Medical and Health Care Applications, Vol. 211, 2026, 79–88
https://doi.org/10.54941/ahfe1007470

Individual Performance Analytics in a Virtual Reality Simulation for Medication and Medical Supply Storage: An Experience Report

Rohan Saxena[1], Andrei Darii[1], Andrei Florea[1], Marian Sorin Nistor[1], Angela Huße[2], Lars Schneidereit[2], and Stefan Pickl[1]

[1]Universität der Bundeswehr München, 85579 Neubiberg, Germany
[2]Sanitätsakademie der Bundeswehr, Neuherbergstr. 11, 80937, München, Germany

ABSTRACT

Safe storage of medications and medical supplies is a core requirement of patient safety. At the Sanitätsakademie der Bundeswehr (Bundeswehr Medical Academy), a deliberately error-seeded "small material distribution unit" (skill lab) is used to illustrate theoretically taught content through practical scenarios. These scenarios address issues such as expired products, improper storage, damaged packaging, and inadequate access control. Inspections are typically conducted in groups, which limits insight into individual competence. A realistic virtual reality (VR) simulation of the storage environment was therefore developed to capture individual performance analytics. The simulation reproduces the skill lab environment in detail and randomises material-related defects and room-level safety checks per run. Its design is informed by findings indicating that high realism in immersive VR enhances learning. An instructor-facing launcher allows training staff to define the difficulty level and configure error types and quantities. At the beginning of a six-week training programme, twelve participants completed a run in the Easy mode, and nine of them returned near the end of the training for a run in the Hard mode. At the end of each run, a report was generated which documented the difficulty level, total number of errors, number and proportion of corrected errors, results of room-configuration checks, and for each item the status "correctly disposed", "incorrectly disposed", or "overlooked". Based on this data, descriptive metrics and illustrative individual performance profiles were derived. This experience report indicates that a VR simulation can meaningfully complement group-based training by providing graded scenarios and individual analytics.

Keywords: Virtual reality, Healthcare technology, Medical education and training, VR evaluation

INTRODUCTION

Virtual reality is commonly defined as the experience of presence in a real or simulated environment mediated by a communication medium (Steuer, 1992). Immersive VR extends this concept by enabling sensor-based interaction and user control, most notably using tracked head-mounted displays (HMDs) (Psotka, 1995).

Received March 30, 2026; Revised April 29, 2026; Accepted May 14, 2026; Available online July 20, 2026

Over the past decade, immersive VR has been increasingly explored as a tool for professional training and education. Several studies have reported that VR-based training environments are easy to use, provide positive learning experiences (Thompson et al., 2020), enhance learning outcomes, improve knowledge and skill acquisition (Fealy et al., 2019), and increase learner motivation (Stepan et al., 2017). The relevance of VR training and evaluation systems has increased further in the post-COVID-19 era. Particularly in the medical context, VR simulations offer certain advantages and address limitations of in-person instruction (Liu and Liu, 2023). They have been successfully used to train and evaluate nursing students and medical personnel on proper procedures for donning and doffing of personal protective equipment (Tsukada et al., 2024; Zikas et al., 2022), caring for patients with infectious diseases (Chang et al., 2025), resuscitation training (Moll-Khosrawi et al., 2022) and COVID-19 swab testing (Zikas et al., 2022), among other applications (Pallavicini et al., 2022). Beyond the medical domain, numerous studies have explored VR in a range of safety-critical domains such as crime scene investigation training (Bures and Lochmannova, 2024), scenario-based training for police officers (Kleygrewe et al., 2024), training paramedics for mass casualty incidents (Lochmannova et al., 2022) and fire safety training (Luimula et al., 2022).

Against this background, VR can be understood not only as a training medium but also as a potential platform for structured evaluation and observation of individual performance in realistic scenarios. In many educational settings, particularly those involving group-based exercises, individual competencies can be difficult to assess transparently. The present work builds on prior research by exploring the use of a realistic VR simulation as an evaluation-oriented complement to existing training practices, with a focus on making individual performance patterns visible within a safety-critical medical logistics context.

