A Virtual Reality Reach Stacker Simulator for Operator Training

Eduardo Germer Pedroso[1], Ewerton Eyre de Morais Alonso[1]

[1]Universidade do Vale do Itajaí (UNIVALI)
Escola Politécnica - Curso de Design de Games
5ª Avenida, 1100 - Municípios, Balneário Camboriú - SC, 88337-300 - Brazil

{ed.pedroso.98@edu.univali.br, ewertonalonso@univali.br}

Abstract. Introduction: *Virtual reality can support training for heavy-equipment operation by reducing exposure to risk and enabling repeatable practice, but studies on reach stacker simulators still provide limited detail about task design and performance logging.* ***Objective:*** *This paper presents the design and exploratory evaluation of a virtual reality reach stacker simulator for operator training, with emphasis on reproducible task conditions, objective telemetry and post-use experience measures.* ***Methodology:*** *Thirteen adults completed a standardized container-stacking task using a head-mounted display and a physical control rig; the simulator logged task time, collisions and final positioning errors, and questionnaires assessed cybersickness, presence, usability, workload and perceived learning.* ***Results:*** *Ten complete logs showed completion times from 56.6 to 447.8 s, final translational errors from 0.052 to 0.391 m and yaw errors below 1.5 degrees. Questionnaire medians indicated low cybersickness, favorable usability (SUS = 85.0), moderate workload and high perceived learning, suggesting that the simulator provides a safe and measurable basis for future studies on port-operator training.*

Keywords *virtual reality; serious games; simulator; reach stacker; simulation-based training; user experience; performance metrics; occupational safety.*

1. Introduction

The Occupational Safety and Health Administration uses the term powered industrial truck to describe mobile, power-driven vehicles used to carry, push, pull, lift, stack or tier materials [Occupational Safety and Health Administration 2024]. Within this broad category, reach stackers are specialized heavy machines designed to handle ISO containers in ports, rail terminals and logistics yards. A reach stacker combines a telescopic boom mounted at the front of the chassis with a spreader that locks onto containers through twistlocks in the four corner castings. Its operation requires attention to the load chart, because stability margins change with boom reach and height. In typical operations, the operator must coordinate driving, braking, steering, boom elevation, boom extension, spreader rotation, side shift and twistlock states while maintaining awareness of visibility, load stability, surrounding stacks and nearby vehicles.

Training exclusively on real equipment is valuable but costly and potentially risky. A trainee's error can damage containers, infrastructure or the machine, and a real reach stacker may need to be removed from production while practice occurs. Simulation-based training addresses part of this problem by allowing instructors to repeat tasks, expose learners to controlled risk situations and collect objective data about actions and outcomes. Immersive virtual reality (VR) adds egocentric perception and natural head movement, which may support spatial awareness during technical tasks and can increase engagement in professional training [Renganayagalu et al. 2021, Radhakrishnan et al. 2021, Stefan et al. 2023].

For training use, however, a simulator should not be evaluated only by its visual realism. The learning task, assessment criteria and feedback data must also be explicit. Technical-training simulators commonly log task time, trajectories, collisions, positioning errors and control inputs, but the definition of these metrics is not always reported in enough detail for comparison or replication [Ernstsen and Nazir 2020, de Oliveira et al. 2022]. This issue is important for reach stacker training because final task success can hide unsafe or inefficient behavior. A learner may place a container correctly after many small collisions, excessive adjustments or poor planning.

This paper presents the design and exploratory evaluation of StackerSim, a VR reach stacker simulator and a standardized container-stacking task for operator training. The simulator combines an HMD, a partial physical rig, a virtual cabin modeled from photogrammetry and an instrumented container-yard scenario. The contribution is a baseline stacking task with task sequence, geometry, telemetry and questionnaire measures described in enough detail to support replication and future instructional use. The study is exploratory and does not claim learning transfer to real machines; instead, it documents whether the task can be completed, whether the experience is comfortable and usable, and whether the collected metrics can support formative feedback.

2. Related Work

Immersive VR has been investigated as a training medium in professional, industrial and safety-critical domains. Systematic reviews report generally promising results for skill acquisition, engagement and safety-relevant training, while also emphasizing heterogeneity in equipment, protocols and outcome measures [Renganayagalu et al. 2021, Radhakrishnan et al. 2021, Stefan et al. 2023]. In training research, this variation makes careful task description essential: the reader needs to know what the trainee practiced, how performance was measured and how subjective experience was assessed.

