Advances in Medical Education and Practice

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Implementation and Analysis of a Fully Immersive Virtual Reality-Based Emergency Training in a Surgical Curriculum

Nader El-Sourani[1,2], Tobias Mühling[3], René Klarmann[2], Susanne Quintes[2], Maximilian Bockhorn[1,2]

[1]Department for General - and Visceral Surgery, University Hospital Oldenburg, Klinikum Oldenburg Aör, Oldenburg, Germany; [2]Faculty for Medical and Health Sciences, Carl von Ossietzky University Oldenburg, Oldenburg, Germany; [3]University Hospital Würzburg, Institute of Medical Teaching and Medical Education Research, Würzburg, Bavaria, Germany

Correspondence: Nader El-Sourani, Email nader.el-sourani@uol.de

Introduction: Due to the changing demands in medical education, there is a necessity to increase the extent and quality of practical teaching. Virtual Reality (VR) enforces learning from simulated experience due to its immersive and interactive environment.

Methods: VR-based training sessions were implemented as the last module before medical students were entering their clerkship. 84 students were enrolled in this study. 24 of them were active users in complex VR-based emergency scenarios (AU) while the other 60 students were observers (OBS). A questionnaire was completed pre - and post intervention to evaluate motion sickness, intuitive use, immersive experience, subjective learning and perceived competence.

Results: Overall, the implementation into the surgical curriculum was feasible. The technical aspects of the program, particularly regarding usability, were generally well-rated by the AU. The degree of immersion and the subjective learning success were reported higher by AU compared to OBS. In the pre/post comparison, a nuanced picture emerged with a significant increase in competence in diagnostic reasoning and initiation of a treatment algorithm, while other competency facets showed no change.

Conclusion: VR can be seen as a good teaching tool in medical education as it improved the subjective learning experience and perceived competence of students. However, for complex clinical emergencies, prior knowledge is usually required, which is why such scenarios are preferably implemented in later stages of the curriculum.

Keywords: virtual reality, medical education, surgery, medical students, virtual emergency

Introduction

The rapid and significant expansion of medical knowledge[1] poses a particular challenge for young doctors at the start of their careers. Given the vast amount of information, the traditional rote learning methods used in medical education are no longer sufficient, underscoring the need for new, meaningful learning approaches.[2,3] This is especially critical in emergency medicine where often immediate action is necessary, leaving no time to consult reference materials. Despite the increasing emphasis on practical training in this area, junior doctors still show deficits in the management of emergency patients.[4–6] However, due to their need for personnel and resources, current simulation approaches are unlikely to adequately address this gap in the near future.

Fortunately, there has been a recent expansion in the methods of delivering medical education. The use of head-mounted Virtual Reality (VR) displays can deliver fully immersive and interactive environments offering the chance for meaningful learning experiences.[7] The use of VR technology in medical education can enhance learning effectiveness and efficiency[8] and learning outcomes may even be superior to traditional teaching methods.[9] Within the literature, various methods are grouped under the term "VR", ranging from 3D representations on screens to 360-degree videos and fully immersive 3D environments on head-mounted displays. Research by Gutiérrez et al indicates that fully immersed VR trainings yield notably higher learning gains compared to partially immersed VR, particularly in medical student

Received: 19 November 2024
Accepted: 14 April 2025
Published: 16 May 2025

education.[10] Due to decreasing hardware prices, scenarios using head-mounted displays are becoming more cost-effective and scalable[11,12] in comparison to analogous simulation settings. Consequently, as an extension of the traditional curriculum, VR trainings offers a way to provide meaningful learning experiences to a large number of students in the future.

Thus far, VR has found successful applications primarily in trainings for robotic surgery, interventional procedures, and hygiene practices[13–15] as well as brain death diagnostic and emergency medicine.[16,17] Especially the latter trainings rely on advanced knowledge and skills (eg, clinical reasoning, task prioritization), making their use particularly beneficial later during the medical curriculum. Recently, Mühling et al implemented a VR-based training approach for complex medical emergencies, exploring perceived stress levels, and subjective learning outcomes among senior students.[18]

Currently, existing studies on the use of complex VR scenarios in medical education take into account, to some extent, the impact of demographic parameters (such as age, gender, and prior experience with 3D/VR applications) on the learning experience.[17,18] However, important user-specific characteristics, such as technological proficiency, as well as hardware- and software-specific factors such as usability, are often overlooked. Yet, poor usability has been proven to be one of the greatest barriers to the curricular implementation of virtual reality.[19,20] Therefore, this study aims to understand not only learning outcomes but also the influence of technological proficiency and usability. Since it is considered good practice, especially in the field of simulation-based training, to take participants' prior knowledge into account, the current study will also assess perceived competence before and after the training.[21] Against that background, the following research questions were formulated:

1. How is the VR-training module perceived by our students with respect to usability (intuitive use, degree of presence and grade of immersion)?
2. Does the VR-training provide a positive learning experience for the students in terms of subjective learning success and perceived competence (the latter measured in a pre-/post design)? Are there any differences between active users and observers?
3. How do technological proficiency, usability (intuitive use, presence) of the VR-training and learning experience correlate with each other?

Materials and Methods

Medical Curriculum

The medical program in Germany is divided into pre-clinical and clinical years. After completing their surgical curriculum consisting of lectures, problem-based-learning and various seminars, students perform a 1-month clinical rotation in surgery. The VR module was implemented as the last teaching unit in our surgical curriculum before students were to take their OSCE (Figure 1) and prior to entering their clinical rotation.

