REGULAR PAPER

Virtual reality in medical emergencies training: benefits, perceived stress, and learning success

Tobias Mühling[1,2] · Isabelle Späth[1] · Joy Backhaus[1] · Nathalie Milke[3] · Sebastian Oberdörfer[3] · Alexander Meining[2] · Marc Erich Latoschik[3] · Sarah König[1]

Abstract

Medical graduates lack procedural skills experience required to manage emergencies. Recent advances in virtual reality (VR) technology enable the creation of highly immersive learning environments representing easy-to-use and affordable solutions for training with simulation. However, the feasibility in compulsory teaching, possible side effects of immersion, perceived stress, and didactic benefits have to be investigated systematically. VR-based training sessions using head-mounted displays alongside a real-time dynamic physiology system were held by student assistants for small groups followed by debriefing with a tutor. In the pilot study, 36 students rated simulation sickness. In the main study, 97 students completed a virtual scenario as active participants (AP) and 130 students as observers (OBS) from the first-person perspective on a monitor. Participants completed questionnaires for evaluation purposes and exploratory factor analysis was performed on the items. The extent of simulation sickness remained low to acceptable among participants of the pilot study. In the main study, students valued the realistic environment and guided practical exercise. AP perceived the degree of immersion as well as the estimated learning success to be greater than OBS and proved to be more motivated post training. With respect to AP, the factor "sense of control" revealed a typical inverse U-shaped relationship to the scales "didactic value" and "individual learning benefit". Summing up, curricular implementation of highly immersive VR-based training of emergencies proved feasible and found a high degree of acceptance among medical students. This study also provides insights into how different conceptions of perceived stress distinctively moderate subjective learning success.

Keywords Dynamic physiology · Emergency medicine · Perceived stress · Simulation sickness · Simulation training · Virtual reality

1 Introduction

Medical emergencies require executive functions and the ability to prioritize tasks. Although emergency medicine is integrated into most curricula of medical degree programs, preparedness among graduates to deal with acute injury or illness is scarce. Recent research has identified the deficits in the early management of medical emergencies, including clinical reasoning/diagnosis, prescribing the appropriate medication, multidisciplinary teamwork, and handover of patients [1–3]. A lack of preparedness was also self-reported by students as well as graduates, and above all observed by clinical educators [2, 4]. As a matter of fact, junior doctors frequently describe feelings of being overstretched when dealing with acutely unwell or critical patients at high risk [5]. Others report that responsibility for severely ill patients

Tobias Mühling and Isabelle Späth have equally contributed to this work.

✉ Tobias Mühling
  muehling_t@ukw.de

1 Institute of Medical Teaching and Medical Education Research, University Hospital Würzburg, Joseph-Schneider-Strasse 2, 97080 Würzburg, Germany

2 Department of Internal Medicine II, University Hospital Würzburg, Würzburg, Germany

3 Human-Computer Interaction, University of Würzburg, Würzburg, Germany

during night shifts is one of the major sources of excessive stress and anxiety during the first year of residency [6].

Educational interventions need to address this lack of competency in the specialty of emergency medicine. To date, different simulation training settings, ranging from simulated patients to full-scale and technology-enhanced analogous trauma rooms have been implemented [7, 8]. Although the outcomes of these physical training environments have been generally positive, its wide-ranging practice was limited owing to the cost in terms of time and money, as well as the availability of room facilities and trained personnel. The development of computer-based simulation has added a new dimension to education and proved to enable the operational learning, e.g., of cardiopulmonary resuscitation [9, 10] and even more complex tasks, such as the management of multiple patients in an emergency room [11, 12]. Unlike the above-mentioned physical simulation facilities, computer-based simulation can be provided without elaborate infrastructure and made available to individuals. Furthermore, users can determine their own pace of learning progress and outcome. Moreover, modern soft and hardware systems are physically capable of calculating multiple physiological parameters, which would provide for a dynamic simulation of diseases covering a wide range of underlying conditions in real time. Accordingly, clinical and laboratory parameters as well as the effects of actions taken by the user (e.g., drug administration, diagnostic and therapeutic interventions) may be programmed, visualized, and direct feedback provided by future systems.

Most recently, virtual reality (VR) technologies have been developed to deliver an immersive, interactive experience that can be conducted privately and inexpensively. Head-mounted displays (HMD) and motion sensors provide the basis for the creation of highly immersive three-dimensional environments. The concept of immersion describes the ability of this technology to create a complete and vivid illusion of reality. While "immersion" refers to representational fidelity from a technical viewpoint, the term "presence" in the context of VR relates to the subjective feeling of being present in the virtual environment [13]. VR-enhanced learning environments with high representational fidelity—and thus a high degree of perceived presence—can increase learner motivation, enhance spatial knowledge representation, and improve contextualization of learning [14–16]. Of note, early VR technology was known to elicit unwanted symptoms (e.g., nausea or disorientation), summed up under the terms "cyber/simulation sickness". These undesirable effects are most likely related to discrepancies between the optical information that the user receives via the HMD and other sensory input such as spatial perception [17]. In the meantime, technical improvements, such as improved HMD resolution and latency, have been implemented to minimize such symptoms.

VR technology without additional devices does still lack the convincing haptic representations of performing surgical procedures and handling tissue for example. Nevertheless, there is substantial evidence supporting VR applications in the training of different surgical and interventional procedures or hygiene measures [18–20]. Although relying less on haptic fidelity, a conclusive concept for internal and emergency medicine has not yet been reported [21].

Here, we implemented and evaluated a novel training concept for medical emergencies that uses up-to-date VR technology alongside a simulated dynamic physiology system. The program "Simulation-based Training of Emergencies for Physicians using Virtual Reality (STEP-VR)" was tested in a pilot phase, during which we evaluated its technical feasibility and the students' tolerance of simulation sickness. Afterward, STEP-VR was implemented as a compulsory training session into the curriculum of a medical degree program.

In order to take advantage of the educational potential offered by VR, we had to develop an understanding of how to maximize the desirable impacts on student learning. We thus aimed to address the following questions as to:

1. Whether the implementation of STEP-VR into the medical curriculum for a semester cohort is feasible and whether students accept the didactic concept;
2. How students perceive VR-based training with respect to simulation sickness and the degree of presence;
3. What level of psychological distress users experience during exposure to the VR content and which factors are associated;
4. How students evaluate the training benefits and whether the active and passive roles adopted during training differ in the estimated learning success.

