Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 59-10 (2025) 2676–2681

Empowering Adaptive Learning in VR Assembly Training Using Real-time Performance Tracking

Eva Geurts * **Arno Verstraete** * **Maarten Wijnants** *

* *UHasselt - Hasselt University, Digital Future Lab, Flanders Make, Diepenbeek, Belgium (e-mail: firstname.lastname@uhasselt.be).*

Abstract: Virtual Reality (VR) provides unique opportunities for creating immersive, personalized and responsive learning environments through advanced features like hand and eye tracking. However, traditional VR training often lacks the flexibility to accommodate diverse learning styles. Personalization of training is crucial in industrial assembly, where user profiles and levels of expertise vary greatly. This paper introduces a novel solution for adaptive learning in VR focused on assembly knowledge training, using hand and eye tracking to deliver real-time feedback and individually adjusted learning paths. We applied our approach to two realistic assembly cases to evaluate its practical application. We hope to inspire future research further to explore and refine this adaptive approach, contributing to developing more flexible and effective VR-based training solutions for the manufacturing industry.

Copyright © 2025 The Authors. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: Industry 4.0; Industry 5.0; Virtual Reality; Assembly Training; Adaptive Learning

1. INTRODUCTION

In the domain of industrial operations, training holds essential importance in performing assemblies, ensuring safety protocols, and navigating digital workflows. Traditionally, on-the-job training, a method where employees learn the skills and knowledge from an expert while actually working, has been the standard for communicating this kind of knowledge to operators (Huang et al. (2021)). Although this training approach can be highly personalized, it brings along a huge cost, is not scalable and reusable, and is not risk-free as it is prone to accidents (Huang et al. (2021); Liu et al. (2022)).

However, the training landscape is rapidly evolving, with the integration of digital technologies offering new opportunities for learning (Palmas et al. (2019)). Digital resources enhance efficiency and facilitate consistent training but lack personalization (Huang et al. (2021)). Virtual Reality (VR) has emerged as a transformative alternative, offering immersive experiences that simulate real-world scenarios (Wolfartsberger et al. (2023)). By replicating assembly processes in a virtual environment, VR training empowers operators to practice without the constraints of physical equipment, enhancing engagement and retention through an interactive, hands-on learning experience (Moreno and Mayer (2002)). Still, traditional VR training solutions often rely on fixed support structures that do not adapt to the user's evolving skill level, which limits their effectiveness in catering to individual learning needs (Ulmer et al. (2022)). Our approach leverages VR's built-in hand and eye tracking to precisely monitor user actions and gaze enabling real-time performance tracking, tailored feedback, and adaptive training. By applying adaptive learning principles, the system continuously personalizes support based on individual progress and past training sessions.

This paper presents our adaptive VR training approach for learning assembly knowledge and sequences, with a key contribution being the integration of dynamic real-time support, feedback mechanisms, and adaptive learning algorithms to enhance training effectiveness. Section 3 describes this tailored approach, while Section 4 demonstrates its application to real-life use cases through collaboration with two manufacturing companies. We discuss its current limitations in Section 5 and conclude with clear action points to advance this work in future research in Section 6.

2. RELATED WORK

We will begin by exploring the latest advancements in digital adaptive learning solutions, which have inspired us to apply these concepts to adaptive learning in VR assembly training. This will be followed by an overview of the state-of-the-art industrial training solutions that use eXtended Reality (XR).

2.1 Adaptive learning from an educational point of view

Taylor et al. (2021) define personalized and adaptive learning as a method of learning that is customized for individual learners. Every learner will follow their own learning path at their own tempo according to the learner's prior knowledge and learning pace. Traditionally, adaptive learning has been implemented through technologies such as adaptive e-learning or Learning Management Systems (LMS) (Vincent-Ruz and Boase (2022)). Adaptive systems often use real-time analysis, continuously tracking a learner's performance to make immediate adjustments,