BACKGROUND

At the Bundeswehr Medical Academy, periodic training is conducted to ensure the correct handling, storage, and inventorying of medications and medical supplies. These materials are stored in a dedicated room that reflects operational storage conditions and regulatory requirements. The training targets medical and support personnel who are expected to identify storage-related deficiencies, assess the usability of medications and supplies, and take appropriate corrective actions to maintain patient safety and compliance with storage protocols.

As part of the training, participants are instructed to identify and correctly dispose of medications and medical supplies that are not fit for use. These include individual medication packages, bottles containing liquid substances such as disinfectants or infusion solutions, and larger cartons used for bulk storage. Defects at the item level may take various forms. Individual medication packages can exhibit physical damage such as crushed or deformed packaging, moisture damage, open boxes, missing or broken seals, discoloration due to external factors such as sun exposure, expired contents, missing Braille tactile markings, counterfeit indicators, or markings identifying the item as training material. Bottles containing liquids may present additional defect types, including crushed containers, missing caps, reduced liquid levels, empty bottles, or visible discoloration of the contents. Large cartons can be physically deformed, show moisture damage, or have missing, torn, or illegible information labels containing critical data such as contents, batch number, and manufacturing or expiry dates. In addition, cartons may lack required symbols, such as orientation arrows or flammability markings, which are essential for correct storage and handling.

Beyond item-level defects, participants are also required to assess the condition and configuration of the storage room itself. Room-level safety checks include verifying the presence of access authorization list at the entrance, inspecting documentation and temperature logs on the refrigerator used for cold-storage materials, and identifying irregular or missing entries. Participants must also assess if direct sunlight enters the room and take corrective action where necessary. A bulletin board within the room may display recall notices issued by manufacturers, identifying defective batches that must be located and removed from storage. The room further contains safes that may store controlled narcotic substances, which must remain locked whenever such substances are present. Another safety-related aspect concerns the storage of flammable materials, where only a limited quantity is permitted in the room.

METHOD

Before a simulation run begins, instructors access a dedicated launcher application that allows them to configure the evaluation scenario. Through this interface, the number and types of item-level defects and room-level safety checks can be defined. While instructors determine the overall configuration, the specific items affected by defects, the type of defect assigned to each item, and their spatial distribution within the room are randomized for each run. The launcher also provides predefined difficulty presets, including Easy, Medium, and Hard modes, for rapid scenario setup. A desktop version of the simulation is also available for participants who experience discomfort when using a head-mounted display.

Figure 1: View while entering the VR evaluation Room.

Virtual Environment and Interaction Flow

The VR evaluation tool is based on a digital replica of the physical storage room located at the Bundeswehr Medical Academy campus in Munich. Each simulation session begins in a virtual corridor leading to the training room, allowing participants to familiarize themselves with the controls and the virtual environment. In this area, participants can adjust display resolution and the height of their virtual avatar.

Figure 2: A participant inspecting defective packages during a simulation: A medicine box with broken seal (left) and a bottle containing discolored liquid (right).

Upon entering the storage room, a 15-minute time limit is activated, accompanied by an audible signal indicating the start of the session. Figure 1 shows a view of the VR Evaluation Room upon entry. Although no fixed sequence of actions is imposed, participants are expected to conduct a systematic inspection of the room. Upon entry, they may verify the access-control signage at the door. A tablet computer placed on the counter serves as a support tool, allowing participants to view their uploaded checklist and access the speech recognition function to verbally record observations. Visual markers labelled as "Fehler" can be placed near objects such as open or unlocked safes to indicate incorrect configurations. Depending on the selected difficulty level, a dynamic number of storage areas is provided, each designated for specific defect categories at the item level. Figure 2 shows a participant inspecting individual medications and medical supplies. Beneath the counter, a medical refrigerator is located for materials requiring cold storage, and its temperature and documentation logs must be inspected. The bulletin board and disposal trolleys for damaged cartons and excess flammable materials are positioned further inside the room. Shelves and cabinets along the walls contain cartons and individual medication packages that must be examined.

