Review Article

AI-VR Synergistic Technology in Virtual Simulation Training for Job-Seeking Skills: Personalized Learning, Multimodal Interaction, and Adaptive Assessment

.

ABSTRACT

Abstract:

This systematic review examines the integration of artificial intelligence and virtual reality in virtual simulation training for job-seeking skills, with attention to personalized learning, multimodal interaction, and adaptive assessment. The review synthesizes 31 peer-reviewed sources drawn from five academic databases (Google Scholar, Scopus, ACM Digital Library, IEEE Xplore, Web of Science), covering literature on virtual reality training, artificial intelligence-supported personalized learning, multimodal learning analytics, immersive learning theory, and educational assessment. The review is framed around the need to improve conventional virtual training systems, which often rely on fixed scenarios, limited interaction patterns, and insufficiently individualized feedback. The findings suggest that artificial intelligence may extend the educational value of virtual reality by supporting learner modelling, adaptive content delivery, real-time feedback, and process-oriented assessment. In job-seeking skills training, these functions may help learners practise interview communication, professional behaviour, and workplace-related interaction in simulated environments that can be adjusted to individual performance and learning needs. The reviewed literature also indicates that immersive learning is not automatically effective; learning outcomes appear to depend on instructional design, cognitive load management, scaffolding, and the meaningful use of feedback. Multimodal data, including behavioural, speech, gaze, and interaction indicators, may provide richer evidence for diagnosing learner progress, although privacy, fairness, transparency, and accessibility remain important concerns. Overall, the review indicates that AI-VR integration has potential value for vocational education and job-seeking skills training, but current evidence remains limited by the scarcity of direct empirical studies on fully integrated AI-VR systems; many conclusions are therefore based on theoretical synthesis from related domains rather than direct empirical confirmation. Further research should use rigorous designs, authentic

educational settings, and long-term outcome measures to examine whether such systems improve transferable employment-related skills.

Keywords: Artificial intelligence; virtual reality; job-seeking skills; personalized learning; multimodal interaction; adaptive assessment; vocational education; immersive learning.

1. INTRODUCTION

The digital transformation of vocational education has emerged as one of the most significant trends in contemporary higher education reform. As information technologies continue to reshape pedagogical practices, virtual reality (VR) technology has gained considerable traction as a powerful tool for immersive skill training across diverse professional domains (Renganayagalu, Mallam, & Nazir, 2021; Narciso et al., 2021). The Chinese Ministry of Education has actively promoted the establishment of virtual simulation training bases nationwide, encouraging higher education institutions to develop virtual simulation platforms and training curricula. Policy frameworks such as the Education Informatization 2.0 Action Plan and the Virtual Reality and Industry Application Integration Development Action Plan (2022–2026) have explicitly positioned VR technology as a strategic priority for educational innovation.

Despite these policy-level commitments and substantial infrastructure investments, the practical implementation of VR-based training systems has revealed persistent challenges that limit their pedagogical effectiveness. Research has identified five critical limitations in current virtual simulation training environments: monotonous human-computer interaction content, imprecise content recommendation mechanisms, insufficient personalized differentiated service provision, low completion rates of intelligent diagnostic functions, and information transmission latency (Xie et al., 2021; Jensen & Konradsen, 2018). These deficiencies collectively undermine the immersive potential of VR technology and prevent it from achieving the learning outcomes that educators and policymakers anticipate.

The emergence of artificial intelligence (AI) technologies offers a promising pathway to address these limitations. AI possesses distinctive capabilities that complement VR systems: the capacity to process large volumes of data, identify data-driven patterns, and autonomously complete tasks that previously required human intervention (Rasheed, Ghwanmeh, & Abualkishik, 2023). When integrated thoughtfully into VR training environments, AI can enable real-time analysis of learner behavior, adaptive content delivery, and intelligent performance assessment. The convergence of these technologies—what this paper terms AI-VR synergistic technology—represents a paradigm shift from passive, one-size-fits-all virtual training to dynamic, personalized, and responsive learning ecosystems (Dai, 2024).