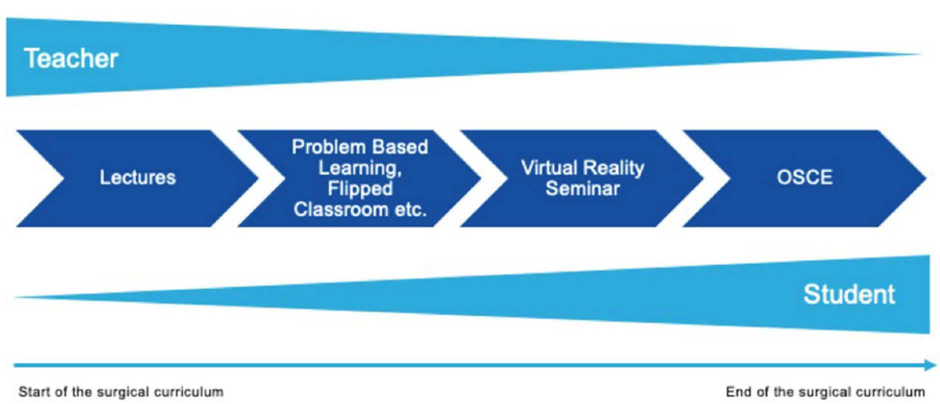

Figure 1 Surgical curriculum before entering the surgical clerkship. The role of the teacher decreases while the role of the student increases. The VR seminar concludes the curriculum.

 https://doi.org/10.2147/AMEP.S507357

Study Design

The VR-based simulation was carried out as the last mandatory teaching session of the surgical curriculum at the European Medical School Oldenburg prior to taking the OSCE and advancing into the four-week surgical clerkship (see Figure 1). Thus, students gradually increased their active role while the teacher's role decreased. All participants were fourth-year medical students. The study was conducted as a pre-post study.

Before participating in the VR seminar, students underwent an inverted classroom, during which they prepared for the virtual seminar using online videos as instructions. The videos explained the handling of the hardware and software. Only the students who successfully completed the inverted classroom were able to participate as active users (AU). Students that completed the inverted classroom participated as AU voluntarily. The remaining students participated as observers (OBS) and were able to view the perspective of the AU on a monitor and provide verbal assistance.

A total of eight VR-based teaching sessions were carried out over a timespan of two hours each. A maximum of 10–15 students were in each group. The VR-based teaching sessions were supervised by a surgical consultant. Three cases for interdisciplinary training, covering both medical and surgical learning objectives, were simulated during each seminar with repeated content: 1. acute biliary pancreatitis with persistent cholestasis, 2. abdominal sepsis due to fistulizing Crohn's disease, 3. oesophageal variceal bleeding due to liver cirrhosis. All medical content of the scenarios was based on established guidelines and reviewed by at least two experienced faculty members with specialization in medical education and/or emergency medicine.[22–24] A questionnaire was provided before and after the seminar. The study was conducted in accordance with the ethics committee of the University of Oldenburg (AZ 2023–130). A flow chart of the study design is shown in Figure 2.

VR-Based Simulation

The STEP-VR program was designed for medical students to simulate the role of physicians encountering various critically ill patients in an emergency department. The training software (version 0.15 beta) was developed by ThreeDee GmbH (Munich, Germany), a startup specializing in 3D visualization. Two head-mounted displays (HMD) were used (Meta Quest 2, META) as seen in Figure 3. The training software and head-mounted displays were running on a high-performance gaming computer and connected by wireless LAN (WLAN). STEP-VR aims to enhance cognitive abilities such as executive functions, situational awareness, time management, and task prioritization. Diagnostic tools provided include medical history, laboratory tests, various imaging techniques, and treatment options, all based on authentic patient data. Certain procedures, such as blood work and intravenous (IV) cannula placement, are simulated in an abstract manner. Once the user determines the most likely diagnosis, therapeutic interventions can be performed, with real-time effects on the simulated patient's physiology. Lastly, a checklist based on German treatment guidelines assessed whether the correct actions have been taken.

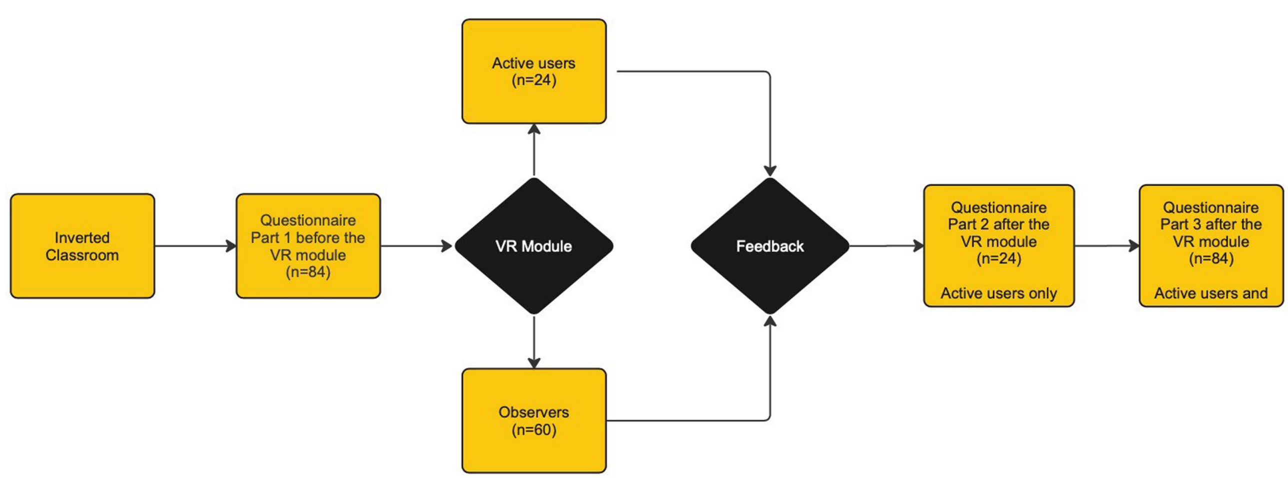

Figure 2 Study protocol displaying the pre-post design including patient-cases and questionnaires.

