the international literature on the effectiveness of VR in medical education, but also provide an empirical basis for the integration of this technology into the medical curriculum in Indonesia [18-21].

2. METHODS

This study used a mixed-methods design with repeated measures, which was designed to assess the feasibility, user experience, and educational impact of a newly developed Virtual Reality (VR) circumcision simulator. The evaluation focused on three main domains: simulator sickness symptoms, user experience, and clinical performance. The VR Simulation Development Study was conducted at the Faculty of Engineering and Vocational Studies, and the trial was conducted at the Faculty of Medicine, Ganesha University of Education (Undiksha), Bali, Indonesia, from December 2024 to January 2025. All participants signed an informed consent form before participating in the research session.

2.1 Participants

The number of research participants was 74, divided into three groups based on clinical experience level to test the construct validity of the simulator, as follows.

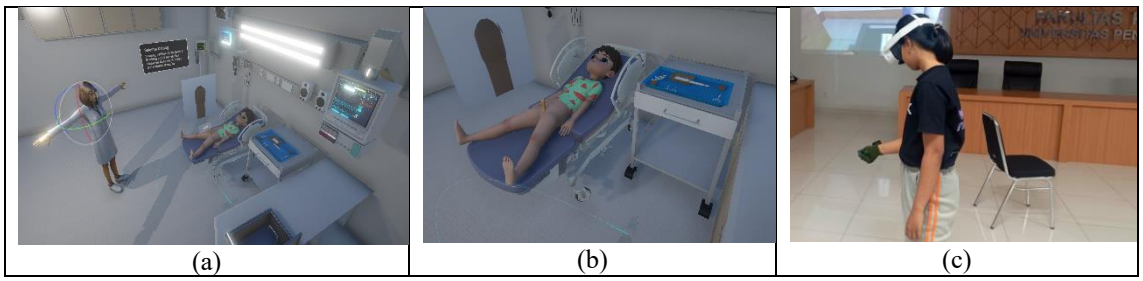
1. Novice (n = 25): Final-year undergraduate medical students (aged 20–23 years) who had never performed circumcision independently.
2. Intermediate (n = 24): Medical students undergoing clinical clerkship or co-assistant doctor training (aged 21–25 years), who had observed circumcision procedures but had limited experience in performing the procedure themselves.
3. Expert (n = 25): Licensed general practitioners (aged 24–50 years) who routinely perform circumcisions in their daily practice.

Gender distribution was balanced, with 37 males and 37 females, to minimize bias related to gender differences in VR tolerance. Recruitment was conducted purposively through the academic networks of the Faculty of Medicine and affiliated teaching hospitals. Inclusion criteria included willingness to participate in the study, normal or well-corrected vision, and no history of severe vestibular disorders. Exclusion criteria included a history of severe motion sickness, epilepsy or seizures, and previous intensive exposure to VR simulations.

2.2 VR Development Equipment

The circumcision simulator was developed using a combination of relatively affordable hardware and software capable of supporting the necessary immersion and interaction. The hardware used was an Oculus Quest 2 standalone VR headset with two hand controllers. This device was chosen for its portability, affordability, and system independence without the need for a high-spec computer.

The development computer, an Intel i7 with 16GB RAM and an NVIDIA RTX 3060 GPU, was used to build 3D models and run the rendering process. The simulator application was developed in Unity 2021 LTS with Oculus SDK and Unity XR Interaction Toolkit integration. The designed 3D model consists of the anatomical stages of the penis, surgical instruments, and the surgical area, visualized using Blender 3.7, resulting in a detailed and realistic anatomical representation, as shown in Figure 1.



(a) (b) (c)

Figure 1 (a) View of the Circumcision Surgery Simulation Room, (b) 3D Animation of a Circumcision Surgery Patient, (c) Participants performing VR-based Circumcision Surgery Simulation

The interaction framework consists of hand controls and object manipulation created with the XR toolkit using the Oculus plugin, while simple vibrations from the controller are integrated as haptic feedback to mark critical stages such as incision, clamping, and suturing. The simulator was developed in early 2024 through collaboration between medical lecturers, software developers, and 3D animators. Design iterations were performed several times based on expert input.

2.3 VR Software Design

This simulator is designed to replicate the standard clinical workflow of circumcision procedures in Indonesia. Three training modes are provided as follows.