2 Materials and methods

2.1 VR-based simulation training

STEP-VR was aimed at advanced medical students who find themselves in the role of physicians being confronted with a variety of critically ill patients in an emergency department (Fig. 1A). The training software (version 0.11b) was provided by ThreeDee GmbH® (Munich, Germany), a startup specializing in 3D visualization. STEP-VR is designed to train cognitive abilities [22] including executive functions, situational awareness, time management, and prioritization of tasks. Relevant diagnostic tools include medical history, laboratory tests, or various imaging modalities, and are based on real patient data (Fig. 1B, C). Specific procedures (e.g., placement of an IV cannula) can only be performed in an abstract
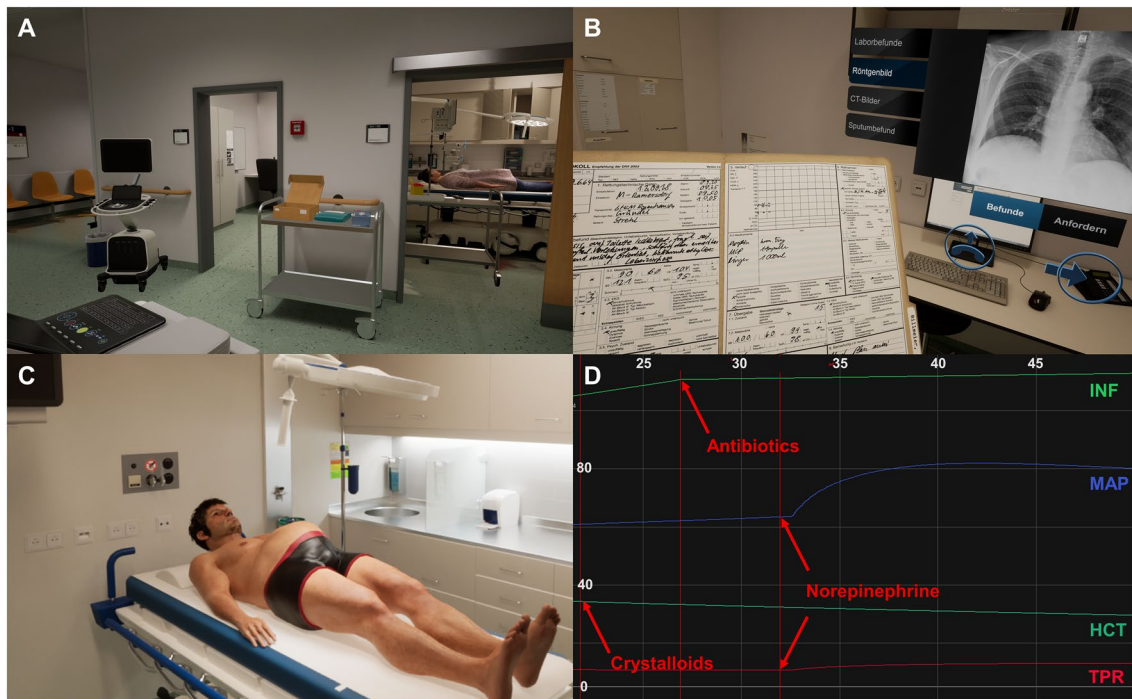
Fig. 1 In-simulation views and parameter monitoring of STEP-VR. The emergency department of a university hospital was transferred to the VR simulation: General interior of the emergency department (**A**). Patient file containing paramedic protocol and computer menu illustrating chest X-ray (**B**). Treatment room with patient (**C**). Parameter monitoring for the scenario sepsis following the administration of i.v. crystalloid fluids, broad-spectrum antibiotics and norepinephrine via syringe pump (**D**). The Y-axis represents the respective units (arbitrary units for infectious load in the bloodstream (INF), mmHg for mean arterial pressure (MAP), % for hematocrit (HCT) and mmHg*min/l for total peripheral resistance (TPR)), the X-axis time indicates the intervals of five minutes

and non-haptic manner of handling. After determining the most likely diagnosis, therapeutic measures must be administered, such as infusions, blood transfusions, application of numerous pharmaceuticals, (non)-invasive ventilation, as well as to arrange endoscopic or cardiologic interventions. The applied measures exert real-time effects on the simulated patients as a pathophysiological continuum. Based on the literature, realistic courses of the underlying conditions are modeled out of > 100 clinical parameters. Accordingly, dose-dependent effects of drugs, transfusions, or interventions provide an immediate response and/or feedback on monitoring parameters (Fig. 1D). Ultimately, a guidelines-based checklist summarizes whether the correct measures have been taken.

The VR hardware setup for this study employed a Schenker XMG Core 15 Laptop (chipset: Intel Core i7-9750H, 6 × 2.6 GHz; graphics adapter: Nvidia GeForce GTX 1650, 4 GB GDDR6 VRAM) and an Oculus Rift S VR HMD. The equipment enabled STEP-VR to run with a constant framerate > 60 frames per second on "high quality" display settings of the HMD.

2.2 Study design and data collection

The study was conducted at a medical school in Germany offering a six-year curriculum. It took place in the fifth year of studies within the framework of short clinical rotations in internal medicine lasting two weeks, including four compulsory afternoon sessions to consolidate clinical reasoning. One of these was STEP-VR, held as small-group training session of eight to ten students followed by debriefing/discussion:

1. VR-based training (120 min) performed in three to four randomly chosen scenarios and guided by a student assistant.
2. Debriefing (60 min) of the emergency cases in the form of a review together with a tutor to summarize and further discuss the issues of problem solving.

In total, five different VR scenarios from the subject of internal medicine were provided, each addressing a specific leading symptom: (1) fever and abdominal pain due

to sepsis, (2) hematemesis due to acute esophageal variceal bleeding, (3) abdominal pain due to acute biliary pancreatitis, (4) chronic cough due to pulmonary tuberculosis, and (5) chest pain due to acute myocardial infarction. A more detailed description of the scenarios and their respective learning objectives is presented in Supplement Table 1. To avoid sequence effects, student groups completed the scenarios in a random order. Each scenario was handled by a different student who volunteered to enter the VR scenario. Owing to the restricted time schedule, a maximum of three to four students out of the group were exposed to the VR environment. They were named the active participants (AP). The other students taking part in the training watched from the first-person perspective on a monitor and were able to assist verbally. They were named the observers (OBS). There was also the specific role of one student to take notes of the key features of the scenario, in order to present them to the tutor in the second part of the session. Students were asked to fill out questionnaires, which were specific to the role of AP (APQ, immediately after the VR exposure) and as OBS (OBSQ, at the end of the session). However, suspecting bias among those OBS who had previous immersive VR experience as AP, these response data were excluded, i.e., cleaned, from the comparative analysis of the measures "presence" and "estimated learning success", as well as the two motivation items (we term these participants "cleaned" OBS).

Figure 2 provides an overview of the study design and data collection. The pilot phase ran from April until September in 2020 with volunteers who attended the VR-based training (approximately 30 min) with the aim of assessing the technical feasibility. Students were asked to fill out pre- and post-intervention questionnaires to rate simulation sickness. The main study was conducted in two subsequent semester cohorts from October 2020 until July 2021. Participants of the study provided written informed consent after being detailed on the study conditions.

2.3 Questionnaires and measures

Questionnaires were developed, of which the measures are listed in Table 1. In the main study, students were asked to enter their matriculation number, in order to match the corresponding data between the questionnaires given to AP and OBS.