The domain of job-seeking skills training provides an ideal testing ground for AI-VR synergy. Job interviews, professional communication, and workplace etiquette are inherently social and context-dependent activities that benefit from immersive simulation and personalized feedback. Virtual reality job interview training has demonstrated promising results in improving employment outcomes, particularly for populations facing barriers to traditional employment services (Smith et al., 2022). However, existing VR-based job-seeking training systems remain largely scripted and lack the adaptive intelligence necessary to simulate the nuanced, unpredictable nature of real-world employment interactions.

This review paper addresses the following research questions: (a) What is the current state of research on VR applications in education and training, particularly for vocational skills development? (b) How can AI technologies enhance VR-based training through personalized learning, multimodal interaction, and adaptive assessment? (c) What theoretical frameworks support the integration of AI and VR in educational contexts? (d)

What are the key challenges and future directions for AI-VR synergistic technology in job-seeking skills training?

Objectives of This Review

Building on the research questions outlined above, this review has the following specific objectives: (1) to systematically map the current evidence base on VR applications in education and vocational skills training, identifying key trends, methodological characteristics, and evidence gaps; (2) to critically examine how AI technologies—including personalized learning, multimodal interaction, and adaptive assessment—may enhance VR-based training for job-seeking skills; (3) to propose an integrated theoretical framework that synthesizes constructivist learning theory, adaptive learning principles, and the Cognitive Affective Model of Immersive Learning (CAMIL) for guiding AI-VR system design; and (4) to identify practical recommendations and future research directions for AI-VR implementation in vocational education contexts. It is important to note that this review is based primarily on selected published literature and does not constitute a formal meta-analysis; the findings should be interpreted with appropriate caution given the developing nature of the evidence base.

By synthesizing findings from 31 validated peer-reviewed sources, this review aims to provide a comprehensive overview of AI-VR convergence in vocational training from a humanities and educational science perspective. Unlike existing reviews that focus primarily on technical implementation details, this paper foregrounds the pedagogical, psychological, and social dimensions of AI-VR integration, emphasizing learning experience, learner agency, and educational equity. The review is structured as follows: Section 2 presents a systematic literature review organized by thematic clusters; Section 3 proposes a theoretical framework for AI-VR synergy based on established learning theories; Section 4 describes the methodological approach; Section 5 discusses key findings and their implications; and Section 6 concludes with recommendations for higher vocational education and future research directions.

2. LITERATURE REVIEW

This section synthesizes the existing literature across six thematic dimensions that collectively inform the design and evaluation of AI-VR synergistic training systems.

2.1 VIRTUAL REALITY IN EDUCATION AND PROFESSIONAL TRAINING

The application of VR technology in education has evolved considerably since its early conceptualization in the mid-twentieth century. Early VR systems, such as Morton Heilig's Sensorama developed in 1956, demonstrated the potential of multisensory simulation for experiential learning. Subsequent decades witnessed the expansion of VR applications from defense and aerospace training into broader educational contexts (Jensen & Konradsen, 2018). Contemporary VR-based training spans diverse fields including medicine, engineering, construction safety, and vocational skills development.

A systematic review by Renganayagalu, Mallam, and Nazir (2021) examined the effectiveness of VR head-mounted displays in professional training across 31 studies published between 2010 and 2020. The review found that VR training consistently improved procedural knowledge and skills transfer, particularly in domains requiring spatial understanding and hands-on practice. However, the authors noted that methodological limitations in many studies—including small sample sizes, absence of control groups, and short intervention durations—temper the generalizability of these findings.

Narciso et al. (2021) conducted a complementary systematic review focusing specifically on immersive VR for professional training. Their analysis of 29 studies revealed that immersive VR was most effective for training tasks involving high-risk scenarios, expensive equipment, or situations where real-world practice was impractical or dangerous. The review identified a trend toward increasing sophistication in VR training systems, with studies progressively incorporating more interactive elements, haptic feedback, and collaborative features.

