

Skepticism towards advancing VR technology – student acceptance of VR as a teaching and assessment tool in medicine

Abstract

Objective: The high didactic potential of Virtual Reality (VR) contrasts with the point of view of students that the technology only has a relatively low significance for current and future teaching. This discrepancy was studied in a differentiated manner in order to gear the further development and implementation of VR towards the target group.

Methods: From January 2020 to July 2020, medical students ($N=318$) were asked to watch ten videos online and rate them on the basis of acceptance indicators (e.g., fun and fairness). Using obstetrics as an example, the videos demonstrated five levels of VR technology functionality (e.g., haptic and adaptive feedback), some of which were visionary, in two use scenarios (teaching and the OSCE). The individual and aggregate indicators were compared with non-parametric testing procedures across application scenarios, functional levels and genders. In addition, correlations between the acceptance and the factors of semester, age, computer affinity, and previous VR experience were analyzed.

Results: Across all functional levels, VR was more likely to be accepted in the classroom than in the OSCE. Comparisons across functional levels also revealed that the VR ready to be marketed was significantly more accepted than the visionary functions. This skepticism toward advancing VR technology was most pronounced with regard to the vision of autonomous VR examinations and among female students with a low computer affinity.

Conclusion: The results suggest that the students' reservations are due to a lack of experience with the VR technology. In order for young physicians to become familiar with the technology and to be able to use it competently in the everyday clinical practice in the future, VR should not only be used as a teaching tool but also be part of the curriculum. Practical examinations using VR, on the other hand, are only recommended once the technology has become established in teaching and has been proven to be reliable.

Keywords: virtual reality, virtual patient, medical training, future, acceptance

Steffen Walter¹
Robert Speidel²
Alexander Hann³
Janine Leitner¹
Lucia Jerg-Bretzke¹
Peter Kropp⁴
Jakob Garbe⁵
Florian Ebner⁶

1 Ulm University Hospital, Department of Psychosomatic Medicine and Psychotherapy, Medical Psychology Section, Ulm, Germany

2 University of Ulm, Faculty of Medicine, Competence Center eEducation in Medicine, Ulm, Germany

3 University Hospital Würzburg, Medical Clinic and Policlinic II, Gastroenterology, Würzburg, Germany

4 Rostock University Medical Center, Institute for Medical Psychology and Medical Sociology, Rostock, Germany

5 Halle University Hospital, Department of Internal Medicine I, Halle, Germany

6 Helios Amper Hospital Dachau, Dachau, Germany

1. The didactic potential of virtual reality

Inert knowledge is not only useless in medicine but life-threatening. Physicians must master specialized knowledge and automate skills in order to make the correct diagnosis and respond adequately in stressful situations. To ensure this level of competency, the National Competency-Based Learning Objective Catalog of Medicine (Nationale Kompetenzbasierte Lernzielkatalog Medizin, NKLM) was introduced in 2015 [<http://www.nklm.de>]. Students can only achieve the practice-oriented graduate profile of the NKLM if they have sufficient opportunity to practice by using authentic problems from everyday clinical practice during their training [1]. For organizational, financial, and ethical reasons, practicing on patients is insufficient in this regard [2]. One in four medical students admits, for example, to at least one error during training that potentially endangered a patient's health [3].

Therefore, simulations are used in medical training with which, for example, clinical decision-making and sensorimotor skills can be trained safely. Simulations are understood as learning environments in which students can influence the course of action of scenarios that are as authentic as possible [4], [5]. The most commonly used simulation method in medicine is role-playing with actor-patients [6] or simulation mannequins. However, since these require space, time, and personnel [7], computer-based simulations are becoming increasingly popular as well since they can be used flexibly and allow for an automatic performance assessment [2]. The subject is usually a "virtual patient" who is presented in an interactive video [8] or animation [9], [10]. Depending on the simulation, students can ask the virtual patients about their medical history, physically examine, or treat them. Since the students are usually seated in front of a conventional screen during the exercise and are limited in the actions they can take, these simulations have a relatively high degree of abstraction.

One way to remedy this deficit is the modern virtual reality (VR) simulation. Using VR goggles, students are immersed in a computer-generated world that gives them the impression of being physically present in the learning environment. This immersion, created in part by the high freedom of movement and interaction [11], makes the individual experience and behavior in VR simulations similar to that in a real situation [12]. This opens up the possibility of experiential learning in VR [13], [14] and thus a lasting learning effect. Although the quality of previous research may raise some issues (including small sample size, lack of control groups, and incomplete study reports), quantitative [15], [16] and qualitative [17] meta-analyses largely attest to the didactic added value of VR. It was found with regard to medical education and training, for example, that learning in virtual reality can be more effective than other digital (e.g., videos and online courses) and traditional (e.g., books and lectures) teaching formats [16].

In view of the competence orientation demanded by the NKLM, VR is of great interest in medicine. Every fourth study in the research field of VR has a medical reference [18]. Its use in medical didactics is increasing as well. At the University of Ulm, for example, students have access to a medical VR lab in which they can diagnose and treat virtual patients. In Germany such opportunities are still the exception [19], [20], whereas in English-speaking countries, VR is more widespread and, in some cases, already an integral part of the curriculum [21]. However, the best technology available today does not fully exploit the potential of this technology. If the technical development continues to advance, the currently predominantly audiovisual stimulation in VR [22] will most likely be supplemented with haptic feedback [23], [24]. This will make it possible to not just practice cognitive skills (e.g., process flows in the operating room) but also sensorimotor skills (e.g., physical examinations). Furthermore, the use of artificial intelligence will, in all likelihood, allow for adaptive performance feedback [25], [26] as well as natural verbal communication with virtual patients [9], [10]. The latter will make it possible to practice social skills (e.g., conversational skills with patients) alone in VR alone. The references cited are pilot projects but already paint a vision of a future medical education in which holistic learning experiences in VR are an integral part of the curriculum.

Whether and in what form this vision can become a reality depends in part on how students accept the technology and its future functions. In recent surveys conducted at German universities, students ascribe only a low to moderate importance to VR for both current and future teaching [19], [27]. The objective of this study is to obtain a more differentiated assessment by medical students in order to explore the discrepancy between the didactic potential and the subjectively perceived importance. The findings may help gear the further technical development and integration of VR into the curriculum toward the target group. The following questions were used to conduct the study and structure the presentation of the results:

1. Does the acceptance of VR differ between the teaching and the OSCE application scenarios?
2. Does the acceptance of VR differ between different levels of functionality, some of which are visionary?
3. Does the acceptance of VR differ between genders?
4. Are there correlations between the acceptance of VR and the particular student's semester, age, computer affinity, and prior VR experience?

Table 1: Sample description

Recruitment	N	Gender		Age		Semester		Computer affinity ^a		VR experience ^b		Own VR goggles	
		♀ N (%)	♂ N (%)	M	SD	M	SD	M	SD	M	SD	Yes N (%)	No N (%)
University Ulm ^c	111	63 (56.8%)	47 (42.3%)	24.49	.78	5.76	.31	3.99	.12	3.32	.18	10 (9.0%)	101 (91.0%)
Other universities ^d	207	139 (67.1%)	68 (32.9%)	24.68	.30	6.42	.22	3.95	.09	2.97	.13	11 (5.3%)	196 (94.7%)
Total	318	202 (63.5%)	115 (36.2%)	24.61	5.98	6.19	3.24	3.96	1.26	3.09	1.84	21 (6.6%)	297 (93.4%)

^a "You love working with computers" - Response scale from 1 ("strongly disagree") to 6 ("strongly agree").

^b "You have already had some form of experience with VR experiences" - Response scale from 1 ("strongly disagree") to 6 ("strongly agree").

^c The response rate was 7.7% in Ulm.

^d Berlin, Bern, Bochum, Bonn, Düsseldorf, Erlangen, Frankfurt, Freiburg, Giessen, Göttingen, Greifswald, Halle, Hamburg, Hanover, Heidelberg, Innsbruck, Jena, Kiel, Cologne, Lübeck, Mainz, Marburg, Munich, Regensburg, Rostock, Tübingen, Würzburg, Zurich, and N/A (M = 6.90, SD = 6.66).

2. Method

In order to answer these questions, medical students from German-speaking countries were asked to participate voluntarily and without remuneration in the online study (see table 1). The majority of the students were enrolled at the University of Ulm and were invited twice via e-mail to participate in the study. Of the students that were contacted, 7.7% responded and participated in the study. At the remaining universities, students were made aware of the study via campus websites and student councils on social media. The anonymous data collection took approximately 15 to 20 minutes for each student and was conducted from January 02, 2020 to July 31, 2020. Upon request, the ethics committee of the University of Ulm decided that the study project did not present a problem and did not require an ethics vote.

An invitation link directed students to the online study where they were asked to watch ten short videos. The videos, which can be accessed on YouTube, demonstrate and explain five levels of functionality of VR technology from the user's perspective (see figure 1) in two different application scenarios (teaching and the OSCE) (see table 2). The level of functionality shown increases successively and cumulatively from market-ready (visual and audiovisual stimulation) to visionary functions (haptics, oral communication, and adaptive feedback/autonomous testing) in both scenarios. Obstetrics was chosen as the application example because students are rarely able to observe and practice in real delivery rooms due to the intimacy surrounding pregnancy and childbirth.

After each video, students were asked to rate the demonstrated use of VR on the basis of acceptance indicators (e.g., presumed learning curve and innovativeness) (see table 3). To do so, students answered items on 6-point Likert-type scales ("strongly disagree" to "strongly agree"). Since the items correlated with each other in a highly significant manner (see attachment 1), they were also aggregated and averaged to obtain an overall acceptance score for each video.

In order to measure the acceptance indicators individually and in an aggregated manner between (1) the application scenarios and (2) the levels of functionality, Wilcoxon signed-rank tests as well as Friedman tests followed by Dunn-Bonferroni post-hoc tests were performed. Mean value comparisons between (3) the genders were addressed with Mann-Whitney U tests, and finally (4) the correlations between the overarching acceptance and the variables semester, age, computer affinity ("You love to work with computers") and prior VR ("You have already had some form of experience with VR experiences") were analyzed using Spearman correlations.