2405-8963 Copyright © 2025 The Authors. This is an open access article under the CC BY-NC-ND license.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2025.09.450

which aligns with the scaffolding teaching technique by providing timely support and guidance (Yu et al. (2024)). In contrast, many traditional LMS assess learners at specific intervals rather than continuously (Strakos et al. (2023)). Another essential component is using algorithms that adapt the learning content based on the learner's unique needs (Ge et al. (2019)). These algorithms consider various factors, including learning styles, preferences, and past performance, enabling dynamic adjustments to the learning path. Aligned with the 'Cognitive load theory', this approach optimizes learning efficiency and engagement by ensuring that cognitive demands remain manageable, allowing learners to focus on essential information without being overwhelmed (Van Merrienboer and Sweller (2005)). Adaptive learning technologies can be used to tailor VR experiences to individual needs, filtering out irrelevant information, and highlighting key content based on real-time performance.

2.2 Personalized XR training: insights and comparisons

Huang et al. (2021) developed an adaptive Augmented Reality (AR) tutoring system that considers the learner's real-time progress, thereby providing different levels of instruction. They found that their proposed solution was more effective and preferred than a non-adaptive variant. Studer et al. (2024) explored using an open-ended system to teach drilling skills. Their solution provides a comprehensive learning experience, including safety protocols, setup procedures, drilling tutorials, and open-ended practice sessions. Users receive real-time feedback on their mistakes and the quality of the drilled geometries. Havard et al. (2021) compared AR tablet guidance to PDF instructions, considering operators' competency levels. They found that advanced operators spent less time waiting for AR information to appear, as they could often rely on text instructions.

Our work builds on the presented findings using real-time hand and eye tracking to personalize VR assembly training. We apply this by dynamically adjusting guidance based on the user's interactions and gaze. The following section describes this conceptual approach.

3. PERSONALIZED LEARNING IN VR USING REAL-TIME PERFORMANCE TRACKING

In this section, we introduce our adaptive learning approach for VR assembly training (see Figure 1), describe the underlying algorithm, and illustrate how real-time performance tracking enables dynamic adjustments to instructional support. By leveraging the adaptive learning techniques from Section 2, this system provides tailored guidance, adapting in real-time based on the learner's actions and performance. We focus on assembly training, which involves numerous pick-and-place actions. Common mistakes are picking the incorrect object, placing the object in an incorrect position, or misunderstanding actions or interactions with objects like clicking, snapping, and screwing. Additionally, the sequence of task execution is often crucial. Therefore, our approach focuses on detecting these mistakes and adapting the support accordingly.

We use the same approach as Huang et al. (2021) by using *Sources of adaptation* and *Targets of adaptation*.

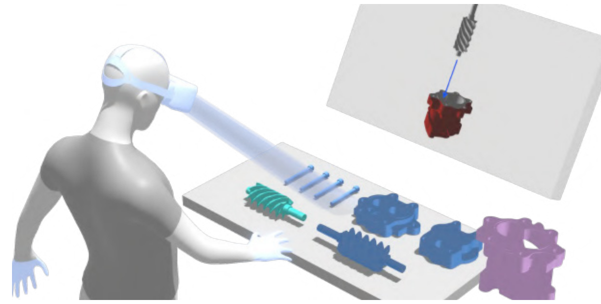


Fig. 1. Illustration of our conceptual VR learning approach. The system recognizes correct objects (teal), incorrect objects (blue), and correct placements (purple), together with hand and eye tracking to track the user's performance in real time and provide personalized feedback.

The factors tracked during training, such as task execution speed, held objects, and eye tracking, are our *Sources of adaptation*. The support systems that instruct the learner, which we adapt to the user, are our *Targets of adaptation*.

3.1 Technical approach

Our adaptive algorithm dynamically manages the activation of instructional support systems based on real-time performance tracking. It identifies user difficulties, referred to as **failure states**, and adjusts the training environment to provide targeted support, referred to as **support systems**. A learner model is continuously updated to track task performance, guiding adaptations over time.

Sources of adaptation. The system gathers key types of data to inform the adaptive algorithm:

Contextual Data : Tracks the assembly task sequence, task duration, and components involved.

Performance/User Data : Monitors task completion time, deviation from target task completion time, triggered failure states, and activated support systems.

Hand Tracking : Detects which objects the user is holding and evaluates correctness based on task context.