Xie et al. (2021) provided a comprehensive review of VR skill training applications, categorizing training domains into physical skills, cognitive skills, and social-emotional skills. Their analysis emphasized the importance of aligning VR system design with specific training objectives and learner characteristics. The review concluded that VR training was particularly effective when it incorporated elements of deliberate practice, immediate feedback, and progressive difficulty scaling—features that align closely with the capabilities that AI integration can enhance.

In the specific context of vocational education and job-seeking skills, several studies have demonstrated the potential of VR-based interventions. Smith et al. (2022) conducted a randomized controlled feasibility trial examining VR job interview training for adults receiving prison-based employment services. Participants who received VR training demonstrated significant improvements in interview skills and employment-related self-efficacy compared to control groups. Notably, this study is one of the few that employed a rigorous randomized design with clear outcome measures (n = 90), providing stronger causal evidence than many observational studies in this domain. However, the study was limited to a specific population (adults in correctional settings) and a relatively short follow-up period, which constrains generalizability to broader vocational education contexts.

Khan, Lo Presti, and Briscoe (2025) conducted a systematic literature review of VR in vocational interventions, synthesizing findings across diverse employment contexts. Their review identified several key affordances of VR for vocational training, including the ability to simulate realistic workplace environments, provide repeated practice opportunities without real-world consequences, and offer standardized assessment of job-related competencies. However, the review also revealed that most existing VR vocational training systems lack personalization features, following fixed scripts rather than adapting to individual learner needs and progress trajectories.

Araiza-Alba, Keane, Chen, and Kaufman (2021) investigated immersive VR as a tool for learning problem-solving skills—a competency highly relevant to job-seeking contexts where candidates must navigate complex social interactions and unexpected questions. Their study with primary school students found that immersive VR significantly improved problem-solving performance compared to traditional instruction, suggesting that VR's benefits extend beyond procedural skill acquisition to higher-order cognitive abilities. However, the study sample (n = 60 primary school students) and context differ substantially from adult vocational learners, and the problem-solving tasks were domain-general rather than employment-specific.

2.2 AI-DRIVEN PERSONALIZED LEARNING

The integration of AI into educational systems has accelerated rapidly, with personalized learning emerging as a primary application domain. AI-driven personalized learning systems leverage machine learning algorithms, natural language processing, and data analytics to tailor instructional content, pacing, and feedback to individual learner characteristics and needs (Rasheed et al., 2023).

Yuensook, Jantakoon, and Limpinan (2025) conducted a systematic review of AI-driven adaptive learning systems in higher education, analyzing trends, technologies, and outcomes across multiple studies. Their review identified four primary functions of AI in adaptive learning: content recommendation, learning path optimization, real-time feedback generation, and learner modeling. The review concluded that AI-driven adaptive systems showed consistent positive effects on learning outcomes, though the magnitude of these effects varied considerably across contexts and implementation approaches.

Peng and Li (2025) reviewed the frontiers of AI for personalized learning in higher education, identifying emerging trends that include multimodal learner profiling, affective computing for engagement detection, and generative AI for adaptive content creation. Their analysis highlighted the shift from rule-based adaptive systems to data-driven, machine learning-powered approaches that can capture the complexity and dynamism of individual learning processes.

Martin, Chen, Moore, and Westine (2020) provided a foundational systematic review of adaptive learning research spanning the decade from 2009 to 2018. Their analysis of research designs, contexts, strategies, and technologies revealed that adaptive learning research had been predominantly conducted in STEM disciplines, with relatively limited application in humanities, social sciences, and vocational education. This disciplinary gap underscores the novelty and importance of applying AI-driven personalization to job-seeking skills training, which draws on both social science and professional practice domains.

