

"HoloMaïa": Augmented Reality gesture simulation system for midwifery clinical training and anxiety analysis

Clara Pérez de los Cobos Cintas

`clara.perez-de-los-cobos-cintas@grenoble-inp`

GSCOP

Nicolas Vuillerme

Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG Sangria

Lionel Di Marco

Department of Midwifery, Université Grenoble Alpes

Guillaume Thomann

GSCOP

Research Article

Keywords: Augmented reality, gesture simulation, educational training, midwifery, anxiety analysis

Posted Date: September 24th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10917081/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Digital technologies increasingly shape modern society, and medical education must adapt by incorporating innovative learning tools such as extended reality (XR). Additionally, in midwifery education, XR simulation has shown promise for preparing students to manage obstetric emergencies.

Methodology

This study aimed to evaluate whether integrating augmented reality (AR) into a practical midwifery course provides immersive learning experiences and affects student anxiety and self-efficacy during high-stress clinical simulations. The AR system, called HoloMaïa, was developed for HoloLens 2 with Unity and Mixed Reality Toolkit. It includes multiple manoeuvres, scenarios, and complications that allow students to practice obstetric emergency management in an immersive environment.

Results

The heart rate increased during AR simulation phases compared with baseline, with specific spikes corresponding to emergency-related cues. Despite this physiological stimulation, the AR group (B) showed no statistically significant increase in overall anxiety compared with the control group (A). Self-confidence levels remained stable across both groups regardless of the intervention method. Furthermore, students showed a positive trend in technology adoption, as reflected by comfort, perceived usefulness, and intention to use.

Conclusion

These findings suggest that AR-based training allows a more immersive learning experience without negatively impacting student anxiety or self-confidence. This supports the integration of AR in obstetric emergency training, and the next steps involve multicentre testing and hardware migration for long-term scalability.

1 Introduction

As digital technologies increasingly shape modern society, younger generations are becoming more accustomed to their integration into everyday life. Consequently, medical education must adapt by incorporating innovative learning tools such as virtual reality (VR) and augmented reality (AR), which can align with students' learning preferences and enrich educational experiences [1]. In midwifery education, evidence suggests that AR simulation may also be an effective strategy for preparing students to manage obstetric emergencies [2].

On the other hand, the mental health of healthcare students is a major concern, as it affects their personal quality of life, academic functioning and quality of patient care, particularly during emergencies

[3]. While practical courses are essential for developing clinical skills, their influence on psychological factors such as performance anxiety and self-confidence remains underexplored. Addressing these aspects of readiness could facilitate students' transition into clinical practice and improve their ability to perform under pressure [4]. Given the significant role that midwives play in reducing maternal and neonatal morbidity and mortality, providing high-quality, low-stress clinical training is essential to the professional development of midwifery students [5].

To address this gap, we must establish whether the implementation of a new midwifery practical course using VR and/or AR provides learning experiences and strategies that improve the course and enhance self-efficacy in practical skills. The need for empirical validation to confirm the efficacy of the new practical course as a pedagogical method and its potential to help manage student anxiety motivated this research.

Specifically, this study evaluated the immersive effect of the new AR midwifery course and its correlation with students' anxiety and self-confidence levels, providing a justification for the formal integration of AR technology into the midwifery curriculum.

1.1 State of the Art

Extended reality (XR) includes VR, AR and mixed reality (MR). XR technologies create immersive and interactive environments with physical and virtual elements that can complement traditional teaching methods. Their immersive features, interactive capabilities, and accessible experiences across all levels make them an attractive pedagogical tool that is becoming increasingly prevalent in education [6].

These technologies allow students to participate in hands-on, experimental learning activities, such as gamified virtual laboratories [7], surgical education [8], [9], or training for high-risk careers as rescue training [10]. They also permit experiments with complex concepts and skills that would be difficult, expensive, or dangerous to replicate in real life. In nursing and midwifery education, studies have shown that XR tools allow students to practice complex clinical scenarios in a safe and controlled environment, improving their preparedness for real-life situations while reducing performance anxiety [11].

To understand the impact of XR on students, measuring psychological and physiological responses that reflect immersion and anxiety levels is also essential. Physiological indicators such as heart rate and heart rate variability [12], alongside self-reported anxiety scales [13], can reveal how immersive simulations influence emotional regulation. Complementary data, such as usability feedback and comfort, contribute to evaluating the overall educational experience. Furthermore, assessing students' adoption of XR technology is important, as acceptance supports sustained use and the effective implementation of XR-based learning in midwifery education [14].

On the basis of these comprehensive evaluation principles, the objective is to assess an AR-based training system, its contribution to a more immersive learning environment, its effect on anxiety and self-

confidence among midwifery students and its technology acceptance levels during simulated clinical emergencies.

1.2 Research Questions

- The goal of this study is to evaluate the impact of AR technologies in midwifery education on immersive learning experiences, student anxiety levels during high-stress clinical simulations, and technology acceptance among students. To guide this investigation, the three following research questions were proposed:
- (RQ1): Does AR-based training allow a more immersive learning environment for students?
- (RQ2): How does anxiety vary with the presence of AR in practical midwifery training?
- (RQ3): What is the level of adoption of the AR learning system among midwifery students following its use?

2 Methodology

2.1 Research Design

The study followed a controlled, randomised, three-session experimental design to assess the impact of the AR intervention. The primary evaluation focused on performance during Session 2, using Session 1 as a common baseline.

All the participants were required to experiment with both conditions: (Session 1) the standard practical session and (Session 2) the AR intervention, as all the students were expected to complete the same curriculum during the academic year. Therefore, a “catch-up” session (Session 3) was included for ethical and pedagogical reasons but was excluded from the primary statistical analysis.

The students were randomly assigned to one of two sequence groups, each following a specific three-session sequence, as presented in Table 1.

Table 1
Experimental design summary.