2.4 Statistical analysis

Analyses were performed with descriptive statistics, including mean (M), minimum (Min), and maximum (Max), standard deviation (SD), and skewness (Skew) of items. Skewness between − 2 and 2 indicated that data were roughly distributed symmetrically. We evaluated the psychometric properties by conducting a maximum likelihood exploratory factor
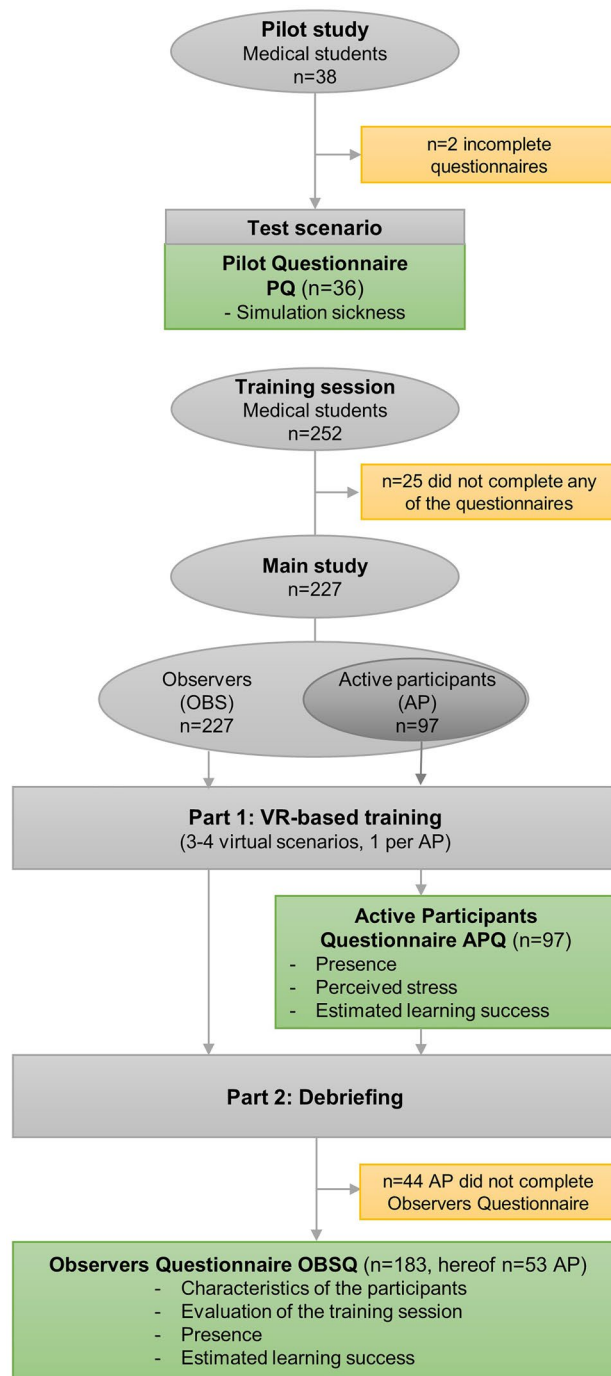

Fig. 2 Schematic design depicting the data collection by questionnaires that were administered to participants. A pilot study on simulation sickness was conducted. For the main study, students fulfilled two roles: all as observers (OBS, $n = 227$) and some as active participants (AP, $n = 97$)

analysis (EFA). The Kaiser criterion was considered for all factors [23]. According to this rule, only the factors that have eigenvalues greater than one are retained for interpretation. Bartlett's test of sphericity and the Kaiser–Meyer–Olkin

Table 1 Overview of the measures addressed in the questionnaires

Sociodemographics and measures	Definition	Adapted questionnaires	Scale values and value labels
Characteristics of participants	Age, gender, career choice (medical/surgical/undecided), number of previous VR experiences, and first-person video games	None	5 items
Simulation sickness	Side effects elicited by use of VR divided into subscales of nausea, disorientation and oculomotor	Simulator sickness questionnaire, ([22], adapted version by [23])	16 items in 3 subsets (Nausea (N), Oculomotor (O), Disorientation(D)), each on a 4-point Likert scale: 1=none, 4=strong. Questionnaires were completed immediately before and after the simulation
Presence	Perception of being physically present in the VR environment	Degree of immersion questionnaire (modified from [24])	5 items on a 5-point Likert scale: 1=strongly disagree, 5=fully agree
Perceived stress	Perception of being stressed and challenged during the management of the virtual emergencies	Perceived stress questionnaire (inspired by [25,26])	15 items on a 5-point Likert scale: 1=strongly disagree, 5=fully agree
Estimated learning success	Self-reported improvement in knowledge and skills	Estimated learning success questionnaire (modified from [24])	9 items on a 5-point Likert scale: 1=strongly disagree, 5=fully agree
Evaluation of the training session	Rating of the didactic quality, difficulty, and motivational effects	Questionnaire comprising items from the practical guidelines for evaluators at faculties and institutes ([27])	17 items, of which 11 items (rating) on a 5-point Likert scale: 1=strongly disagree, 5=fully agree. 2 items (difficulty) on a 5-point Likert scale: 1=far too easy, 5=far too high. 2 items (motivation pre- and post-training) on a 5-point Likert scale: 1=very low, 5=very high

coefficient (KMO coefficient) were used to evaluate whether we could subject data to EFA. The factor solution had to meet the following criteria [24] of a Bartlett's test of sphericity ($p < 0.05$) and a KMO-coefficient > 0.50. Consistency was assessed by employing Cronbach's alpha (α). A value exceeding 0.7 was considered as acceptable, greater than 0.8 as good.

The Welch test was used to compare between groups. Nominally scaled variables were compared with the Chi-square test. The Wilcoxon test was used with the level of significance set to $p < 0.05$ to compare simulation sickness in the pre- and post-intervention questionnaires as two matched samples. Correlation according to Pearson was investigated [25]. Additionally, curvilinear relationships were analyzed [26]. Furthermore, we applied Little's Missing Completely at Random (MCAR) test [27] to check whether values were systematically missing.

2.5 Ethics approval

The local institutional review and ethics board judged the project as not representing medical or epidemiological research on human subjects and as such adopted a simplified assessment protocol. The project was approved without any reservation under the proposal number 20201012-01. Survey data from the questionnaires were retrieved anonymously using the EvaSys® platform (Lüneburg, Germany). A student's decision to participate or not, as well as the results of the questionnaire, had no consequences on the students' academic progress. Data were processed and stored in accordance with the local data protection laws.

3 Results

3.1 Characteristics of the participants and evaluation of the training session

A total of 38 students participated in the pilot study, of whom 36 answered the questionnaire assessing simulation sickness. Two hundred fifty two students were enrolled in the VR-based training sessions, of whom 227 took part in the main study and adopted the role of OBS at some time during the training sessions. The OBSQ was completed by 183 participants, of whom 130 were so-called "cleaned" OBS, meaning that they had not assumed the role of an AP before. The APQ was completed by all 97 AP.

Characteristics of all participants who filled out the OBSQ are listed in Table 2. More than 60% of the participants were aged between 22 and 25 years. Sixty-three students were male (34.5%) and 119 were female (65.0%). The result of Little's MCAR test for the key variables provided

Table 2 Characteristics of the participants in the main study ($n = 183$)

Parameter	Subgroup	n	Percentage (%)
Age	< 22	6	3.3
	22–25	112	61.2
	26–30	39	21.3
	> 30	23	12.6
	No response	3	1.6
Gender	Female	119	65.0
	Male	63	34.5
	Other	0	0.0
	No response	1	0.5
Career choice	Medical	83	45.4
	Surgical	56	30.6
	Undecided	42	23.0
	No response	2	1.0
Experience with first-person video games (FPG)	None	117	63.9
	Rarely	28	15.3
	Sometimes	19	10.4
	Frequently	8	4.4
	Often	8	4.4
	No response	3	1.6
Experience with VR applications	None	119	65.0
	< 1 h	34	18.6
	1–5 h	18	9.8
	5–10 h	4	2.2
	> 10 h	5	2.7
	No response	3	1.6

an χ^2 distance of 1689.6 ($df = 1706$; $p = 0.607$.), indicating that there was no systematic accumulation of missing data.

The training session was evaluated as part of the OBSQ (Supplement Table 2). Students rated the quality of the training session highly and as well suited to conveying the learning content. They assessed the didactic concept of the training session as a good mixture of practical exercise and discussion. They emphasized the good working atmosphere. Moreover, the results highlighted the importance of the helpful comments by the student assistant and additional value of the tutor to clarify uncertainties. The overall difficulty of the training session was perceived as appropriate, as was the density of learning input.

