Medical and Health Care Applications, Vol. 211, 2026, 133–142
https://doi.org/10.54941/ahfe1007476

Comparing Human- and Machine-Guided Virtual Reality Training: The Role of Physiological Stress in Learning Outcomes

Dilan Cabuk-Colak[1], Mehmet Emin Aksoy[2,3], Dilek Kitapcioglu[3,4], Tuba Usseli[5], and Hayrettin Can Südor[2]

[1]Department of Psychology, Faculty of Humanities and Social Sciences, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey
[2]Department of Biomedical Device Technology, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey
[3]CASE (Center of Advanced Simulation and Education), Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey
[4]Department of Medical Education, School of Medicine, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey
[5]Program of Anesthesia, Vocational School of Health Services, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey

ABSTRACT

Virtual reality (VR) is increasingly used in medical education due to its capacity to provide immersive, standardized, and scalable training environments, yet direct comparisons between human-guided and machine-guided VR instruction remain limited, particularly regarding the role of learners' physiological responses. The present study compared human-guided (IG) and machine-guided (MG) VR-based Basic Life Support (BLS) training and examined whether physiological stress responses during training and examination phases moderated learning outcomes. Fifty-five undergraduate students completed a VR BLS session under either IG (N = 25) or MG (N = 30), followed by a VR-based exam assessing learning outcomes. Electrodermal activity (EDA) was recorded continuously as an index of physiological stress, with mean EDA values computed separately for the training and exam phases, and participants also reported their sense of presence in the virtual environment. Independent samples t-tests indicated no significant group differences in physiological stress during either the training or exam phases, suggesting comparable levels of physiological activation across instructional modalities. In contrast, participants in IG reported a significantly higher sense of presence than those in MG. An ANCOVA controlling for presence and stress levels during both phases revealed a significant main effect of instructional group on exam performance, with participants in IG achieving higher scores than those in MG, while none of the covariates significantly predicted performance. These findings indicate that the benefits of human-guided VR instruction extend beyond differences in average physiological arousal or subjective presence.

Keywords: Basic life support, Virtual reality, Stress, Learning

Received February 6, 2026; Revised April 2, 2026; Accepted April 18, 2026; Available online July 20, 2026

INTRODUCTION

Learning complex procedural skills requires repeated practice, timely feedback, and opportunities for error adjustment (Burgess et al., 2020; Metcalfe, 2017). Educational contexts that involve high-stakes procedures, such as medical emergencies, therefore face the challenge of balancing effective learning with safety, instructional standardization, and practical constraints. Although theory- and simulation-based training methods in medical education remain foundational (van der Vlugt et al., 2002), opportunities for hands-on practice are often limited, especially for emergency procedures that occur infrequently yet demand rapid and accurate responses when they do arise (Schonnop et al., 2020). From a learning perspective, these constraints may restrict effective skill acquisation due to limited opportunities for feedback-driven refinement, and consolidation of procedural knowledge through repeated exposure to the learning environment.

Virtual reality (VR) has emerged as a promising response to these challenges for effective learning in medical training and is now increasingly integrated into educational contexts that require experiential and applied learning. By providing immersive and interactive environments, VR simulations allow learners to repeatedly practice complex skills without risk to real patients while preserving key perceptual, temporal, and contextual features of real-world settings (Weller et al., 2012). Studies suggest that VR can effectively improve procedural accuracy (Logishetty et al., 2019) and learner confidence (Tay el al., 2025). Beyond performance outcomes, it is possible to offer a high degree of instructional standardization with VR-based training, ensuring that all learners are exposed to identical scenarios, task demands, and assessment criteria. This level of control is particularly valuable, as it reduces variability in instructional input that can otherwise obscure causal relationships between learning conditions and outcomes.

Among the various applications of VR-based training, Basic Life Support (BLS) represents a compelling case for studying learning under pressure. BLS procedures require maintenance of airway clearance, support of breathing and circulation, in the absence of advanced medical equipment. VR-based BLS training inherently enables repeated rehearsal of these procedures within realistic emergency scenarios, providing an ideal context for examining how VR-design influences learning, retention, and performance under stressful conditions.