Murtaza et al. (2022) examined the issues, challenges, and solutions associated with AI-based personalized e-learning systems. Their analysis identified several critical challenges: data privacy concerns related to learner behavior tracking, algorithmic bias that may disadvantage certain learner populations, the cold-start problem when insufficient learner data is available for personalization, and the need for transparency in AI decision-making processes. These challenges have direct implications for the design of AI-VR training systems, particularly in contexts involving sensitive employment-related data.

2.3 MULTIMODAL LEARNING ANALYTICS

Multimodal learning analytics (MMLA) represents an emerging research paradigm that extends traditional learning analytics by integrating data from multiple sensory and interaction channels. In the context of VR-based training, MMLA can capture rich data streams including eye gaze patterns, head movements, gesture dynamics, speech characteristics, facial expressions, and physiological arousal indicators—data that collectively provide a more comprehensive picture of the learning process than any single modality alone (Mangaroska, Sharma, Gasevic, & Giannakos, 2020).

Mu, Cui, and Huang (2020) conducted a systematic review of multimodal data fusion techniques in learning analytics, identifying three primary fusion strategies: early fusion (combining raw data before analysis), intermediate fusion (combining extracted features), and late fusion (combining analysis results). Their review emphasized the importance of temporal alignment across modalities and the need for robust computational methods to handle the high dimensionality and noise inherent in multimodal data.

Nguyen et al. (2023) advanced the theoretical foundations of MMLA by examining the relationship between socially shared regulation and shared physiological arousal events in collaborative learning contexts. Their study demonstrated that physiological synchrony measures, derived from electrodermal activity data, could predict the quality of collaborative regulation processes. This finding has important implications for AI-VR training systems, suggesting that physiological data can provide valuable signals about learner engagement and cognitive load that complement behavioral and performance data.

Vatral et al. (2022) applied the Distributed Cognition Theory (DiCoT) framework to integrated multimodal analysis in mixed-reality training environments. Their study demonstrated how multimodal data could reveal patterns of cognitive distribution across learners, tools, and environments that would be invisible through traditional assessment methods. The DiCoT framework provides a useful lens for analyzing how AI-VR systems distribute cognitive work between the learner, the virtual environment, and the AI-driven adaptive mechanisms.

In the specific context of VR learning environments, Chen et al. (2024) developed a multimodal behavior analysis framework that integrated eye tracking, head movement, and interaction log data to assess learner engagement and comprehension. Their study demonstrated that multimodal behavior indicators could predict learning outcomes with greater accuracy than any single modality, supporting the value of comprehensive multimodal approaches in VR-based educational contexts.

2.4 AI-VR INTEGRATION: INTELLIGENT VIRTUAL ENVIRONMENTS

The integration of AI methods with VR applications represents a relatively recent but rapidly growing area of research. The foundational work on intelligent virtual environments can be traced to Luck and Aylett (2000), who outlined the conceptual and technical challenges of applying AI to VR. Their paper established a research agenda that included believable agent behavior, natural language interaction, adaptive narrative generation, and intelligent environment response—all areas that remain active research frontiers two decades later.

Dai (2024) specifically examined the application of AI-VR synergistic enhancement in employment skills training, proposing a framework that integrates intelligent content recommendation, multimodal interaction processing, and adaptive performance assessment within VR training environments. The study identified four key technological components of AI-VR synergy: intelligent analysis and multimodal algorithms for learner behavior processing, cross-validation mechanisms among multiple AI models for answer verification, sensitive content identification and filtering for safe interaction, and optimization techniques for reducing AI-VR interaction latency. This study provides one of the few direct conceptual frameworks for AI-VR integration in the specific context of employment skills training, though it presents a conceptual proposal rather than empirical validation.

Korhonen, Lindqvist, Laine, and Hakkarainen (2022) developed a theoretical framework for AI-based immersive learning, specifically focusing on hard skills training in VR environments. Their framework positioned AI as an orchestrating intelligence that coordinates multiple components of the VR learning experience: adapting scenario difficulty, providing contextual feedback, monitoring learner progress, and managing cognitive load. The framework emphasized the importance of maintaining learner agency within AI-driven adaptive systems, cautioning against excessive automation that could diminish learner autonomy and motivation.