1. Autonomous Mode: Participants perform the entire procedure independently with minimal visual guidance.
2. Guided Mode: Step-by-step instructions are displayed within the VR environment, including highlights on important anatomical landmarks.
3. Haptic-Enhanced Mode: Although Oculus Quest 2 does not support advanced force feedback, simple vibrations are used to signal when critical actions are taking place.

The procedure is divided into main stages: patient preparation, local anesthesia, prepuce incision, dissection, hemostasis, suturing, and dressing application. Each stage is accompanied by layered anatomical visualization, realistic tissue textures, and sound effects.

2.4 Sampling Strategy

The participant recruitment strategy used purposive sampling to ensure representation of each skill level. The number of participants in each group was determined based on previous studies on the validation of surgical VR simulators, which typically recruited 20–30 people per group. The target of 25 novices, 24 intermediates, and 25 experts was considered sufficient to ensure the statistical power of the repeated measures design. Students were recruited through class announcements and university emails, while general practitioners were recruited through affiliated hospitals and network clinics.

2.5 Study Procedures

The study took place at the Undiksha Clinical Skills Laboratory, which was adapted for VR simulation purposes. Each participant underwent the following procedure:

1. Orientation: Explanation of the research objectives, use of the headset, and VR controllers. Participants were given 5 minutes to familiarize themselves with the equipment to minimize the novelty effect.
2. Initial measurement: Participants filled out a demographic questionnaire and the Sickness Simulator Questionnaire (SSQ) prior to the session.
3. Simulation session: Participants undergo the three modes (Autonomous, Guided, Haptic) in a randomized order to reduce the learning sequence effect. Each mode lasts 15–20 minutes.
4. Post-session measurements: After completing each mode, participants completed the SSQ (post), FMS, and VRNQ (VRISE and UX subscales).
5. Performance assessment: During the session, the examiner used a 20-item procedural checklist and the OSATS (Objective Structured Assessment of Technical Skills) scale.
6. Break: A 5–10 minute rest is provided between modes to prevent VR fatigue.
7. Retention test: The checklist is repeated one week and one month after training to assess skill retention.

All activities were supervised by research assistants and medical lecturers who ensured participant safety and assessment consistency. The research procedure flowchart is illustrated in Figure 2. The flowchart is as follows.

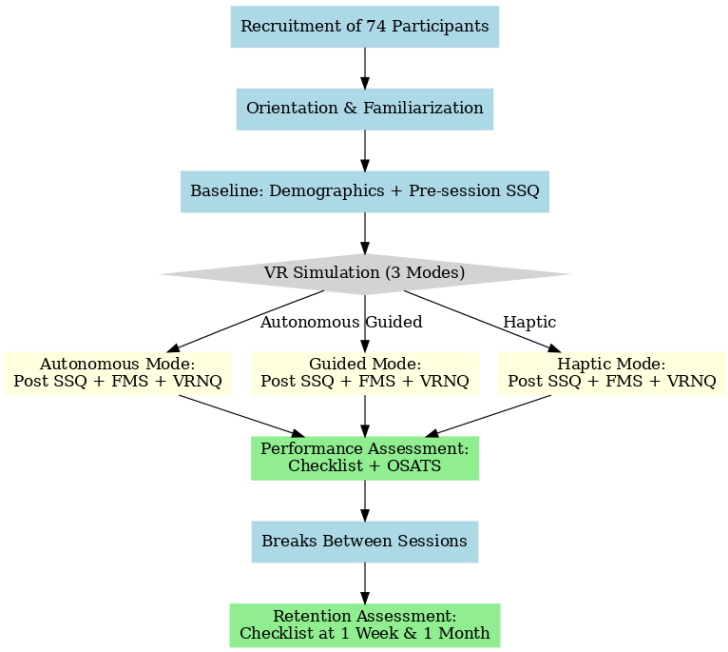


Figure 2 Research Flowchart.

2.6 Instruments

Research instruments are grouped into three categories as follows:

1. Simulator Sickness consists of the Simulator Sickness Questionnaire (SSQ) using 16 international standard items to assess symptoms of nausea, ocular discomfort, and disorientation. The Fast Motion Sickness Scale (FMS) consists of a single 20-point scale, which is quick to assess the perception of momentary sickness. The VRNQ-VRISE subscale is a brief tool for VR-induced symptoms.
2. User Experience consists of the VRNQ-UX subscale, which aims to measure enjoyment, presence, and usability, and the User Experience Questionnaire–Short (UEQ-S), which aims to assess pragmatic and hedonic quality.
3. Clinical Performance consists of a procedural checklist using 20 binary items (true/false), as well as OSATS using five domains (tissue handling, instrument handling, flow, time and motion, procedural knowledge).