Figure 3 (**A**) Head-mounted displays by META (Quest 2). (**B**) In-App view of the virtual emergency department displaying a patient connected to a monitor measuring vital signs. (**C**) Active (AU) and passive (OBS) participants during the VR-based simulation.

Questionnaire

The questionnaire was developed with the department of academic affairs, quality management and evaluation. The software used was "QuestPro 5 by Blubbsoft GmbH". Questionnaires that were adapted or modified are marked in Table 1. It was divided into three parts. Part 1 was answered by all participants and included demographics, experience with VR as well as technological proficiency and perceived clinical competence before the seminar. Part 2 was answered by AU only and included the performed case, motion sickness, presence, intuitive use and perceived stress. Part 3 was answered by all participants and included degree of immersion, subjective learning success, perceived clinical competence after the seminar and a field for comments. A simplified overview can be seen in Table 1.

Statistical Analysis

The analysis was conducted using descriptive statistics, including mean (M), minimum (Min), maximum (Max) and standard deviation (SD). The Little's Missing Completely at Random Test (MCAR) was applied to the entire dataset to

Table 1 Overview of the Questionnaire Used

	Part 1 (Pre-Session)	**Part 2 (Post-Session)**	**Part 3 (End of Seminar)**
Student Group	**All Students**	**Only AU**	**All Students**
Theme:	Demographics	VR-scenario	Grade of immersion
Questions:	*Age* *Sex* *Previous use of VR and console-gaming*	*One-choice-only*	*Four questions with Likert-scale 1–5*
Theme	Technological proficiency	Motion Sickness*	Subjective learning success*
Questions:	*Four questions with Likert-scale 1–6*	*One-choice-only and open questions to symptoms (based on*	*Five questions with Likert-scale 1–5*
Theme:	Perceived competence (before)	Presence*	Perceived competence (after)
Questions:	*Seven questions with Likert-scale 1–5*	*Two questions with Likert-scale 1–5*	*Seven questions with Likert-scale 1–5 (identical to item 4)*
Theme:		Intuitive use	Open comments

(*Continued*)

Table 1 (Continued).

	Part 1 (Pre-Session)	Part 2 (Post-Session)	Part 3 (End of Seminar)
Student Group	**All Students**	**Only AU**	**All Students**
Questions:		*Eight questions with Likert-scale 1–5*	
Theme:		Perceived stress*	
Questions:		*Nine questions with Likert-scale 1–5*	

Notes: *Data from these studies.[18,25–29]

assess whether values were missing systematically or completely randomly. As most of the parameters were not distributed normally, we chose non-parametric tests: Mann–Whitney Test was used for comparisons between groups (eg AU/OBS). Perceived competence before and after the seminar was analysed using Wilcoxon matched pairs signed rank tests. Spearman's rank test was used for calculating correlation coefficients (r). For the correlation analysis, we calculated the difference in perceived competence between values collected before and after the VR training, labelled as "perceived competence gain". Data analysis was performed using IBM SPSS Statistics 29.0.1 (International Business Machines Corporation, Armonk, NY, USA), and table creation was done using Excel or Word (Microsoft 2010, Redmond, WA, USA).

Results

Demographics

Out of all 84 students enrolled in the study, 60 students (71.4%) were OBS and 24 students (28.6%) were AU. A vast majority of the students were female (81.9%). 52 students (61.9%) were between the ages of 18–25, 25 students (29.8%) between the ages of 26–30 and seven students (8.3%) were over 30 years old. No significant differences were found regarding age (p=0.655) and gender (p=0.591) when comparing AU and OBS. 57 students (65.9%) never played any sort of PC or console-gaming and 79 students (94%) denied any previous use of VR or AR. Students' demographics are displayed in Table 2.

Table 2 Students Demographics

Variable	Sub-Variable	N (%)	AU N (%)	OBS N (%)
Age in years	<18	0 (0.0)	0 (0.0)	0 (0.0)
	18–25	52 (61.9)	15 (62.5)	37 (61.7)
	26–30	25 (29.8)	8 (33.3)	17 (28.3)
	>30	7 (8.3)	1 (4.2)	6 (10.0)
Sex	Male	15 (17.9)	5 (20.8)	10 (16.7)
	Female	68 (81.9)	18 (75)	50 (83.3)
	Diverse	0 (0.0)	0 (0.0)	0 (0.0)
	Not specified	1 (1.2)	1 (4.2)	0 (0.0)
Use of PC or console gaming	Never	57 (67.9)	16 (66.7)	41 (68.3)
	Once a month	18 (21.4)	3 (12.5)	15 (25.0)

(*Continued*)

 https://doi.org/10.2147/AMEP.S507357

Table 2 (Continued).

Variable	Sub-Variable	N (%)	AU N (%)	OBS N (%)
	Once a week	6 (7.1)	4 /16.7)	2 (3.3)
	Frequently	2 (2.4)	1 (4.2)	1 (1.7)
	Daily	1 (1.2)	0 (0.0)	1 (1.7)
	Not specified	0 (0.0)	0 (0.0)	0 (0.0)
Use of VR or AR (gaming)	Never	79 (94.0)	23 (95.8)	56 (93.3)
	Once a month	3 (3.6)	1 (4.2)	2 (3.3)
	Once a week	0 (0.0)	0 (0.0)	0 (0.0)
	Frequently	0 (0.0)	0 (0.0)	0 (0.0)
	Daily	0 (0.0)	0 (0.0)	0 (0.0)
	Not specified	2 (2.4)	0 (0.0)	2 (3.3)

Technological Proficiency Before the VR-Module

The analysis of "technological proficiency" showed no significant statistical differences for all items questioned. While the enjoyment of dealing with technical problems was below average (2.95 ± 1.35), participants felt confident in resolving them independently (3.94 ± 1.34). Results of the comparison between AU and OBS for "technological proficiency" are displayed in Table 3.