3.2 Simulation sickness

The pilot study served to assess the feasibility of the training session and to rule out disturbing side effects such as simulation sickness. Overall mean values for simulation sickness were significantly higher after the exposure in the subscales of nausea and disorientation, but still fairly low to acceptable after the training session (Table 3).

Table 3 Descriptive data of simulation sickness sub-scores among participants of the pilot study ($n=36$). Items were rated on a four-point Likert scale with a maximum value of 4

	n	Pre		Post		
		Mean	SD	Mean	SD	P
Nausea	36	1.12	0.23	1.19	0.25	< 0.05
Oculomotor	36	1.18	0.25	1.20	0.23	Not significant
Disorientation	36	1.08	0.19	1.17	0.21	< 0.05

Table 4 Comparison between the group of "cleaned" OBS and AP in the measures "presence" and "estimated learning success"

	"Cleaned" OBS ($n=130$)		AP ($n=97$)		
	Mean	SD	Mean	SD	P
Presence	3.26	0.74	3.81	0.98	< 0.001
Estimated learning success	3.72	0.77	3.87	0.75	Not significant
Motivation*					
Pre-training	3.23	0.76	3.26	0.89	Not significant
Post-training	3.61	0.84	3.97	0.88	< 0.01

Depicted are the averages of all items of the respective measure. Motivation was rated by both groups at two points in time as a single item post-training and retrospectively pre-training

*Note that "motivation", part of the OBSQ, was calculated from $n=53$ AP, as $n=44$ OBS from AP were missing

3.3 "Presence", "estimated learning success", and motivation rated by the group of AP and OBS

As the participation of students in the VR pilot study was voluntary, we first compared both groups of AP and OBS with respect to demographic characteristics. Using the Welch test, no differences were found with respect to age, career choice, or experience with VR. Interestingly, both the chi-square test for the nominally distributed gender variable ($p < 0.05$) and the Welch test for the question on the frequency of playing games from the first-person perspective (FPG) ($p < 0.05$) demonstrated that AP were more likely to be male and with previous gaming experience. In accordance, a highly significant correlation was found between male students and the frequency of playing FPG ($r=-0.52$).

Students' ratings of the measures "presence", "estimated learning success", and for the items of motivation are summarized in Table 4. As expected, AP rated the degree of presence significantly higher than "cleaned" OBS ($M=3.81$, SD$=0.98$, $n=97$ and $M=3.26$, SD$=0.74$, $n=130$ respectively, $p < 0.001$). However, the rating of "cleaned" OBS was still above average. Looking into more detail on item level Supplement Table 2), there was large agreement that simulation created a realistic environment, while interaction with the virtual patient was only rated above average, as was the feeling of being in a real emergency. The "estimated learning

success" was rated higher in the group of AP than "cleaned" OBS ($M=3.87$, SD$=0.75$, $n=97$ and $M=3.72$, SD$=0.77$, $n=130$ respectively, $p=0.05$), however with only marginal significance. The usefulness of the learning tool was generally rated highly, and students would strongly embrace more opportunities to train in VR simulations. However, the effect on the improvement in confidence or in the ability to prioritize tasks was only rated just above average. Furthermore, students were asked to rate their motivation before (pre-training) and after the session (post-training). Baseline motivation in both groups was similar. However, a significantly higher motivation post training was reported by AP in comparison to "cleaned" OBS ($M=3.97$, SD$=0.88$, $n=53$ and $M=3.61$, SD$=0.84$, $n=130$ respectively, $p < 0.01$).

3.4 Impact of "perceived stress" on "estimated learning success"

The group of AP ($n=97$) rated the "perceived stress" induced by the VR training. We subjected the respective items to an EFA with maximum likelihood estimation and promax rotation. The KMO-coefficient measure of sampling adequacy was very good (0.83); Bartlett's test of sphericity was significant ($\chi^2=465.91$, $df=36$, $p < 0.001$). A three-factor solution provided the best fit and good internal consistency in accordance with the criteria outlined in the statistical analysis section. We defined the factors (1) "sense of control" (the self-assessed competence to manage the patient cases, with extremes, such as under- and over-estimation being deleterious, number of items$=6$, $\alpha=0.755$), (2) "challenge to perform" (the challenge to make correct decisions based on specific factual and procedural knowledge, number of items$=5$, $\alpha=0.698$) and (3) "social interaction to learn" (the perception of being observed and commented on by others, number of items$=4$, $\alpha=0.669$) (Table 5). Altogether, the three factors explained 52% of the variance. As an indication of discriminant validity, correlation between the three scales was only moderate ($r=0.322$–0.406, $p < 0.05$).

Discriminant validity was analyzed by inspecting group differences. We found no significant differences for gender in relation to all three factors (data not illustrated). A direct comparison of age groups revealed that the younger participants up to 22 years of age experienced "social interaction to learn" to a greater extent than participants older than 29 years ($M=1.75$, SD$=0.0$, $n=6$ and $M=1.28$, SD$=0.28$,

Table 5 Exploratory factor analysis for items of the "perceived stress" measure for the AP ($n = 97$)

	Sense of control	Challenge to perform	Social interaction to learn	Mean	SD	Skew
2. I had to fulfill several tasks simultaneously	**0.55**			3.61	1.07	− 0.58
3. I was afraid of being unable to manage the medical emergency appropriately	**0.56**			3.03	1.39	− 0.03
11. I felt stressed while managing the emergency	**0.72**	− 0.46		2.86	1.23	− 0.06
12. Others watching me during the scenario exerted pressure on me	0.57		**0.48**	2.58	1.42	0.43
13. I was stressed because several aspects were out of my control	**0.74**			2.52	1.11	0.34
14. Initiating the appropriate diagnostics challenged me	**0.53**		− 0.45	3.31	1.03	− 0.48
15. I found it challenging to initiate the appropriate intervention/therapy	**0.73**			3.36	1.22	− 0.51
1. I felt I was able to handle the scenario*	0.37	**0.49**		2.35	1.01	0.42
4. I felt motivated by the clinical decisions required*		**0.58**	0.35	1.96	0.92	0.57
5. My knowledge was sufficient to handle the emergency cases*		**0.67**		2.36	0.97	0.47
6. It was easy for me to find a potential diagnosis*		**0.64**		1.77	0.92	0.97
7. I could easily draw the right conclusions giving the data provided*		**0.85**		2.32	0.88	− 125
8. Guidance by the tutor made me feel safe*			**0.62**	1.59	0.85	1.64
9. The assistance of my fellow students made me feel safe*			**0.68**	1.84	1.02	1.30
10. Altogether, the VR simulation was fun*			**0.78**	1.29	0.63	2.25
Cronbach's α	**0.76**	**0.70**	**0.67**			
Explained variance	**25.47**	**13.93**	**12.68**			
Eigenvalues after rotation	**3.82**	**2.09**	**1.90**			

Bold values indicate the item loadings assigned to respective scale

*Reverse-coded item. Loadings between − 0.3 and 0.3 not displayed

$n = 23$, respectively, $p < 0.01$). Marginally significant differences among two factors were found, depending on the frequency participants indicated that they play FPG. "Sense of control" was rated higher by those who responded that they never play FPG than by those playing occasionally ($M = 3.21$, SD $= 0.63$, $n = 117$ and $M = 2.92$, SD $= 0.67$, $n = 63$ respectively, $p < 0.1$. Items in the factor "challenge to perform" were rated more highly by those who never played games (compared to those playing occasionally ($M = 2.14$, SD $= 0.62$, $n = 117$ and $M = 2.02$ SD $= 0.59$, $n = 63$, respectively, $p < 0.1$). Additionally, the three factors of the "perceived stress" measure were found not to correlate with the five items from the "presence" measure.