Studies close to container-handling training illustrate this diversity. VR has been used for ship-handling simulations with presence and cybersickness measures [Bassano et al. 2019], repeated forklift-driving practice with novices [Islam et al. 2023], construction-equipment hazard recognition [Shringi et al. 2022], robotic teleoperation training [Adami et al. 2021] and shipboard-crane operation with anti-sway control [Martin and Irani 2024]. Comparative work in occupational safety and industrial contexts has also examined VR against more conventional training strategies, although conclusions depend on task, sample, exposure duration and assessment method [Stark et al. 2023, Pětvaldský et al. 2025].

The literature specifically focused on reach stackers is smaller. Prior work includes a visual reach stacker simulator [Ribeiro Neto 2014], accident reconstruction using a virtual reach stacker [Estork 2022], phase-and-task modules for operator training [Hermes 2022], transshipment simulators for port personnel [Haller et al. 2015] and extended-reality approaches for terminal and port operations [Bruzzone and Shinelschikov 2024]. These studies support the relevance of simulation in port training, but many reports provide limited information about the task layout, completion criteria, collision thresholds, final positioning metrics or questionnaire batteries.

The present work complements that literature by focusing on a single, reproducible training task. The task is intentionally simple: pick up one 40-foot container from the ground and stack it on another 40-foot container. Its simplicity is a methodological choice, because a controlled baseline is useful for future pre/post studies, instructor feedback experiments and progressive training curricula for reach stacker operators.

3. Simulator and Training Task

3.1. Hardware, virtual model and interaction

StackerSim ran on a Windows 11 PC with an NVIDIA RTX 4080 GPU, 32 GB of RAM and a Meta Quest 2 HMD connected through Link at 90 Hz. The physical station combined a USB steering wheel with three pedals and a USB joystick. The steering wheel controlled vehicle direction; the pedals controlled longitudinal motion; and the joystick controlled boom elevation, boom extension, spreader rotation, lateral displacement and twistlock opening or locking. The simulation was fully rendered in the HMD; it was not a projection system or a mixed-reality overlay. The user interacted with the physical devices in the rig, while the cabin, controls, reach stacker, containers and yard were represented virtually.

Figure 1. Physical control rig with steering wheel, pedals and joystick.

The virtual cabin was modeled from photographs of a real reach stacker cabin using photogrammetry as geometric reference. The photogrammetry mesh was retopologized and adjusted into a lighter model with proportions close to the physical cabin. The virtual steering wheel, pedals and joystick were positioned to match the physical devices after a user-triggered calibration step at the beginning of the session. This calibration repositioned only the virtual control objects relative to the seated user and physical rig, while also resetting the camera's vertical orientation to recenter the view. As a result, the virtual controls appeared in approximately the same direction as the corresponding physical devices. This allowed participants to locate the controls mainly by touch while looking through the virtual windshield or mirrors. Hand tracking from the HMD was used to represent the user's hands on the virtual wheel and joystick, reinforcing the relation between physical actions and the simulated machine.

Figure 2. Virtual cabin model and photogrammetry-derived reference mesh.

Figure 3. Tracked virtual hands aligned with the steering wheel and joystick.

The boom and spreader were modeled with rigid bodies and joints in Unity. The simulator included 20-foot and 40-foot ISO container models, with attention to the corner regions where twistlocks engage. The spreader had a passive vertical degree of freedom so that, when lowered onto a container, it could settle before locking. In real equipment, sensors near the twistlocks indicate whether the spreader is resting on the container and whether the locks are engaged. The simulator reproduced this operational logic through visual states for unlocked, seated and locked conditions. Active spreader tilt was not implemented in this version; tilt emerged passively from gravity, contacts and forces transmitted through the boom and load.

3.2. Task sequence and telemetry

The evaluated scenario consisted of a single stacking task. At the start, the reach stacker faced a 40-foot container resting on the ground. The participant had to approach the container, align the spreader, lower and seat it, lock the twistlocks, lift the container, move to the target position, place the container on a second 40-foot container, unlock the twistlocks and reverse away from the stack. A static stack of three 40-foot containers was placed behind the target container to create a controlled opportunity for mild contacts during approach and final positioning. A world-space interface inside the cabin displayed the main task steps and marked completed steps.