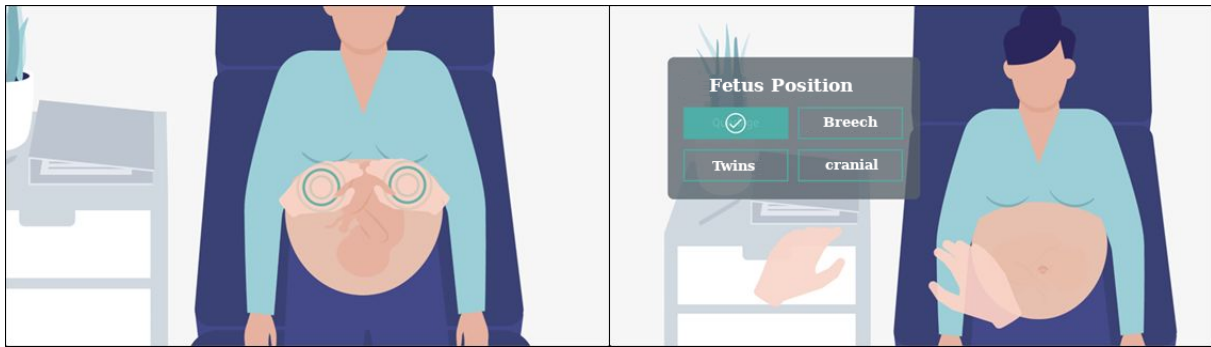


Figure 1: Video screenshot of the haptics function in the teaching (left) and OSCE scenario (right)

Table 2: YouTube links and duration of the videos

Level of functionality in the application scenarios	YouTube link	Duration in seconds
Visual stimulation		
Teaching	https://youtu.be/l4488FovEjU	63
OSCE	https://youtu.be/Mi6ocJg_rOc	37
Audiovisual stimulation		
Teaching	https://youtu.be/tvXegfsa-R4	39
OSCE	https://youtu.be/M4nV1oy5gnY	44
+ Haptics		
Teaching	https://youtu.be/YNU4JhRIAMk	57
OSCE	https://youtu.be/lvaSj_asDB0	41
+ Oral communication		
Teaching	https://youtu.be/EHcRSaVksSg	69
OSCE	https://youtu.be/qplw-ckhccl	32
+ Adaptive feedback, or autonomous testing		
Teaching	https://youtu.be/F4RjPSg73Ao	44
OSCE	https://youtu.be/ZNxTxyvGiRs	34

Table 3: Acceptance indicators

Acceptance indicator	Operational scenario	Item ^a
Learning curve	Teaching	"I believe that my learning curve could be improved if such a vision became a reality"
Fun	Teaching	"I think I would enjoy learning more if such a vision became a reality"
Fairness	OSCE	"I believe that the examination would be more objective and fairer than an examination involving a lecturer"
Simulation quality ^b	Teaching & OSCE	"I believe that the FUNCTION can simulate the patient contact well"
Innovative strength	Teaching & OSCE	"I think the development of the FUNCTION in teaching is innovative"
Usefulness	Teaching & OSCE	"I consider the FUNCTION to be useful in the application as a whole"
Feasibility	Teaching & OSCE	"I think the implementation of the FUNCTION in teaching is realistic"

^a "FUNCTION" is representative for the respective complexity level of the technology.

^b The simulation quality indicator was omitted from the adaptive teaching scenario.

Table 4: Overarching acceptance of levels of functionality between teaching and the OSCE

Application scenario/test	Statistics	Degree of function				
		Visual stimulation	Audiovisual stimulation	+ Haptics	+ Oral communication	+ Adaptive feedback/ autonomous testing
Teaching	<i>M</i>	4.14	4.45	4.02	3.94	3.90
	<i>SD</i>	1.10	1.08	1.24	1.28	1.31
OSCE	<i>M</i>	3.95	4.13	3.68	3.65	3.45
	<i>SD</i>	1.06	1.08	1.24	1.31	1.30
Wilcoxon signed-rank test	<i>Z</i>	4.27	8.35	8.83	6.70	7.16
	<i>p</i>	.000***	.000***	.000***	.000***	.000***
	<i>r</i>	.24	.47	.50	.38	.40

Note: The response scale ranged from 1 ("strongly disagree") to 6 ("strongly agree").

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 5: Dunn-Bonferroni post-hoc tests on the overall acceptance between levels of functionality

Teaching			OSCE		
Comparison of levels of functionality	Test statistics	p^a	Comparison of levels of functionality	Test statistics	p^a
F1 - F2	-.81	.000***	F1 - F2	-.39	.020*
F1 - F3	-.15	1.000	F1 - F3	-.50	.001**
F1 - F4	-.14	1.000	F1 - F4	-.42	.009**
F1 - F5	-.17	1.000	F1 - F5	-.89	.000***
F2 - F3	-.96	.000***	F2 - F3	-.89	.000***
F2 - F4	-.95	.000***	F2 - F4	-.80	.000***
F2 - F5	-.98	.000***	F2 - F5	-1.27	.000***
F3 - F4	-.01	1.000	F3 - F4	-.09	1.000
F3 - F5	-.02	1.000	F3 - F5	-.38	.022*
F4 - F5	-.03	1.000	F4 - F5	-.47	.002**

Note: Visual stimulation (F1), audiovisual stimulation (F2), + haptics (F3), + oral communication (F4), + adaptive feedback/ autonomous testing (F5).

^aTo avoid an alpha error accumulation, the p -value was corrected by using the Bonferroni method.

* $p < .05$. ** $p < .01$. *** $p < .001$

3. Results

3.1. Acceptance of the level of functionality regarding teaching and the OSCE

The use of VR depicted in the videos was significantly more accepted in teaching than in the OSCE across all levels of functionality (see table 4).

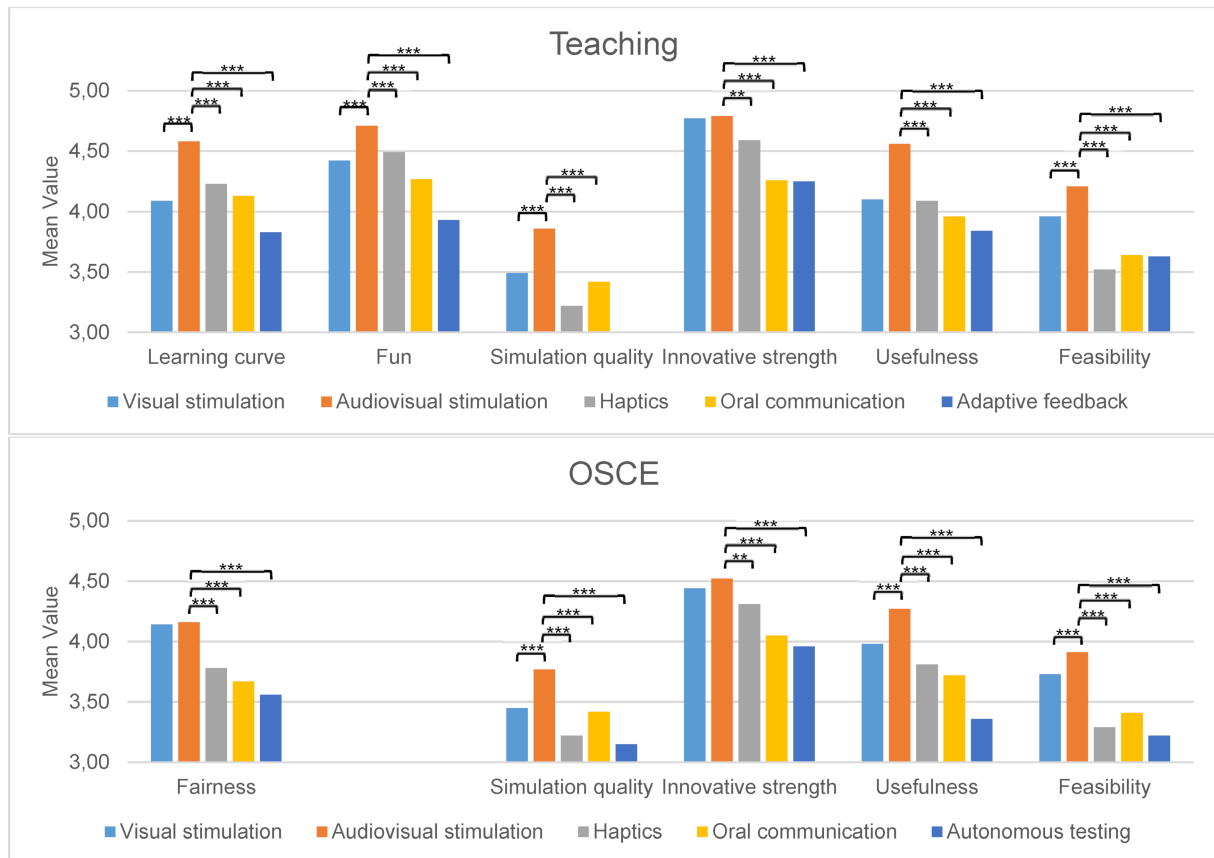
3.2. Acceptance differences between the levels of functionality

In both teaching ($\chi^2(4)=97.41$, $p<.001$, $n=318$) and the OSCE ($\chi^2(4)=138.07$, $p<.001$, $n=318$), VR acceptance differed significantly between the levels of functionality. The post-hoc testing (see table 5) specified that the audiovisual degree of functionality was significantly best accepted in both application scenarios. Visual stimulation was the second preference but was only rated significantly better than the more complex haptic, oral communication and autonomous levels of functionality in the OSCE. Students rated the latter the lowest in the OSCE with a signi-

ficant margin. A differentiated look at the individual acceptance indicators shows that all items scored significantly highest for the audiovisual stimulation (see figure 2, see attachment 1, table A1 and table A2). The exception is the visual stimulation, which was only descriptively rated lower for the innovativeness and fairness indicators.