Eye Tracking : Analyzes user gaze to identify interactions with instructional materials, target objects, wrong objects or distractors. If eye tracking is unavailable, head tracking is used as a fallback, albeit with reduced accuracy.

Failure states & Targets of adaptation. The system classifies user difficulties into specific failure states, each linked to corresponding support systems:

Object Failure : Occurs when the user selects the wrong object. Adaptive responses include highlighting the correct object or differentiating similar objects.

Target Failure : Triggered when the user struggles to identify where to place an object. Support systems may highlight target locations or provide directional animations.

Action Failure : Results from unclear task interactions, such as missing a required action like screwing. Additional instructions or animations can help in such situations.

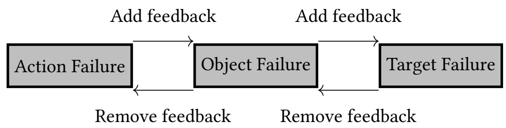


Fig. 2. Default hierarchy for failure states. If the Default Failure is reached (see Figure 3), support systems are incrementally added from left to right. This means that when the Default Failure is reached, we check if a support system from the Action Failure can be activated. If all of these systems are already active, we check if a system from Object Failure can be activated, and lastly from Target Failure. Removing support happens in the same manner but reversed.

Default Failure : A fallback state that activates support incrementally, following a predefined hierarchy (see Figure 2).

The *Targets of adaptation* (or support systems) can be defined in terms of the assembly use case. We will provide concrete examples in Section 4.

Figures 2 and 3 summarize the default hierarchy for failure states and the algorithm's decision-making process. Support systems are activated progressively based on detected failure states, with deactivation following the reverse order to promote user autonomy. In doing so, we apply adaptive learning principles in training sessions by reducing the number of support systems by one for each task in each iteration. By gradually reducing the amount of provided information, we draw on the 'Cognitive load theory', which suggests that introducing a manageable level of "desirable difficulty" can enhance engagement and improve learning outcomes (Van Merrienboer and Sweller (2005)).

3.2 Workflow of the adaptive algorithm

Tasks are timed independently, and failure states are detected based on real-time data as illustrated in Figure 3. The algorithm automatically detects when to activate support when thresholds are exceeded. Conversely, during subsequent training iterations, support is reduced to encourage independent task completion and learning progress.

Illustrative example. Consider a VR task to assemble a stool with a set of failure states and support systems (see Table 1).

When the user is new to the training, they start with only the first level of *action failure* active, which is probably an instruction (text or image). Based on eye and hand tracking, we decide, as explained before, what other support systems should be activated during the execution of the training. When assembling a stool leg, the user must select a bolt (correct object), align it with a hole (correct target), and screw it in (correct action). If the user picks the wrong part, which will eventually trigger an *object failure*, the system highlights the correct object. If they struggle to locate the assembly point, it will trigger a *target failure* and the system highlights the target area. Finally, if the screwing action is missed, an *action failure* will be triggered and text instructions will explain the