Figure 4. Side view of the baseline stacking scenario.

Figure 5. In-cabin step interface used during the stacking task.

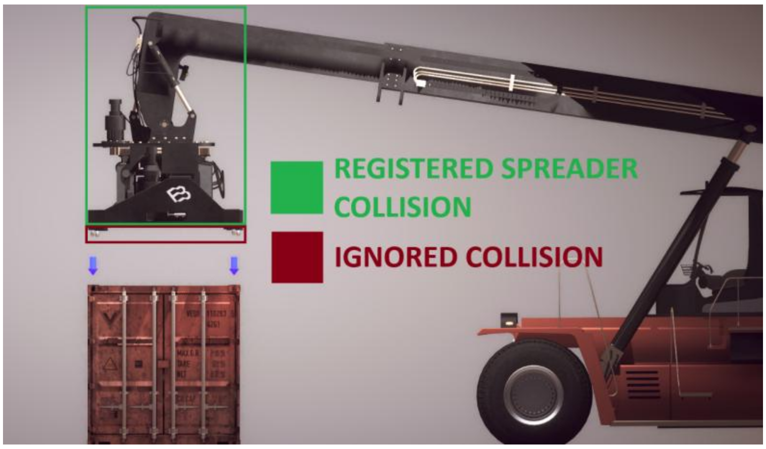

Figure 6. Spreader contact logic distinguishing expected seating contact from counted spreader collisions.

The simulator used a right-handed world frame, with the XZ plane as the horizontal plane and Y as the vertical axis. World units corresponded to meters. The center of the target container was positioned 10.83 m ahead of the center of the ground container in the XZ plane. The rear static stack was positioned with a 0.063 m horizontal gap behind the target container and was modeled as immovable.

Task time did not start when the scene loaded. It started when any reach stacker collider entered a cylindrical trigger region with a 2.0 m radius around the ground container. The task was complete only when two conditions were met: the lifted container was resting stably on the target container, and the reach stacker was at least 15.0 m from the target container in the XZ plane. Metrics were saved to CSV files at the end of each logged attempt. Table 1 summarizes the main variables.

Table 1. Telemetry metrics recorded by the simulator.

Metric	Symbol / unit	Description
Task time	t (s)	Total time from the beginning of the approach to the final retreat from the target stack.
Other contacts	n_other	Low-severity contacts in the logger's Other category, mainly involving the chassis or boom.
Spreader collisions	n_spreader	Lateral, frontal, rear or top collisions involving the spreader, excluding expected bottom contact during seating.
Translational error	e_pos (m)	Planar XZ distance between centers of the lifted and target containers at task completion.
Yaw error	e_ang (°)	Absolute yaw difference between the lifted and target containers at task completion.

Collisions were classified using the impulse magnitude reported by the physics engine. Contacts with impulse below 30 were classified as low-impact, whereas medium and high impacts were grouped as impulse values equal to or above 30. Only low-impact contacts are reported in Table 3. Contacts on the bottom face of the spreader in the seating region were ignored so that normal placement on a container would not be counted as an error. The translational error was computed as the planar XZ distance between the centers of the lifted and base containers in the local frame of the base container. The angular error was computed as the absolute yaw difference between the same containers.

4. Evaluation Method

4.1. Study design, participants and procedure

The study followed an exploratory, descriptive, single-group design [Gil 2008]. Thirteen adults were recruited by convenience among people able to use an HMD. Inclusion criteria were age of 18 years or older, no known contraindication to HMD use and agreement to participate through informed consent. The group was intentionally heterogeneous, including participants with different levels of experience with VR, games, simulators and material-handling equipment.

Each session started with an in-person briefing about the controls without the HMD. After wearing the HMD, the participant received the task instructions and could complete the scenario at most twice: a familiarization attempt without logged performance metrics and a second attempt with telemetry recording. Sessions could be interrupted if participants reported persistent discomfort. Comfort was supported by a fixed seat, stable frame rate, low perceived latency and pauses on demand, consistent with recommendations in immersive training literature [Renganayagalu et al. 2021].

4.2. Questionnaires and analysis

After the simulator run, participants answered a post-use questionnaire. The first block collected profile and prior-experience information, including use of glasses or contact lenses, professional experience with material-handling equipment, current role, previous use of this simulator, previous use of other simulators, previous HMD experience and tendency to experience motion sickness in VR.