3.3. Acceptance differences between the sexes

Male students exhibited a higher overarching acceptance than female students in all levels of functionality and application scenarios. Regarding the OSCE scenario videos, this gender difference became significant with regard to the functional levels of audiovisual stimulation ($U=10048.00$, $Z=-2.00$, $p=.045$, $d_{Cohen}=.23$) and autonomous examination ($U=9950.50$, $Z=-2.12$, $p=.034$, $d_{Cohen}=.24$). In addition, the computer affinity was also higher among male students ($U=7206.50$, $Z=-5.79$, $p<.000$, $d_{Cohen}=.67$).



Note: The marked results of the Wilcoxon signed-rank tests between the audiovisual stimulation and the other levels of functionality are shown in the attached tables A1 and A2. The response scale ranged from 1 ("strongly disagree") to 6 ("strongly agree"). * $p < .05$. ** $p < .01$. *** $p < .001$

Figure 2: Acceptance indicators by level of functionality and application scenario

3.4. Relationship between overarching acceptance and control factors

The acceptance scores across the application scenarios correlated strongly between the levels of functionality ($r = .493$ to $.851$) and moderately with regard to the individual computer affinity ($r = .229$ to $.372$) (see attachment 1, table A3). In contrast, the control factors of age, semester and prior VR experience did not show a significant relationship with student acceptance.

4. Discussion

The high didactic potential of VR contrasts with the point of view of students that the technology only has a relatively low significance for current and future teaching. This discrepancy was examined in a differentiated manner in the present study in order to gear the further technical development and the implementation of VR into the curriculum towards the target group. With this goal "in mind", the student acceptance between different (1) application scenarios and (2) levels of functionality was compared. In addition, it was checked whether, with regard to the acceptance, (3) gender differences or (4) correlations between the factors of semester, age, computer affinity and previous VR experience exist.

As in previous surveys [19], [27], the acceptance of the use of VR was generally moderate.

Referencing the comparisons made, however, this general statement is not specific enough. Across all levels of functionality, for example, VR was found to be more accepted in teaching than in the OSCE. This suggests that students have doubts about whether the novel VR technology can reliably support practical performance assessments or even automate them in the future. With regard to the Technology Acceptance Model [28], the low acceptance of autonomous VR examinations can probably be attributed to the lack of control over the system. Comparisons between levels of functionality reveal further reservations about the technology. The market-ready status quo of VR (visual and audiovisual stimulation) was significantly better accepted than the features (haptics, oral communication and adaptive feedback/autonomous testing) that are still under development in both application scenarios. Medical students thus appear to be skeptical about technological innovations that have not yet been proven in practice. This skepticism is most pronounced for autonomous VR examinations and among female students.

Since there are only few empirical arguments for the visionary VR features, this reluctance is especially understandable when it comes to examination scenarios that are directly relevant to a student's grading. In detail,

however, the results are counterintuitive. As the level of functionality increased, not only did the assumed feasibility decrease but so did the innovativeness, simulation quality, meaningfulness, and learning curve. The same is true for the application-specific factors of fun, learning curve, and fairness. These statements contradict the potential attributed to the technical innovations. While haptic feedback is primarily intended to improve the immersion and quality of the simulation, the integration of artificial intelligence in VR promises individual learning and examination experiences with lifelike virtual patients. These potentials are not reflected in the assessment of the medical students. One reason for this counterintuitive assessment may be the human tendency to embrace the ordinary (audiovisual stimulation) and to avoid the unknown (e.g., autonomous performance assessment) [29]. This basic attitude is substantially reflected in the positive correlation between acceptance and computer affinity. Since the use of haptic feedback and AI is still in development and thus generally unknown, student skepticism towards visionary levels of functionality can possibly be attributed to a lack of experience.

The restrained acceptance of the market-ready status quo can possibly be explained in a similar manner. In a cross-curriculum survey by Weisflog and Böckel [19], students considered VR more important for their studies if the technology was available to them at their university. In 2019, however, only about 4% of the students in Germany had VR equipment at their disposal on campus. Even in the private sector, the use of VR – despite increasing hardware and software sales – is currently still limited to a technology-savvy minority [<https://www2.deloitte.com/de/de/pages/presse/contents/zukunftsperspektiven-fuer-virtual-augmented-reality.html>]. Consequently, only a small percentage of students has had the opportunity to make practical use of the current VR technology and overcome possible reservations. However, this explanatory approach is not supported by the prior VR experience measured in the study. The corresponding item presumably fails to express the individual level of knowledge of VR technology since the quality and quantity of the VR experiences were not surveyed.

Also noteworthy is the consistently high correlation between the acceptance indicators (see attachment 1, table A3). The high internal consistency suggests that many medical students have an established opinion about the use of VR that has conditioned their responses across application scenarios and levels of functionality. The transferability of these findings needs to be critically discussed, however. The low response rate of 7.7% in Ulm limits the representativeness of the sample. In addition, the video demonstrations referred exclusively to obstetrics, which means that the acceptance values can be concluded for other fields of application only indirectly. Finally, when interpreting the results, it must also be taken into account that the students only evaluated videos on the use of VR since the higher levels of functionality have not yet reached market maturity. In line with the assumption that the acceptance of VR increases

as awareness rises, the assessment would probably be more positive if students had had a practical demonstration.

Regardless, however, the study results make it clear that the concerns and technical knowledge of students must be taken into account in the further development and the implementation of VR in the curriculum. In order for real learning effects to result from the potential of VR, the technology should be introduced not only as a teaching tool but also be part of the medical studies curriculum. The opportunity to learn about the application of VR in a guided manner and to discuss it critically also promotes the media competence of medical students who will possibly work with VR in their future daily clinical routine (e.g., simulation-based planning of surgical interventions). Whether VR will also be suitable for conducting practical examinations or whether student skepticism will continue to be justified in the future should be scientifically examined as soon as the technology has become established in teaching and proven to be reliable.

5. Conclusion

The differentiated assessment of the acceptance of VR by students revealed that the subjectively low perceived importance of VR is due to a skepticism towards emerging technologies. Medical students seem to have too little knowledge about VR to adequately assess its didactic potential. In order for medical students to become familiar with the technology and to be able to use it competently in their everyday clinical practice in the future, VR should be introduced not only as a teaching tool but also integrated into the curriculum. The implementation of practical examinations in VR, on the other hand, is only recommended once the technology has proven itself to be reliable in teaching.

Authors

The authors S. Walter and R. Speidel share the first authorship.

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from

<https://www.egms.de/en/journals/zma/2021-38/zma001496.shtml>

- Attachment_1.pdf (108 KB)
Attachment 1

References

1. Boshuizen H, Schmidt H. The development of clinical reasoning expertise. *Clin Reason Health Prof.* 2008;3:113-121.
2. Keifenheim KE, Teufel M, Ip, J, Speiser N, Leehr EJ, Zipfel S, Herrmann-Werner A. Teaching history taking to medical students: A systematic review. *BMC Med Educ.* 2015;15:159. DOI: 10.1186/s12909-015-0443-x
3. Kiesewetter J, Kager M, Lux R, Zwissler B, Fischer MR, Dietz I. German undergraduate medical students' attitudes and needs regarding medical errors and patient safety - a national survey in Germany. *Med Teach.* 2014;36(6):505-510. DOI: 10.3109/0142159X.2014.891008
4. Cook DA. How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. *Med Educ.* 2014;48(8):750-760. DOI: 10.1111/medu.12473
5. Kononowicz AA, Woodham LA, Edelbring S, Stathakarou N, Davies D, Saxena N, Tudor Car L, Carlsted-Duke J, Car J, Zary N. Virtual patient simulations in health professions education: Systematic review and meta-analysis by the digital health education collaboration. *J Med Internet Res.* 2019;21(7):e14676. DOI: 10.2196/14676
6. Huwendiek S, De Ieng BA, Zary N, Fischer MR, Ruiz JG, Ellaway R. Towards a typology of virtual patients. *Med Teach.* 2009;31(8):743-748. DOI: 10.1080/01421590903124708
7. Bokken L, Rethans JJ, Scherpier AJ, van der Vleuten CP. Strengths and weaknesses of simulated and real patients in the teaching of skills to medical students: a review. *Simul Healthc.* 2008;3(3):161-169. DOI: 10.1097/SIH.0b013e318182fc56
8. Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hatala R. Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Med Teach.* 2013;35(1):867-898. DOI: 10.3109/0142159X.2012.714886
9. Shorey S, Ang E, Yap J, Ng ED, Lau ST, Chui CK. A Virtual Counseling Application Using Artificial Intelligence for Communication Skills Training in Nursing Education: Development Study. *J Med Internet Res.* 2019;21(10):e14658. DOI: 10.2196/14658
10. Talbot TB, Rizzo AS, Dubbels BR. Virtual standardized patients for interactive conversational training: A grand experiment and new approach. *Hershey, PA: IGI Global;* 2019. DOI: 10.4018/978-1-5225-7461-3.ch003
11. Slater M, Wilbur S. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence.* 1997;6(6):603-616. DOI: 10.1162/pres.1997.6.6.603
12. Chirico A, Gaggioli A. When Virtual Feels Real: Comparing Emotional Responses and Presence in Virtual and Natural Environments. *Cyberpsychol Behav Soc Netw.* 2019;22(3):220-226. DOI: 10.1089/cyber.2018.0393
13. Kolb DA. *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, N.J.: Prentice-Hall; 1984.
14. Kwon C. Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR technologies. *Virt Reality.* 2019;23(1):101-118. DOI: 10.1007/s10055-018-0364-1
15. Chavez B, Bayona S. Virtual Reality in the Learning Process. In: Rocha Á, Adeli H, Reis LP, Costanzo S, editors. *Advances in Intelligent Systems and Computing. Trends and Advances in Information Systems and Technologies.* Cham: Springer International Publishing; 2018. p.1345-1356. DOI: 10.1007/978-3-319-77712-2_129
16. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, Divakar U, Masiello I, Kononowicz AA, Zary N, Tudor Car L. Virtual Reality for Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. *J Med Internet Res.* 2019;21(1):e12959. DOI: 10.2196/12959
17. Pourmand A, Davis S, Lee D, Barber S, Sikka N. Emerging Utility of Virtual Reality as a Multidisciplinary Tool in Clinical Medicine. *Games Health J.* 2017;6(5):263-270. DOI: 10.1089/g4h.2017.0046
18. Cipresso P, Giglioli IAC, Raya MA, Riva G. The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature. *Front Psychol.* 2018;9:2086. DOI: 10.3389/fpsyg.2018.02086
19. Weisflog W, Böckel A. Ein studentischer Blick auf den Digital Turn: Auswertung einer bundesweiten Befragung von Studierenden für Studierende. Arbeitspapier Nr. 54. Berlin: Stifterverband für die Deutsche Wissenschaft, Hochschulforum Digitalisierung; 2020. Zugänglich unter/available from: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_54_Studierendenbefragung.pdf
20. Kuhn S, Frankenhauser S, Tolks D. Digitale Lehr- und Lernangebote in der medizinischen Ausbildung: Schon am Ziel oder noch am Anfang? [Digital learning and teaching in medical education : Already there or still at the beginning?]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2018;61(2):201-209. DOI: 10.1007/s00103-017-2673-z
21. Pottle J. Virtual reality and the transformation of medical education. *Future Healthc J.* 2019;6(3):181-185. DOI: 10.7861/fhj.2019-0036
22. Zender R, Weise M, von der Heyde M, Söbke H. Lehren und Lernen mit VR und AR - Was wird erwartet? Was funktioniert? In: Krömker D, Schroeder U, editors. *DeLFI 2018 - Die 16. E-Learning Fachtagung Informatik.* Bonn: Gesellschaft für Informatik e.V.; 2018. p.275-276.
23. Frithioff A, Frendø M, Mikkelsen PT, Sørensen MS, Andersen SA. Ultra-high-fidelity virtual reality mastoidectomy simulation training: a randomized, controlled trial. *Eur Arch Otorhinolaryngol.* 2020;277(5):1335-1341. DOI: 10.1007/s00405-020-05858-3
24. Rangarajan K, Davis H, Pucher PH. Systematic Review of Virtual Haptics in Surgical Simulation: A Valid Educational Tool? *J Surg Educ.* 2020;77(2):337-347. DOI: 10.1016/j.j Surg Educ. 2019.09.006
25. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS ONE.* 2020;15(2):e0229596. DOI: 10.1371/journal.pone.0229596
26. Veneziano D, Cacciamani G, Rivas JG, Marino NS, Bhaskar K. VR and machine learning: novel pathways in surgical hands-on training. *Curr Opin Urol.* 2020;30(6):817-822. DOI: 10.1097/MOU.0000000000000824
27. Speidel R, Schneider A, Koerner J, Grab-Kroll C, Oechsner W. Did video kill the XR star? Digital trends in medical education before and after the COVID-19 outbreak from the perspective of students and lecturers from the faculty of medicine at the University of Ulm. *GMS J Med Educ.* 2021;38(6):Doc101. DOI: 10.3205/zma001497
28. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. *J Biomed Inform.* 2010;43(1):159-172. DOI: 10.1016/j.jbi.2009.07.002
29. Castro D, McQuinn A. *The Privacy Panic Cycle: A Guide to Public Fears About New Technologies.* Washington, DC: The Information Technology & Innovation Foundation; 2015. p.39. Zugänglich unter/available from: <http://www2.itif.org/2015-privacy-panic.pdf>