This study did not include physiological measurements such as HRV, EDA, or eye-tracking because these facilities were not available at Undiksha at the time of the study. The focus of measurement was on simulator sickness symptoms, user experience, and procedural performance relevant to educational objectives.

2.7 Data Analysis

Data were analyzed using SPSS version 28. Descriptive analysis used mean, standard deviation, median, and frequency distribution. Simulator sickness used the Friedman test for SSQ, FMS, and VRNQ between modes, with post-hoc Wilcoxon (Bonferroni correction). User experience used UX score comparisons between modes with repeated measures ANOVA (or Friedman if assumptions were violated). Clinical performance used Mixed ANOVA with VR conditions as a within-subject factor and skill groups as a between-subject factor; the non-parametric Kruskal–Wallis alternative was also used. Retention was assessed using the Friedman test for Checklist scores at four time points (Baseline, Post, 1w, 1m). Correlations were assessed using Spearman's rho analysis between sickness, UX, and performance scores. The significance level was set at $p < 0.05$. Effect sizes were reported as partial η^2 or rho values for correlations.

3. RESULTS AND DISCUSSION

A total of 74 participants were enrolled in this study, consisting of 37 males and 37 females distributed evenly. Participants were grouped into three levels of clinical expertise. The Beginner group consisted of 25 final-year undergraduate medical students with a mean age of 21.3 years (SD = 1.4). The Intermediate group consisted of 24 medical students undergoing clinical clerkship, with an average age of 22.8 years (SD = 1.3). The Expert group consisted of 25 licensed general practitioners who routinely performed circumcisions, with an average age of 36.5 years (SD = 7.7). The age distribution in each group was consistent with the established stratification criteria outlined in the research protocol.

3.1 Descriptive Statistics

Table 1 shows the main descriptive statistics (mean ± SD, median, range) for the variables Simulator Sickness, User Experience, and clinical performance. As can be seen, SSQ and FMS scores tend to decrease as skill levels increase, while performance scores (Checklist, OSATS) increase consistently in the Expert group.

Table 1 Descriptive Statistics per Group

Group	Age	SSQ_ Auto	SSQ_ Guided	SSQ_ Haptic	FMS_Auto	FMS_Guided	FMS_Haptic	VRISE	UX	Checklist_Post	OSATS
Novice	21.3 ± 1.4	24.7 ± 9.2	20.5 ± 8.3	21.9 ± 7.9	7.4 ± 3.6	5.9 ± 3.2	6.4 ± 3.5	3.9 ± 0.6	3.2 ± 0.5	68.3 ± 9.7	14.5 ± 1.6
Intermediate	22.8 ± 1.3	20.7 ± 8.2	17.9 ± 7.2	18.1 ± 6.9	6.2 ± 3.0	4.7 ± 2.8	5.1 ± 2.9	3.4 ± 0.7	3.6 ± 0.6	75.2 ± 8.8	17.4 ± 1.5
Expert	36.5 ± 7.7	14.1 ± 7.6	11.2 ± 6.3	12.0 ± 6.8	4.3 ± 2.5	3.1 ± 2.2	3.4 ± 2.1	2.8 ± 0.5	4.2 ± 0.4	86.9 ± 6.1	21.7 ± 1.0
Overall	26.7 ± 8.6	19.9 ± 9.1	16.6 ± 8.0	17.4 ± 8.1	6.0 ± 3.2	4.6 ± 2.9	5.0 ± 3.0	3.4 ± 0.7	3.7 ± 0.6	76.7 ± 10.8	17.9 ± 3.1

3.2 Simulator Sickness Outcomes

Nonparametric analysis using the Friedman test showed significant differences in the level of simulator sickness symptoms between the three VR training modes shown in Table 2. For the Simulator Sickness Questionnaire (SSQ), the test results showed significant differences ($\chi^2(2) = 28.1$, $p < 0.001$). Similarly, Fast Motion Sickness (FMS) scores also differed significantly between modes ($\chi^2(2) = 21.7$, $p < 0.001$).