The "estimated learning success" measure rated by AP ($n = 97$) was also subjected to EFA. The 2-factor solution was the best-fitting model that also had conceptual coherence as (1) "didactic value" (the general potential attributed to VR simulation by participants, number of items $= 4$, $\alpha = 0.83$) and (2) "individual learning benefit" (the positive effects on theoretical/practical skills for the respective individual, number of items $= 4$, $\alpha = 0.82$) (Table 6). The internal consistency of the two factors was considered good. They explained 71% of the variance ($\lambda_1 = 57\%$; $\lambda_2 = 14\%$). In this case, there was a high degree of correlation between the two factors ($r = 0.71$, $p < 0.01$). According to the quality criteria of factor solutions, one item (number 4) of the "estimated learning success"

measure had to be eliminated. The following statistical analysis was performed with this shorter version as the final inventory.

Finally, we investigated the relationship between the measures of "perceived stress" and "estimated learning success". The factor "sense of control" was associated quadratically with "didactic value" and "individual learning benefit" (Fig. 3). In more detail, both scales increased as the "sense of control" improved from low up to medium levels. However, when reaching higher levels of "sense of control", the relationship reversed and the scale value declined, as described by the Yerkes–Dodson law [28]. Optimal levels of "didactic value" and "individual learning benefit" were experienced when "sense of control" was in the medium range, which we saw as acceptable. The factor "challenge to perform" correlated negatively with "didactic value" ($r = − 0.26$) and "individual learning benefit" ($r = − 0.24$, $p < 0.05$). There was some correlation between "social interaction to learn" and "didactic value" as well as "individual learning benefit", although it was not statistically significant.

4 Discussion

Advances in VR technology enabling the production of high-quality simulation environments that train the management of patient cases are dramatically affecting educational

Table 6 Exploratory factor analysis for items of the "estimated learning success" measure, which was assessed by AP ($n = 97$)

	Didactic value	Individual learning benefit	Mean	SD	skew
5. I would favor more opportunities to train/learn in VR simulations	**0.89**		4.57	0.83	− 2.75
3. The VR simulation is suitable to acquire practical skills for medical emergencies	**0.82**		4.26	0.87	− 1.21
6. The VR simulation is a good learning tool	**0.80**		4.57	0.77	− 2.00
1. The VR simulation helps me to react better in medical emergencies	**0.42**		3.61	1.25	− 0.99
9. The VR simulation is useful for my future work as a medical doctor		**0.43**	3.74	1.29	− 1.16
8. My ability to prioritize actions in medical emergencies has improved		**1.05**	3.19	1.13	− 0.29
7. I have more confidence now to react appropriately in medical emergencies		**0.79**	3.31	1.13	− 0.36
2. The VR simulation helps me understand the physiological background		**0.42**	3.63	1.10	− 0.98
Cronbach's α	**0.83**	**0.82**			
Explained variance	**56.70**	**14.34**			
Eigenvalues after rotation	**4.53**	**1.15**			

Bold values indicate the item loadings assigned to the respective scale

*Reverse-coded item. Loadings between − 0.3 and 0.3 not displayed

concepts, teaching activities, and the outcome of student learning. In the present study, we developed and implemented a novel VR-based training environment for medical emergencies, evaluated its acceptance by students, and explored selected psychological and didactic aspects that were associated with the benefits of the VR method and teaching intervention.

A primary objective of the study was to demonstrate that implementation of a VR-based training session as a compulsory course in the medical degree curriculum is feasible and well accepted by students. With the chosen format, we did not experience any administrative or technical difficulties and thus were able to enroll 252 participants, of whom 227 took part in our study. Compared to other recent work examining the feasibility and outcomes of VR-based training [20, 29], the sample size of our study was considerably large. Investigating sociodemographic parameters, we experienced that the willingness to act as AP was greater among male participants, of whom a higher percentage had prior experience in FPG. This is in line with recent publications demonstrating that twice as many men compared to women play computer games several times a week [30]. Possibly as a consequence, female students, especially those with less computer affinity, exhibit greater skepticism toward VR technology for teaching purposes [31]. For any broad application in future, the reservations of these individuals in particular should therefore be identified and addressed specifically. However, considering the whole teaching approach, evaluation of the VR-based training session revealed a wide acceptance among both male and female participants. Of note, debriefing and the discussion of open questions with a tutor following the exposure to VR were seen as an integral part of the course. This speaks in favor of embedding VR into a sound didactic concept rather than having students experimenting on their own (e.g., in a "VR lab" or at home). Beyond any didactic principles, the guided application of VR-based teaching is mandatory to promote the necessary media competence among medical students [31].

Our second aim was to investigate the technical aspects of the simulation, namely simulation sickness and presence. In our pilot study, VR exposure caused simulation sickness. However, reported values in the three elements nausea, oculomotor, and disorientation did not indicate high levels of discomfort in our sample. Referring to a comprehensive literature review including 55 studies recording simulator sickness scores among > 3000 VR participants [32], values obtained in our study even appeared to be relatively low in comparison. In light of the fact that only few participants had any active experience with VR prior to the study, STEP-VR as a simulation-based training environment can clearly be offered to a wider audience without any concern relating to adverse events such as vegetative reactions. Completion of actions only possible in one ground plane, which was congruent with the actual bodily position of the participant, might be one reason for the good tolerance. Additionally, the restrictive use of content-rich text boxes and a high average frame rate per second attained by the equipment we employed reduced the perception and perhaps effects of simulation sickness. It is worth noting that research in healthy participants and vestibular patients revealed a decline in motion sickness susceptibility with increasing age [33]. In our study, the majority of participants were 22–25 years old and therefore fairly young.