Group	Session 1 (Baseline)	Session 2 (Primary evaluation)	Session 3 (Catch-up)
Group A (Control)	Standard Session	Standard Session	AR Session
Group B (Experimental)	Standard Session	AR Session	Standard Session

2.2 Scenarios and teaching content

Midwifery training follows a progressive and comprehensive approach designed to prepare students for the challenges of clinical practice, particularly in managing obstetrical emergencies. Before progressing to practical sessions, students complete a theoretical course explaining shoulder dystocia birth complications and the necessary management procedures. The standard shoulder dystocia practical course consists of an emergency gesture simulation using a low-fidelity birthing manikin (bottom half of the mother and a baby). In this simulation, the teacher initiates the birth, and the student is required to perform the necessary manoeuvres, as shown in Fig. 1.

In contrast, the AR practical course (see Fig. 2) also includes the birthing manikin but introduces scenario variability and immersive elements, such as delivery room sounds, social interaction or visual indicators. These digital elements are designed to bridge the gap between static manikin gesture simulation practices and the reality of a live clinical environment.

2.3 Materials

The hardware materials used in this study included devices and tools essential for implementing and monitoring mixed reality simulations and physiological data collection.

- Microsoft HoloLens 2 headset: creates an immersive learning environment that includes real-world elements, such as the student's hands or the mannequin, with interactive digital components, such as the baby's heart rate (HR) [15].
- Polar H9 sensors: To monitor the participants' HR throughout the experimental (AR) session [16].

The software materials included applications and platforms used for designing the AR experience, managing data collection, and analysing participant performance and responses.

- HoloMaia program: This program was developed with Unity, a cross-platform game engine widely used for real-time 3D applications, in combination with the Mixed Reality Toolkit (MRTK).
- HR tracking software: A custom Python program was developed to capture and record heart rate signals from the Polar sensors on a second-by-second basis. The connectivity issues identified during preliminary testing were resolved to ensure reliable data acquisition.
- Python for data analysis: Different libraries such as SciPy, NumPy, Pandas, and matplotlib were used.

Different questionnaires used to evaluate the experimental sessions. A set of five general questions about the course, the State and Trait Anxiety Inventory (STAI) [17], the Nursing Anxiety and Self-Confidence with Clinical Decision Making (NASC-CDM) [13], the Instructional Materials Motivation Survey (IMMS) [18], the NASA Task Load Index (TLX) [19], the Comfort Questionnaire [16], the Technology Acceptance Model (TAM) [17], the Net Promoter Score (NPS) [18] and an open question concerning the users' experience, was used. These include assessments of students' anxiety levels

throughout the course, as well as measures of comfort, usability, and other aspects related to the use of AR.

2.4 System description: HoloMaïa

HoloMaïa is an AR application developed for the Microsoft HoloLens 2 device, which is designed to make the simulation of obstetric emergencies and labour management more immersive. The system architecture is divided into three primary components:

- Onboarding and Calibration: To minimise the "novelty effect" and technical barriers, the application includes an integrated video tutorial that teaches the user gesture control and holographic spatial alignment.
- Simulation environment: The system features three clinical scenarios that are chosen by the teacher via a "blind" selection process. This approach allows the teacher to decide the specific manoeuvres and clinical complications without the student's prior knowledge, testing real-time decision-making.
- Immersive environment features: The simulation incorporates the following:

Visual cues: A dynamic foetal heart rate monitor that responds to complications, an information panel displaying completed tasks and other information, and a timer indicating the elapsed time since the start of the simulation. As shown in Fig. 3.

Auditory cues: High-fidelity sounds, including heart rate, maternal vocalisations and neonatal crying.

Interactive elements: social interaction, such as a responsive mother who behaves on the basis of the scenario, or other medical staff.

- Data Logging and Performance Analysis: The application collects data, featuring an automated information panel that registers every action taken by the student. After the scenario is finished, the system generates a summative report including a chronological log of actions, facilitating objective debriefing and assessment.

2.5 Participants

All fourth-year midwifery students from Grenoble Alpes University were invited to participate in the study. Participation took place during a mandatory course that forms part of the midwifery curriculum.

The group of students was evenly distributed into two groups for the experiments, with 18 participants each. The experimental groups were well balanced in terms of population type. No significant differences were found across the groups for average age ($p = 0.76$), the number of students repeating the year ($p = 0.46$), or gender ($p = 1.00$).

2.6 Data collection and analysis

Data were collected from questionnaires and biometric devices. In addition, the entire simulation session was recorded on video, along with a first-person video recording from the HoloLens perspective for each participant while they performed the manoeuvres.

Both quantitative and qualitative analyses were conducted. The quantitative data included physiological measurements and questionnaire responses, which were analysed via Python. Qualitative data were derived from the open-ended questions and the observational notes from the experiments and other comments.

2.7 Ethics consideration

As the research project was classified as a “noninterventional study”, it did not require review by an ethics committee [20]. Nevertheless, all participants provided informed consent. The consent form details the experimental session methodology, participants’ rights (including withdrawal), potential benefits and risks, the right to ask questions, and the dissemination of results.

To ensure fairness and maintain the integrity of the experimental design, all the participating students followed an identical curriculum, including the same courses and a “catch-up” session. All questionnaires and data collected during the study were strictly anonymized to prevent potential bias.

3 Results

3.1 Event-related physiological responses during AR simulation

HR was continuously recorded throughout the entire two-and-a-half-hour session 2 for each student. For the purpose of this analysis, HR data corresponding specifically to the AR simulation phase (for example, 200 seconds in Fig. 4) were isolated and compared with HRs measured outside the simulation period (during the same session).

In addition, video recordings captured by the HoloLens during the session were used to identify the timing of key events occurring within the AR scenario. These events included elements such as the appearance of a red emergency panel, the moment when the student was required to call the obstetrician for assistance, or the end of the simulation marked by the new-born’s cry. The temporal alignment of these events with the HR signal allowed event-related segments of physiological data to be extracted.

Figure 4 illustrates a representative heart rate segment during the AR simulation, with annotated events overlaid on the HR time series. The periods corresponding to the AR simulation have higher average HR values than the periods outside the simulation. Variations in HR can be observed in temporal proximity to specific simulation events, including increases following emergency-related cues and a decrease toward the end of the scenario.

Since heart rate variations detected during the AR simulation were temporally aligned with key scenario events, this suggests that participants were more engaged during the learning activity. In simulation-based education, these physical responses are often used to show cognitive and emotional involvement. They reflect how the learner is actively processing the situation and responding to the changing demands of the task [21].