As VR technologies continue to evolve, new questions arise as to how VR-based instruction should be designed to maximize learning outcomes. While early VR applications can rely on instructors to guide learners through virtual scenarios, recent technological advances have enabled the development of automated, machine-guided instruction within the virtual environment (Cabuk-Colak et al., 2026; Kitapcioglu et al., 2024). These systems can deliver prompts, feedback, and performance evaluations without the presence of a human instructor, raising important questions about the equivalence of different guidance modalities in terms of learning outcomes.

Despite the rapid expansion of VR-based education with machine-guidance, direct experimental comparisons between human-guided and

machine-guided VR instruction remain limited, and the mechanisms through which different guidance modalities influence learning are not yet well understood. Human instructors may provide dynamic interpersonal cues during the learning phase. Moreover, they can dynamically adapt instruction by modifying feedback, pacing, or explanations in response to learners' momentary needs. Such adaptivity may be especially important in immersive VR environments, where learners must coordinate perceptual, cognitive, and motor processes in real time. In contrast, relying on predefined instructional rules, machine-guided systems may fail to provide these advantages.

Learning in high-stakes contexts is also closely intertwined with stress. Stress can influence learning in complex ways: through impairing working memory processes, affecting cognitive flexibility, or prompting intuitive thinking (Goldfarb et al., 2017; Schoofs et al., 2008; Starcke & Brand, 2016). Moreover, continuous exposure to stress may obscure learning (LeBlanc, 2009). In immersive training environments, understanding how stress unfolds across instructional phases is therefore essential for interpreting learning outcomes.

In addition to physiological responses, subjective experience within VR environments may also be an important factor shaping learning outcomes. One possible experience is presence, defined as the subjective sense of "being there" in a virtual environment (Witmer & Singer, 1998). While the relationship between presence and learning outcomes remains an open question in the literature, social elements within human-guided trainings in VR settings, even when subtle, may enhance perceived presence by increasing realism and interpersonal connection. Examining presence alongside physiological measures, therefore, may therefore clarify its role in immersive learning.

The present study examined whether instructional guidance modality influences learning outcomes in VR-based training and whether physiological stress responses moderate this relationship. Undergraduate participants completed a VR-based BLS training under either human-guided or machine-guided instruction, followed by a VR-based standardized exam. Electrodermal activity was recorded continuously to assess physiological stress during both training and exam phases, and participants reported their sense of presence following the session. By integrating behavioural performance, subjective experience, and physiological measures, this study adopts a multimodal approach to comparing human- and machine-guided VR instruction.

METHODS

Participants

56 participants completed a VR-based BLS training session under either a human-guided (instructor-guided; IG) or a machine-guided (MG) group. Participants were randomly assigned to groups using a coin-toss procedure.

All participants were second-year anesthesiology students who had received formal theoretical instruction in BLS protocols as part of their curriculum but had no prior experience with VR-based training. Eligibility criteria required participants to be currently enrolled in the anesthesiology

program, fluent in Turkish—the language used for instruction and within the VR interface—and physically able to use head-mounted VR equipment. Participants were excluded if they had previously completed VR-based BLS training, reported a history of pronounced motion sickness related to VR, had significant visual or motor impairments, or had medical conditions (e.g., vertigo) or medication use that could interfere with safe participation in immersive virtual environments.

One participant in the IG group did not complete the post-training measures and was therefore excluded from the analyses. The final sample consisted of 25 participants in the IG group and 30 participants in the MG group. All participants took part voluntarily, and written informed consent was obtained prior to participation.

Materials

VR-Based Training Module

The VR-based training module was designed to support the acquisition of core BLS skills, including chest compressions, airway management, and rescue breathing. The module presented a single-patient emergency scenario in which participants were required to perform BLS procedures in a stepwise manner, following internationally established resuscitation guidelines. The simulation was developed through collaboration with the Center of Advanced Simulation and Education (CASE) at Acıbadem Mehmet Ali Aydınlar University and has been used and validated in prior research employing similar instructional contexts (Aksoy 2019, 2020; Kitapcioglu et al., 2025).