The subjective-experience block included instruments commonly used in VR and simulator evaluation. Cybersickness was measured with CSQ-VR 6, a six-item questionnaire designed for cybersickness assessment in VR and validated in comparison with SSQ and VRSQ [Kourtesis et al. 2023]; the SSQ was also considered as the historical basis for simulator-sickness assessment [Kennedy et al. 1993]. Presence was assessed with a Single-Item Presence measure [Bouchard et al. 2004]. Usability was measured with the System Usability Scale (SUS) [Brooke 1996]. Workload was assessed with NASA-TLX [Hart and Staveland 1988]. In this study, the six NASA-TLX items were averaged on the 1-7 scale and also normalized to 0-100. Custom perceived learning items assessed reach stacker parts, driving and locomotion, stability, boom and spreader operation, load handling and container positioning. A separate item about safe operation was collected but was not included in the aggregate score.

Analyses were descriptive. For telemetry, the focus was on task time, contacts and final positioning errors. Questionnaire scores were calculated as described above and summarized with medians and observed ranges. No inferential comparison was performed because the study was not designed to compare training methods, experience levels or repeated-learning effects.

5. Results

All thirteen participants completed the session and the post-use questionnaire, and all questionnaire responses were included in the analysis. Ten logged attempts had complete telemetry and were included in the performance analysis. A CSV-saving bug affected three sessions and prevented their telemetry files from being saved correctly. The bug was promptly corrected after it was identified, but the affected telemetry could not be recovered. Across all questionnaires, two participants reported reach stacker operator experience, including one instructor/operator. Only the instructor/operator had complete telemetry. Among the ten participants with complete logs, five had a professional link to reach stackers or related equipment.

5.1. Questionnaire scores

Table 2 summarizes the questionnaire scores. Cybersickness was close to zero for most respondents. Presence and perceived session smoothness tended to be high. SUS indicated favorable perceived usability, while NASA-TLX suggested a moderate workload. Perceived learning ratings were concentrated toward the upper end of the scale, indicating that participants felt the experience helped them understand basic aspects of reach stacker operation.

Table 2. Descriptive measures of questionnaire scores (n = 13).

Instrument / dimension	Scale	Median (range)	Normalized 0-100
CSQ-VR 6 (cybersickness)	0-6	0.0 (0.0-1.2)	0.0
Presence (SIP)	0-6	5.0 (1.0-6.0)	83.3
Perceived smoothness	1-5	5.0 (3.0-5.0)	100.0
SUS usability	0-100	85.0 (60.0-92.5)	85.0
NASA-TLX workload	1-7	3.3 (1.8-5.2)	38.9
Perceived learning	1-5	4.3 (2.5-5.0)	83.3

5.2. Telemetry and participant-level results

Table 3 combines the main telemetry values with workload, perceived learning, direct reach stacker operator experience and professional links to reach stackers or related equipment. Completion times ranged from 56.6 to 447.8 s. Despite this variation, final translational errors remained between 0.052 and 0.391 m, and final yaw errors remained below approximately 1.5 degrees. Low-severity contacts were common for some users, especially during approach and fine adjustments.

Table 3. Logged task results ordered by total task time (n = 10).

Code	t (s)	e_pos (m)	e_ang (°)	Other	Spreader	TLX	P. learning	RS operator	Link
URKT2	56.6	0.089	1.273	8	7	2.8	2.5	No	Yes
NNSQ5	89.1	0.205	0.285	15	16	3.2	5.0	No	Yes
LHCN0	93.3	0.093	0.646	5	6	2.5	5.0	Yes	Yes
SOGC4	200.8	0.280	0.668	6	2	4.2	4.7	No	Yes
VSJL9	205.2	0.076	0.329	7	7	5.0	4.3	No	No
YDJS3	223.4	0.333	1.058	16	7	4.2	4.2	No	No
UTLH6	240.0	0.058	0.469	32	22	3.0	3.5	No	No
MGAA1	277.6	0.391	0.093	24	7	4.0	4.8	No	No
AKYV4	336.9	0.278	1.406	31	11	5.2	4.0	No	Yes
FAGX3	447.8	0.052	0.428	33	39	3.3	4.3	No	No

In Table 3, RS operator indicates direct reach stacker operator experience among participants with complete logs. Link indicates a professional connection to reach stackers or related equipment and may include instruction, maintenance, development or other relevant work. Other is the logger's non-spreader category and mainly represents chassis or boom contacts. Four of the five fastest logs had Link = Yes, although one participant with Link = Yes was much slower. The sample was too small for group comparisons.