However, gender-specific differences were found in the self-assessment of "technological proficiency" as well as in the "use of PC or console games". Among all male students, 46.7% stated that they use PC or console games at least once a week. Among female students, 81% denied any use. Supplement Table 1 presents a gender-specific comparison of "technological proficiency". Male participants especially rated their abilities higher when it came to "independently solving technical problems" ($p=0.002$) and "enjoying dealing with technical problems and solving them" ($p=0.003$). Two statements were not statistically significant.

AU Results for Presence, Intuitive Use and Perceived Stress and Motion Sickness

The questions answered for the items presence, intuitive use, perceived stress by AU after the simulation are displayed in Table 4. Regarding the item presence, the students fully agreed with the statement "I had the feeling of being in the room with the patient" (4.42 ± 0.50) and "during the simulation I had the feeling to be fully immersed" (4.29 ± 0.69). Most items of the measure "intuitive use" were answered around the average value, with moderate agreement that the use of

Table 3 Comparison Between AU and OBS for the Item "Technological Proficiency" Before the VR-Module

Technological Proficiency	AU + OBS Mean ± SD	AU Mean ± SD	OBS Mean ± SD	p-value
1. I can solve a technological problem that I encounter independently	3.94 ± 1.34	4.08 ± 1.28	3.88 ± 1.37	0.53
2. Technical devices are often opaque and difficult to master*	2.76 ± 1.32	2.54 ± 1.44	2.85 ± 1.27	0.21
3. I enjoy dealing with technical problems and solving them	2.95 ± 1.35	2.92 ± 1.28	2.97 ± 1.39	0.90
4. Even when encountering resistance, I continue to work on a technical problem	3.71 ± 1.37	3.96 ± 1.27	3.60 ± 1.40	0.25

Note: *inverted item.

 https://doi.org/10.2147/AMEP.S507357

Table 4 AU Results for Presence, Intuitive Use and Perceived Stress.

Topic / Item	Mean ± SD
Presence	
I had the feeling of being in the room with the patient	4.42 ± 0.50
During the simulation I had the feeling to be fully immersed	4.29 ± 0.69
Intuitive use	
The operation of the VR-simulation felt familiar to me	2.83 ± 1.24
No difficulties were encountered while using the VR-simulation	2.71 ± 1.08
Using the VR-simulation was straightforward	3.50 ± 0.98
I succeeded in achieving my goals as I had envisioned	3.54 ± 0.88
Using the VR-simulation was effortless from the beginning	3.25 ± 1.11
I had a clear understanding of what steps to take when operating the VR-simulation	3.17 ± 1.17
The use of the VR-simulation went smoothly	3.13 ± 1.23
I enjoyed the VR-simulation	4.79 ± 0.51
Perceived stress	
I felt capable of meeting the demands	3.33 ± 1.05
I knew enough to handle the case	3.34 ± 1.14
It was easy for me to formulate a diagnosis	3.75 ± 1.15
I found it easy to formulate a course of action from the findings	3.33 ± 0.76
I had to manage many tasks simultaneously*	4.00 ± 0.98
I was concerned about not handling the emergency appropriately*	3.50 ± 1.38
I felt tense while handling the emergency*	3.52 ± 1.20
Completing the VR-simulation in the presence of observing students put me under pressure*	2.48 ± 1.44
I was stressed because I could not control many things*	2.75 ± 1.15

Note: *inverted item.

the VR simulation was "straightforward" (3.50 ± 0.98). In contrast, there was very strong confirmation of the statement "I enjoyed the VR simulation" (4.79 ± 0.51). Regarding the items of perceived stress, the students had mostly issues with managing several tasks at once (4.00 ± 0.98) and some were concerned not being able to address the emergency appropriately (3.50 ± 1.38). The feeling of being observed by fellow students exerted only little pressure on participants (2.48 ± 1.44). Five out of 24 AU reported symptoms of motion sickness such as nausea, headache and slight dizziness.

Comparison Between AU and OBS Regarding Grade of Immersion, Subjective Learning Success and Perceived Competence

Grade of Immersion

While the AU were nearly "completely immersed" in the VR simulation (4.58 ± 0.83), the OBS reported only average agreement for this question (3.02 ± 0.93, $p<0.001$). Interaction with the patient was rated as less immersive compared to other items by the AU, but still above average (3.58 ± 1.06). All results are listed in Table 5. In general, AU rated the degree of immersion significantly higher (AU: 4.05 ± 0.72, OBS: 3.28 ± 0.70, $p<0.001$).

Subjective Learning Success

The outcomes of the subjective learning success after completion of the VR training for both AU and OBS are displayed in Table 5. While learning success across all items was rated high by both groups, the AU's ratings for some items were significantly higher. For instance, the VR simulation was seen as an effective learning tool (AU: 4.79 ± 0.42, OBS: 4.20 ± 0.78, p<0.001). Items related to individual learning benefit were consistently rated more reserved but still above average (eg, "I now have more confidence in being able to respond appropriately in emergency situations", AU: 3.50 ± 0.91, OBS: 3.22 ± 0.90, p=0.26). On average, subjective learning success was also rated higher by AU compared to OBS (AU: 4.15 ± 0.55, OBS: 3.69 ± 0.72, p = 0.005).