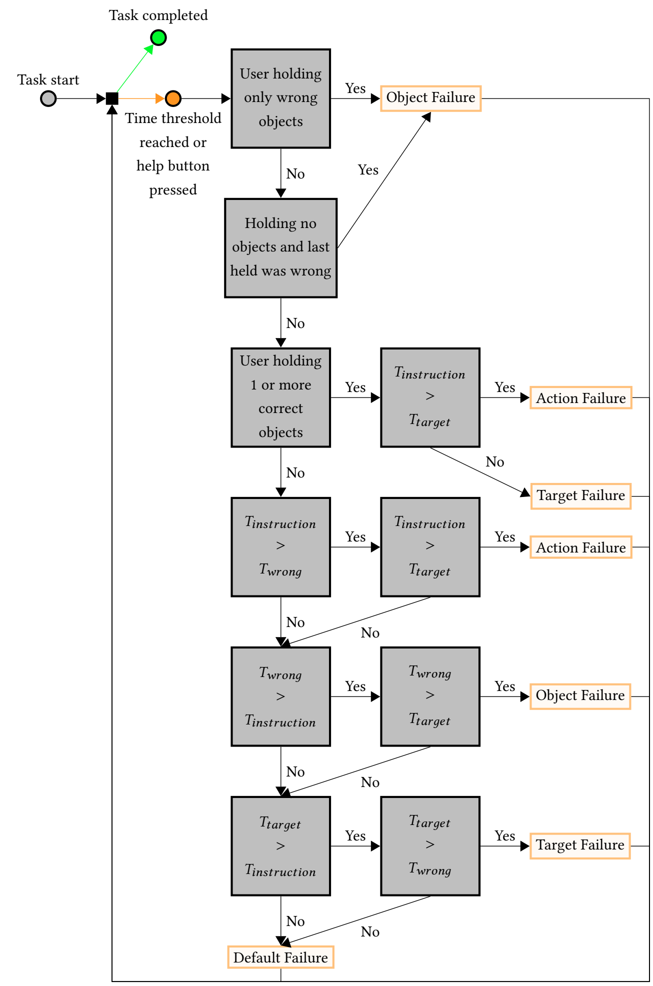


Fig. 3. Support is adaptively activated during the training depending on the failure state outcome of the algorithm. The algorithm determines the most suitable failure state each time a time threshold is reached (the user did not complete a task in time) or when the help button is pressed. Firstly, the objects the user is holding are taken into consideration. If these do not trigger a failure state, the eye tracking data will be compared. $T_{instructions}$, T_{target} and T_{wrong} refer to the time the user has been looking at elements tagged as Instructions, Target, and Wrong Objects (as described in Section 3.1) respectively during the current task.

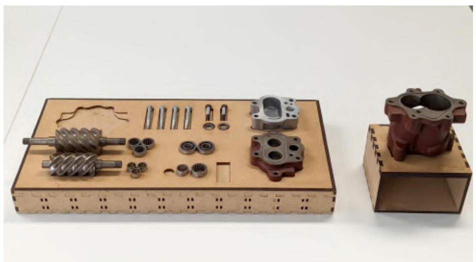
required action. In the next training iteration, the number of support systems is reduced by one for each task following the proposed hierarchy (see Figure 2). Following our illustrative example, if the step of applying the bolt to the hole resulted in additional support for an *action failure* and *object failure*. The next training iteration will only show support for the *action failure* for that specific step.

3.3 Fallback and adaptability

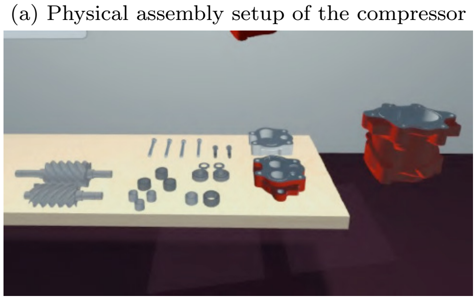
The system is designed for compatibility across VR platforms, with head tracking as an alternative to eye tracking. This ensures broader accessibility without compromising the adaptability of the training experience.

Table 1. Failure states and their corresponding support systems.

	Support systems
Object Failure: Wrong object selection	Incorrect object selections trigger warnings, and correct objects are highlighted.
Target Failure: Wrong object placement	Placement areas are highlighted with visual cues like circles.
Action Failure: Incorrect/No action execution	Text-based instructions explain the required actions.



(a) Physical assembly setup of the compressor



(b) VR replica of the compressor

Fig. 4. Physical and virtual setup of the compressor.

4. APPLICATION TO REAL-LIFE COMPANY CASES

In this section, we describe how we applied our concept to two real use cases from local industrial companies to demonstrate the applicability and generalizability of our approach.

VR training for compressor assembly. In this use case, trainees learn how to assemble a compressor in VR. The compressor consists of several components such as bearings, bolts, and rotors that must be assembled into a housing, as seen in Figure 4a. Figure 4b shows the virtual replica in the VR training. All assembly steps are performed without tools. In this case, the adaptive algorithm of Figure 3 reaches *task completed* each time the user places the correct component in the correct position with the correct orientation.