Critical moments in the scenario, such as emergency cues or the need to request assistance, were associated with more pronounced heart rate changes. These events require rapid evaluation and action, which are known to increase cognitive load and emotional arousal [22] [23] [24]. The physiological responses observed during these phases match an increased level of situational awareness, as participants had to monitor clinical signs, anticipate consequences, and adapt their actions.

The presence of AR may have contributed to these responses by increasing the realism of the simulated environment. By including auditory and visual cues in the learner's field of view, AR can support the perception of being involved in a more realistic clinical situation.

3.2 Anxiety scores between sessions and groups

This section examines the differences in anxiety scores after the simulations and compares anxiety levels between the non-AR group (Group A) and the AR group (Group B). Anxiety was assessed after each experimental session via the same questionnaires for both groups. The STAI-6 was used to evaluate anxiety after the first session (S1), in which all participants completed the standard course. The NASC-CDM was administered after the second session (S2), where Group A repeated the standard course and Group B completed the AR course.

Session 1 (S1): Baseline anxiety (STAI-6)

As S1 served as a standardised baseline for all 36 participants, an analysis of the STAI-6 scores was conducted to ensure group comparability prior to the AR intervention. The STAI-6 results revealed that both groups had similar levels of calmness ($M_A = 3.12$ vs. $M_B = 3.00$), relaxation ($M_A = 3.06$ vs. $M_B = 2.61$), and comfort ($M_A = 2.75$ vs. $M_B = 2.72$). Specifically, Group B was slightly more tense ($M_A = 2.00$ vs. $M_B = 2.61$) and worried about making mistakes ($M_A = 2.31$ vs. $M_B = 2.72$). However, Group A was more upset about the scenario itself than Group B was ($M_A = 2.12$ vs. $M_B = 1.44$). The results are represented in Fig. 5 and summarised in Table 2.

Session 2 (S2): AR Intervention (NASC-CDM)

In S2, Group B used AR, whereas Group A continued with the traditional simulation. Despite Group B's greater initial tension and performance concerns in S1, the introduction of AR did not lead to a statistically significant increase in overall anxiety (Q1 to Q15) compared with the control group in S2 ($M_A = 45.59$ vs. $M_B = 50.52$; $p > 0.05$). Furthermore, self-confidence (Q16 to Q27) remained stable across both groups ($M_A = 46.94$ vs. $M_B = 44.52$), as shown in Fig. 6 and summarised in Table 2.

Table 2

Comparison of average anxiety and self-confidence levels in Session 1 (S1), following the standard course, and in Session 2 (S2) with (Group B) and without AR (Group A).

Metric	Group A	Group B	Difference (ΔM)
Tense (S1) (1 to 5)	2.00 $\pm$ 0.84	2.61 $\pm$ 0.91	+ 0.61
Comfortable facing SD (S1) (1 to 5)	2.75 $\pm$ 0.77	2.72 $\pm$ 0.75	-0.03
Anxiety (S2) (from 14 to 84)	45.59 $\pm$ 12.10	50.52 $\pm$ 13.70	+ 4.93
Self-Confidence (S2) (from 13 to 78)	46.94 $\pm$ 5.40	44.52 $\pm$ 9.88	-2.42

3.3 Technology adoption

Technology adoption was measured via the comfort questionnaire, the TAM (technology acceptance model) and the NPS (net promoter score). All participants completed the full protocol. Data were collected from Group B during S2 (individual simulation) and Group A during S3 (catch-up session with the whole group in the practical classroom).

Across all questionnaires, only two items had isolated missing values (TAM4 for one participant in Group B; TAM6 and TAM7 for one participant in Group A), which were handled through pairwise exclusion in the subscale calculations.

Physical Comfort

Discomfort scores were low across all participants ($M = 1.85$, $SD = 0.54$ on a 1–6 scale). The item-level analysis revealed that eye fatigue (CO4: $M = 2.08$) and difficulty appreciating fine details (CO5: $M = 2.97$) were the most frequently reported symptoms, whereas headaches (CO3: $M = 1.19$), nausea (CO6: $M = 1.08$), and general discomfort (CO1: $M = 2.05$) remained near the floor of the scale, as shown in Fig. 7. No significant difference in discomfort was found between the groups ($t(34) = 0.14$, $p = .887$).

Technology acceptance (TAM)

Technology acceptance was assessed via an 11-item questionnaire structured around four TAM subscales: perceived usefulness (PU; items 1–3), perceived ease of use (PEOU; items 4–6), enjoyment/engagement (ENJ; items 7–8, with item 8 reverse-coded), and behavioural intention (BI; items 9–11). The internal consistency was satisfactory for PU ($\alpha = .89$) and BI ($\alpha = .93$) and acceptable for PEOU ($\alpha = .67$).

The mean scores and standard deviations for each subscale are reported in Table 3. Overall, all subscales exceeded the scale midpoint (3/5), with ENJ ($M = 4.33$, $SD = 0.55$) and BI ($M = 4.13$, $SD = 0.77$) having the highest scores. PU was also positive ($M = 3.75$, $SD = 0.63$). The PEOU was the lowest subscale ($M = 3.04$, $SD = 0.65$), with the highest variability, as shown in Fig. 8.

Table 3

TAM results across all students. Note. Scale 1–5. Item 8 reverse-coded (r). Welch's t tests, two-tailed. No comparisons reached statistical significance. α for ENJ is left blank because it has only 2 items.

Subscale	Items	α	Overall M (SD)	Group B M (SD)	Group A M (SD)	t	p
Perceived Usefulness (PU)	1–3	.89	3.75 (0.63)	3.58 (0.69)	3.94 (0.52)	-1.79	.083
Ease of Use (PEOU)	4–6	.67	3.04 (0.65)	3.02 (0.63)	3.07 (0.68)	-0.23	.818
Enjoyment (ENJ)	7–8r	—	4.33 (0.55)	4.39 (0.39)	4.26 (0.69)	0.69	.499
Behavioural Intention (BI)	9–11	.93	4.13 (0.77)	3.96 (0.87)	4.31 (0.62)	-1.39	.173
TAM Global	1–11	—	3.81 (0.44)	3.74 (0.46)	3.90 (0.41)	-1.09	.283

Welch t tests revealed no statistically significant differences between groups on any subscale (all $p > .05$), indicating consistent acceptance across participants. A nonsignificant trend toward higher PU was found in Group A ($p = .083$).