6. Discussion

The first relevant finding is that the standardized task was feasible for a heterogeneous group. Even participants without direct reach stacker operation experience completed the full sequence and placed the container close to the target. In a training context, this suggests that the task can serve as an introductory exercise: learners practice the relation between vehicle positioning, boom movement, spreader seating, twistlock state and final alignment without the complexity of a full port yard.

Open-ended feedback was treated only as contextual evidence, not as a formal qualitative analysis. Comments from experienced participants suggested refinements to machine behavior and task design, reinforcing that simulator sessions can also be used to discuss operational procedures and safe practices.

The second finding concerns comfort and usability. Low cybersickness, high session smoothness and favorable SUS scores indicate that the simulator was acceptable for the tested sample. This matters because a training tool that causes discomfort or imposes excessive interface friction may fail even if its visual model is detailed. The physical controls and calibrated virtual cabin appear to have helped participants connect hand and foot actions with machine behavior.

The third finding concerns assessment. Final positioning errors alone would suggest that most participants performed well, but collision counts and completion times reveal different task-execution patterns. Some participants achieved precise final placement only after many low-severity contacts or a long series of adjustments. For instructors, this distinction is useful: telemetry can support feedback about clearance, planning, speed control and unnecessary contact, instead of judging only whether the container eventually reached the target.

The implementation details also support replication: cabin modeling, calibration, task geometry, start and finish criteria, collision logic and participant-level metrics can be reused or adapted in future studies.

The study also reinforces the need to separate perceived learning from demonstrated learning. Participants reported that the session helped them understand reach stacker operation, but there was no pre-test, post-test, comparison group or assessment on real equipment. Therefore, the results indicate usability, comfort and perceived training potential, not learning transfer. Future studies should connect simulator metrics with formal knowledge tests, repeated-practice curves and, when feasible, supervised assessments on real or higher-fidelity equipment.

Several limitations remain. The sample was small and recruited by convenience, and only ten participants had complete telemetry because the data from three sessions affected by the CSV-saving bug could not be recovered. The scenario was a simplified baseline task; real yards include higher stacks, different loads, restricted visibility, weather, traffic, time pressure and local procedures. The perceived learning items were custom and should not be treated as a validated training scale. Finally, the implementation depends on a specific game engine, HMD, control rig and physics configuration, so replication may require calibration of collisions, rigid bodies and HMD performance targets.

7. Conclusion

This paper presented the design and exploratory evaluation of a VR reach stacker simulator for operator training. The simulator combined an HMD, physical steering/pedal/joystick controls, a virtual cabin modeled with photogrammetry and an instrumented container-stacking task. The exploratory evaluation with thirteen participants indicated low cybersickness, favorable usability, moderate workload and positive perceived learning, while telemetry from ten complete logs captured task time, collisions and final positioning errors.

The main contribution is not a claim that VR replaces real-equipment training or demonstrates learning transfer, but a reproducible basis for training research on reach stacker operation and related container-handling equipment. By specifying the task sequence, scenario geometry, completion criteria, instruments and metrics, the work supports future studies that can test learning progression, compare instructional strategies and investigate how simulator data can inform actionable feedback for port operator training.

Ethics Statement

All participants were adults, agreed to participate through informed consent and were represented by recurring codes rather than personal identification in the dataset. Sessions could be interrupted on demand in case of discomfort.

Use of Generative AI Assistance

The authors used ChatGPT by OpenAI to help revise the text and translate parts from Portuguese to English.

References

Adami, P., Rodrigues, P. B., Woods, P. J., Becerik-Gerber, B., Soibelman, L., Copur-Gencturk, Y., and Lucas, G. (2021). Effectiveness of VR-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation. Advanced Engineering Informatics, 50:101431.

Bassano, C., Chessa, M., Fengone, L., Isgrò, L., Solari, F., Spallarossa, G., Tozzi, D., and Zini, A. (2019). Evaluation of a virtual reality system for ship handling simulations. In Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019), Volume 2: HUCAPP, pages 62-73.

Bouchard, S., Robillard, G., St-Jacques, J., Dumoulin, S., Patry, M.-A., and Renaud, P. (2004). Reliability and validity of a single-item measure of presence in VR. In IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications (HAVE 2004), pages 59-61.