Electrodermal Activity

Electrodermal activity (EDA) was employed as a physiological indicator of stress during the simulation task. EDA captures changes in skin conductance driven by sympathetic nervous system activity and is widely regarded as a sensitive measure of stress responses (Boucsein, 2012). Its suitability for evaluating stress in immersive and technology-supported learning environments, including VR and simulation-based training, has also been demonstrated in prior research (Chiossi et al., 2022).

EDA data were collected using a Shimmer3 GSR+ sensor (Shimmer, Dublin, Ireland), with electrodes attached to the palmar surface of the index and middle fingers of the participant's left hand. Recordings were obtained in a controlled environment at a sampling rate of 256 Hz and streamed wirelessly via Consensys software. To capture tonic stress-related activity, the raw signal was low-pass filtered and downsampled, yielding skin conductance level (SCL) measures for subsequent analyses.

Presence Questionnaire

Participants' perceived sense of presence in the virtual environment was assessed using the Turkish version of the Presence Questionnaire (PQ). The PQ is a well-established self-report instrument for measuring subjective presence in VR experiences (Witmer & Singer, 1998). The validated Turkish

adaptation consists of 29 items rated on a 7-point Likert scale (Gökoğlu & Çakıroğlu, 2019). Overall presence scores were calculated by averaging responses across all items.

Data Analyses

All statistical analyses were conducted to examine group differences in physiological stress, subjective presence, and learning outcomes, as well as to test whether stress-related measures moderated the relationship between instructional guidance and performance.

To compare physiological stress levels between groups, independent-samples t-tests were performed on mean EDA values computed separately for the training phase and the exam phase. Mean EDA values were calculated by averaging continuous EDA recordings within each phase for each participant. An additional independent-samples t-test was conducted to compare self-reported presence scores between the IG and MG groups.

To examine learning outcomes while accounting for individual differences in subjective experience and physiological activation, an analysis of covariance (ANCOVA) was conducted with Exam Score as the dependent variable and Group (IG vs. MG) as the between-subjects factor. Presence, mean EDA during the training phase, and mean EDA during the exam phase were included as covariates to assess whether these variables were associated with performance and whether group differences in exam scores persisted after statistically controlling for them.

All statistical tests were two-tailed, and the alpha level was set at .05.

Ethical Considerations

The study was conducted in compliance with established ethical principles governing research with human participants. Ethical approval was granted by the Scientific Ethics Committee of Acıbadem Mehmet Ali Aydınlar University (Approval No. 2024-17/647) prior to data collection. All participants were informed about the study procedures, their rights as participants, and the voluntary nature of their involvement, and written informed consent was obtained before participation. Participant anonymity was protected by assigning coded identifiers and removing any personally identifiable information from the dataset. Participation was voluntary, and no monetary or academic incentives were offered.

Procedure

The study consisted of four sequential phases conducted within a single session: orientation, training, assessment, and post-experience measures.

Orientation Phase: All participants first completed a brief orientation session designed to familiarize them with the VR equipment and interaction mechanics. Using a separate introductory module, participants were introduced to the head-mounted display, hand controllers, and basic navigation within the virtual environment. This phase aimed to minimize novelty effects and ensure that any possible performance differences were not attributable to unfamiliarity with the technology.

Training Phase: Following orientation, participants entered the training phase, where they completed the same VR-based cardiac arrest simulation. Participants were randomly assigned to either the IG group or the MG group. The simulation content, task sequence, and difficulty level were identical across groups; the only difference concerned the source and format of instructional guidance.

In the MG group, guidance was delivered entirely by the VR system through automated prompts and feedback, providing step-by-step instructions without any human involvement. This setup was designed to ensure maximal standardization of instruction across participants.