Perceived Competence Before and After the VR-Module

All students rated their own subjective perceived competence in dealing with emergency situations, identifying physiological relationships, coming to a diagnosis and initiating a subsequent therapy both before and after the VR-module as seen in Table 6. Three of the seven items were graded significantly better after the VR-module as depicted in Figure 4. Students felt well-prepared (before: 2.76 ± 0.80, after: 3.00 ± 0.84, p=0.001), deducing a diagnosis (before: 3.21 ± 0.73, after: 3.41 ± 0.74, p=0.005) and initiating an appropriate treatment (before: 2.51 ± 0.82, after: 2.89 ± 0.94, p<0.001). The items "I am confident to respond appropriately to emergency situations", "I feel confident in managing medical emergencies" and "I am able to understand the underlying physiology in emergency situations" did not statistically differ before/after the VR-module while the item "I feel confident in understanding connections and deriving consequences" was graded worse (before: 3.80 ± 0.76, after: 3.60 ± 0.73, p=0.023).

Table 5 Results of the Items of "Grade of Immersion" and "Subjective Learning Success" After the VR-Module for All Students (AU +OBS) and Separated in AU and OBS

	AU+OBS Mean ± SD	AU Mean ± SD	OBS Mean ± SD	p-value
Grade of immersion				
The VR-simulation created a realistic learning environment	3.93 ± 0.76	4.50 ± 0.66	3.70 ± 0.67	**<0.001**
I was completely immersed in the VR-simulation	3.48 ± 1.15	4.58 ± 0.83	3.02 ± 0.93	**<0.001**
I found the interaction with the patient realistic	3.24 ± 0.99	3.58 ± 1.06	3.10 ± 0.93	**0.03**
I feel as I am in a real emergency situation	3.35 ± 1.06	3.54 ± 1.28	3.27 ± 0.95	0.18
Total average	3.50 ± 0.78	4.05 ± 0.72	3.28 ± 0.70	**<0.001**
Subjective learning success				
The VR-simulation is a valuable tool for acquiring clinical competence in emergency situations	4.21 ± 0.75	4.50 ± 0.78	4.10 ± 0.71	**0.008**
I consider the VR-simulation to be an effective learning tool	4.37 ± 0.74	4.79 ± 0.42	4.20 ± 0.78	**<0.001**
I now have more confidence in being able to respond appropriately in emergency situations	3.30 ± 0.91	3.50 ± 0.91	3.22 ± 0.90	0.26
I am now better equipped to prioritise tasks effectively in emergency situations	3.37 ± 0.90	3.65 ± 0.94	3.26 ± 0.87	0.14
Practicing emergency situations in the VR-simulation will assist me in my future career as a physician	3.77 ± 0.97	4.22 ± 0.74	3.59 ± 1.01	**0.009**
Total average	3.82 ± 0.71	4.15 ± 0.55	3.69 ± 0.72	**0.005**

Notes: statistically significant p-values (<0.05) are shown in bold.

Table 6 Results of the Item "Perceived Competence" Before and After the VR-Module

Perceived Competence	Before/After	AU + OBS		AU		OBS		
		Mean ± SD	p-value bef./after	Mean ± SD	p-value bef./after	Mean ± SD	p-value bef./after	p-value AU/OBS
1. I feel confident in understanding connections and deriving consequences	Before	3.80 ± 0.76	**0.03**	3.92 ± 0.78	**0.048**	3.75 ± 0.76	0.30	0.54
	After	3.60 ± 0.73		3.54 ± 0.78		3.62 ± 0.72		0.64
2. I am confident to respond appropriately to emergency situations	Before	3.29 ± 0.89	0.75	3.46 ± 1.00	0.39	3.22 ± 0.83	0.95	0.37
	After	3.27 ± 0.81		3.29 ± 1.08		3.25 ± 0.69		0.86
3. I feel confident in deducting a diagnosis from symptoms	Before	3.21 ± 0.73	**0.005**	3.21 ± 0.66	0.99	3.22 ± 0.76	**0.005**	0.83
	After	3.41 ± 0.74		3.25 ± 0.85		3.47 ± 0.70		0.24
4. I feel well-prepared to treat medical emergencies	Before	2.76 ± 0.80	**0.001**	2.96 ± 0.91	0.99	2.68 ± 0.75	**0.005**	0.30
	After	3.00 ± 0.84		3.00 ± 1.06		3.00 ± 0.74		0.94
5. I am able to understand the underlying physiology in emergency situations	Before	2.96 ± 0.68	0.07	2.92 ± 0.72	0.34	2.98 ± 0.66	0.17	0.75
	After	3.13 ± 0.71		3.08 ± 0.72		3.15 ± 0.72		0.80
6. I feel confident in initiating an appropriate treatment algorithm	Before	2.51 ± 0.82	**<0.001**	2.46 ± 0.88	0.08	2.53 ± 0.80	**0.002**	0.87
	After	2.89 ± 0.94		2.83 ± 1.17		2.92 ± 0.84		0.58
7. I feel confident in managing medical emergencies	Before	2.76 ± 0.83	0.47	2.92 ± 1.02	0.61	2.70 ± 0.74	0.19	0.30
	After	2.83 ± 0.88		2.79 ± 1.14		2.85 ± 0.76		0.54
Total average	Before	3.04 ± 0.56	0.09	3.12 ± 0.66	0.94	3.01 ± 0.52	**0.04**	0.71
	After	3.16 ± 0.66		3.11 ± 0.84		3.18 ± 0.57		0.45

Notes: statistically significant p-values (<0.05) are shown in bold.

Correlation Analysis

To identify factors influencing the perceived competence gain (Δpost – pre) among the AU, correlation analyses were conducted. Other measures considered in these analyses were the results of the questionnaires on technological proficiency, presence, intuitive use and subjective learning success.

The perceived competence gain correlated moderately and significantly ($r = 0.46$, $p = 0.03$) with the subjective learning success. A slight, yet non-significant correlation was found between the perceived competence gain and the intuitive use ($r = 0.33$, $p = 0.11$). The connection between intuitive use and subjective learning success was more pronounced ($r = 0.51$, $p = 0.01$).