Table 2 Simulator Sickness (Friedman and Post-hoc)

Outcome	Friedman χ^2(df)	p	Comparison	p_adj
SSQ_Post	28.1 (2)	<0.001	Auto > Guided	<0.001
			Auto > Haptic	0.003
			Guided < Haptic	0.045
FMS	21.7 (2)	<0.001	Auto > Guided	0.001
			Auto > Haptic	0.005
			Guided vs. Haptic	0.212

Further testing with the Wilcoxon signed-rank test corrected using Bonferroni showed a consistent pattern of differences. On the SSQ, the Autonomous mode produced higher scores than both Guided ($p < 0.001$) and Haptic ($p = 0.003$), indicating a greater level of physiological discomfort. In addition, the comparison between Guided and Haptic also showed a significant difference, with Guided having a lower score ($p = 0.045$). The FMS results confirmed these

findings, with Autonomous being significantly higher than Guided ($p = 0.001$) and Haptic ($p = 0.005$). However, the difference between Guided and Haptic did not reach statistical significance ($p = 0.212$).

This study shows that simulator sickness symptoms measured using the Simulator Sickness Questionnaire (SSQ) and Fast Motion Sickness Scale (FMS) were highest in Autonomous mode compared to Guided and Haptic modes. This pattern is consistent with recent reports confirming that the higher the level of unguided interaction, the greater the cognitive load and potential visual disorientation experienced by users [1], [22]. Recent systematic research highlights that exposure duration, rendering quality, and interactivity level are the main determinants of virtual reality–induced symptoms and effects (VRISE) [23]. In this context, the use of Guided Mode has been shown to significantly reduce sickness symptoms. Instructional support appears to function as scaffolding that reduces excessive cognitive load, allowing users to adapt better. These findings support a gradual pedagogical approach, in which VR-based medical training should begin with instructional modes before transitioning to autonomous modes. Additionally, the integration of simple haptics through controller vibrations acts as a multisensory cue that aids spatial orientation. Recent evidence demonstrates that multisensory cues are effective in reducing vestibular discomfort during VR exposure [2], [3].

3.3 User Experience Outcomes

Kruskal–Wallis analysis in Table 3 shows a significant difference in VRNQ–VRISE scores among the three participant groups ($H = 15.3$, p < 0.001). The observed pattern shows a decrease in symptom intensity with increasing levels of clinical experience. Additionally, User Experience (UX) scores also differed significantly between groups ($H = 12.9$, $p = 0.002$), with the expert group reporting the most positive user experience compared to the other groups.

Table 3. User Experience (Kruskal–Wallis)

Variable	H	df	p
VRNQ-VRISE	15.3	2	<0.001
VRNQ-UX	12.9	2	0.002

The increase in UX scores in the Expert group demonstrates the construct validity of the simulator: individuals with more mature clinical experience are able to appreciate procedural realism and assess pedagogical usefulness more highly. Conversely, the high VRISE scores in the Novice group highlight the need for stronger instruction for beginners. Recent literature encourages the implementation of adaptive VR systems that adjust the level of instruction based on user experience, thereby improving the learning curve while reducing the risk of sickness [6] [8][7]

3.4 Performance Outcomes

Assessment using Kruskal–Wallis in Table 4 shows significant performance variations between groups, both in the post-training Checklist ($H = 20.7$, p < 0.001) and in the OSATS ($H = 24.1$, p < 0.001). Overall, the Expert group consistently scored higher than the Intermediate group, while the Novice group scored the lowest. This indicates that the simulator is able to differentiate skills according to clinical experience level.

Table 4. Performance Outcomes (Kruskal–Wallis)

Variable	H	df	p
Checklist_Post	20.7	2	<0.001
OSATS_score	24.1	2	<0.001

3.5 Retention Analysis

Analysis using the Friedman test in Table 5 on the Checklist scores at four points in time (Baseline, Post, 1 week, and 1 month) showed statistically significant differences between

measurements ($\chi^2(3) = 54.6$, p < 0.001). These results indicate that participants' procedural skills changed significantly over time after the VR-based training sessions.

Table 5. Retention Analysis (Friedman)

Outcome	χ^2(df)	p	Comparison	p_adj
Checklist	54.6 (3)	<0.001	Baseline < Post	<0.001
			Baseline < 1 week	<0.001
			Baseline < 1m	<0.001
			Post vs 1 week	0.216
			1 week vs 1 month	0.341

Post-hoc tests with Bonferroni correction showed that scores at Post-training were significantly higher than Baseline scores ($p < 0.001$). This confirms that VR training sessions directly improved participants' performance in the circumcision procedure. The comparison between Post-training and 1 week did not show a significant difference ($p = 0.216$), nor did the comparison between 1 week and 1 month ($p = 0.341$). This means that the skills acquired after training can be maintained stably for at least one month after the intervention. Furthermore, scores at the three post-training time points (Post, 1 week, and 1 month) all remained significantly higher than the Baseline scores ($p < 0.001$). These findings confirm that VR training not only produces immediate skill improvements after the session but also supports skill retention in the medium term.