Net Promoter Score (NPS)

Students' likelihood of recommending the HoloMaia tool was evaluated through two NPS items rated on a 0–10 scale. The mean ratings were 7.14 (SD = 1.74) for the general recommendation of the tool and 7.42 (SD = 1.61) for the recommendation as a method for learning emergency procedures such as shoulder dystocia. Using standard NPS classification (promoters: 9–10; passives: 7–8; detractors: 0–6), the global NPS was -13.9 for general use and -8.3 for emergency procedure learning.

These scores reflect a predominantly passive response pattern rather than active dissatisfaction. The majority of participants gave scores of 7 or 8, indicating conditional approval, as shown in Fig. 9. Group A showed a more favourable profile, with an NPS of -5.9 for the general tool and 0.0 for the emergency procedures, the second indicating an equal balance of promoters and detractors.

4 Discussion

4.1 Result interpretation

The physiological responses should not be interpreted only as stress markers. Instead, they could indicate engagement within a complex learning environment. In the context of obstetric emergency training, this type of engagement is a key part of learning although experience, helps the development of situational awareness and prepares students for real-life clinical situations [25].

These interpretations are further supported by the psychological measures collected during the study. Baseline levels of confidence and anxiety (S1) were comparable between the groups. Although Group B

showed slightly greater anxiety after S1 (which was conducted without AR for all participants), the introduction of AR in S2 did not result in a statistically significant increase in overall anxiety compared with the control group. Furthermore, self-confidence remained stable across both groups.

Taken together, these findings suggest that the increased physiological responses observed in the AR group may reflect engagement and immersion rather than stress exclusively within a complex learning environment [26]. However, this immersion may be attributable not only to the technology itself but also to the presence of simulated scenarios in S2 compared with the simple gesture simulation of S1.

If anything, AR may slightly challenge learners' perceived confidence, possibly because increased realism is coupled with increased emotional and sensory information. The participants were required to process multiple elements at once, such as visual cues, clinical signs, and procedural decisions, similar to the demands of real-life obstetric emergencies. This complexity may temporarily increase arousal while facilitating experiential learning [27]. In addition, the heightened realism provided by AR may make learners more aware of the challenges of real clinical situations, which could reduce their perceived confidence while simultaneously preparing them more effectively for future professional practice, as we see after S2 in Group B compared with Group A.

In the technology adoption assessment, there was no significant difference between groups on any measure, suggesting that positive reception was consistent across contexts. If anything, Group A tended toward higher usefulness scores. This group performed the AR session after completing two standard courses and performed the AR simulation in a group format rather than individually as Group B, which may have influenced their perception. This topic may need exploration in future studies with larger samples.

Discomfort scores were low across all participants, indicating that the AR system imposed minimal physical burden during the simulation session. These results suggest that the headset is well tolerated for short clinical training simulations, which is an important safety and usability precondition for wearable technology, particularly in environments where learners already face high cognitive loads [28].

With respect to technology acceptance, all the subscales surpassed the scale midpoint, with enjoyment and behavioural intention having the highest scores, indicating strong engagement and willingness to reuse the tool. Perceived usefulness was also positive with good internal consistency, reflecting students' perception that AR could enhance their learning of shoulder dystocia procedures. Perceived ease of use was the lowest subscale, suggesting that ease of use, while above the midpoint, remains an area for design improvement.

Finally, the negative NPS data reflect the strict promoter/detractor methodology, where only 9–10 s count as promoters, rather than genuine student dissatisfaction, placing most students solidly in "passive" territory rather than "detractor." This pattern is consistent with first-exposure adoption curves in health education [29].

4.2 Limitations

Several limitations should be considered when these findings are interpreted. First, the sample size was relatively small, which may have limited the statistical power to detect subtle differences in anxiety or self-confidence between groups. A larger cohort would strengthen the robustness and generalizability of the results [30].

Second, physiological activation was assessed primarily through HR variations. While HR is a sensitive indicator of stimulation, it does not allow differentiation between positive engagement, cognitive load, and stress-related responses. The inclusion of additional physiological markers (such as heart rate variability, electrodermal activity, or cortisol levels) would provide a deeper understanding of the emotional and cognitive mechanisms of AR-based learning [31].

Third, the study was conducted within a single institutional and educational context. As simulation practices, student characteristics, and curricular structures may vary across institutions, caution should be exercised when these findings are generalised to other midwifery or healthcare training programs. Nevertheless, the next steps in this research will include multicentre testing and evaluation.

Additionally, the short-term nature of the evaluation limits conclusions regarding long-term educational outcomes. While AR appeared to enhance immersion without increasing anxiety during the sessions, its impact on knowledge retention, skill transfer, and clinical performance over time remains to be determined.

Finally, the novelty effect of AR technology may have influenced participants' engagement and physiological responses. Further studies are needed to determine whether the effects remain consistent once they become familiar with the technology.

5 Conclusion

This study aimed to evaluate whether AR-based training allows immersion (RQ1), how anxiety varies with the presence of AR in midwifery practical training (RQ2), and the level of adoption among students (RQ3). The findings indicate that AR contributes to a more immersive learning environment. This is shown by the close temporal alignment between scenario events and students' physical responses, which suggests greater engagement.

In terms of anxiety, the introduction of AR did not result in a significant increase in self-reported anxiety or a decrease in self-confidence. Although physiological arousal was greater during AR-enhanced simulation, this could reflect more focus and emotional involvement. The more realistic and stimuli-rich environment could be the reason for the slightly higher anxiety and lower levels of the experimental group (Group B).

Overall, the system was considered acceptable for use in midwifery training. The students reported low physical discomfort, positive perceived usefulness, high engagement, and the intention to reuse the system. The consistency of these results across groups strengthens confidence in their reliability.

Together, these results support the use of AR in obstetric emergency training to provide a more immersive learning experience without negatively affecting student anxiety.