Regarding the interplay of user-specific factors with hardware-/software specific factors, there were no correlations with the technological proficiency of participants or the experience of presence in the VR-based scenario. However, participants' technological proficiency correlated moderately to strongly with Intuitive Use ($r = 0.62$, $p = 0.001$). Lastly, prior experience with computers or VR technology did not correlate with any of the learning outcomes, intuitive use, or immersive experience. Only between prior experience with computers and technological proficiency, a moderate, significant correlation was found ($r = 0.44$, $p < 0.001$).

 https://doi.org/10.2147/AMEP.S507357

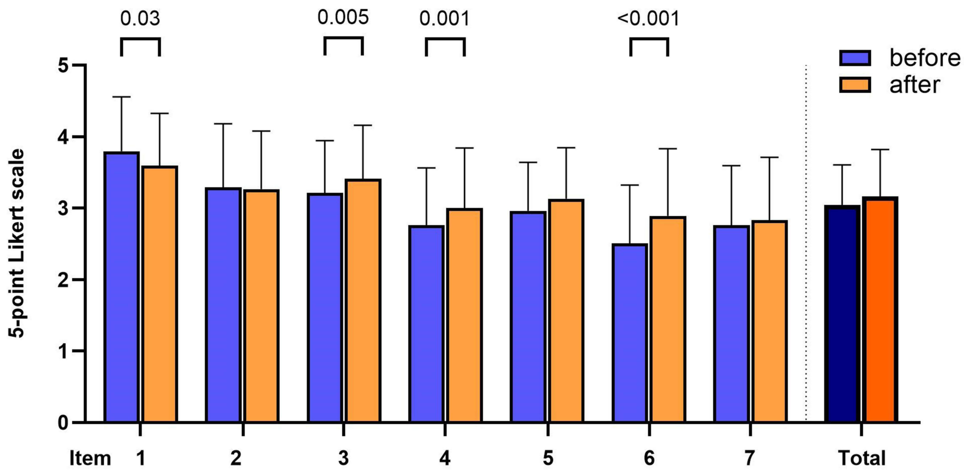

Figure 4 Comparing the results for perceived competence before and after the VR-module.

Discussion

In this study, we integrated a VR-based training program for complex interdisciplinary emergencies into the surgical curriculum and compared the usability and subjective learning experiences between fully immersed AUs and partially immersed OBS. Another focus was on examining the impact of technological proficiency and usability measures on the learning experience.

Regarding our primary objective of assessing various usability aspects, participants rated the intuitiveness of use as slightly above average. Comparable results are reported by other studies: Users of a VR-based training for paramedics reported a Usability Score (SUS) of 65/100.[24] Similarly, participants in a VR-based emergency training for medical students (conducted in a non-native language) also reported average values for intuitive use.[18] However, subjective experiences with the intuitive use of VR training are currently being shaped by the novelty of VR applications as most participants are using such technology for the first time. Considering this, it is difficult to determine whether such values reflect the actual usability of the programs or rather the fact that users are generally not familiar with the technology. In our specific population, a particularly low prior experience with 3D or VR applications, compared to other studies,[17,18] may have additionally influenced the results. The relatively high correlation of intuitive use with subjective learning outcomes in our study, however indicates, that VR learning applications should generally undergo rigorous optimization of accessibility and usability to maximize the potential for an optimal learning experience.

Concurrently, our participants assessed the level of immersion and presence as high to very high. Similar to a previous study with a comparable concept, but focusing on an internal medicine perspective, the substantial difference in immersion experience between AU and OBS could be confirmed.[18] A higher degree of realism in the case of fully immersed VR training can enhance satisfaction among learners[30] and also objective learning outcomes.[10] In our study, however, a stronger feeling of presence among AU was not correlated with learning outcomes. Given the high degree of reported presence in our study, this could imply that a ceiling effect has been reached, beyond which greater immersion no longer has a subjective impact on learning success.

As a second objective, this study aimed to analyse two subjective learning outcomes: The first one, subjective learning success, was rated by the participants as substantially above average, with AU giving higher ratings than OBS. Retrospective subjective reports of learning success are often used to assess (simulated) learning experiences. However, the items of such reports frequently capture aspects of the individual learning experience together with a general assessment of VR-based learning.[17,18,31] Moreover, the retrospective nature of the assessment also raises doubts about its validity.

Therefore, as a second learning outcome, a measure of perceived competence before and after the VR module was introduced. A similar approach was already successfully taken for analogous simulation training.[32] Interestingly, the increases in perceived competence were more heterogeneous than would have been expected from the high subjective learning success. While specific competency facets such as diagnostic reasoning and initiation of a treatment algorithm demonstrated significant improvements, others remained unchanged or even worsened (such as understanding of connections and deriving of consequences). Mühling et al already suggested that the potential students generally see in the method may not necessarily translate into personal competence improvement within relatively short training duration.[18] Furthermore, participants may have been confronted with their own professional deficiencies due to the complexity of the cases, leading them to worse (but more realistic) self-assessments.

Of note, all significant differences between groups were exploratory assessed without controlling for type-1 error caused by multiple testing.[33] Given the moderate sample size of AU in combination with predominantly highly significant results, findings are considered likely to be valid.[34] Moreover, it must be emphasized that also the individual competency gain, collected in a pre/post design is still a subjective measure. While some authors suggest a poor correlation between subjective and objective knowledge outcomes,[35] others found that high self-reported preparedness translates into good performance[36] Although these considerations highlight the need for objective, randomized-controlled data, there is currently no evidence on learning outcomes of VR-based complex emergency medical trainings.[9,37] Admittedly, a randomized-controlled setup for examining objective endpoints (skills or knowledge), eg within the context of an objective structured clinical examination (OSCE) is significantly more elaborate and may initially lead to technical issues. Therefore, as part of the implementation, we opted for a questionnaire-based study design as a first step. However, a corresponding approach should be undertaken to objectively investigate the translatability of the results into the clinical-practical context.