At the end of each simulation run, an automated report is generated. The report includes the participant's unique ID, start and end times, and the results of room-level safety checks. It further contains three structured tables: one listing items that were correctly identified and disposed of appropriately, one listing items that were correctly identified as defective but placed in incorrect storage areas, and one listing defects that were not identified. Summary metrics, including the proportion of corrected errors, are presented alongside an illustrative pie chart to support rapid interpretation.

Technical Implementation

The VR evaluation tool was developed using Unreal Engine 5.5, a physics-based game engine widely used for interactive simulations in both desktop and VR contexts. The accompanying launcher application was implemented in Python. The system was tested using Meta Quest 3 and Meta Quest Pro head-mounted display devices.

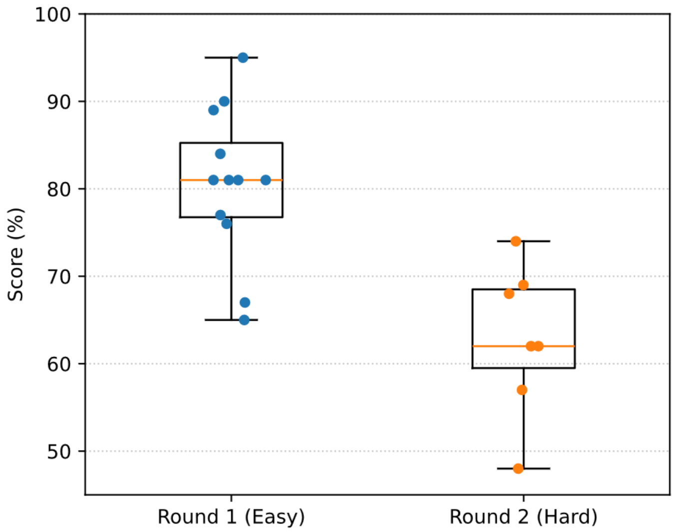

Figure 3: Distribution of participant scores in the easy and hard simulation rounds shown as boxplots. The figure summarizes the spread and central tendency of scores for each round, highlighting the overall shift in performance with increased difficulty.

RESULTS AND DISCUSSION

The VR-based evaluation tool was deployed during a six-week training program at the Bundeswehr Medical Academy. At the beginning of this training period, twelve participants completed an individual simulation run configured in the Easy difficulty mode. Before the start of the evaluation session, the participants were given an overview of the entire system. Data privacy was ensured by the instructor by assigning IDs instead of using their actual names as input into the launcher. For each participant, before starting the simulation, they were given time to adjust to the HMD. They were familiarized with the controls and given the opportunity to adjust the straps on the HMD, to ensure that they were completely comfortable before they began the simulation. They also got the opportunity to adjust their in-simulation height and the resolution of the HMD display. At the end of each run, a structured evaluation report was generated, marked by the ID assigned to the participant at the beginning. This assignment ensured that each participant's performance could be compared with their second run, while maintaining privacy. This report documented the selected difficulty level, the total number of errors present in the scenario, and the number and proportion of errors that were correctly identified and addressed by the participant. In addition, the report recorded the outcomes of room-level safety checks and, for each individual item, whether it was handled correctly, disposed of incorrectly, or overlooked entirely. This data formed the basis for the derivation of descriptive metrics and illustrative individual performance profiles, allowing for a systematic overview of observed performance patterns across participants and difficulty levels.