Corresponding author:

Steffen Walter
Ulm University Hospital, Department of Psychosomatic
Medicine and Psychotherapy, Medical Psychology Section,
Frauensteige 6, D-89075 Ulm, Germany
steffen.walter@uni-ulm.de

Please cite as

Walter S, Speidel R, Hann A, Leitner J, Jerg-Bretzke L, Kropp P, Garbe J,
Ebner F. Skepticism towards advancing VR technology – student
acceptance of VR as a teaching and assessment tool in medicine. *GMS
J Med Educ.* 2021;38(6):Doc100.
DOI: 10.3205/zma001496, URN: urn:nbn:de:0183-zma0014967

This article is freely available from

<https://www.egms.de/en/journals/zma/2021-38/zma001496.shtml>

Received: 2021-01-27

Revised: 2021-05-17

Accepted: 2021-06-25

Published: 2021-09-15

Copyright

©2021 Walter et al. This is an Open Access article distributed under
the terms of the Creative Commons Attribution 4.0 License. See license
information at <http://creativecommons.org/licenses/by/4.0/>.

Skepsis gegenüber fortschreitender VR-Technologie – die studentische Akzeptanz von VR als Lehr- und Prüfungsmedium in der Medizin

Zusammenfassung

Zielsetzung: Das hohe didaktische Potenzial von VR steht im Kontrast zur studentischen Einschätzung, dass die Technologie nur eine relativ geringe Bedeutung für die gegenwärtige und zukünftige Lehre hat. Diese Diskrepanz wurde differenziert untersucht, um die Weiterentwicklung und Implementierung von VR zielgruppenorientiert auszurichten.

Methodik: Von Januar bis Juli 2020 wurden Medizinstudierende (N=318) gebeten, zehn Videos online anzusehen und anhand von Akzeptanzindikatoren (z.B. Spaß und Fairness) zu bewerten. Am Beispiel der Geburtshilfe demonstrierten die Videos fünf teils visionäre Funktionsgrade der VR-Technologie (z.B. haptisches und adaptives Feedback) in zwei Einsatzszenarien (Lehre und OSCE-Prüfung). Die einzelnen und aggregierten Indikatoren wurden mit nicht-parametrischen Testverfahren zwischen den Einsatzszenarien, Funktionsgraden und Geschlechtern verglichen. Darüber hinaus wurden Zusammenhänge zwischen der Akzeptanz und den Faktoren Fachsemester, Alter, Computeraffinität und VR-Vorerfahrung analysiert.

Ergebnisse: Über alle Funktionsgrade hinweg wurde VR in der Lehre eher akzeptiert als in der OSCE-Prüfung. Die Vergleiche zwischen den Funktionsgraden offenbarten außerdem, dass der marktreife Status Quo von VR signifikant besser angenommen wurde als die visionären Funktionen. Diese Skepsis gegenüber der fortschreitenden VR-Technologie war bei der Vision autonomer VR-Prüfungen und unter weiblichen Studierenden mit geringer Computeraffinität am stärksten ausgeprägt.

Schlussfolgerung: Die Ergebnisse legen nahe, dass die Vorbehalte der Studierenden auf mangelnde Erfahrungswerte mit der VR-Technologie zurückzuführen sind. Damit die angehenden Ärzte die Technologie kennenlernen und zukünftig kompetent im Klinikalltag anwenden können, sollte VR nicht nur als Lehrmedium sondern auch als Lerngegenstand curricular eingeführt werden. Praktische Prüfungen in VR empfehlen sich dagegen erst, wenn sich die Technologie in der Lehre etabliert und als verlässlich erwiesen hat.

Schlüsselwörter: Virtuelle Realität, virtueller Patient, medizinische Ausbildung, Zukunft, Akzeptanz

Steffen Walter¹
Robert Speidel²
Alexander Hann³
Janine Leitner¹
Lucia Jerg-Bretzke¹
Peter Kropp⁴
Jakob Garbe⁵
Florian Ebner⁶

1 Universitätsklinikum Ulm, Klinik für Psychosomatische Medizin und Psychotherapie, Sektion Medizinische Psychologie, Ulm, Deutschland

2 Universität Ulm, Medizinische Fakultät, Kompetenzzentrum eEducation in der Medizin, Ulm, Deutschland

3 Universitätsklinikum Würzburg, Medizinische Klinik und Poliklinik II, Gastroenterologie, Würzburg, Deutschland

4 Universitätsmedizin Rostock, Institut für Medizinische Psychologie und Medizinische Soziologie, Rostock, Deutschland

5 Universitätsklinikum Halle, Klinik für Innere Medizin I, Halle, Deutschland

6 Helios Amper-Klinikum Dachau, Dachau, Deutschland

1. Das didaktische Potenzial der virtuellen Realität

Träges Wissen ist in der Medizin nicht nur unbrauchbar, sondern lebensbedrohlich. Ärzte müssen Fachwissen durchdringen und Fertigkeiten automatisieren, um in Belastungssituationen die richtige Diagnose zu stellen und adäquat zu reagieren. Um dieses Kompetenzniveau sicherzustellen, wurde 2015 der Nationale Kompetenzbasierte Lernzielkatalog Medizin (NKLM) eingeführt [<http://www.nklm.de>]. Für Studierende ist das praxisorientierte Absolventenprofil des NKLM nur erreichbar, wenn sie in ihrer Ausbildung ausreichend Gelegenheit haben, anhand authentischer Problemstellungen aus dem Klinikalltag praktisch zu üben [1]. Das Üben am Patienten ist dafür aus organisatorischen, finanziellen und ethischen Gründen ungenügend [2]. Beispielsweise räumt jeder vierte Medizinstudierende mindestens einen Fehler in der Ausbildung ein, der potenziell die Gesundheit eines Patienten gefährdete [3].

In der medizinischen Ausbildung werden deshalb Simulationen eingesetzt, mit denen beispielsweise die klinische Entscheidungsfindung und sensomotorische Fertigkeiten gefahrlos trainiert werden können. Unter Simulationen werden Lernumgebungen verstanden, in denen Studierende den Handlungsverlauf von möglichst authentischen Szenarien beeinflussen können [4], [5]. Die in der Medizin am häufigsten angewendete Simulationsmethode sind Rollenspiele mit Schauspielpatienten [6] oder Simulationsspinnen. Da diese jedoch räumlich, zeitlich und personell gebunden sind [7], werden vermehrt auch computerbasierte Simulationen verwendet, die flexibel einsetzbar sind und eine automatische Leistungsbeurteilung ermöglichen [2]. Der Lerngegenstand ist hierbei in der Regel ein „virtueller Patient“, der als interaktives Video [8] oder Animation [9], [10] dargestellt wird. Die Studierenden können die virtuellen Patienten je nach Simulation zur ihrer Krankheitsgeschichte befragen, körperlich untersuchen oder behandeln. Da sie bei der Übung meist vor einem konventionellen Bildschirm sitzen und in ihren Aktionsmöglichkeiten beschränkt sind, weisen diese Simulationen einen relativ hohen Abstraktionsgrad auf.