Retention analysis indicates that the improvement in post-training performance lasts up to one month, reinforcing the argument that VR can facilitate medium-term skill retention. Recent studies confirm that VR training not only improves technical skills but also has transferability to real clinical practice [4], [7]. In the Indonesian context, where opportunities to encounter elective cases such as circumcision are limited, the existence of this simulator provides strategic added value. Although limitations such as the absence of physiological data (HRV, EDA, eye-tracking) reduce the depth of interpretation of biological mechanisms, behavioral evidence from performance scores is strong enough to support pedagogical effectiveness [22] [24].

3.6 Correlation Analyses

Spearman's correlation analysis in Table 6 shows a consistent pattern between virtual reality-induced symptoms and effects (VRISE), user experience (UX), and clinical performance indicators. VRNQ-VRISE scores showed a strong positive correlation with the Simulator Sickness Questionnaire (SSQ) ($\rho = 0.52–0.61$; $p < 0.001$) and with Fast Motion Sickness (FMS) ($\rho = 0.44–0.55$; $p < 0.01$). This relationship indicates that the higher the reported VRISE symptoms, the greater the level of simulator sickness discomfort measured by standard instruments. In other words, the convergent validity between VRNQ-VRISE and SSQ and FMS is confirmed, as all instruments measure similar phenomena despite their different formats.

Table 6. Spearman Correlations

Anchor	Variable	ρ	p
VRISE	SSQ_Auto	0.55	<0.001
VRISE	SSQ_Guided	0.52	<0.001
VRISE	SSQ_Haptic	0.61	<0.001
VRISE	FMS_Auto	0.48	<0.001
VRISE	FMS_Guided	0.44	0.002
VRISE	FMS_Haptic	0.55	<0.001
VRISE	Checklist_Post	-0.32	0.008
VRISE	OSATS	-0.36	0.004
UX	SSQ_Auto	-0.51	<0.001
UX	SSQ_Guided	-0.43	0.002
UX	SSQ_Haptic	-0.58	<0.001
UX	FMS_Auto	-0.47	0.001
UX	FMS_Guided	-0.41	0.003
UX	FMS_Haptic	-0.52	<0.001

UX	Checklist_Post	0.39	0.004
UX	OSATS	0.43	0.002

Conversely, VRISE scores had a negative correlation with clinical performance indicators. Checklist_Post scores were negatively correlated (ρ = −0.32; p < 0.01) and OSATS scores also decreased as VRISE symptoms increased (ρ = −0.36; p < 0.01). This pattern confirms that the discomfort caused by VR not only affects subjective comfort, but also has a direct impact on participants' ability to complete procedures correctly. This is consistent with cognitive load theory, in which increased physiological load due to nausea or disorientation can reduce cognitive capacity to perform technical skills optimally.

On the other hand, UX scores showed the opposite pattern. UX was negatively correlated with SSQ and FMS (ρ = −0.41 to −0.58; $p < 0.01$), indicating that participants who reported a more positive user experience tended to experience lower levels of simulator sickness symptoms. This relationship shows that perceptions of comfort, ease of use, and level of presence in VR play a protective role against the emergence of unwanted physical symptoms. Additionally, UX showed a positive correlation with clinical performance indicators (Checklist_Post: ρ = 0.39; OSATS: ρ = 0.43; both $p < 0.01$). Thus, the higher the perceived quality of the user experience, the better the technical skills achieved.

This combination of findings supports the argument that the success of VR implementation in medical education is not only determined by anatomical validity or technical realism, but also by the extent to which the simulator is able to minimize VRISE symptoms while improving UX. The positive correlation of VRISE with sickness and the negative correlation with performance underscore the risks of VR, while the opposite pattern in UX confirms its potential. Thus, simulator design strategies need to consider the balance between technical immersion and user comfort, as both aspects directly affect clinical learning outcomes. Correlation analysis shows that VRISE scores are positively related to SSQ/FMS and negatively related to performance. Conversely, UX is negatively correlated with sickness and positively correlated with performance. This pattern is consistent with *cognitive load* theory*,* in which sensory discomfort interferes with working memory capacity, thereby reducing skill acquisition [5], [11]. Recent literature confirms the importance of integrating ergonomic interface design, high *frame rates*, low latency, and gradual instructional modes to minimize sickness [12], [13]. The practical implication of these findings is that medical simulator development should not only focus on anatomical and procedural validity, but also on optimizing the user experience. This approach is relevant for middle-income countries, where affordable devices such as the Oculus Quest 2 can be used to produce a learning experience equivalent to high-cost simulators in developed countries [14], [15].