In contrast, participants in the IG group were guided by an experienced human instructor with extensive expertise in medical education. The instructor observed participants' actions in real time through a shared virtual interface and provided verbal feedback during the simulation. To maintain consistency, the instructor followed a predefined guidance protocol aligned with the automated prompts used in the MG group.

Throughout the training phase, EDA was continuously recorded to assess physiological stress responses during learning.

Assessment Phase: Immediately after completing the training, participants undertook a VR-based exam within the same virtual environment. The exam assessed procedural knowledge and task execution related to the training content. Performance was automatically scored by the VR system on a 100-point scale, ensuring objective and consistent evaluation across groups. EDA recording continued during the exam phase to capture stress responses during task retrieval and performance under evaluation.

Post-Experience Measures: Upon completion of the exam, participants immediately completed the Presence Questionnaire via Qualtrics to assess their subjective sense of presence during the virtual experience. The questionnaire was administered without delay to reduce memory-related biases. Afterward, participants were debriefed, thanked for their participation, and given the opportunity to ask questions about the study procedures.

RESULTS

Stress Levels During Training and Examination

Independent-samples t-tests were conducted to examine whether physiological stress levels differed between the instructor-guided (IG) and machine-guided (MG) groups during the training and exam phases. The analyses revealed no significant group differences in mean stress levels during the training phase, $t(53) = -0.58$, $p = .57$, nor during the exam phase, $t(53) = -0.74$, $p = .46$. These results indicate that participants in the IG and MG groups experienced comparable levels of physiological activation across both phases of the VR task.

Presence

Group differences in perceived presence were examined using an independent-samples t-test. The analysis showed that participants in the IG group reported

significantly higher presence scores than those in the MG group, $t(53) = 2.11$, $p = .04$, indicating a stronger subjective sense of immersion and engagement when instruction was provided by a human instructor.

Exam Performance

To examine whether instructional guidance influenced learning outcomes while accounting for individual differences in stress and presence, an ANCOVA was conducted with exam score as the dependent variable, Group (IG vs. MG) as the between-subjects factor, and Presence, Stress During Training, and Stress During Exam as covariates. The ANCOVA revealed a significant main effect of Group, $F(1, 50) = 6.18$, $p = .02$, indicating that participants in the IG group achieved higher exam scores than those in the MG group.

None of the covariates significantly predicted exam performance (Presence: $p = .16$; Stress During Training: $p = .97$; Stress During Exam: $p = .98$). Thus, physiological stress levels during either the learning or assessment phase were not associated with exam performance, and the instructional advantage of the human-guided condition remained significant after statistically controlling for these factors.

DISCUSSION

The present study investigated whether the modality of guidance in immersive VR—human-guided versus machine-guided—affects learning outcomes in BLS, and whether physiological stress responses and presence contribute to or moderate these effects. Three main findings emerged. First, participants trained with a human instructor outperformed those trained with machine-guidance on the VR-based exam. Second, although the human-guided group reported a higher sense of presence, presence did not statistically predict exam performance. Third, physiological stress levels, indexed by EDA, did not differ between groups and were not associated with learning outcomes during either the training or exam phases.

The observed performance advantage for the instructor-guided group suggests that human guidance can confer learning benefits even within highly standardized VR environments. Notably, this advantage remained after statistically controlling for presence and physiological stress, indicating that it cannot be attributed to greater immersion or differences in arousal alone. Because instructional content and sequencing were held constant across groups, the advantage of human guidance is unlikely to reflect differences in instructional strategy or adaptivity. Instead, it may arise from subtle social and paralinguistic cues inherent to human instruction, such as vocal tone, timing, emphasis, or perceived responsiveness, which were not experimentally controlled. Indeed, contemporary perspectives in educational psychology emphasize that interpersonal features of teaching can meaningfully shape learners' engagement and processing of information (Collie & Martin, 2017). These findings suggest that human instructors may influence learning through mechanisms that extend beyond explicit instruction, highlighting the need for future research to isolate and systematically manipulate social cues within VR-based learning environments.