Ein Lösungsansatz für dieses Defizit sind moderne Simulationen in der virtuellen Realität (VR). Mit einer VR-Brille tauchen Studierende in eine computergenerierte Welt ein, die ihnen den Eindruck vermittelt, physisch in der Lernumgebung präsent zu sein. Durch diese Immersion, die unter anderem durch die hohe Bewegungs- und Interaktionsfreiheit entsteht [11], ähnelt das individuelle Erleben und Verhalten in VR-Simulationen dem in einer realen Situation [12]. Das eröffnet die Möglichkeit, in VR erfahrungsbasiert zu lernen [13], [14] und somit einen nachhaltigen Lerneffekt zu erzielen. Zwar ist die Qualität der bisherigen Forschung zu diskutieren (u.a. geringe Stichprobengröße, fehlende Kontrollgruppen und lückenhafte Studienberichte), doch quantitative [15], [16] und qualitative [17] Metaanalysen bescheinigen den didaktischen Mehrwert von VR weitgehend. Mit Blick auf gesundheits-

bezogene Aus- und Weiterbildungen zeigt sich etwa, dass das Lernen in der virtuellen Realität effektiver sein kann als in anderen digitalen (z.B. Videos und Online-Kurse) und traditionellen (z.B. Bücher und Vorlesungen) Lehrformaten [16].

Angesichts der vom NKLM geforderten Kompetenzorientierung trifft VR in der Medizin auf großes Interesse. Bereits jede vierte Studie im Forschungsfeld VR weist einen medizinischen Bezug auf [18]. Auch der medizinischdidaktische Einsatz nimmt zu. An der Universität Ulm steht den Studierenden beispielsweise ein medizinisches VR-Lab zur Verfügung, in dem sie virtuelle Patienten diagnostizieren und behandeln können. In Deutschland sind solche Angebote noch die Ausnahme [19], [20], im englischsprachigen Raum ist VR dagegen weiter verbreitet und teils schon fester Bestandteil des Curriculums [21]. Dabei reizt der aktuelle Stand der Technik nicht das Potenzial der Technologie aus. Schreitet die technische Entwicklung weiter voran, wird die bisher vorwiegend audiovisuelle Stimulation in VR [22] voraussichtlich mit haptischem Feedback [23], [24] ergänzt. Dadurch können nicht mehr nur kognitive (z.B. Prozessabläufe im OP-Bereich) sondern auch sensomotorische Fertigkeiten (z.B. körperliche Untersuchungen) geschult werden. Außerdem wird der Einsatz von künstlicher Intelligenz aller Wahrscheinlichkeit nach adaptives Feedback zur erbrachten Leistung [25], [26] sowie eine natürliche, mündliche Kommunikation mit virtuellen Patienten [9], [10] ermöglichen. Letzteres würde es schließlich erlauben, auch soziale Fertigkeiten (z.B. Gesprächsführung mit Patienten) alleine in VR zu trainieren. Die angeführten Referenzen sind Pilotprojekte, zeichnen jedoch bereits eine Vision der zukünftigen medizinischen Ausbildung, in der holistische Lernerfahrungen in VR fester Bestandteil des Curriculums sind.

Ob und in welcher Form diese Vision realisiert wird, ist nicht zuletzt davon abhängig, wie die Studierenden die Technologie und ihre zukünftigen Funktionen annehmen. In aktuellen, an deutschen Hochschulen durchgeführten Umfragen bescheinigen sie VR nur eine geringe bis moderate Bedeutung für sowohl die gegenwärtige als auch die zukünftige Lehre [19], [27]. Das Ziel dieser Studie ist es, eine differenziertere Einschätzung der Medizinstudierenden zu erhalten, um die Diskrepanz zwischen didaktischem Potenzial und subjektiv wahrgenommener Bedeutung zu ergründen. Die Erkenntnisse sollen dazu beitragen, die technische Weiterentwicklung und curriculare Implementierung von VR zielgruppenorientiert auszurichten. Die folgenden Fragen leiteten das Studienvorhaben und gliedern die Darstellung der Ergebnisse:

1. Unterscheidet sich die Akzeptanz von VR zwischen den Einsatzszenarien Lehre und OSCE-Prüfung?
2. Unterscheidet sich die Akzeptanz von VR zwischen unterschiedlichen, teils visionären Funktionsgraden?
3. Unterscheidet sich die Akzeptanz von VR zwischen den Geschlechtern?
4. Bestehen Zusammenhänge zwischen der Akzeptanz von VR und dem Fachsemester, dem Alter, der Computeraffinität und der VR-Vorerfahrung?

Tabelle 1: Stichprobenbeschreibung

Rekrutierung	N	Geschlecht		Alter		Fachsemester		Computer-Affinität ^a		VR-Erfahrung ^b		Eigene VR-Brille	
		♀ N (%)	♂ N (%)	M	SD	M	SD	M	SD	M	SD	Ja N (%)	Nein N (%)
Universität Ulm ^c	111	63 (56.8%)	47 (42.3%)	24.49	.78	5.76	.31	3.99	.12	3.32	.18	10 (9.0%)	101 (91.0%)
Restliche Universitäten ^d	207	139 (67.1%)	68 (32.9%)	24.68	.30	6.42	.22	3.95	.09	2.97	.13	11 (5.3%)	196 (94.7%)
Total	318	202 (63.5%)	115 (36.2%)	24.61	5.98	6.19	3.24	3.96	1.26	3.09	1.84	21 (6.6%)	297 (93.4%)

^a „Sie lieben es, mit Computern umzugehen.“ - Antwortskala von 1 („trifft nicht zu“) bis 6 („trifft maximal zu“).

^b „Sie haben bereits in irgendeiner Form Erfahrungen mit VR-Erlebnissen gemacht.“ - Antwortskala von 1 („trifft nicht zu“) bis 6 („trifft maximal zu“).

^c Die Rücklaufquote betrug in Ulm 7,7%.

^d Berlin, Bern, Bochum, Bonn, Düsseldorf, Erlangen, Frankfurt, Freiburg, Gießen, Göttingen, Greifswald, Halle, Hamburg, Hannover, Heidelberg, Innsbruck, Jena, Kiel, Köln, Lübeck, Mainz, Marburg, München, Regensburg, Rostock, Tübingen, Würzburg, Zürich und k.A. (M = 6.90, SD = 6.66).

2. Methode

Um diese Fragestellungen zu untersuchen, wurden Medizinstudierende aus dem deutschsprachigen Raum gebeten, freiwillig und ohne Entlohnung an der online durchgeführten Studie teilzunehmen (siehe Tabelle 1). Der Großteil der Studierenden war an der Universität Ulm eingeschrieben und wurde zweimalig via E-Mail zur Studie eingeladen. 7.7% der Adressaten sind dem Aufruf gefolgt und haben an der Studie teilgenommen. An den restlichen Universitäten wurden die Studierenden über Campus-Webseiten und Fachschaften in sozialen Medien auf die Studie aufmerksam gemacht. Die anonyme Datenerhebung dauerte jeweils ca. 15 bis 20 Minuten und wurde vom 02. Januar bis zum 31. Juli 2020 durchgeführt. Auf Anfrage entschied die Ethikkommission der Universität Ulm, dass das Studienvorhaben unbedenklich sei und keinem Ethikvotum bedürfe.

Ein Einladungslink leitete die Studierenden zur Online Studie, bei der sie aufgefordert wurden, zehn kurze Videos anzusehen. Die auf YouTube abrufbaren Videos demonstrieren und erklären fünf Funktionsgrade der VR-Technologie aus der Perspektive des Anwenders (siehe Abbildung 1) in zwei verschiedenen Einsatzszenarien (Lehre und OSCE-Prüfung) (siehe Tabelle 2). Der dargestellte Funktionsgrad steigert sich in beiden Szenarien sukzessive und kumulativ von marktreifen (visuelle und audiovisuelle Stimulation) bis hin zu visionären Funktionen (Haptik, mündliche Kommunikation und adaptives Feedback/autonome Prüfung). Als Anwendungsbeispiel wurde die Geburtshilfe ausgewählt, da Studierende wegen der Intimität rund um Schwangerschaft und Geburt selten im realen Kreißsaal hospitieren und üben können.

Nach jedem Video wurden die Studierenden angehalten, den demonstrierten Einsatz von VR anhand von Akzeptanzindikatoren (z.B. vermutete Lernkurve und Innovationskraft) zu beurteilen (siehe Tabelle 3). Dazu beantworteten die Studierenden Items auf 6-stufigen-Skalen des Likert-Typs („trifft nicht zu“ bis „trifft maximal zu“). Da die Items untereinander hochsignifikant korrelierten (siehe Anhang 1), wurden sie zusätzlich aggregiert und gemittelt, um für jedes Video einen übergreifenden Akzeptanzwert zu erhalten.

Um die Akzeptanzindikatoren einzeln und aggregiert zwischen den (1) Einsatzszenarien und zwischen den (2) Funktionsgraden zu vergleichen, wurden Wilcoxon-Rang-Tests sowie Friedman-Tests mit anschließenden Dunn-Bonferroni Post-hoc-Tests durchgeführt. Mittelwertvergleiche zwischen den (3) Geschlechtern wurden mit Mann-Whitney-U-Tests angestellt. Abschließend wurden die (4) Zusammenhänge zwischen der übergreifenden Akzeptanz und den Variablen Fachsemester, Alter, Computeraffinität („Sie lieben es mit Computern umzugehen.“) und VR-Vorerfahrung („Sie haben bereits in irgendeiner Form Erfahrungen mit VR-Erlebnissen gemacht.“) anhand von Spearman-Korrelationen analysiert.