6 Contribution of the study

This study proposes a structured framework for the development of an AR system for midwifery education, with the potential to be extended to other midwifery contexts and healthcare education more broadly. The approach is reproducible and flexible and can be adapted across various clinical scenarios, training programs, and cultural environments.

From a practical perspective, this study offers insights into how an AR system can be effectively integrated into existing curricula, providing a more immersive and experiential learning experience without negatively affecting students' emotional state. It also helps identify the main advantages and limitations of AR as a teaching tool and guides future implementation strategies in educational institutions.

Finally, this work has strong international development potential through ongoing collaborations, particularly in Brazil. These partnerships will support the deployment and evaluation of the system across new educational and cultural contexts, enabling assessment of its adaptability, scalability, and broader relevance.

7 Future research directions

Future research directions will focus on extending the applicability and scalability of the HoloMaia system.

First, the system will be adapted and deployed in an international context by modifying and testing the tool within the obstetric nursing training programs of the National Institute of Telecommunications (Inatel) and the Centro de Simulação em Saúde at the Federal University of Paraíba (UFPB), both of which are in Brazil. This will facilitate the evaluation of its usability, cultural adaptability, and educational effectiveness across different healthcare and pedagogical settings while also establishing a standardised deployment protocol for global implementation.

Second, to ensure the long-term viability of the system following the discontinuation of Microsoft HoloLens 2, the HoloMaia system is migrated to Meta Quest 3. This transition will allow continued use of the AR program while supporting broader implementation within educational courses.

Abbreviations

XR: Extended reality

AR: Augmented reality

VR: Virtual reality

MR: Mixed reality

RQ: Research question

HR: Heart rate

NASC-CDM: Nursing Anxiety and Self-Confidence with Clinical Decision Making

NPS: Net Promoter Score

TAM: Technology acceptance model

Declarations

9.1 Ethics approval and consent to participate

See 2.7.

9.2 Consent for publication

Written informed consent for publication was obtained from all participants.

9.3 Availability of data and materials

The dataset generated during this study is available from the corresponding author upon reasonable request.

9.4 Competing interests

The authors declare that they have no competing interests.

9.5 Funding

This work was funded by the French National Research Agency (ANR) in the framework of the Investissements d'avenir program (ANR-10-AIRT-05) and the MIAI Cluster (ANR-23-IACL-0006).

9.6 Author contributions

All the authors contributed equally.

9.7 Acknowledgements

This work was funded by the French National Research Agency (ANR) in the framework of the Investissements d'avenir program (ANR-10-AIRT-05) and the MIAI Cluster (ANR-23-IACL-0006).

References

1. Visan IG, Toma CV, Petca R, Petrescu GED, Noditi A-R, Petca A. Exploring the Efficacy of Virtual Reality Training in Obstetric Procedures and Patient Care—A Systematic Review. *Healthcare*. Apr. 2025;13(7):784. 10.3390/healthcare13070784.
2. Vogel K, Bernloehr A, Willmeroth T, Blattgerste J, Hellmers C, Bauer NH. Augmented reality simulation-based training for midwifery students and its impact on perceived knowledge, confidence and skills for managing critical incidents, *Midwifery*, vol. 136, p. 104064, Sep. 2024, 10.1016/j.midw.2024.104064
3. Frajerman A, Colle R, Jollant F, Corruble E. Mental health in midwifery students: A French nationwide survey. *Midwifery*. Dec. 2024;139:104165. 10.1016/j.midw.2024.104165.
4. Ang WHD, et al. Evaluation of a psychological readiness program and final clinical practicum among final year nursing students: A mixed methods study. *Nurse Educ Today*. Oct. 2024;141:106317. 10.1016/j.nedt.2024.106317.
5. Seyedi R, Dousti S, Shabani F, Hajian S. Psychological stressors affecting midwifery students in clinical education: a systematic review. *BMC Med Educ*. Oct. 2025;25(1):1464. 10.1186/s12909-025-08051-4.
6. Al-Harthy AS, Al Balushi MY, Al Badi AH. Exploration of Educational Possibilities of Extended Reality(XR) in Higher Education Society, in 2023 *International Conference on Intelligent Metaverse Technologies & Applications (iMETA)*, Tartu, Estonia: IEEE, Sep. 2023, pp. 1–6. 10.1109/iMETA59369.2023.10294659
7. Vahdatikhaki F, Friso-van Den I, Bos S, Mowlaei, Kollöffel B. Application of gamified virtual laboratories as a preparation tool for civil engineering students, *Eur. J. Eng. Educ.*, vol. 49, no. 1, pp. 164–191, Jan. 2024, 10.1080/03043797.2023.2265306
8. Ntakakis G, Plomariti C, Frantzidis C, Antoniou PE, Bamidis PD, Tsoulfas G. Exploring the use of virtual reality in surgical education. *World J Transpl*. Feb. 2023;13(2):36–43. 10.5500/wjt.v13.i2.36.
9. Wan T, Liu K, Li B, Wang X. Validity of an immersive virtual reality training system for orthognathic surgical education. *Front Pediatr*. Mar. 2023;11:1133456. 10.3389/fped.2023.1133456.
10. Pedram S, Palmisano S, Skarbez R, Perez P, Farrelly M. Investigating the process of mine rescuers' safety training with immersive virtual reality: A structural equation modelling approach. *Comput Educ*. Aug. 2020;153:103891. 10.1016/j.compedu.2020.103891.
11. Pérez C, De Cintas LC, Vuillerme N, Thomann G, Di Marco L. Effect of an Extended Reality Simulation Intervention on Midwifery Students' Anxiety: Systematic Review. *JMIR Nurs*. Jun. 2025;8:e68984–68984. 10.2196/68984.
12. Kim H-G, Cheon E-J, Bai D-S, Lee YH, Koo B-H. Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature, *Psychiatry Investig.*, vol. 15, no. 3, pp. 235–245, Mar. 2018, 10.30773/pi.2017.08.17