We observed a high degree of presence among AP, who perceived the VR simulation as very realistic. This was confirmed by OBS, but to a significantly lower degree. However,
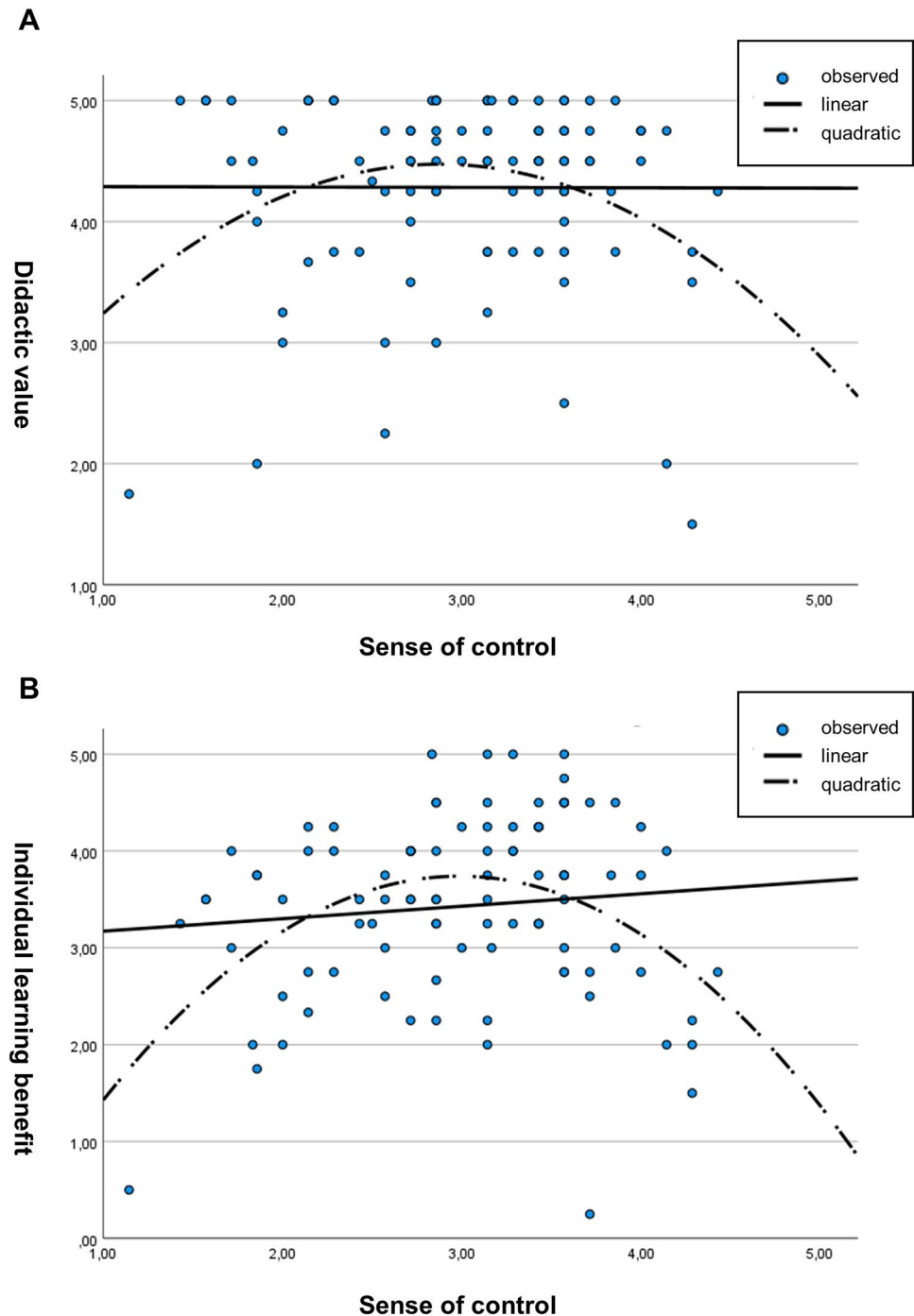
when exploring the interaction with patients on item level, authenticity was only rated slightly above average by AP and average by OBS. One explanation for this finding may be attributed to the fact that the virtual patients were portrayed in a rather motionless manner and communication could be performed only via a dialog menu. Owing to the focus lying on training procedural skills and acceptance of the lack of haptic handling in the VR environment, many procedures were only depicted in a simplified manner, e.g., the insertion of an IV cannula as "drag & drop" to the patient's arm. Therefore, future development of the software should aim to include more realistic patient modeling and characteristics including responsive and life-like gestures. Another possible way to enhance authenticity could be to incorporate physical objects into the virtual environment, providing a more realistic and haptic experience for students. This approach, described by the terms augmented reality (AR) or mixed reality (MR), can also facilitate interaction with tutors and

student observers, as the objects are visible during the simulation. However, the downside of using AR/MR technology is that the required HMD is more expensive, and additional equipment (e.g., phantoms) is needed, which can hinder its widespread implementation.

More research is also needed to clarify how real-time communication affects the perception of high-level patient interaction in VR. However, this requires the integration of sophisticated applications to enable speech recognition. Recent developments in natural language processing in virtual standardized patients only demonstrated a response accuracy of 74–92% in dialogs [34, 35]. Especially in the management of emergencies, even higher levels of precision might prove mandatory, in order to prevent the misconception of information under time pressure.

Third, an aim was to evaluate the extent and quality of psychological distress among the participants of the VR simulation. On average, AP reported being moderately challenged during the management of virtual patients. More specifically, EFA identified three distinct factors corresponding to subtypes of perceived stress. The first two, "sense of control" and "challenge to perform" represent more classic concepts of stress and have been reported as appearing in various contexts of medical education [36, 37]. Both parameters correlated negatively with the use of FPG, which is in line with enhanced spatial navigation and visuomotor coordination in this subgroup [38, 39], presumably facilitating the use of VR-based simulations and decreasing perceived stress. A third identified factor of perceived stress entitled "social interaction to learn" is based on the assumption that the possibility of publicly unmasking knowledge deficits, as well as the feeling of being observed, also exert discomfort on AP. This might be a particular concern, as it is not possible to notice the audience around when wearing an HMD. Interestingly, by comparing age groups, we revealed that young participants experienced this factor significantly to a greater extent than their older counterparts. This suggests that age might have a moderating/mediating effect on the social experience in virtual reality, a correlation that has not yet been investigated in more detail. Future studies are needed to elucidate the role of (social) stress in VR-based simulations by comparing different settings (with and without observers).

Finally, we were also interested in the perceived learning benefits and the relation to putative associated factors. Estimated learning success was certainly high for both AP and OBS at the end of the training session. In our comparative analysis, AP rated learning benefit significantly higher. This might be attributable to the fact that a strong feeling of presence has been shown to enhance skill acquisition and factual memory [29, 40, 41]. Higher motivation scores among the AP compared to OBS, possibly owing to higher vigilance and increased activity while training,

might serve as another explanation for the better learning success for AP. The promotion of motivation and self-efficacy are key concepts of simulation-based learning and have already been reproduced in the setting of VR-based approaches [14, 42].

Performing EFA on the "estimated learning success" measure revealed that items contributing to the "didactic value" of the used VR simulation were rated very highly among AP, while the items comprising translation into an "individual learning benefit" were rated somewhat lower. A possible explanation for this difference is the limited time in the training session, which was approximately 30 min per AP. Furthermore, the version of the VR software in use did not provide any immediate and specific feedback interface (such as a text-based performance report) to the students currently active. Instead, feedback was only provided during debriefing by the tutor, which might lead to a lower rating for the learning benefit acquired through the training scenario.

Investigating the potential influence of perceived stress on estimated learning success among AP, we were able to demonstrate that the factor "sense of control" revealed a typical inverse U-shaped correlation to both subtypes of "estimated learning success", resembling the relationship that Yerkes and Dodson described for arousal and performance in challenging tasks [28]. According to one study we found in the literature, stress responsiveness was a predictor of good performance [43], while most work highlights the impairment of cognitive skills under excessive stress in high-fidelity surgical or medical emergency simulations [44–46]. The results of the aforementioned study are also concordant with the negative correlation that we discovered for the second factor "challenge to perform" and "estimated learning success". These findings indicate that support should be provided, e.g., by means of a student assistant or a digital tutor to mitigate excessive stress caused by the demands of this teaching method. Interestingly, on the other side of the inverse U shape, students with a high "sense of control" also reported poor learning. These participants may have had previous experience in dealing with medical emergencies, and/or the theoretical content presented was already familiar to them such that they found no gain in competence. However, owing to the lack of objective performance data, we cannot conclude whether this self-estimation was justified or misleading. In contrast, "social interaction to learn" was the only factor of "perceived stress", which exhibited no correlation to the two factors of "estimated learning success". Presence of an audience is described as a stressor in conventional healthcare simulation approaches [47, 48] and correlated with signs of increased stress and a trend towards worse learning outcomes in a physical intubation model [49]. However, the audience in our training session was generally perceived as supportive, as indicated by the results on item level, which might have somehow balanced

positive and negative effects, thus making it difficult to establish any meaningful correlation.