Although participants in the instructor-guided group reported higher presence, presence did not predict exam performance, nor did it account for the group difference in learning outcomes. This pattern is consistent with prior work suggesting that presence may enhance engagement (see Wei et al., 2025) without necessarily translating into superior performance (Smith & Mulligan, 2021). In other words, learners can feel highly immersed without optimally encoding procedural rules or decision sequences. The current findings therefore contribute to ongoing debates about the role of presence in VR-based learning, suggesting that instructional factors may be more critical determinants of learning than immersion alone.

The absence of group differences in physiological stress and the lack of association between stress and exam performance suggest that both instructor-guided and machine-guided VR training elicited comparable levels of autonomic activation during both learning and retrieval. One possible interpretation is that the immersive and time-critical nature of the BLS simulation itself was sufficient to induce a moderate level of physiological arousal across participants, leaving limited scope for additional modulation by the source of guidance. It is also plausible that learners' subjective experience of stress, rather than objective physiological activation, plays a more central role in retention or transfer to real-world performance—outcomes that were not directly assessed in the present study. Future research would therefore benefit from jointly examining subjective stress appraisals alongside physiological measures to clarify how perceived and physiological stress interact to shape learning in immersive training environments.

Taken together, the findings underscore that the educational value of VR does not reside solely in technological immersion or automation. While machine-guided systems offer clear advantages in scalability and standardization, the current results suggest that human instructors provide added value for learning in high-stakes procedural medical contexts. This has important implications for the design of VR-based training programs. Rather than viewing human and machine-guidance as mutually exclusive, hybrid instructional models may be particularly promising—leveraging automation for standardization while preserving human adaptivity at critical moments in the learning process.

Several limitations should be acknowledged. The sample consisted of undergraduate students with similar educational backgrounds, which may limit generalizability to expert populations or practicing clinicians. Additionally, learning outcomes were assessed immediately after training, precluding conclusions about long-term retention. Future research should examine delayed testing, transfer to physical task performance, and the potential role of adaptive machine-guided systems that respond to learners' physiological or behavioral signals in real time.

In conclusion, this study contributes to the understanding of learning in immersive VR environments by demonstrating that instructional guidance modality influences performance independently of presence and physiological stress. This finding highlights the importance of socially mediated instruction—even in technologically advanced training environments—and emphasize that effective VR-based learning depends not only on immersion, but on how instructional support is designed to align with fundamental principles of human learning.

REFERENCES

Aksoy, E. (2019). Comparing the Effects on Learning Outcomes of Tablet-Based and Virtual Reality–Based Serious Gaming Modules for Basic Life Support Training: Randomized Trial. *JMIR Serious Games, 7*(2), e13442. https://doi.org/10.2196/13442

Aksoy, M. E. (2020). Comparing Basic Life Support Serious Gaming Scores With Hands-on Training Platform Performance Scores: Pilot Simulation Study for Basic Life Support Training. *JMIR Serious Games, 8*(4), e24166. https://doi.org/10.2196/24166

Boucsein, W. (2012). Electrodermal Activity, Second Edition. *Electrodermal Activity, Second Edition*, 1-+. https://doi.org/10.1007/978-1-4614-1126-0

Burgess, A., Van Diggele, C., Roberts, C., & Mellis, C. (2020). *Tips for teaching procedural skills*. BMC Medical Education, *20*. https://doi.org/10.1186/s12909-020-02284-1.

Cabuk-Colak, D., Aksoy, M. E., Kitapcioglu, D., Usseli, T., Ozkan, A. E., Sudor, H. C., & Yoner, S. I. (2025, June). Pulse of Learning: Complexity, Guidance, and Performance in VR-Based Serious Gaming for Medical Training. In *International Conference on Human-Computer Interaction* (pp. 276–284). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-12808-9_17

Chiossi, F., Welsch, R., Villa, S., Chuang, L., & Mayer, S. (2022). Virtual Reality Adaptation Using Electrodermal Activity to Support the User Experience. *Big Data and Cognitive Computing, 6*(2).