3.7 Correlation Matrix

Figure 3 shows the Spearman correlation matrix (ρ) between the Virtual Reality Neuroscience Questionnaire – VRISE scores and simulator sickness symptoms (SSQ, FMS) as well as clinical performance (Checklist, OSATS). The color pattern on the heatmap shows that positive correlations (marked with red gradations) dominate the relationship between VRISE and sickness measures. This indicates that the higher the VR-induced symptoms, the higher the SSQ and FMS scores reported by participants. Conversely, the relationship between VRISE and clinical performance is marked by blue colors, indicating a negative correlation, suggesting that higher sickness symptoms correlate with decreased performance on the procedural Checklist and OSATS scores. These findings are consistent with the literature emphasizing that physiological intolerance to the VR environment can interfere with concentration and reduce clinical skill learning outcomes. Thus, VRISE can serve as a sensitive indicator that predicts both discomfort and potential performance decline during VR-based training.

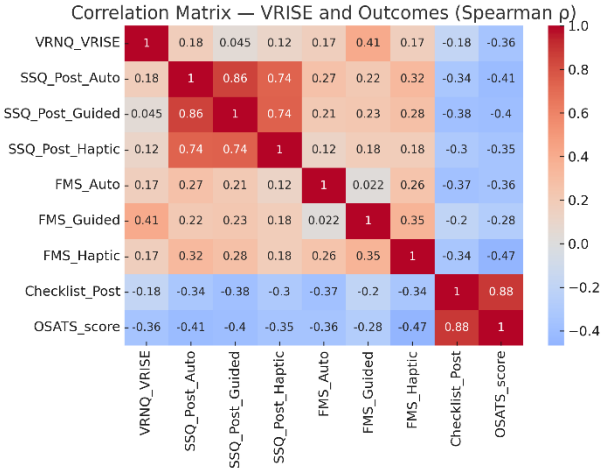


Figure 3. Correlation matrix between VRNQ-VRISE and outcomes (Spearman ρ).

Figure 4 shows the Spearman correlation matrix (ρ) between user experience (UX) scores and the same variables. The pattern is opposite to that in Figure 1. The blue color in the relationship between UX and SSQ and FMS indicates a negative correlation, meaning that the more positive the user experience, the lower the perceived symptoms of simulator sickness. Conversely, the relationship between UX and performance scores (Checklist and OSATS) appears in red, indicating a positive correlation. This means that the better the user's perception of the enjoyment, presence, and usability aspects of the simulator, the higher the clinical performance achieved. These results reinforce the argument that an optimal user experience not only has implications for comfort, but also directly supports technical skill learning. This correlation is consistent with previous research showing that user-centered design in VR simulators plays an important role in improving the efficiency of skill transfer to real-world practice.

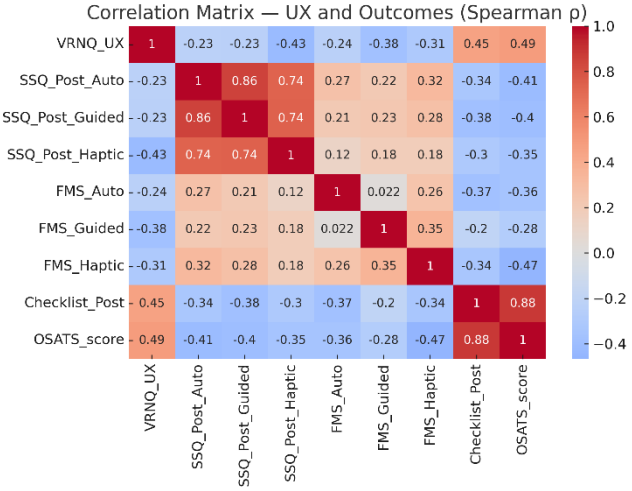


Figure 4. Correlation matrix between UX and outcomes (Spearman ρ).