Lastly, this study addressed the relationships between technological proficiency, usability, and learning experience to identify potential barriers to the widespread implementation of VR technology. We found no significant correlation between technological proficiency and learning outcomes, suggesting that all students benefited from VR training regardless of technical skills. The minimization of the operational complexity of the software STEP-VR is likely one of the factors responsible for this. All interactions occur with a single controller button (plus an additional one for occasional teleports) and the interaction elements within the software have been continuously refined to be more intuitive since 2018. However, the homogeneous and low prior experience in PC and VR usage of the sample must also be mentioned as a possible (additional) factor. Perhaps a correlation would emerge more clearly with a more varied distribution of prior experience.

While literature reports some correlations between acceptance, learning success, and computer affinity or prior 3D/VR experience,[18] many studies overlook the link between these parameters and the actual (self-assessed) technical abilities.[15,17] In the correlation between the attribute "technological proficiency" and the degree of intuitive use, we see support for the validity of this construct. In contrast to that, demographic factors on previous use of console and VR gaming exhibited no influence on the degree of intuitive use (group comparisons, data not shown). This might suggest that the measure of technological proficiency is more suitable for examining the usability of a new VR program based on the individual prerequisites of the users.

While subjective learning success correlated significantly with intuitive use, this was only marginally the case for perceived competence gain. We interpreted this to mean that the retrospective assessment of subjective learning success is more strongly influenced by the positive usability experience. Therefore, determining perceived competence gain could be a more independent and hence more meaningful parameter.

Strengths

The study documents the implementation of a VR training program in the surgical curriculum. While most studies in a surgical context analyse VR technology used to train practical skills, our study included the development of clinical competencies in managing interdisciplinary emergencies. The relatively large sample size of this mandatory course also included students with critical attitudes towards new technologies. In addition to a comprehensive set of user-related and technical parameters, learning outcomes were measured retrospectively as well as in a pre/post design. This approach provided further insights into the complex interplay of individual-specific, technical factors and learning experience in immersive settings.

 https://doi.org/10.2147/AMEP.S507357

Limitations

The study was conducted at a single medical school in Germany and restricted to the final academic year, limiting generalizability of the results. As only students that successfully completed the inverted classroom were able to participate as AU, a selection bias cannot be ruled out. This discrepancy in interaction may have influenced their perceptions of the effectiveness of VR. Future studies should consider evaluating learning outcomes separately for each group to assess the differential impact of active and passive participation. Moreover, the study relied solely on subjective responses from participants, and the effects of the training were not compared with a control group. Future research ideally should assess objective performance (eg in the framework of clinical-practical examinations), resulting from VR training and control conditions.

Conclusion

The curricular implementation of a VR training program for interdisciplinary emergencies during the later stages of surgical education is organizationally and technically feasible. Student acceptance and enthusiasm are high, and most competency facets can be positively influenced. Further research such as randomised-controlled-trials comparing OSCE results for VR and non-VR sessions are necessary to evaluate the objective learning outcomes.

Acknowledgments

We thank Joy Backhaus (University of Würzburg) for the statistical consultation.

Disclosure

T.M. was involved in the development of STEP-VR. S.Q. reports grants from Stiftung Innovation in der Hochschullehre, during the conduct of the study. All other authors declare that they have no conflicts of interest in this work.

References

1. Schilling DR. Knowledge doubling every 12 months, soon to be every 12 hours. 2013. Available from: http://www.industrytap.com/knowledge-doubling-every-12-months-soon-to-be-every-12-hours/3950. Accessed February 10, 2024.
2. Mayer RE. Rote versus meaningful learning. *Theory Into Pract*. 2002;41(4):226–232. doi:10.1207/s15430421tip4104_4
3. Sayma M, Williams HR. A new method for teaching physical examination to junior medical students. *Adv Med Educ Pract*. 2016;7:91–97. doi:10.2147/AMEP.S100509
4. Monrouxe LV, Grundy L, Mann M, et al. How prepared are UK medical graduates for practice? A rapid review of the literature 2009–2014. *BMJ Open*. 2017;7(1):e013656. doi:10.1136/bmjopen-2016-013656
5. Bugaj TJ, Nikendei C, Groener JB, et al. Ready to run the wards? - A descriptive follow-up study assessing future doctors' clinical skills. *BMC Med Educ*. 2018;18(1):257. doi:10.1186/s12909-018-1370-4
6. Burridge S, Shanmugalingam T, Nawrozzadeh F, Leedham-Green K, Sharif A. A qualitative analysis of junior doctors' journeys to preparedness in acute care. *BMC Med Educ*. 2020;20(1):12. doi:10.1186/s12909-020-1929-8
7. Bailenson J. *Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do*. W. W. Norton & Company; 2018.
8. Schuir J, Behne A, Teuteberg F. Chancen und Herausforderungen von Virtual Reality in der Aus- und Weiterbildung im Gesundheitswesen. In: *Informatik 2019*. Bonn; 2019: S.676f.
9. Kim HY, Kim EY. Effects of medical education program using virtual reality: a systematic review and meta-analysis. *Int J Environ Res Public Health*. 2023;20(5):3895. doi:10.3390/ijerph20053895
10. Gutiérrez F, Pierce J, Vergara VM, et al. The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. *Stud Health Technol Inform*. 2007;125:155–160.
11. Farra SL, Gneuhs M, Hodgson E, et al. Comparative cost of virtual reality training and live exercises for training hospital workers for evacuation. *Comput Inform Nurs*. 2019;37(9):446–454. doi:10.1097/CIN.0000000000000540
12. Pottle J. Virtual reality and the transformation of medical education. *Future Healthc J*. 2019;6(3):181–185. doi:10.7861/fhj.2019-0036
13. Alaker M, Wynn GR, Arulampalam T. Virtual reality training in laparoscopic surgery: a systematic review & meta-analysis. *Int J Surg*. 2016;29:85–94. doi:10.1016/j.ijsu.2016.03.034
14. Bernardo A. Virtual reality and simulation in neurosurgical training. *World Neurosurg*. 2017;106:1015–1029. doi:10.1016/j.wneu.2017.06.140
15. Bric JD, Lumbard DC, Frelich MJ, Gould JC. Current state of virtual reality simulation in robotic surgery training: a review. *Surg Endosc*. 2016;30(6):2169–2178. doi:10.1007/s00464-015-4517-y
16. Junga A, Kockwelp P, Valkov D, et al. Teach the unteachable with a virtual reality (VR) brain death scenario - 800 students and 3 years of experience. *Perspect Med Educ*. 2025;14(1):44–54. doi:10.5334/pme.1427
17. Mahling M, Wunderlich R, Steiner D, Gorgati E, Festl-Wietek T, Herrmann-Werner A. Virtual reality for emergency medicine training in medical school: prospective, large-cohort implementation study. *J Med Internet Res*. 2023;25:e43649. doi:10.2196/43649
18. Mühling T, Späth I, Backhaus J, et al. Virtual reality in medical emergencies training: benefits, perceived stress, and learning success. *Multimedia Syst*. 2023;29(4):2239–2252. doi:10.1007/s00530-023-01102-0