4.1 Strengths of the study

In our study, we used state-of-the-art technology to enable the best possible learning success by high-presence VR training and mitigate major side effects to a large extent. The number of participants was comparably large, coming from two unbiased semesters, and was thus a statistically meaningful dataset. In comparison with voluntary or incentive-driven studies to recruit students, this compulsory training session is also likely to include participants with more critical attitudes towards VR technology. In this way, the participants of our study may well prove to be representative of the heterogeneity of medical students in general.

4.2 Limitations of the study

One limitation of the study is that the sample was restricted to undergraduate students in one academic year and at one institution. Performed at a single medical school in Germany, further studies at other universities and with different student cohorts will be required to assess the validity of the results. In addition, as the course was performed exclusively in small groups under guidance, the additional influence of both the student assistant and the observing peers could not be clarified. As participation in the role of AP was voluntary, we are clearly unable to exclude selection bias in this group.

A second limitation is the use of adapted or newly created questionnaires. As the combination of the teaching method (VR-based simulation) and setting (small group teaching) is a relatively new approach, the currently available questionnaires, especially to evaluate perceived stress, were not considered applicable and had to be adapted. Of course, these strongly modified/newly created questionnaires await re-validation and currently restrict our ability to compare with other work.

Most importantly, all the results presented are subjective data recorded using surveys. While, at least to our knowledge, no concept exists for the objective measurement of simulation sickness or feeling of presence (based on physiologic parameters, for example), the reported "estimated learning success" certainly needs confirmation through a randomized controlled trial and test performance data derived from a pre- and post-intervention study design. Additionally, future work might place emphasis on objective measurements of stress, such as heart rate variation or electrodermal activity [50, 51], and investigate any possible habituation as a result of the repeated use of VR-based simulation.

5 Conclusions

Curricular implementation of a highly immersive VR-based training session of medical emergencies is feasible. The chosen teaching modality including guidance by a student assistant and subsequent debriefing by a tutor enjoys a high level of acceptance among medical students. Self-rated high level of learning success was reported by both AP and OBS, the effect on learning being more pronounced in students taking the role as AP.

Moreover, this study provides insights into how different conceptions of perceived stress distinctively moderate subjective learning success. Whether peer observers are perceived as helpful or seen as stressors could not be elucidated entirely and warrants further studies comparing group to individual teaching settings. Obviously, as the learning benefit in this study was reported subjectively by the participants, there is the need for an objective evaluation of the learning outcomes achieved by STEP-VR in a controlled assessment setting, ideally with different levels of explicit and implicit feedback.

Declarations

References

1. McEvoy, M.D., Dewaay, D.J., Vanderbilt, A., et al.: Are fourth-year medical students as prepared to manage unstable patients as they are to manage stable patients? Acad. Med. **89**, 618–624 (2014). https://doi.org/10.1097/ACM.0000000000000192

2. Monrouxe, L.V., Grundy, L., Mann, M., et al.: How prepared are UK medical graduates for practice? A rapid review of the literature 2009-2014. BMJ Open **7**, e013656 (2017). https://doi.org/10.1136/bmjopen-2016-013656

3. Bugaj, T.J., Nikendei, C., Groener, J.B., et al.: Ready to run the wards? A descriptive follow-up study assessing future doctors' clinical skills. BMC Med. Educ. **18**, 257 (2018). https://doi.org/10.1186/s12909-018-1370-4

4. Ochsmann, E.B., Zier, U., Drexler, H., et al.: Well prepared for work? Junior doctors' self-assessment after medical education. BMC Med. Educ. **11**, 99 (2011). https://doi.org/10.1186/1472-6920-11-99

5. Tallentire, V.R., Smith, S.E., Skinner, J., et al.: Understanding the behaviour of newly qualified doctors in acute care contexts. Med. Educ. **45**, 995–1005 (2011). https://doi.org/10.1111/j.1365-2923.2011.04024.x

6. Goodyear, H.M.: First year doctors experience of work related wellbeing and implications for educational provision. Int. J. Med. Educ. **5**, 103–109 (2014). https://doi.org/10.5116/ijme.5380.6ef1

7. Ilgen, J.S., Sherbino, J., Cook, D.A.: Technology-enhanced simulation in emergency medicine: a systematic review and meta-analysis. Acad. Emerg. Med. **20**, 117–127 (2013). https://doi.org/10.1111/acem.12076

8. ten Eyck, R.P.: Simulation in emergency medicine training. Pediatr. Emerg. Care **27**, 333–341 (2011). https://doi.org/10.1097/PEC.0b013e3182131fe0

9. Schwid, H.A., Rooke, G.A., Ross, B.K., et al.: Use of a computerized advanced cardiac life support simulator improves retention of advanced cardiac life support guidelines better than a textbook review. Crit. Care Med. **27**, 821–824 (1999). https://doi.org/10.1097/00003246-199904000-00045

10. Roppolo, L.P., Heymann, R., Pepe, P., et al.: A randomized controlled trial comparing traditional training in cardiopulmonary resuscitation (CPR) to self-directed CPR learning in first year medical students: the two-person CPR study. Resuscitation **82**, 319–325 (2011). https://doi.org/10.1016/j.resuscitation.2010.10.025

11. Heinrichs, W.L., Youngblood, P., Harter, P., et al.: Training healthcare personnel for mass-casualty incidents in a virtual emergency department: VED II. Prehosp. Disaster. Med. **25**, 424–432 (2010). https://doi.org/10.1017/s1049023x00008505

12. Middeke, A., Anders, S., Schuelper, M., et al.: Training of clinical reasoning with a Serious Game versus small-group problem-based learning: a prospective study. PLoS One **13**, e0203851 (2018). https://doi.org/10.1371/journal.pone.0203851

13. Slater, M., Wilbur, S.: A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. Presence **6**, 603–616 (1997). https://doi.org/10.1162/pres.1997.6.6.603

14. McGrath, J.L., Taekman, J.M., Dev, P., et al.: Using virtual reality simulation environments to assess competence for emergency medicine learners. Acad. Emerg. Med. **25**, 186–195 (2018). https://doi.org/10.1111/acem.13308

15. Kyaw, B.M., Saxena, N., Posadzki, P., et al.: Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. J. Med. Internet Res. **21**, e12959 (2019). https://doi.org/10.2196/12959

16. Gutiérrez, F., Pierce, J., Vergara, V.M., et al.: The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. Stud. Health Technol. Inform. **125**, 155–160 (2007)

17. Davis S, Nesbitt K, Nalivaiko E: A Systematic Review of Cybersickness. In: Blackmore K, Nesbitt K, Smith SP (eds) Proceedings of the 2014 Conference on Interactive Entertainment. ACM, New York, NY. pp 1–9 (2014)

18. Alaker, M., Wynn, G.R., Arulampalam, T.: Virtual reality training in laparoscopic surgery: a systematic review & meta-analysis. Int. J. Surg. **29**, 85–94 (2016). https://doi.org/10.1016/j.ijsu.2016.03.034

19. Bernardo, A.: Virtual reality and simulation in neurosurgical training. World Neurosurg. **106**, 1015–1029 (2017). https://doi.org/10.1016/j.wneu.2017.06.140