Scenarios in the Easy mode contained on average approximately 21 errors. Across these runs, a mean detection or correction rate of around 80% was observed. Analysis of item-level outcomes indicated that overlooked defects occurred more frequently than misclassifications. Participants rarely removed intact items incorrectly, whereas defects related to labelling and information were more often missed. Typical examples included missing symbols, expired products, and broken or absent seals. Regarding room-level safety checks, performance was generally strong across most categories. However, repeated omissions were observed for access-control signage at the entrance to the room, suggesting that this check was less consistently attended to during the inspection process.

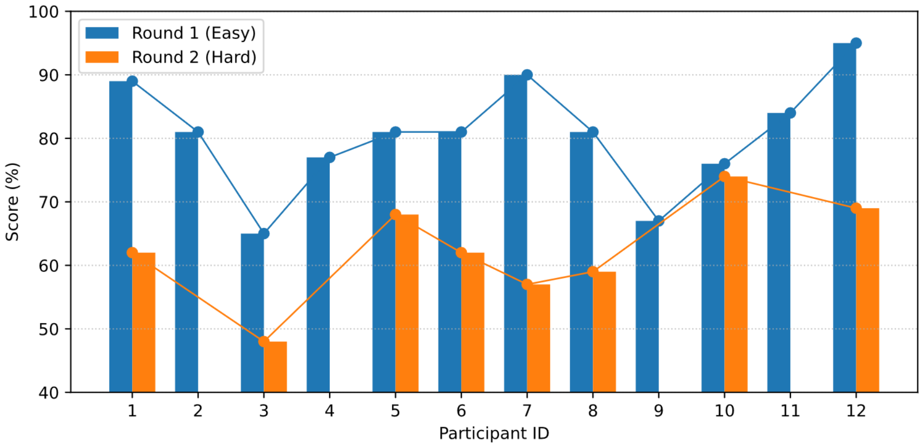

Figure 4: Grouped bar chart of participant scores in the easy and hard simulation rounds, with overlaid lines indicating scores per round as a visual guide.

Near the end of the training program, nine of the original participants returned for a second run in the Hard difficulty. Out of these nine, two participants used the desktop mode, one of which felt difficulties getting accustomed to the controls and had to abandon (see Limitations). In the Hard mode, scenarios contained on average approximately 36 errors, reflecting both an increased number of defects and the inclusion of additional error categories. Under these conditions, the mean detection or correction rate decreased to approximately 62%. Figure 3 shows the distribution of the participant scores across the two simulation rounds. Overall performance declined relative to the Easy mode, although some participants showed comparatively stable results across both runs. This pattern illustrates how increased scenario complexity and error density can differentiate levels of robustness in individual performance. Examination of item-level logs revealed recurring weaknesses, particularly for recall-related defects, missing Braille labelling, and markings identifying items as training material. In contrast, defects involving clearly visible physical damage or pronounced changes in liquid appearance remained reliably detected across difficulty levels.

Taken together, these observations indicate that the VR-based evaluation tool can make individual performance patterns visible in a structured and repeatable manner, as shown in Figure 4. By providing graded and randomized scenarios and automatically generated performance reports, the system enables transparent inspection of overall detection rates, defect-type specific strengths and weaknesses, and systematic omissions in room-level safety checks. This supports targeted, participant-specific feedback within the existing training framework and complements group-based inspection exercises by offering an additional perspective on individual task performance.

Limitations

The use of immersive VR systems is associated with practical constraints that need to be considered. Factors such as headset weight, thermal comfort, limited freedom of movement, and discrepancies between physical motion and visual perception may contribute to VR-related discomfort and increased cognitive load, particularly for inexperienced users (Oun et al., 2024). Individual differences in digital literacy and interaction preferences can further affect usability. During our study, one participant experienced difficulties in both VR and desktop modes, which may be related to them being left-handed and a control scheme not adapted for this user group. These observations highlight the importance of considering accessibility and interaction variability when deploying VR-based evaluation tools.