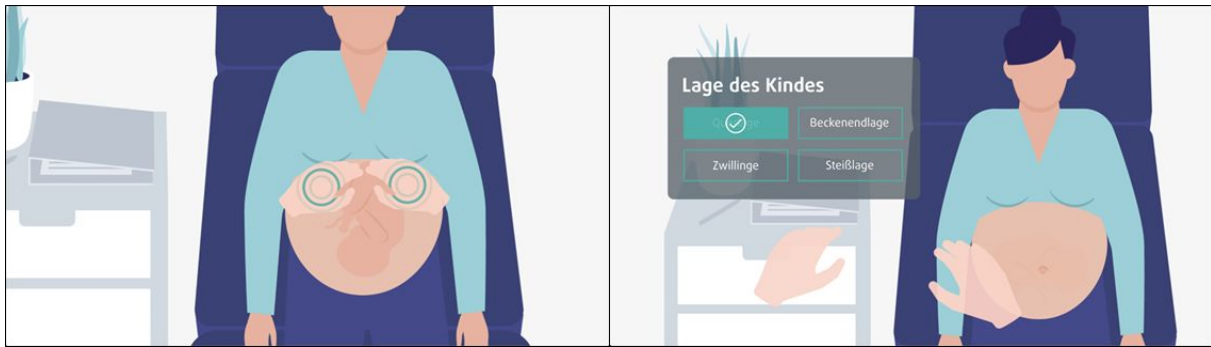


Abbildung 1: Videoscreenshot zum Funktionsgrad Haptik im Lehr- (links) und OSCE-Prüfungsszenario (rechts)

Tabelle 2: YouTube-Links und Dauer der Videos

Funktionsgrad in Einsatzszenarien	YouTube-Link	Dauer in Sekunden
Visuelle Stimulation		
Lehre	https://youtu.be/l4488FovEjU	63
OSCE-Prüfung	https://youtu.be/Mi6ocJg_rOc	37
Audiovisuelle Stimulation		
Lehre	https://youtu.be/tvXegfsa-R4	39
OSCE-Prüfung	https://youtu.be/M4nV1oy5gnY	44
+ Haptik		
Lehre	https://youtu.be/YNU4JhRIAMk	57
OSCE-Prüfung	https://youtu.be/lvaSj_asDB0	41
+ Mündliche Kommunikation		
Lehre	https://youtu.be/EHcRSaVksGg	69
OSCE-Prüfung	https://youtu.be/qplw-ckhccl	32
+ Adaptives Feedback, bzw. Autonome Prüfung		
Lehre	https://youtu.be/F4RjPSg73Ao	44
OSCE-Prüfung	https://youtu.be/ZNxTxyvGiRs	34

Tabelle 3: Indikatoren der Akzeptanz

Akzeptanzindikator	Einsatzszenario	Item ^a
Lernkurve	Lehre	„Ich glaube, dass meine Lernkurve durch die Verwirklichung einer solchen Vision verbessert werden könnte.“
Spaß	Lehre	„Ich glaube mit der Verwirklichung einer solchen Vision hätte ich mehr Spaß am Lernen.“
Fairness	OSCE-Prüfung	„Ich empfinde die Prüfungssituation als objektiver und gerechter als jene in einer Prüfung mit einem Dozenten.“
Simulationsqualität ^b	Lehre & OSCE-Prüfung	„Ich glaube, dass die FUNKTION den Patientenkontakt gut simulieren kann.“
Innovationskraft	Lehre & OSCE-Prüfung	„Ich halte die Entwicklung der FUNKTION in der Lehre für innovativ.“
Sinnhaftigkeit	Lehre & OSCE-Prüfung	„Ich erachte die FUNKTION in der Anwendung insgesamt für sinnvoll.“
Umsetzbarkeit	Lehre & OSCE-Prüfung	„Ich halte die Umsetzung der FUNKTION in der Lehre für realistisch.“

^a „FUNKTION“ ist stellvertretend für den jeweiligen Funktionsgrad.

^b Der Indikator Simulationsqualität wurde beim adaptiven Lehrszenario ausgelassen.

Tabelle 4: Übergreifende Akzeptanz der Funktionsgrade zwischen Lehre und OSCE-Prüfung

Einsatzszenario/ Test	Statistik	Funktionsgrad				
		Visuelle Stimulation	Audiovisuelle Stimulation	+ Haptik	+ mündliche Kommunikation	+ adaptives Feedback/ autonome Prüfung
Lehre	M	4.14	4.45	4.02	3.94	3.90
	SD	1.10	1.08	1.24	1.28	1.31
OSCE-Prüfung	M	3.95	4.13	3.68	3.65	3.45
	SD	1.06	1.08	1.24	1.31	1.30
Wilcoxon-Rang-Test	Z	4.27	8.35	8.83	6.70	7.16
	p	.000***	.000***	.000***	.000***	.000***
	r	.24	.47	.50	.38	.40

Notiz. Die Antwortskala reichte von 1 („trifft nicht zu“) bis 6 („trifft maximal zu“).

*p < .05. **p < .01. ***p < .001

Tabelle 5: Dunn-Bonferroni Post-hoc-Tests zur übergreifenden Akzeptanz zwischen den Funktionsgraden

Lehre			OSCE-Prüfung		
Vergleich von Funktionsgraden	Teststatistik	p ^a	Vergleich von Funktionsgraden	Teststatistik	p ^a
F1 – F2	-.81	.000***	F1 – F2	-.39	.020*
F1 – F3	-.15	1.000	F1 – F3	-.50	.001**
F1 – F4	-.14	1.000	F1 – F4	-.42	.009**
F1 – F5	-.17	1.000	F1 – F5	-.89	.000***
F2 – F3	-.96	.000***	F2 – F3	-.89	.000***
F2 – F4	-.95	.000***	F2 – F4	-.80	.000***
F2 – F5	-.98	.000***	F2 – F5	-1.27	.000***
F3 – F4	-.01	1.000	F3 – F4	-.09	1.000
F3 – F5	-.02	1.000	F3 – F5	-.38	.022*
F4 – F5	-.03	1.000	F4 – F5	-.47	.002**

Notiz. Visuelle Stimulation (F1), audiovisuelle Stimulation (F2), + Haptik (F3), + mündliche Kommunikation (F4), + adaptives Feedback/ autonome Prüfung (F5)

^a Um eine Alphafehler-Kumulierung zu vermeiden, wurde der p-Wert mit der Bonferroni-Methode korrigiert.

*p < .05. **p < .01. ***p < .001

3. Ergebnisse

3.1. Akzeptanz der Funktionsgrade zwischen Lehre und OSCE-Prüfung

Der in den Videos dargestellte Einsatz von VR wurde über alle Funktionsgrade hinweg signifikant mehr in der Lehre als in der OSCE-Prüfung akzeptiert (siehe Tabelle 4).

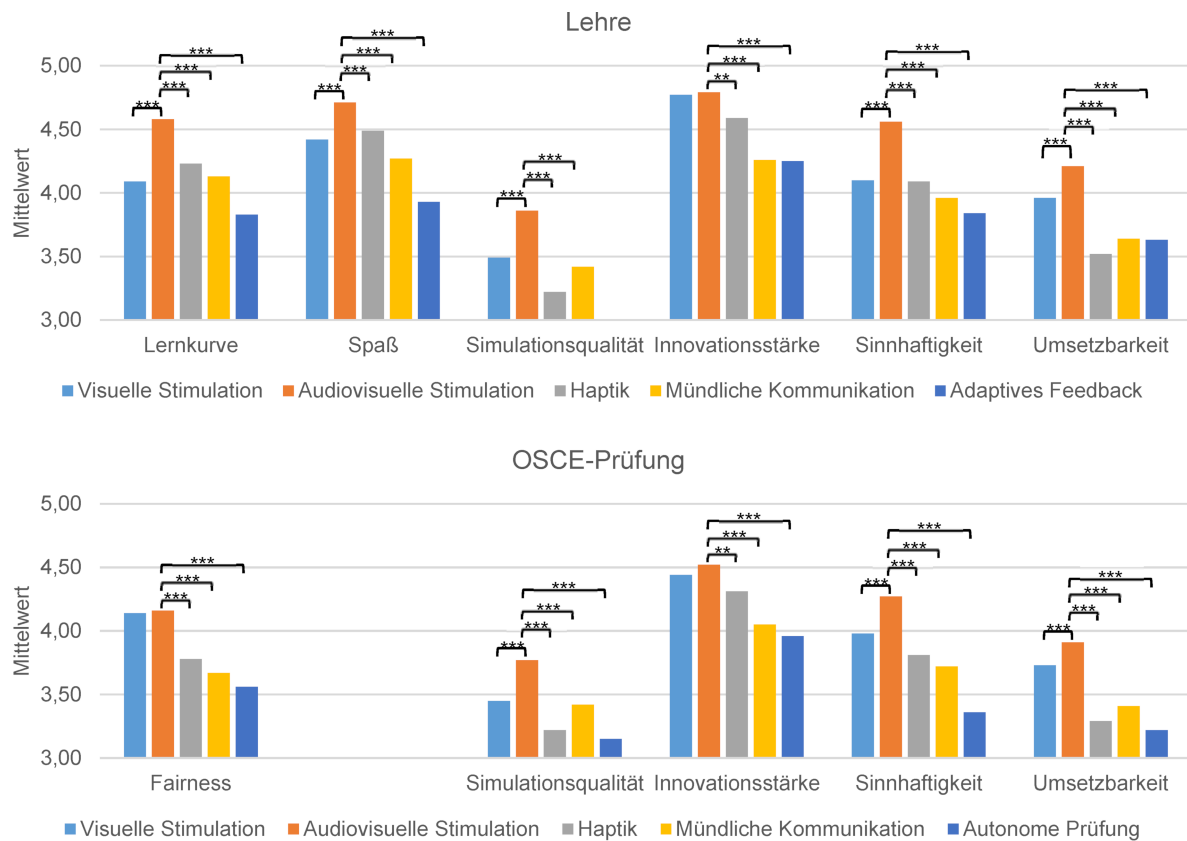
3.2. Akzeptanzunterschiede zwischen den Funktionsgraden

Sowohl in der Lehre ($\chi^2(4)=97.41$, $p<.001$, $n=318$) als auch in der OSCE-Prüfung ($\chi^2(4)=138.07$, $p<.001$, $n=318$) unterschied sich die Akzeptanz von VR signifikant zwischen den Funktionsgraden. Die Post-hoc-Testung (siehe Tabelle 5) präziserte, dass in beiden Einsatzszenarien der audiovisuelle Funktionsgrad signifikant am besten angenommen wurde. Als zweite Präferenz stellte sich die visuelle Stimulation heraus, die jedoch nur in der OSCE-Prüfung signifikant besser beurteilt wurde als die komplexeren Funktionsgrade Haptik, mündliche Kommunikation

und autonome Prüfung. Letzteren haben die Studierenden im Rahmen der OSCE-Prüfung signifikant am geringsten bewertet. Ein differenzierter Blick auf die einzelnen Akzeptanzindikatoren zeigt, dass alle Items bei der audiovisuellen Stimulation signifikant am höchsten ausfielen (siehe Abbildung 2, siehe Anhang 1, Tabelle A1 und Tabelle A2). Die Ausnahme ist die visuelle Stimulation, die bei den Indikatoren Innovationskraft und Fairness nur deskriptiv geringer bewertet wurde.