4. CONCLUSIONS

This study successfully developed and validated a Virtual Reality (VR)-based circumcision simulator using Oculus Quest 2, Unity, and Blender, and tested it on 74 participants with varying levels of experience. The results show that this simulator has good construct validity: the Expert group consistently showed higher performance than the Intermediate and Novice groups. User experience (UX) was positively correlated with performance outcomes and negatively correlated with simulator sickness symptoms, while VRISE scores were negatively

correlated with performance. This confirms that subjective comfort and physiological tolerance are important determinants of VR learning effectiveness. Another key finding shows that Guided Mode and the integration of simple haptic feedback significantly reduce simulator sickness symptoms compared to Autonomous Mode. Additionally, clinical skill retention persists up to one month after the training session, confirming VR's potential as a medium-term training tool.

These results support the integration of VR as a circumcision training medium in Indonesia, especially in medical education institutions with limited patient access or clinical resources. The relatively affordable Oculus Quest 2 can be a low-cost solution to expand the use of VR in developing countries. Although this study provides significant empirical contributions, several limitations need to be addressed in future studies. First, physiological measurements such as *heart rate variability* (HRV), *electrodermal activity* (EDA), and *eye-tracking* should be added to understand the biological mechanisms of simulator sickness more comprehensively. Second, further research needs to evaluate *the transfer of training* to real clinical practice through field tests on patients. Third, the development of AI-based adaptive modes that can adjust instruction levels to user experience will improve learning efficiency. Finally, multi-center studies with larger samples from various institutions in Indonesia are needed to strengthen the generalization of results.

ACKNOWLEDGEMENTS

This research was financially supported by the Institute for Research and Community Service (LPPM) of Ganesha University of Education through an internal research funding scheme based on Number: SP DIPA-023.17.2.677530/2024. The author would like to express their appreciation to the Faculty of Engineering and Vocational Studies and the Faculty of Medicine at Ganesha University of Education for providing laboratory facilities, as well as to all research participants who voluntarily participated in the simulator testing process.

REFERENCES

[1] R. Firdaus, A. R. Tantri, and S. K. Manggala, "Factors Influencing Virtual Reality Sickness in Emergency Simulation Training," *Med Sci Educ,* vol. 34, no. 6, pp. 1309–1315, Dec. 2024, doi: 10.1007/s40670-024-02102-z.

[2] M. Mergen, M. Meyerheim, and N. Graf, "Reviewing the current state of virtual reality integration in medical education – a scoping review protocol," *Syst Rev*, vol. 12, no. 1, 2023, doi: 10.1186/s13643-023-02266-6.

[3] H. Jiang, S. Vimalesvaran, J. K. Wang, K. B. Lim, S. R. Mogali, and L. T. Car, "Virtual Reality in Medical Students' Education: Scoping Review," *JMIR Med Educ*, vol. 8, no. 1, 2022, doi: 10.2196/34860.

[4] D. Mistry, C. A. Brock, and T. Lindsey, "The Present and Future of Virtual Reality in Medical Education: A Narrative Review," *Cureus*, 2023, doi: 10.7759/cureus.51124.

[5] C. Elendu *et al.*, "The impact of simulation-based training in medical education: A review," Jul. 05, 2024, *Lippincott Williams and Wilkins*. doi: 10.1097/MD.0000000000038813.

[6] C. J. Clay, J. R. Budde, A. Q. Hoang, and A. Gushchyan, "An evaluation of the effectiveness of immersive virtual reality training in non-specialized medical procedures for caregivers and students: a brief literature review," 2024, *Frontiers Media SA*. doi: 10.3389/frvir.2024.1402093.

[7] S. S. Lie, N. Helle, N. V. Sletteland, M. D. Vikman, and T. Bonsaksen, "Implementation of Virtual Reality in Health Professions Education: Scoping Review," 2023. doi: 10.2196/41589.

[8] M. A. Minouei, A. Omid, A. Mirzaie, H. Mahdavifard, and A. Rahimi, "Effectiveness of virtual reality on medical students' academic achievement in anatomy: systematic review," *BMC Med Educ*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12909-024-06402-1.

[9] Y. M. Tang, K. Y. Chau, A. P. K. Kwok, T. Zhu, and X. Ma, "A systematic review of immersive technology applications for medical practice and education - Trends, application areas, recipients, teaching contents, evaluation methods, and performance," *Educ Res Rev,* vol. 35, no. November 2021, p. 100429, 2022, doi: 10.1016/j.edurev.2021.100429.