19. Halbig A, Babu SK, Gatter S, Latoschik ME, Brukamp K, von Mammen S. Opportunities and challenges of virtual reality in healthcare – a domain experts inquiry. *Front Virtual Real*. 2022;3:837616. doi:10.3389/frvir.2022.837616
20. Lie SS, Helle N, Sletteland NV, Vikman MD, Bonsaksen T. Implementation of virtual reality in health professions education: scoping review. *JMIR Med Educ*. 2023;9:e41589. doi:10.2196/41589
21. Ross JG. Simulation and psychomotor skill acquisition: a review of the literature. *Clin Simulation Nursing*. 2012;8(9):e429–e435. ISSN 1876-1399. doi:10.1016/j.ecns.2011.04.004
22. Götz M, Anders M, Biecker E, et al. S2k-leitlinie gastrointestinale blutung [S2k Guideline Gastrointestinal Bleeding - Guideline of the German Society of Gastroenterology DGVS]. *Z Gastroenterol*. 2017;55(9):883–936. German. doi:10.1055/s-0043-116856
23. Beyer G, Hoffmeister A, Michl P, et al. S3-Leitlinie Pankreatitis – leitlinie der Deutschen Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten (DGVS) – September 2021 – AWMF Registernummer 021-003. *Z Gastroenterol*. 2022;60(3):419–521. German. doi:10.1055/a-1735-3864
24. Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med*. 2021;47(11):1181–1247. doi:10.1007/s00134-021-06506-y
25. Skagerlund K, Forsblad M, Tinghög G, et al. Decision-making competence and cognitive abilities: which abilities matter? *Behav Decis Mak*. 2022;35(1). doi:10.1002/bdm.2242
26. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *PARE*. 2005;10:7. doi:10.7275/jyj1-4868
27. Klemmt C, Backhaus J, Jeske D, et al. Psychometric properties and calibration of the SPOREEM (students' perception of the operating room educational environment measure). *J Surg Educ*. 2021;78(4):1151–1163. doi:10.1016/j.jsurg.2020.10.015
28. Benesty J, Chen J, Huang Y, et al. Pearson correlation coefficient. In: Cohen I, Huang Y, Chen J, et al. editors. *Noise Reduction in Speech Processing*. Vol. 2, Berlin, Heidelberg: Springer; 2009:1–4.
29. Ezekiel M. A method of handling curvilinear correlation for any number of variables. *J Am Stat Assoc*. 1924;19(148):431–453. doi:10.2307/2281561
30. Mütterlein J. The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. In: Proceedings of the 51st Hawaii International Conference on System Sciences: Hawaii International Conference on System Sciences; 2018. doi: 10.24251/HICSS.2018.174.
31. Lerner D, Mohr S, Schild J, Göring M, Luiz T. An immersive multi-user virtual reality for emergency simulation training: usability study. *JMIR Serious Games*. 2020;8(3):e18822. PMID: 32735548. doi:10.2196/18822
32. Stroben F, Schröder T, Dannenberg KA, Thomas A, Exadaktylos A, Hautz WE. A simulated night shift in the emergency room increases students' self-efficacy independent of role taking over during simulation. *BMC Med Educ*. 2016;16(1):177. doi:10.1186/s12909-016-0699-9
33. Rice TK, Schork NJ, Rao DC. Methods for handling multiple testing. *Adv Genet*. 2008;60:293–308. doi:10.1016/S0065-2660(07)00412-9
34. Royall RM. The effect of sample size on the meaning of significance tests. *Am Stat*. 1986;40(4):313–315. doi:10.1080/00031305.1986.10475424
35. Sitzmann T, Ely K, Brown KG, Bauer KN. Self-assessment of knowledge: a cognitive learning or affective measure. *Acad Manag*. 2017;9:169–191. doi:10.5465/amle.9.2.zqr169
36. Bloch SA, Bloch AJ. Simulation training based on observation with minimal participation improves paediatric emergency medicine knowledge, skills and confidence. *Emerg Med J*. 2015;32(3):195–202. doi:10.1136/emermed-2013-202995
37. Liu JYW, Yin YH, Kor PPK, et al. The effects of immersive virtual reality applications on enhancing the learning outcomes of undergraduate health care students: systematic review with meta-synthesis. *J Med Internet Res*. 2023;25:e39989. doi:10.2196/39989