3.3. Akzeptanzunterschiede zwischen den Geschlechtern

Die männlichen Studierenden wiesen in allen Funktionsgraden und Einsatzszenarien deskriptiv eine höhere, übergreifende Akzeptanz auf als die weiblichen Studierenden. Bei den Videos zum OSCE-Prüfungsszenario wurde dieser Geschlechtsunterschied bei den Funktionsgraden audiovisuelle Stimulation ($U=10048.00$, $Z=-2.00$, $p=.045$, $d_{Cohen}=.23$) und autonome Prüfung ($U=9950.50$, $Z=-2.12$, $p=.034$, $d_{Cohen}=.24$) signifikant. Darüber hinaus fiel auch die Computeraffinität bei den männlichen Stu-



Notiz. Die markierten Ergebnisse der Wilcoxon-Rang-Tests zwischen der audiovisuellen Stimulation und den restlichen Funktionsgraden sind in den angehängten Tabellen A1 und A2 aufgeführt. Die Antwortskala reichte von 1 („trifft nicht zu“) bis 6 („trifft maximal zu“). * $p < .05$. ** $p < .01$. *** $p < .001$

Abbildung 2: Akzeptanzindikatoren nach Funktionsgrad und Einsatzszenario

dierenden höher aus ($U=7206.50$, $Z=-5.79$, $p<.000$, $d_{cohen}=.67$).

3.4. Zusammenhang zwischen übergreifender Akzeptanz und Kontrollfaktoren

Die Akzeptanzwerte korrelieren über die Einsatzszenarien hinweg stark zwischen den Funktionsgraden ($r=.493$ bis $.851$) und moderat mit der individuellen Computeraffinität ($r=.229$ bis $.372$) (siehe Anhang 1, Tabelle A3). Die Kontrollfaktoren Alter, Fachsemester und VR-Vorerfahrung wiesen hingegen keinen signifikanten Zusammenhang mit der studentischen Akzeptanz auf.

4. Diskussion

Das hohe didaktische Potenzial von VR steht im Kontrast zur studentischen Einschätzung, dass die Technologie nur eine relativ geringe Bedeutung für die gegenwärtige und zukünftige Lehre hat. Diese Diskrepanz wurde in der vorliegenden Studie differenziert untersucht, um die technische Weiterentwicklung und curriculare Implementierung von VR zielgruppenorientiert auszurichten. Mit diesem Ziel „vor Augen“ wurde die studentische Akzeptanz zwischen unterschiedlichen (1) Einsatzszenarien und (2) Funktionsgraden verglichen. Außerdem wurde

überprüft, ob bezüglich der Akzeptanz (3) Geschlechterunterschiede oder (4) Zusammenhänge mit den Faktoren Fachsemester, Alter, Computeraffinität und VR-Vorerfahrung bestehen.

Wie in vorangegangenen Umfragen [19], [27] fiel die Akzeptanz für den Einsatz von VR allgemein moderat aus. Mit Verweis auf die angestellten Vergleiche greift diese pauschale Aussage jedoch zu kurz. So zeigte sich über alle Funktionsgrade hinweg, dass VR eher in der Lehre akzeptiert wird als in der OSCE-Prüfung. Dieser Umstand lässt vermuten, dass die Studierenden Zweifel daran haben, ob die neuartige VR-Technologie praktische Leistungserhebungen verlässlich unterstützen oder in Zukunft gar automatisieren kann. Mit Verweis auf das Technology Acceptance Modell [28] lässt sich die geringe Akzeptanz von autonomen VR-Prüfungen vermutlich auf die fehlende Kontrolle über das System zurückführen. Die Vergleiche zwischen den Funktionsgraden offenbaren weitere Vorbehalte gegenüber der Technik. Der marktreife Status Quo von VR (visuelle und audiovisuelle Stimulation) wurde in beiden Einsatzszenarien signifikant besser angenommen als die Funktionen, die sich noch in der Entwicklung befinden (Haptik, mündliche Kommunikation und adaptives Feedback/autonome Prüfung). Medizinstudierende scheinen demnach technologischen Innovationen, die sich noch nicht praktisch bewährt haben, skeptisch gegenüberzustehen. Diese Skepsis ist bei autonomen VR-

Prüfungen und unter weiblichen Studierenden am stärksten ausgeprägt.

Da es bisher wenig empirische Argumente für die visionären VR-Funktionen gibt, ist diese Zurückhaltung besonders in Prüfungsszenarien nachvollziehbar, die unmittelbar für den Studienerfolg relevant sind. Im Detail sind die Ergebnisse jedoch kontraintuitiv. Mit steigendem Funktionsgrad nahmen bei den Studierenden nicht nur die vermutete Umsetzbarkeit, sondern auch die Innovationskraft, Simulationsqualität, Sinnhaftigkeit und Lernkurve ab. Gleiches gilt für die einsatzspezifischen Faktoren Spaß, Lernkurve und Fairness. Diese Angaben stehen im Widerspruch zum Potenzial, das den technischen Neuerungen zugesprochen wird. Während das haptische Feedback vor allem die Immersion und Qualität der Simulation verbessern soll, verspricht die Integration von künstlicher Intelligenz in VR individuelle Lern- und Prüfungserfahrungen mit natürlich anmutenden virtuellen Patienten. Diese Potenziale spiegeln sich nicht in der Einschätzung der Medizinstudierenden wider. Ein Grund für diese kontraintuitive Bewertung mag die menschliche Tendenz sein, das Gewöhnliche (audiovisuelle Stimulation) anzunehmen und das Unbekannte (z.B. autonome Leistungsbeurteilung) zu meiden [29]. Diese Grundeinstellung zeigt sich bereits wesentlich im positiven Zusammenhang zwischen Akzeptanz und Computeraffinität. Da sich der Einsatz von haptischem Feedback und KI noch in der Entwicklung befindet und somit grundsätzlich unbekannt ist, lässt sich die studentische Skepsis gegenüber den visionären Funktionsgraden möglicherweise auf fehlende Erfahrungswerte zurückführen.

Im gleichen Sinne lässt sich womöglich auch die verhaltene Akzeptanz des marktreifen Status Quo erklären. In einer studiengangübergreifenden Umfrage von Weisflog und Böckel [19] erachteten Studierende VR für ihr Studium wichtiger, wenn ihnen die Technologie an ihrer Hochschule zur Verfügung stand. 2019 traf das in Deutschland jedoch nur auf etwa 4% der Studierenden zu. Auch im privaten Bereich beschränkt sich der Einsatz von VR – trotz steigendem Hard- und Software-Absatz – derzeit noch auf eine technikaffine Minderheit [<https://www2.deloitte.com/de/de/pages/presse/contents/zukunftsperspektiven-fuer-virtual-augmented-reality.html>]. Folglich hatte bisher nur ein kleiner Teil der Studierenden die Möglichkeit, den gegenwärtigen Stand der VR-Technik praktisch anzuwenden und mögliche Vorbehalte abzubauen. Dieser Erklärungsansatz wird jedoch nicht von der in der Studie gemessenen VR-Vorerfahrung gestützt. Das entsprechende Item hat vermutlich zu wenig Aussagekraft über den individuellen Kenntnisstand der VR-Technologie, da die Qualität und Quantität der abgefragten VR-Erlebnisse nicht erhoben wurden.

Bemerkenswert ist auch die durchweg hohe Korrelation zwischen den Akzeptanzindikatoren (siehe Anhang 1, Tabelle A3). Die hohe interne Konsistenz legt nahe, dass viele Medizinstudierende eine gefestigte Meinung zum VR-Einsatz haben, die die Antworten über die Einsatzszenarien und Funktionsgrade hinweg bedingt hat. Die Übertragbarkeit dieser Erkenntnisse muss jedoch kritisch

diskutiert werden. Die geringe Rücklaufquote von 7,7% in Ulm schränkt die Repräsentativität der Stichprobe ein. Außerdem bezogen sich die Videodemonstrationen ausschließlich auf die Geburtshilfe, sodass die Akzeptanzwerte nur indirekt auf andere Anwendungsfelder schließen lassen. Letztlich muss bei der Interpretation der Ergebnisse auch berücksichtigt werden, dass die Studierenden nur Videos über den Einsatz von VR bewerteten, da die höheren Funktionsgrade noch keine Marktreife erreicht haben. Entsprechend der Annahme, dass die Akzeptanz von VR mit steigendem Bekanntheitsgrad zunimmt, sollte die Bewertung bei einer praktischen Demonstration positiver ausfallen.