[10] B. M. Kyaw *et al.*, "Virtual reality for health professions education: Systematic review and meta-analysis by the digital health education collaboration," 2019. doi: 10.2196/12959.

[11] H. Sung, M. Kim, J. Park, N. Shin, and Y. Han, "Effectiveness of Virtual Reality in Healthcare Education: Systematic Review and Meta-Analysis," Oct. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/su16198520.

[12] A. Rubio-López, R. G. Carmona, L. Z. Román, A. R. Navas, A. González-Pinto, and P. Cardinal-Fernández, "Measuring Stress and Perceptions for a Virtual Reality–Based Pericardiocentesis Procedure Simulation for Medical Training: Usability Study," *JMIR Serious Games,* vol. 13, Jan. 2025, doi: 10.2196/68515.

[13] E. Jochmann, T. Jochmann, M. Weber, K. Weigel, and C. Klingner, "Impact of sensorimotor mismatch on virtual reality sickness and user experience: age-related differences in a randomized trial," J *Neuroeng Rehabil,* vol. 22, no. 1, Dec. 2025, doi: 10.1186/s12984-025-01677-x.

[14] J. H. Won, H. C. Na, and Y. S. Kim, "A New Training Method for VR Sickness Reduction," *Applied Sciences (Switzerland)*, vol. 14, no. 8, Apr. 2024, doi: 10.3390/app14083485.

[15] T. Abu Selo, Z. Hussain, Q. Hussain, W. Alhajyaseen, S. Al-Quradaghi, and M. Y. Alqaradawi, "Exploring Simulation Sickness in Virtual Reality Pedestrian Scenarios: Effects of Gender, Exposure, and User Perceptions," *Safety*, vol. 11, no. 3, p. 63, Jul. 2025, doi: 10.3390/safety11030063.

[16] P. Kourtesis, J. Linnell, R. Amir, F. Argelaguet, and S. E. MacPherson, "Cybersickness in Virtual Reality Questionnaire (CSQ-VR): A Validation and Comparison against SSQ and VRSQ," *Virtual Worlds,* vol. 2, no. 1, 2023, doi: 10.3390/virtualworlds2010002.

[17] I. Arshad, P. De Mello, M. Ender, J. D. McEwen, and E. R. Ferre, "Reducing Cybersickness in 360-Degree Virtual Reality," *Multisens Res*, vol. 29, 2021, doi: 10.1163/22134808-bja10066.

[18] J. R. Abbas *et al.*, "Virtual reality in simulation-based emergency skills training: A systematic review with a narrative synthesis," *Resusc Plus*, vol. 16, p. 100484, Dec. 2023, doi: 10.1016/J.RESPLU.2023.100484.

[19] A. Erbaş, E. Akyüz, and Ş. Ergöl, "The Effects of Virtual Reality Used in Healthcare Education on Cybersickness and Sense of Presence: A Systematic Review," *PRESENCE: Virtual and Augmented Reality*, vol. 33, pp. 161–178, Jul. 2024, doi: 10.1162/pres_a_00414.

[20] R. Walls *et al.*, "Virtual reality as an engaging and enjoyable method for delivering emergency clinical simulation training: a prospective, interventional study of medical undergraduates," *BMC Med*, vol. 22, no. 1, Dec. 2024, doi: 10.1186/s12916-024-03433-9.

[21] D. Saredakis, A. Szpak, B. Birckhead, H. A. D. Keage, A. Rizzo, and T. Loetscher, "Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis," *Front Hum Neurosci,* vol. 14, 2020, doi: 10.3389/fnhum.2020.00096.

[22] S. Cossio *et al.*, "Cybersickness and discomfort from head-mounted displays delivering fully immersive virtual reality: A systematic review," *Nurse Educ Pract*, vol. 85, p. 104376, May 2025, doi: 10.1016/J.NEPR.2025.104376.

[23] L. Simón-Vicente, S. Rodríguez-Cano, V. Delgado-Benito, V. Ausín-Villaverde, and E. Cubo Delgado, "Cybersickness. A systematic literature review of adverse effects related to virtual reality," *Neurología (English Edition),* vol. 39, no. 8, pp. 701–709, Oct. 2024, doi: 10.1016/J.NRLENG.2022.04.007.

[24] T. Tene, D. F. Vique López, P. E. Valverde Aguirre, L. M. Orna Puente, and C. Vacacela Gomez, "Virtual reality and augmented reality in medical education: an umbrella review," 2024, *Frontiers Media SA*. doi: 10.3389/fdgth.2024.1365345.