Our adaptive learning approach for this case consists of four support systems, inspired by prior work underpinning the advantages of using object coloring (Jasche et al. (2021); Vosinakis and Koutsabasis (2017)). They are also completely language-independent. Table 2 shows how these support systems were divided among the appropriate failure states. *Object Failure* has two support systems in this case: highlighting wrong objects and highlighting correct objects. Table 2 also shows the hierarchy between the two. This means that if an *Object Failure* is triggered, the highlighting of wrong objects will be activated before the highlighting of correct objects. Thus, the highlighting of correct objects will only get activated if the highlighting of wrong objects was already active

Table 2. Assignment of support systems to failure states for **compressor case**.

	Support system hierarchy	
Object Failure	Color wrong objects red	Color correct objects green
Target Failure	Hologram on target location	
Action Failure	Visual instructions	

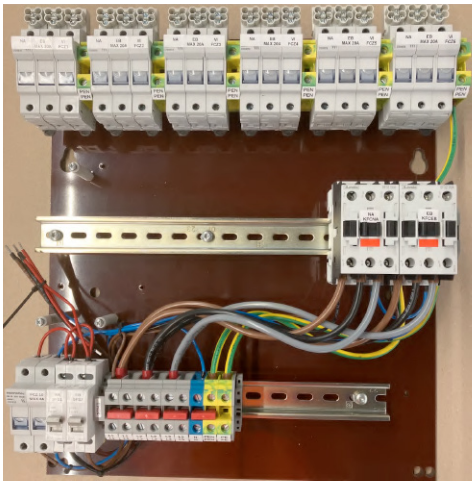
Table 3. Assignment of support systems to failure states for **electrical cabinet case**.

	Support system hierarchy	
Object Failure	Color grabbed object	Color object on picking box
Target Failure	Hologram + outline on target location	
Action Failure	Visual instructions	Animations

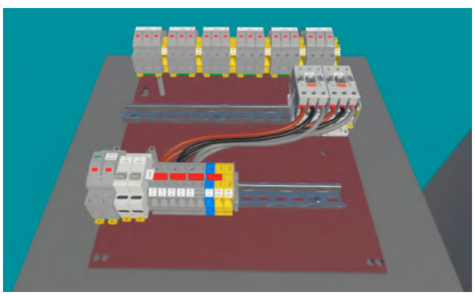
when this failure state was triggered. The order in the Object Failure support hierarchy is dictated by the fact that wrong objects only get highlighted in red when the user picks them up, while correct objects are highlighted in green regardless of whether the user has picked them up. Highlighting the correct objects in green will quickly show the user which object is correct, without overloading them with information. Thus, highlighting objects in this way inherently means that highlighting correct object will give the user more information about picking than only highlighting the wrong object the user is currently holding. This is also our reasoning behind the hierarchy between these two support systems. When support is deactivated, the order is reversed.

VR training for assembly of electrical cabinet. Our second use case is a VR tool that teaches the assembly procedure of an electrical cabinet for operators in a sheltered workplace (see Figure 5). This is still an assembly case, but it is quite different from the compressor. Other than picking and placing components inside an assembly housing, the training also focuses on wiring and screwing. Because of this, screwing motions have to be performed, sometimes by hand and sometimes with a specific screwdriver, depending on the component or wire. Other than screwing interactions, some components need to click into a rail and components that have a tab that needs to be pushed after placing it on the board. All these interactions come with a motion that the user has to perform in the virtual environment. This assembly procedure also has more than double the number of tasks and components compared to the compressor case. The adaptive algorithm of Figure 3 reaches *task completed* whenever the user places the correct component in the correct position with the correct orientation for picking steps. For steps that require interactions (screwing, clicking), it reaches *task completed* when the user successfully completes the entire interaction.