In addition, the findings of this experience report are limited by the single-center setting and the relatively small number of participants. The system represents a prototype-level implementation and was evaluated within a specific training context, which restricts the generalizability of the observed performance patterns. Hardware-related limitations of the commercially available VR headsets used in this project may also have influenced user experience, although several participants did not perceive these constraints as problematic.

Finally, ethical considerations related to immersive VR, including potential long-term effects of repeated exposure and differences between behavior in virtual and real-world settings (Madary & Metzinger, 2016) should be taken into account when interpreting the results. Moreover, behavior exhibited in virtual settings may differ from real-world conduct. For instance, actions such as casually handling or discarding medications in a virtual environment carry no immediate consequences, whereas such behavior would be unacceptable in practice.

CONCLUSION

This experience report presented the deployment of a virtual reality–based evaluation tool for medication and medical supply storage within an existing training programme at the Bundeswehr Medical Academy. The observations illustrate how randomized, graded simulation scenarios and automated reporting can make individual performance patterns visible in tasks that are otherwise assessed in group settings. By capturing item-level defects, room-level safety checks, and differences in performance across difficulty levels, the system provides a structured basis for participant-specific feedback without replacing established training formats.

While the scope of the deployment was limited, the results of this experience report indicate that the approach is suitable for further investigation in a larger and more systematic study. Such a study could examine the impact of VR-based evaluation on performance assessment, user acceptance, and practical feasibility across repeated training cycles and a broader participant base. If these future investigations demonstrate sustained value, VR-based evaluation tools could be considered for more formal integration into training curricula as a complementary assessment component in safety-critical educational contexts.

ACKNOWLEDGEMENT

The authors would like to thank dtec.bw — Digitalization and Technology Research Center of the Bundeswehr for funding the project “Competencies for the Digital Working World (KoDiA) - Empowerment for Digitalization”, and the Center for Technology Based Education and Training (ZtB) at Helmut Schmidt University/Universität der Bundeswehr Hamburg as lead partner. dtec.bw is funded by the European Union – NextGenerationEU.

AI DISCLOSURE

The authors used generative artificial intelligence tools to assist with language editing, rephrasing, and improving readability of the manuscript. The AI tools were not used to generate scientific content, data, results, or conclusions. All responsibility for the content remains with the authors.

REFERENCES

Bures, M., Lochmannová, A. (2024). Use of virtual reality for crime scene investigation training by security forces. In: Tareq Ahram and Christianne Falcão (eds) Human Factors in Virtual Environments and Game Design. AHFE (2024) International Conference. AHFE Open Access, vol 137. AHFE International, USA. http://doi.org/10.54941/ahfe1004983

Chang W, Lin CC, Crilly J, Lee HL, Chen LC, Han CY. Virtual Reality Simulation for Undergraduate Nursing Students for Care of Patients With Infectious Diseases: Mixed Methods Study. JMIR Med Educ. 2025;11:e64780. Published 2025 Feb 11. doi:10.2196/64780

Fealy S, Jones D, Hutton A, et al. The integration of immersive virtual reality in tertiary nursing and midwifery education: A scoping review. Nurse Educ Today. 2019;79:14-19. doi:10.1016/j.nedt.2019.05.002

Kleygrewe, L., Hutter, R. I. V., Koedijk, M., & Oudejans, R. R. D. (2024). Virtual reality training for police officers: a comparison of training responses in VR and real-life training. Police Practice and Research, 25(1), 18–37. https://doi.org/10.1080/15614263.2023.2176307

Liu, Y., Liu, Y. (2023). Scenario innovation of virtual reality in medical education: Possibility Advantages and Barriers. In: Tareq Ahram and Waldemar Karwowski (eds) Augmented, Virtual and Mixed Reality Simulation. AHFE (2023) International Conference. AHFE Open Access, vol. 118. AHFE International, USA. http://doi.org/10.54941/ahfe1004439