20. Birrenbach, T., Zbinden, J., Papagiannakis, G., et al.: Effectiveness and utility of virtual reality simulation as an educational tool for safe performance of COVID-19 diagnostics: prospective randomized pilot trial. JMIR Serious Games **9**, e29586 (2021). https://doi.org/10.2196/29586

21. Pantelidis, P., Chorti, A., Papagiouvanni, I., et al.: Virtual and augmented reality in medical education. In: Tsoulfas, G. (ed.) Medical and surgical education—past present and future. InTech, London (2018)

22. Skagerlund, K., Forsblad, M., Tinghög, G., et al.: Decision-making competence and cognitive abilities: which abilities matter? Behav. Decis. Mak. (2022). https://doi.org/10.1002/bdm.2242

23. Costello AB, Osborne J: Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. PARE.2005;10:7. https://doi.org/10.7275/jyj1-4868

24. 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. **78**, 1151–1163 (2021). https://doi.org/10.1016/j.jsurg.2020.10.015

25. Benesty, J., Chen, J., Huang, Y., et al.: Pearson correlation coefficient. In: Cohen, I., Huang, Y., Chen, J., et al. (eds.) Noise reduction in speech processing, vol. 2, pp. 1–4. Springer, Berlin, Heidelberg (2009)

26. Ezekiel, M.: A method of handling curvilinear correlation for any number of variables. J. Am. Stat. Assoc. **19**, 431 (1924). https://doi.org/10.2307/2281561

27. Little, R.J.A.: Statistical analysis with missing data. In: Wiley series in probability and statistics and statistics Ser, v.333, 2nd edn. John Wiley & Sons Incorporated, New York (2002)

28. Yerkes, R.M., Dodson, J.D.: The relation of strength of stimulus to rapidity of habit-formation. J. Comp. Neurol. Psychol. **18**, 459–482 (1908). https://doi.org/10.1002/cne.920180503

29. Lerner, D., Mohr, S., Schild, J., et al.: An immersive multi-user virtual reality for emergency simulation training: usability study. JMIR Serious Games **8**, e18822 (2020). https://doi.org/10.2196/18822

30. Pawlik V (2022) Umfrage unter Frauen und Männern zur Häufigkeit des Spielens von Computerspielen 2021. https://de.statista.com/statistik/daten/studie/828207/umfrage/umfrage-unter-frauen-und-maennern-zur-haeufigkeit-des-spielens-von-computerspielen/

31. Walter, S., Speidel, R., Hann, A., et al.: Skepticism towards advancing VR technology—student acceptance of VR as a teaching and assessment tool in medicine. GMS J. Med. Educ. **38**, Doc100 (2021). https://doi.org/10.3205/zma001496

32. Saredakis, D., Szpak, A., Birckhead, B., et al.: Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Front. Hum. Neurosci. **14**, 96 (2020). https://doi.org/10.3389/fnhum.2020.00096

33. Paillard, A.C., Quarck, G., Paolino, F., et al.: Motion sickness susceptibility in healthy subjects and vestibular patients: effects of gender, age and trait-anxiety. J. Vestib. Res. **23**, 203–209 (2013). https://doi.org/10.3233/VES-130501

34. Campillos-Llanos, L., Thomas, C., Bilinski, É., et al.: Lessons learned from the usability evaluation of a simulated patient dialogue system. J. Med. Syst. **45**, 69 (2021). https://doi.org/10.1007/s10916-021-01737-4

35. Maicher, K., Danforth, D., Price, A., et al.: Developing a conversational virtual standardized patient to enable students to practice history-taking skills. Simul. Healthc. **12**, 124–131 (2017). https://doi.org/10.1097/SIH.0000000000000195

36. Sanders, A.E., Lushington, K.: Effect of perceived stress on student performance in dental school. J. Dent. Educ. **66**, 75–81 (2002)

37. Alzahem, A.M., van der Molen, H.T., Alaujan, A.H., et al.: Stress amongst dental students: a systematic review. Eur. J. Dent. Educ. **15**, 8–18 (2011). https://doi.org/10.1111/j.1600-0579.2010.00640.x

38. Momi, D., Smeralda, C., Sprugnoli, G., et al.: Acute and long-lasting cortical thickness changes following intensive first-person action videogame practice. Behav. Brain. Res. **353**, 62–73 (2018). https://doi.org/10.1016/j.bbr.2018.06.013

39. Wu, S., Spence, I.: Playing shooter and driving videogames improves top-down guidance in visual search. Atten. Percept. Psychophys. **75**, 673–686 (2013). https://doi.org/10.3758/s13414-013-0440-2

40. Piccione, J., Collett, J., de Foe, A.: Virtual skills training: the role of presence and agency. Heliyon **5**, e02583 (2019). https://doi.org/10.1016/j.heliyon.2019.e02583

41. Makowski, D., Sperduti, M., Nicolas, S., et al.: "Being there" and remembering it: presence improves memory encoding. Conscious Cogn. **53**, 194–202 (2017). https://doi.org/10.1016/j.concog.2017.06.015

42. Bracq, M.-S., Michinov, E., Jannin, P.: Virtual reality simulation in nontechnical skills training for healthcare professionals: a systematic review. Simul. Healthc. **14**, 188–194 (2019). https://doi.org/10.1097/SIH.0000000000000347

43. Keitel, A., Ringleb, M., Schwartges, I., et al.: Endocrine and psychological stress responses in a simulated emergency situation. Psychoneuroendocrinology **36**, 98–108 (2011). https://doi.org/10.1016/j.psyneuen.2010.06.011

44. McGraw, L.K., Out, D., Hammermeister, J.J., et al.: Nature, correlates, and consequences of stress-related biological reactivity and regulation in Army nurses during combat casualty simulation. Psychoneuroendocrinology **38**, 135–144 (2013). https://doi.org/10.1016/j.psyneuen.2012.05.009

45. Wetzel, C.M., Black, S.A., Hanna, G.B., et al.: The effects of stress and coping on surgical performance during simulations. Ann. Surg. **251**, 171–176 (2010). https://doi.org/10.1097/SLA.0b013e3181b3b2be

46. Mauriz, E., Caloca-Amber, S., Córdoba-Murga, L., et al.: Effect of psychophysiological stress and socio-emotional competencies on the clinical performance of nursing students during a simulation practice. Int. J. Environ. Res. Public Health (2021). https://doi.org/10.3390/ijerph18105448

47. Shearer, J.N.: Anxiety, nursing students, and simulation: state of the science. J. Nurs. Educ. **55**, 551–554 (2016). https://doi.org/10.3928/01484834-20160914-02

48. Taylor, S.E., Seeman, T.E., Eisenberger, N.I., et al.: Effects of a supportive or an unsupportive audience on biological and psychological responses to stress. J. Pers. Soc. Psychol. **98**, 47–56 (2010). https://doi.org/10.1037/a0016563

49. Bensouda, B., Mandel, R., Mejri, A., et al.: Effect of an audience on trainee stress and performance during simulated neonatal intubation: a randomized crossover trial. BMC Med. Educ. **18**, 230 (2018). https://doi.org/10.1186/s12909-018-1338-4

50. Kyriakou, K., Resch, B., Sagl, G., et al.: Detecting moments of stress from measurements of wearable physiological sensors. Sens. (Basel) (2019). https://doi.org/10.3390/s19173805

51. Bhoja, R., Guttman, O.T., Fox, A.A., et al.: Psychophysiological stress indicators of heart rate variability and electrodermal activity with application in healthcare simulation research. Simul Healthc **15**, 39–45 (2020). https://doi.org/10.1097/SIH.0000000000000402

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.