13. White KA. Development and Validation of a Tool to Measure Self-Confidence and Anxiety in Nursing Students During Clinical Decision Making, *J. Nurs. Educ.*, vol. 53, no. 1, pp. 14–22, Jan. 2014, 10.3928/01484834-20131118-05
14. Welfare M, Crowther S, Chapman L, Frost J. Exploring virtual reality in undergraduate midwifery education: A qualitative systematic review. *Collegian*. Jun. 2026;33(3):232–9. 10.1016/j.colegn.2026.03.007.
15. Palumbo A. Microsoft HoloLens 2 in Medical and Healthcare Context: State of the Art and Future Prospects. *Sensors*. Oct. 2022;22:7709. 10.3390/s22207709.
16. Gomes N, Pato M, Lourenço AR, Datia N. A Survey on Wearable Sensors for Mental Health Monitoring. *Sensors*. Jan. 2023;23(3):1330. 10.3390/s23031330.
17. Zsido AN, Teleki SA, Csokasi K, Rozsa S, Bandi SA. Development of the short version of the spielberger state–trait anxiety inventory. *Psychiatry Res*. Sep. 2020;291:113223. 10.1016/j.psychres.2020.113223.
18. Loorbach N, Peters O, Karreman J, Steehouder M. Validation of the Instructional Materials Motivation Survey (IMMS) in a self-directed instructional setting aimed at working with technology. *Br J Educ Technol*. Jan. 2015;46(1):204–18. 10.1111/bjet.12138.
19. Hart SG, Field M. NASA-Task Load Index (NASA-TLX): 20 Years Later.
20. Flochlay A-C. Le comité d'éthique pour la recherche Grenoble Alpes (CERGA), Université Grenoble Alpes. Accessed: Feb. 25, 2026. [Online]. Available: <https://www.univ-grenoble-alpes.fr/universite/engagements/ethique-et-deontologie/comite-d-ethique-pour-la-recherche-grenoble-alpes/le-comite-d-ethique-pour-la-recherche-grenoble-alpes-cerga-1240768.kjsp?RH=1683886433067>
21. Ayres P, Lee JY, Paas F, Van JJG, Merriënboer. The Validity of Physiological Measures to Identify Differences in Intrinsic Cognitive Load. *Front Psychol*. Sep. 2021;12:702538. 10.3389/fpsyg.2021.702538.
22. Fernández-Ayuso D, et al. The Modification of Vital Signs According to Nursing Students' Experiences Undergoing Cardiopulmonary Resuscitation Training via High-Fidelity Simulation: Quasi-Experimental Study. *JMIR Serious Games*. Aug. 2018;6(3):e11061. 10.2196/11061.
23. Kozlowski D, Hutchinson M, Hurley J, Rowley J, Sutherland J. The role of emotion in clinical decision making: an integrative literature review. *BMC Med Educ*. Dec. 2017;17(1):255. 10.1186/s12909-017-1089-7.
24. Que Y, Zheng Y, Hsiao JH, Hu X. Using eye movements, electrodermal activities, and heart rates to predict different types of cognitive load during reading with background music. *Sci Rep*. Sep. 2025;15(1):32635. 10.1038/s41598-025-03052-1.
25. Morgan P et al. Jan., Using a situational awareness global assessment technique for interprofessional obstetrical team training with high fidelity simulation, *J. Interprof. Care*, vol. 29, no. 1, pp. 13–19, 2015, 10.3109/13561820.2014.936371

26. Thomas V, Hönning A, Curiel RC, Harhaus-Wähner L, Prahm C. Physiological responses to virtual sensory stimuli in immersive XR: a systematic review. *Front Virtual Real.* May 2026;7:1810042. 10.3389/frvir.2026.1810042.
27. Pico J, et al. From Realism to Learner Engagement: Rethinking Fidelity in Simulation-Based Education. *JMIR Med Educ.* Feb. 2026;12:e84684–84684. 10.2196/84684.
28. Ryan G, et al. Virtual reality in midwifery education: A mixed methods study to assess learning and understanding. *Nurse Educ Today.* Dec. 2022;119:105573. 10.1016/j.nedt.2022.105573.
29. Alvarado-Robles B, Rodriguez-Ramirez AG, Luviano-Cruz D, Ortiz-Muñoz D, Alonso-Mendoza VM, Garcia-Luna F. Technology Assessment Models in Healthcare Education: An Integrative Review and Future Perspectives in the Era of AI and VR. *Appl Sci.* Jan. 2026;16(3):1213. 10.3390/app16031213.
30. Stuart EA, Bradshaw CP, Leaf PJ. Assessing the Generalizability of Randomised Trial Results to Target Populations. *Prev Sci.* Apr. 2015;16(3):475–85. 10.1007/s11121-014-0513-z.
31. Vage A, Gormley GJ, McKeown G, Hamilton PK. The evaluation of stress during medical simulation: a feasibility study involving final-year medical students. *Int J Healthc Simul.* Jun. 2023;psla1602. 10.54531/psla1602.

Figures



Figure 1

Standard shoulder dystocia practical course (without scenario).



Figure 2

AR shoulder dystocia practical course (with scenario variability).



Figure 3

Screenshot from HoloLens 2 of the HoloMaïa system.