Brooke, J. (1996). SUS: A quick and dirty usability scale. In Jordan, P. W., Thomas, B., Weerdmeester, B. A., and McClelland, A. L., editors, Usability Evaluation in Industry, pages 189-194. Taylor & Francis, London.

Bruzzone, A. G. and Shinelschikov, K. (2024). Immersive XR simulation for enhanced terminal & port operations. In 15th International Workshop on Applied Modelling and Simulation (WAMS 2024), Shaoxing.

de Oliveira, R. P., Carim Junior, G., Pereira, B., Hunter, D., Drummond, J., and Andre, M. (2022). Systematic literature review on the fidelity of maritime simulator training. Education Sciences, 12(11):817.

Ernstsen, J. and Nazir, S. (2020). Performance assessment in full-scale simulators: A case of maritime pilotage operations. Safety Science, 129:104775.

Estork, N. A. de N. (2022). Uso de simulador virtual 3D de máquina tipo reach stacker para reconstrução de acidentes. Undergraduate final work, Instituto Federal de Santa Catarina, Itajaí, Brazil.

Gil, A. C. (2008). Métodos e Técnicas de Pesquisa Social. Atlas, São Paulo, Brazil, 6th edition.

Haller, A., Putz, L. M., and Schauer, O. (2015). Transhipment simulators for training of ports' personnel. Advanced Engineering Forum, 13:277-281.

Hart, S. G. and Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In Hancock, P. A. and Meshkati, N., editors, Human Mental Workload, volume 52 of Advances in Psychology, pages 139-183. North-Holland, Amsterdam.

Hermes, F. S. F. (2022). Implementação de um sistema de fases e tarefas utilizando o simulador 3D de empilhadeira do tipo reach stacker para treinamento e reciclagem de operadores. Undergraduate final work, Instituto Federal de Santa Catarina, Itajaí, Brazil.

Islam, M. S., Jamshid Nezhad Zahabi, S., Kim, S., Lau, N., Nussbaum, M. A., and Lim, S. (2023). Forklift driving performance of novices with repeated VR-based training. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, volume 67, pages 1480-1481.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology, 3(3):203-220.

Kourtesis, P., Linnell, J., Amir, R., Argelaguet, F., and MacPherson, S. E. (2023). Cybersickness in virtual reality questionnaire (CSQ-VR): A validation and comparison against SSQ and VRSQ. Virtual Worlds, 2(1):44-84.

Martin, I. A. and Irani, R. A. (2024). The examination of operator performance when controlling a shipboard crane anti-sway control system within a virtual-reality simulator. Ocean Engineering, 298:117164.

Occupational Safety and Health Administration (2024). Powered industrial trucks (forklifts): OSHA standard 1910.178 and related interpretations. Accessed: 2026-04-27.

Pětvaldský, T., Kočkár, S., Lepík, P., Hollá, K., and Kuricová, A. (2025). A comparative analysis of OSH training: Evaluating traditional methods versus interactive and virtual reality approaches in the context of sustainability. Sustainability, 17(12):5570.

Radhakrishnan, U., Koumaditis, K., and Chinello, F. (2021). A systematic review of immersive virtual reality for industrial skills training. Behaviour & Information Technology, 40(12):1310-1339.

Renganayagalu, S. K., Mallam, S. C., and Nazir, S. (2021). Effectiveness of VR head mounted displays in professional training: A systematic review. Technology, Knowledge and Learning, 26(4):999-1041.

Ribeiro Neto, A. (2014). Simulador visual de reach stacker utilizando recursos de realidade virtual. Master's dissertation, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

Shringi, A., Arashpour, M., Golafshani, E. M., Rajabifard, A., Dwyer, T., and Li, H. (2022). Efficiency of VR-based safety training for construction equipment: Hazard recognition in heavy machinery operations. Buildings, 12(12):2084.

Stark, C., Wiener, S., and Chin, J. F. (2023). Comparing industry training using virtual reality against conventional training: A case study. In Technological Advancement in Instrumentation & Human Engineering, volume 882 of Lecture Notes in Electrical Engineering, pages 161-176. Springer, Singapore.

Stefan, H., Mortimer, M., and Horan, B. (2023). Evaluating the effectiveness of virtual reality for safety-relevant training: A systematic review. Virtual Reality, 27(4):2839-2869.