Collie, R. J., & Martin, A. J. (2017). Teachers' sense of adaptability: Examining links with perceived autonomy support, teachers' psychological functioning, and students' numeracy achievement. *Learning and Individual Differences, 55*, 29–39. https://doi.org/https://doi.org/10.1016/j.lindif.2017.03.003

Goldfarb, E. V., Froböse, M. I., Cools, R., & Phelps, E. A. (2017). Stress and Cognitive Flexibility: Cortisol Increases Are Associated with Enhanced Updating but Impaired Switching. *Journal of Cognitive Neuroscience, 29*(1), 14–24. https://doi.org/10.1162/jocn_a_01029

Gökoğlu, S., & Çakıroğlu, Ü. (2019). SANAL GERÇEKLİK TEMELLİ ÖĞRENME ORTAMLARINDA BULUNUŞLUK HİSSİNİN ÖLÇÜLMESİ: BULUNUŞLUK ÖLÇEĞİNİN TÜRKÇE'YE UYARLANMASI. *Eğitim Teknolojisi Kuram ve Uygulama, 9*(1), 169–188. https://doi.org/10.17943/etku.441497

Kitapcioglu, D., Aksoy, M. E., Ozkan, A. E., & Usseli, T. (2024). Comparing Learning Outcomes of Machine-Guided Virtual Reality–Based Training With Educator-Guided Training in a Metaverse Environment: Randomized Controlled Trial. *JMIR Serious Games, 12(1)*, e58654. https://doi.org/10.2196/58654

LeBlanc, V. R. (2009). The Effects of Acute Stress on Performance: Implications for Health Professions Education. *Academic Medicine, 84*(10).

Logishetty, K., Rudran, B., & Cobb, J. (2019). Virtual reality training improves trainee performance in total hip arthroplasty: a randomized controlled trial. *The bone & joint journal, 101* (12), 1585–1592 . https://doi.org/10.1302/0301-620x.101b12.bjj-2019-0643.r1.

Metcalfe, J. (2017). *Learning from Errors*. Annual Review of Psychology, 68, 465–489. https://doi.org/10.1146/annurev-psych-010416-044022.

Schonnop, R., Stauffer, B., Gauri, A., & Ha, D. (2020). Procedural Skills Training in Emergency Medicine Physicians Within the Edmonton Zone: A Needs Assessment. *AEM Education and Training, 5*, 1–11. https://doi.org/10.1002/aet2.10495.

Schoofs, D., Preuss, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology, 33*(5), 643–653. https://doi.org/10.1016/j.psyneuen.2008.02.004

Smith, S. A., & Mulligan, N. W. (2021). Immersion, presence, and episodic memory in virtual reality environments. *Memory, 29*(8), 983–1005. https://doi.org/10.1080/09658211.2021.1953535

Starcke, K., & Brand, M. (2016). Effects of stress on decisions under uncertainty: A meta-analysis. *Psychological Bulletin, 142*(9), 909–933. https://doi.org/10.1037/bul0000060

Tay, G., Tong, M., Yap, J., Mak, H., Goh, S., & Ho, C. (2025). Virtual reality for experiential learning: enhancing agitation management skills, confidence, and empathy in healthcare students. *Medical Education Online, 30*. https://doi.org/10.1080/10872981.2025.2542809.

Tussyadiah, I. P., Wang, D., Jung, T. H., & Tom Dieck, M. C. (2018). Virtual reality, presence, and attitude change: Empirical evidence from tourism. *Tourism management, 66*, 140–154. https://doi.org/10.1016/j.tourman.2017.12.003

van der Vlugt, T. M., & Harter, P. M. (2002). Teaching procedural skills to medical students: one institution's experience with an emergency procedures course. *Annals of emergency medicine, 40*(1), 41–49. https://doi.org/10.1067/mem.2002.125613

Weller, J. M., Nestel, D., Marshall, S. D., Brooks, P. M., & Conn, J. J. (2012). Simulation in clinical teaching and learning. Medical Journal of Australia, 196(9), 594–594. https://doi.org/10.5694/mja12.11015

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments, 7*(3), 225–240. https://doi.org/10.1162/105474698565686