Table 3 illustrates the design of support systems tailored for this case. Similar to the approach used for the com-



(a) Physical assembly setup of the electrical cabinet



(b) VR replica of the electrical cabinet

Fig. 5. Physical and virtual setup of the electrical cabinet.

pressor, the support systems were adapted to meet the specific requirements of this scenario. For *Object Failure*, the grabbed object will be highlighted in green to confirm whether the user is holding the correct object or highlighted in red to inform that it is the wrong object. At a secondary level of support, the correct area in the environment where the object can be picked is also highlighted, similar to traditional picking boxes in assembly setups (due to the higher number of components needed for this assembly). For *Target Failure*, the same principles applied to the compressor case are used, with the addition of outlining the target location. This enhancement is necessary because the target location spans a larger area compared to the compressor, making the outline crucial for better visibility and user orientation. Regarding *Action Failure*, visual instructions are provided similarly to the compressor. However, a second level of support with animations is introduced for actions involving motion interactions. These animations serve to demonstrate the expected operations, such as screwing animations for screwing actions, pushing animations for components with tabs, and clicking animations for components that need to be clicked onto rails. This visual guidance supports users to clearly understand the required actions.

5. DISCUSSION

The application of the adaptive learning concept to two realistic assembly use cases demonstrated both the feasibility and flexibility of our approach. A key aspect of our work is the real-time adaptation enabled by built-in VR features, such as hand and eye tracking, which allow for the detection of failure states and their linkage to targeted support systems. While our concept provides a structured approach to adaptation, it remains flexible—developers can determine which support systems are best suited for different failure states based on the specific requirements of each use case. Unlike predefined feedback levels commonly found in the literature (Ulmer et al. (2022); Huang et al. (2021)), our approach employed a fully adaptive system that continuously tailors feedback to the trainee's specific needs. This dynamic approach ensures a more personalized and responsive learning experience, potentially improving training effectiveness across diverse learner profiles.

One limitation of our current work is that the evaluation was conducted on two realistic assembly use cases, which are more limited in scope as they primarily involve basic skill requirements and pick-and-place operations. While these scenarios reflect common tasks in industrial settings, future research should explore the application of our adaptive VR approach to more complex industrial use cases that involve advanced skills, decision-making, or collaborative tasks. Broadening the scope of evaluation will help assess the versatility of the approach and its adaptability to diverse industrial environments.

6. FUTURE WORK

One important direction for future research is the evaluation of our adaptive VR approach in comparison to traditional non-adaptive methods. Conducting a study that directly contrasts these two approaches would offer valuable insight into the efficiency of adaptive learning in VR, as well as its impact on user experience and perception, similar to the study approach of Huang et al. (2021). Understanding whether personalized learning paths enhance task performance and retention compared to a more generic approach will help guide the adoption of adaptive learning in industrial training environments. Secondly, the adaptive VR approach we described here has a large focus on adding additional support adaptively. However, removing support is performed based on a predetermined hierarchy of support systems. Future work should investigate how the information we are collecting during VR training can be used to remove support systems in a more personalized and adaptive way as well. We recommend a similar approach to adding feedback, where user behavior is analyzed to determine which support systems can be excluded. This can be achieved by tracking factors such as the duration each support system was active in an iteration and using this data to decide which one to remove in the next iteration.

Beyond evaluating the VR experience, it is crucial to assess how well physical skills learned in VR transfer to real-world tasks (Mayer (2011)). While VR training enhances cognitive and procedural knowledge, its real-world effectiveness remains key. We will evaluate both in a follow-up study.

7. CONCLUSION

In this paper, we have demonstrated the concept of supporting adaptive learning in VR assembly training by providing real-time guidance using VR's built-in eye and hand tracking. Our approach provides personalized support when the user is having difficulties performing tasks and removes support dynamically when the user is more proficient. Based on a collaboration with two manufacturing companies, we demonstrated the feasibility of applying our concept to a realistic use case within each business. Future research should evaluate the effectiveness of such adaptive VR training by comparing it to traditional non-adaptive approaches, focusing not only on user experience and efficiency but also on the transfer of physical skills to real-world performance. Additionally, exploring the potential of similar adaptive learning approaches in Augmented Reality and digital work instructions presents an exciting avenue, especially in environments where physical interaction is central, though challenges related to real-time object tracking remain.

ACKNOWLEDGEMENTS

This research was supported by Flanders Make, the strategic research centre for the manufacturing industry, in the project SKILLEDWORKFORCE.