Die Studienergebnisse verdeutlichen nichtsdestotrotz, dass bei der Weiterentwicklung und curricularen Implementierung von VR die Bedenken und Technikenkenntnisse der Studierenden berücksichtigt werden müssen. Damit aus dem Potenzial von VR reale Lerneffekte resultieren, sollte die Technologie nicht nur als Lehrmedium sondern auch als Lerngegenstand im Medizinstudium eingeführt werden. Die Möglichkeit, die Anwendung von VR angeleitet zu erlernen und kritisch zu diskutieren, fördert auch die Medienkompetenz angehender Ärzte, die im zukünftigen Klinikalltag womöglich mit VR arbeiten werden (z.B. simulationsbasierte Planung von chirurgischen Eingriffen). Ob sich VR in Zukunft auch für die Durchführung praktischer Prüfungen anbietet oder die studentische Skepsis künftig weiterhin begründet ist, sollte wissenschaftlich geprüft werden, sobald sich die Technologie in der Lehre etabliert und als verlässlich erwiesen hat.

5. Schlussfolgerung

Die differenzierte Untersuchung der studentischen VR-Akzeptanz offenbarte, dass sich die subjektiv gering wahrgenommene Bedeutung von VR in einer Skepsis gegenüber aufkommenden Technologien begründet. Medizinstudierende scheinen zu wenig Kenntnisse über VR zu haben, um dessen didaktisches Potenzial adäquat einzuschätzen. Damit die angehenden Ärzte die Technologie kennenlernen und zukünftig kompetent im Klinikalltag anwenden können, sollte VR nicht nur als Lehrmedium sondern auch als Lerngegenstand curricular eingeführt werden. Die Durchführung praktischer Prüfungen in VR empfiehlt sich dagegen erst, wenn sich die Technologie in der Lehre als verlässlich erwiesen hat.

Autoren

Die Autoren S. Walter und R. Speidel teilen sich die Erstautorenschaft.

Interessenkonflikt

Die Autor*innen erklären, dass sie keinen Interessenkonflikt im Zusammenhang mit diesem Artikel haben.

Anhänge

Verfügbar unter

<https://www.egms.de/de/journals/zma/2021-38/zma001496.shtml>

1. Anhang_1.pdf (110 KB)
Anhang 1

Literatur

1. Boshuizen H, Schmidt H. The development of clinical reasoning expertise. *Clin Reason Health Prof.* 2008;3:113-121.
2. Keifenheim KE, Teufel M, Ip, J, Speiser N, Leehr EJ, Zipfel S, Herrmann-Werner A. Teaching history taking to medical students: A systematic review. *BMC Med Educ.* 2015;15:159. DOI: 10.1186/s12909-015-0443-x
3. Kiesewetter J, Kager M, Lux R, Zwissler B, Fischer MR, Dietz I. German undergraduate medical students' attitudes and needs regarding medical errors and patient safety - a national survey in Germany. *Med Teach.* 2014;36(6):505-510. DOI: 10.3109/0142159X.2014.891008
4. Cook DA. How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. *Med Educ.* 2014;48(8):750-760. DOI: 10.1111/medu.12473
5. Kononowicz AA, Woodham LA, Edelbring S, Stathakarou N, Davies D, Saxena N, Tudor Car L, Carlsted-Duke J, Car J, Zary N. Virtual patient simulations in health professions education: Systematic review and meta-analysis by the digital health education collaboration. *J Med Internet Res.* 2019;21(7):e14676. DOI: 10.2196/14676
6. Huwendiek S, De Ieng BA, Zary N, Fischer MR, Ruiz JG, Ellaway R. Towards a typology of virtual patients. *Med Teach.* 2009;31(8):743-748. DOI: 10.1080/01421590903124708
7. Bokken L, Rethans JJ, Scherpier AJ, van der Vleuten CP. Strengths and weaknesses of simulated and real patients in the teaching of skills to medical students: a review. *Simul Healthc.* 2008;3(3):161-169. DOI: 10.1097/SIH.0b013e318182fc56
8. Cook DA, Hamstra SJ, Brydges R, Zendejas B, Szostek JH, Wang AT, Erwin PJ, Hatala R. Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Med Teach.* 2013;35(1):867-898. DOI: 10.3109/0142159X.2012.714886
9. Shorey S, Ang E, Yap J, Ng ED, Lau ST, Chui CK. A Virtual Counseling Application Using Artificial Intelligence for Communication Skills Training in Nursing Education: Development Study. *J Med Internet Res.* 2019;21(10):e14658. DOI: 10.2196/14658
10. Talbot TB, Rizzo AS, Dubbels BR. Virtual standardized patients for interactive conversational training: A grand experiment and new approach. *Hershey, PA: IGI Global;* 2019. DOI: 10.4018/978-1-5225-7461-3.ch003
11. Slater M, Wilbur S. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence.* 1997;6(6):603-616. DOI: 10.1162/pres.1997.6.6.603
12. Chirico A, Gaggioli A. When Virtual Feels Real: Comparing Emotional Responses and Presence in Virtual and Natural Environments. *Cyberpsychol Behav Soc Netw.* 2019;22(3):220-226. DOI: 10.1089/cyber.2018.0393
13. Kolb DA. *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, N.J.: Prentice-Hall; 1984.
14. Kwon C. Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR technologies. *Virt Reality.* 2019;23(1):101-118. DOI: 10.1007/s10055-018-0364-1
15. Chavez B, Bayona S. Virtual Reality in the Learning Process. In: Rocha Á, Adeli H, Reis LP, Costanzo S, editors. *Advances in Intelligent Systems and Computing. Trends and Advances in Information Systems and Technologies.* Cham: Springer International Publishing; 2018. p.1345-1356. DOI: 10.1007/978-3-319-77712-2_129
16. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, Divakar U, Masiello I, Kononowicz AA, Zary N, Tudor Car L. Virtual Reality for Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. *J Med Internet Res.* 2019;21(1):e12959. DOI: 10.2196/12959
17. Pourmand A, Davis S, Lee D, Barber S, Sikka N. Emerging Utility of Virtual Reality as a Multidisciplinary Tool in Clinical Medicine. *Games Health J.* 2017;6(5):263-270. DOI: 10.1089/g4h.2017.0046
18. Cipresso P, Giglioli IAC, Raya MA, Riva G. The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature. *Front Psychol.* 2018;9:2086. DOI: 10.3389/fpsyg.2018.02086
19. Weisflog W, Böckel A. Ein studentischer Blick auf den Digital Turn: Auswertung einer bundesweiten Befragung von Studierenden für Studierende. Arbeitspapier Nr. 54. Berlin: Stifterverband für die Deutsche Wissenschaft, Hochschulforum Digitalisierung; 2020. Zugänglich unter/available from: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_54_Studierendenbefragung.pdf
20. Kuhn S, Frankenhauser S, Tolks D. Digitale Lehr- und Lernangebote in der medizinischen Ausbildung: Schon am Ziel oder noch am Anfang? [Digital learning and teaching in medical education : Already there or still at the beginning?]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2018;61(2):201-209. DOI: 10.1007/s00103-017-2673-z
21. Pottle J. Virtual reality and the transformation of medical education. *Future Healthc J.* 2019;6(3):181-185. DOI: 10.7861/fhj.2019-0036
22. Zender R, Weise M, von der Heyde M, Söbke H. Lehren und Lernen mit VR und AR - Was wird erwartet? Was funktioniert? In: Krömker D, Schroeder U, editors. *DeLFI 2018 - Die 16. E-Learning Fachtagung Informatik.* Bonn: Gesellschaft für Informatik e.V.; 2018. p.275-276.
23. Frithioff A, Freundø M, Mikkelsen PT, Sørensen MS, Andersen SA. Ultra-high-fidelity virtual reality mastoidectomy simulation training: a randomized, controlled trial. *Eur Arch Otorhinolaryngol.* 2020;277(5):1335-1341. DOI: 10.1007/s00405-020-05858-3
24. Rangarajan K, Davis H, Pucher PH. Systematic Review of Virtual Haptics in Surgical Simulation: A Valid Educational Tool? *J Surg Educ.* 2020;77(2):337-347. DOI: 10.1016/j.jsurg.2019.09.006
25. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS ONE.* 2020;15(2):e0229596. DOI: 10.1371/journal.pone.0229596

26. Veneziano D, Cacciamani G, Rivas JG, Marino NS, Bhaskar K. VR and machine learning: novel pathways in surgical hands-on training. *Curr Opin Urol.* 2020;30(6):817-822. DOI: 10.1097/MOU.0000000000000824
27. Speidel R, Schneider A, Koerner J, Grab-Kroll C, Oechsner W. Did video kill the XR star? Digital trends in medical education before and after the COVID-19 outbreak from the perspective of students and lecturers from the faculty of medicine at the University of Ulm. *GMS J Med Educ.* 2021;38(6):Doc101. DOI: 10.3205/zma001497
28. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. *J Biomed Inform.* 2010;43(1):159-172. DOI: 10.1016/j.jbi.2009.07.002
29. Castro D, McQuinn A. The Privacy Panic Cycle: A Guide to Public Fears About New Technologies. Washington, DC: The Information Technology & Innovation Foundation; 2015. p.39. Zugänglich unter/available from: <http://www2.itif.org/2015-privacy-panic.pdf>

Bitte zitieren als

Walter S, Speidel R, Hann A, Leitner J, Jerg-Bretzke L, Kropp P, Garbe J, Ebner F. Skepticism towards advancing VR technology – student acceptance of VR as a teaching and assessment tool in medicine. *GMS J Med Educ.* 2021;38(6):Doc100.
DOI: 10.3205/zma001496, URN: urn:nbn:de:0183-zma0014967

Artikel online frei zugänglich unter

<https://www.egms.de/en/journals/zma/2021-38/zma001496.shtml>

Eingereicht: 27.01.2021

Überarbeitet: 17.05.2021

Angenommen: 25.06.2021

Veröffentlicht: 15.09.2021

Copyright

©2021 Walter et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.

Korrespondenzadresse:

Steffen Walter
Universitätsklinikum Ulm, Klinik für Psychosomatische
Medizin und Psychotherapie, Sektion Medizinische
Psychologie, Frauensteige 6, 89075 Ulm, Deutschland
steffen.walter@uni-ulm.de