Lochmannová, A.; Šimon, M.; Hořejší, P.; Bárdy, M.; Reichertová, S.; Gillernová, K. The Use of Virtual Reality in Training Paramedics for a Mass Casualty Incident. Appl. Sci. 2022, 12, 11740. https://doi.org/10.3390/app122211740

Luimula, M., Majapuro, J., Bellalouna, F., Jedidi, A., Somerkoski, B., Haavisto, T. (2022). Hazardous Training Scenarios in Virtual Reality - A Preliminary Study of Training Scenarios for Massive Disasters in Metaverse. In: Tareq Ahram and Christianne Falcão (eds) Human Factors in Virtual Environments and Game Design. AHFE (2022) International Conference. AHFE Open Access, vol. 50. AHFE International, USA. http://doi.org/10.54941/ahfe1002062

Madary, Michael & Metzinger, Thomas. (2016). Real Virtuality: A Code of Ethical Conduct. Recommendations for Good Scientific Practice and the Consumers of VR-Technology. Frontiers in Robotics and AI. 3. 10.3389/frobt.2016.00003.

Moll-Khosrawi, P., Falb, A., Pinnschmidt, H. et al. Virtual reality as a teaching method for resuscitation training in undergraduate first year medical students during COVID-19 pandemic: A randomised controlled trial. BMC Med Educ 22, 483 (2022). https://doi.org/10.1186/s12909-022-03533-1

Oun, Ahmed & Hagerdorn, Nathan & Scheideger, Caleb & Cheng, Xiangyi. (2024). Mobile Devices or Head-Mounted Displays: A Comparative Review and Analysis of Augmented Reality in Healthcare. IEEE Access. PP. 1–1. 10.1109/ACCESS.2024.3361833.

Pallavicini F, Pepe A, Clerici M, Mantovani F. Virtual Reality Applications in Medicine During the COVID-19 Pandemic: Systematic Review. JMIR Serious Games. 2022;10(4):e35000. Published 2022 Oct 25. doi:10.2196/35000

Psotka, J. Immersive training systems: Virtual reality and education and training. Instr Sci 23, 405–431 (1995). https://doi.org/10.1007/BF00896880

Stepan K, Zeiger J, Hanchuk S, et al. Immersive virtual reality as a teaching tool for neuroanatomy. Int Forum Allergy Rhinol. 2017;7(10):1006-1013. doi:10.1002/alr.21986

Steuer, J., 1992. Defining virtual reality: Dimensions determining telepresence. Journal of communication, 42(4), pp. 73–93.

Thompson DS, Thompson AP, McConnell K. Nursing students' engagement and experiences with virtual reality in an undergraduate bioscience course. Int J Nurs Educ Scholarsh. 2020;17(1):/j/ijnes.2020.17.issue-1/ijnes-2019-0081/ijnes-2019-0081.xml. Published 2020 Sep 16. doi:10.1515/ijnes-2019-0081

Tsukada K, Yasui Y, Miyata S, et al. Effectiveness of Virtual Reality Training in Teaching Personal Protective Equipment Skills: A Randomized Clinical Trial. JAMA Netw Open. 2024;7(2):e2355358. Published 2024 Feb 5. doi:10.1001/jamanetworkopen.2023.55358

Tusher, Hasan Mahbub & Nazir, Salman & Mallam, Steven. (2022). Ethical Considerations for the Application of Virtual Reality in Education and Training. 10.54941/ahfe1002396.

Zikas, Paul & Kateros, Steve & Lydatakis, Nick & Kentros, Mike & Geronikolakis, Efstratios & Kamarianakis, Manos & Evangelou, Giannis & Kartsonaki, Ioanna & Apostolou, Achilles & Birrenbach, Tanja & Exadaktylos, Aristomenis & Sauter, Thomas & Papapagiannakis, George. (2022). Virtual Reality Medical Training for COVID-19 Swab Testing and Proper Handling of Personal Protective Equipment: Development and Usability. Frontiers in Virtual Reality. 2. 10.3389/frvir.2021.740197.