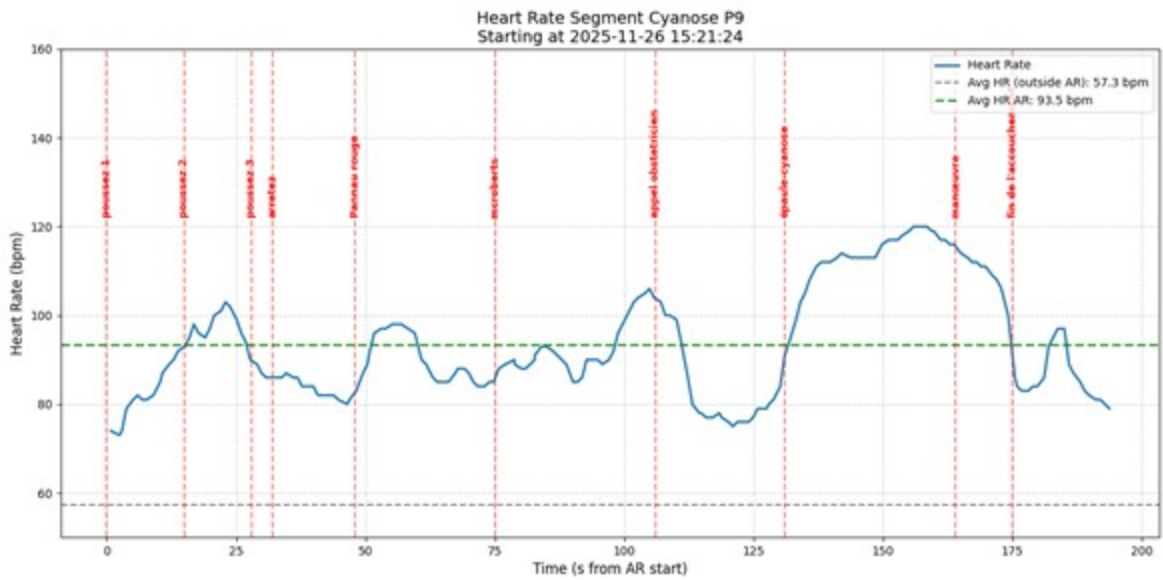


Figure 4

HR during simulation with annotated events of one of the participants.

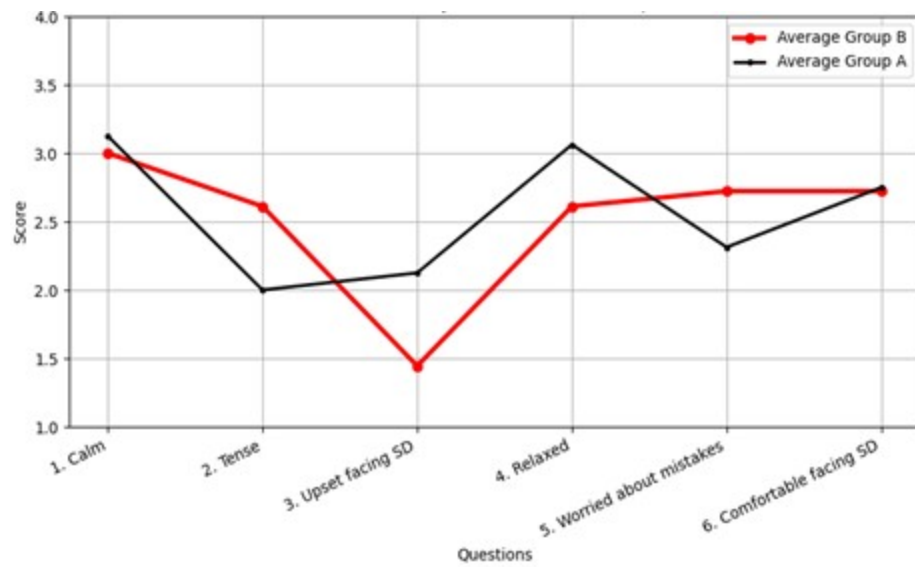


Figure 5

STAI-6 average results of Groups A and B after the first experimental session (S1).

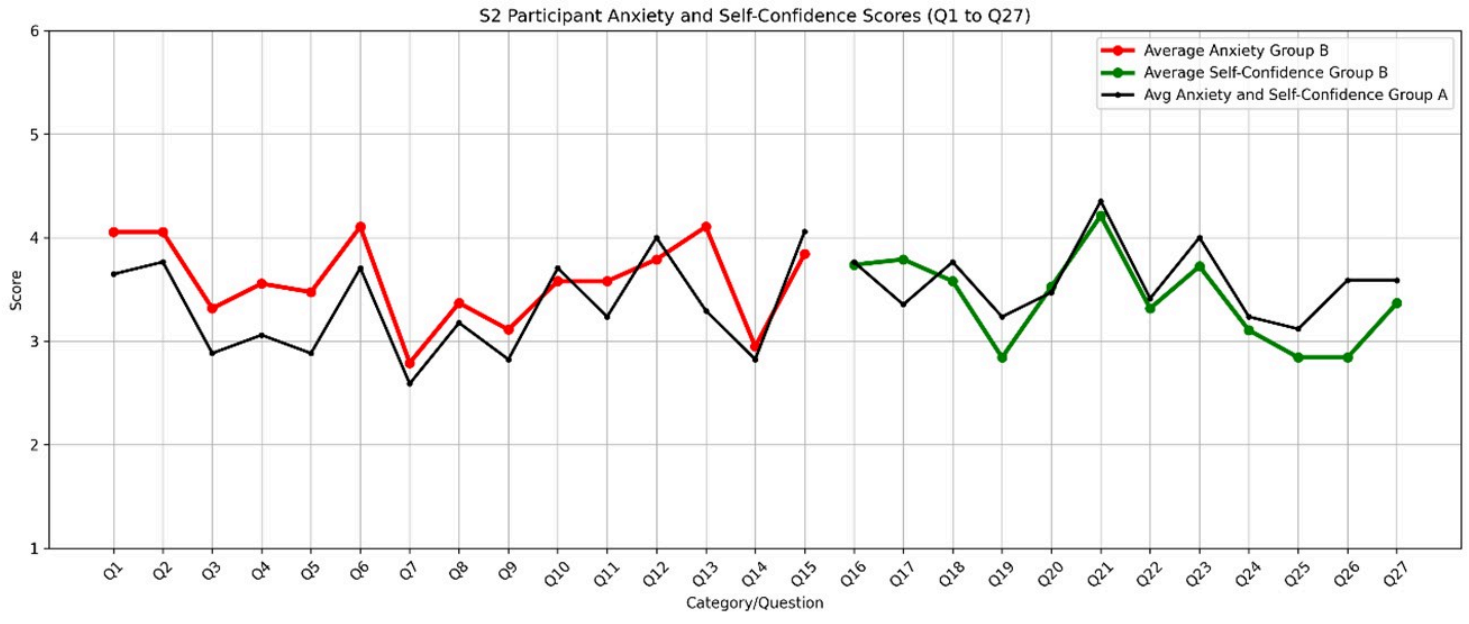


Figure 6

NASC-CDM, average anxiety (Q1-Q15) and self-confidence (Q16-Q27), results of Groups A and B after the second experimental session (S2).