REFERENCES

Ge, Z., Xi, M., and Li, Y. (2019). A literature review of the adaptive algorithms adopted in adaptive learning systems. In *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*, 254–258. IEEE.

Havard, V., Baudry, D., Jeanne, B., Louis, A., and Savatier, X. (2021). A use case study comparing augmented reality (ar) and electronic document-based maintenance instructions considering tasks complexity and operator competency level. *Virtual Reality*, 25(4), 999–1014. doi:10.1007/s10055-020-00493-z.

Huang, G., Qian, X., Wang, T., Patel, F., Sreeram, M., Cao, Y., Ramani, K., and Quinn, A.J. (2021). AdapTutAR: An adaptive tutoring system for machine tasks in augmented reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM. doi:10.1145/3411764.3445283.

Jasche, F., Hoffmann, S., Ludwig, T., and Wulf, V. (2021). Comparison of different types of augmented reality visualizations for instructions. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA. doi:10.1145/3411764.3445724. URL `https://doi.org/10.1145/3411764.3445724`.

Liu, X.W., Li, C.Y., Dang, S., Wang, W., Qu, J., Chen, T., and Wang, Q.L. (2022). Research on training effectiveness of professional maintenance personnel based on virtual reality and augmented reality technology. *Sustainability*, 14(21), 14351. doi:10.3390/su142114351.

Mayer, R.E. (2011). Applying the science of learning to multimedia instruction. In *Psychology of learning and motivation*, volume 55, 77–108. Elsevier. doi: https://doi.org/10.1016/B978-0-12-387691-1.00003-X.

Moreno, R. and Mayer, R.E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of educational psychology*, 94(3), 598. doi:10.1037/0022-0663.94.3.598.

Palmas, F., Labode, D., Plecher, D.A., and Klinker, G. (2019). Comparison of a gamified and non-gamified virtual reality training assembly task. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE. doi: 10.1109/vs-games.2019.8864583.

Strakos, J.K., Douglas, M.A., McCormick, B., and Wright, M. (2023). A learning management system-based approach to assess learning outcomes in operations management courses. *The International Journal of Management Education*, 21(2), 100802.

Studer, K., Lie, H., Zhao, Z., Thomson, B., Turakhia, D.G., and Liu, J. (2024). An open-ended system in virtual reality for training machining skills. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI EA '24. Association for Computing Machinery, New York, NY, USA. doi: 10.1145/3613905.3648666.

Taylor, D.L., Yeung, M., and Bashet, A.Z. (2021). *Personalized and Adaptive Learning*, 17–34. Springer International Publishing, Cham. doi:10.1007/978-3-030-58948-6_2.

Ulmer, J., Braun, S., Cheng, C.T., Dowey, S., and Wollert, J. (2022). Gamification of virtual reality assembly training: Effects of a combined point and level system on motivation and training results. *International Journal of Human-Computer Studies*, 165, 102854. doi: 10.1016/j.ijhcs.2022.102854.

Van Merrienboer, J.J. and Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational psychology review*, 17, 147–177.

Vincent-Ruz, P. and Boase, N.R.B. (2022). Activating discipline specific thinking with adaptive learning: A digital tool to enhance learning in chemistry. *PLOS ONE*, 17(11), 1–21. doi:10.1371/journal.pone.0276086. URL `https://doi.org/10.1371/journal.pone.0276086`.

Vosinakis, S. and Koutsabasis, P. (2017). Evaluation of visual feedback techniques for virtual grasping with bare hands using leap motion and oculus rift. *Virtual Reality*, 22(1), 47–62. doi:10.1007/s10055-017-0313-4.

Wolfartsberger, J., Zimmermann, R., Obermeier, G., and Niedermayr, D. (2023). Analyzing the potential of virtual reality-supported training for industrial assembly tasks. *Computers in Industry*, 147, 103838. doi: https://doi.org/10.1016/j.compind.2022.103838.

Yu, J., Kim, H., Zheng, X., Li, Z., and Zhu, X. (2024). Effects of scaffolding and inner speech on learning motivation, flexible thinking and academic achievement in the technology-enhanced learning environment. *Learning and Motivation*, 86, 101982.