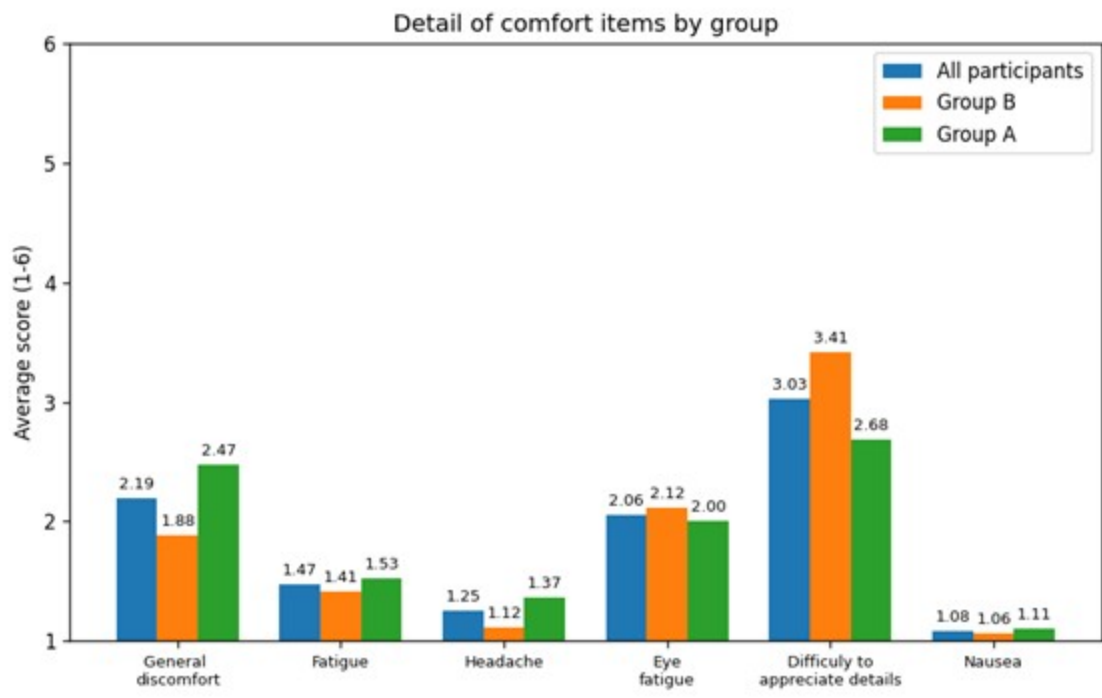


Figure 7

Bar chart of mean scores per comfort item (C01–C06), by group and overall. Scale 1-5.

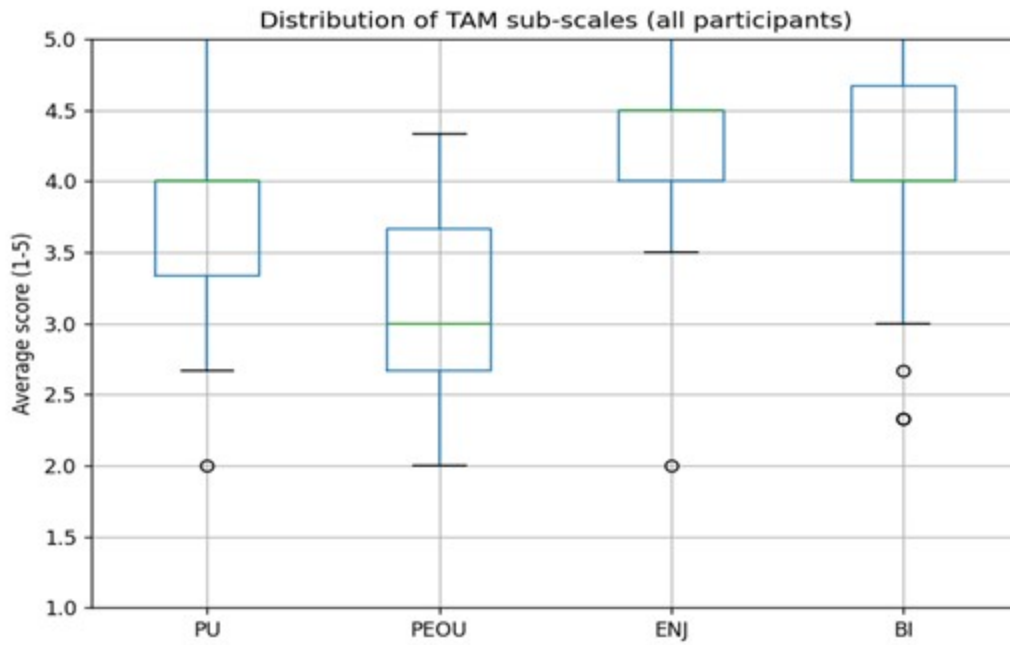


Figure 8

Boxplot of TAM subscales overall. Scale 1-5.

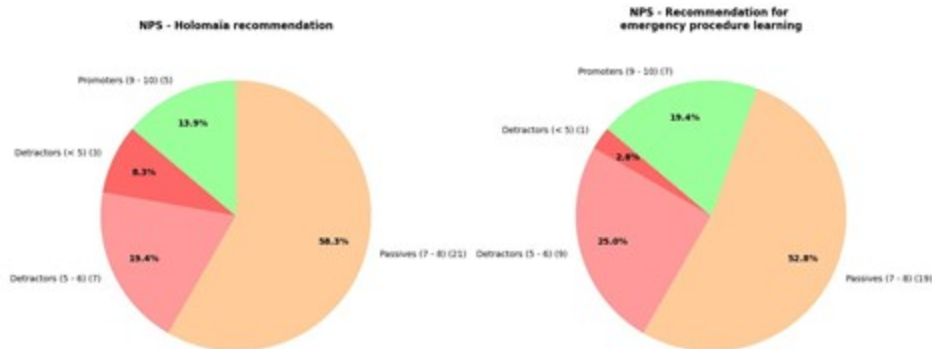


Figure 9

Promoter/passive/detractor breakdown. Scale 1-10.