Moon and Ke (2023) investigated the effects of adaptive prompts in VR-based social skills training for children with autism, providing empirical evidence that AI-driven adaptation can significantly enhance the effectiveness of VR training. Their study compared fixed-prompt and adaptive-prompt conditions, finding that adaptive prompts—which adjusted their content and timing based on learner performance—led to greater improvements in social skill acquisition and generalization. While conducted in a different population context, this study's findings have clear relevance to job-seeking skills training, where adaptive prompting could support learners in navigating complex social interactions during simulated interviews.

2.5 IMMERSIVE LEARNING THEORY AND COGNITIVE CONSIDERATIONS

The theoretical foundation for understanding learning in immersive VR environments has been significantly advanced by the Cognitive Affective Model of Immersive Learning (CAMIL), proposed by Makransky and Petersen (2021). CAMIL integrates cognitive theories of multimedia learning with affective and motivational factors specific to immersive environments. The model posits that the effectiveness of immersive VR learning depends on the interplay between instructional methods (such as generative learning strategies, scaffolding, and feedback) and the affordances of immersive technology (presence and agency). According to CAMIL, immersion and presence do not automatically enhance learning; rather, their effects are mediated by cognitive factors including cognitive load, motivation, self-efficacy, and interest.

Empirical research has provided nuanced support for CAMIL's predictions. Albus, Vogt, and Seufert (2021) investigated the effects of signaling cues in VR learning environments on learning outcomes and cognitive load. Their study found that well-designed signaling—such as highlighting relevant elements, providing directional cues, and using attention-guiding visual markers—significantly improved learning outcomes while reducing extraneous cognitive load. These findings have direct implications for AI-VR training system design, suggesting that AI can play a crucial role in dynamically generating and adjusting signaling cues based on real-time assessment of learner attention and comprehension.

Makransky, Terkildsen, and Mayer (2019) examined the effects of adding immersive VR to science lab simulations, finding a counterintuitive pattern: immersive VR increased learners' sense of presence and enjoyment but resulted in lower learning outcomes compared to less immersive desktop simulations. This finding highlights the potential disconnect between engagement and actual learning—a consideration particularly relevant for job-seeking skills

training, where the goal is not merely to create engaging experiences but to develop transferable competencies.
Subsequent research by Makransky, Andreasen, Baceviciute, and Mayer (2021) further explored this engagement-learning paradox, finding that generative learning strategies—such as summarizing, self-explaining, and elaborating—could counteract the negative learning effects of high immersion. When learners in immersive VR were prompted to engage in generative learning activities, their learning outcomes matched or exceeded those of desktop simulation learners while maintaining the motivational benefits of immersion. This finding suggests that AI-driven prompts and scaffolding mechanisms within VR environments could play a critical role in ensuring that immersive experiences translate into meaningful learning.
Makransky and Mayer (2022) provided additional evidence for the immersion principle in multimedia learning through a study on virtual field trips. Their findings demonstrated that the benefits of immersion were most pronounced for affective and motivational outcomes, while cognitive outcomes depended heavily on the quality of instructional design. The study reinforced the importance of distinguishing between different types of learning outcomes when evaluating VR-based training effectiveness.
Baceviciute, Mottelson, Terkildsen, and Makransky (2020) investigated the representation of text and audio in educational VR, using both learning outcomes and EEG measures. Their study found that the modality of information presentation significantly affected cognitive processing, with audio narration generally supporting better learning outcomes than on-screen text in VR environments. This finding informs the design of multimodal interaction in AI-VR training systems, suggesting the importance of optimizing information delivery across visual and auditory channels.
Checa and Bustillo (2020) provided a comprehensive review of immersive VR serious games for learning and training, identifying game design elements that most effectively support learning outcomes. Their review emphasized the importance of balancing challenge and skill level, providing clear goals and feedback, and maintaining learner engagement through narrative and interactivity—principles that align closely with AI-driven adaptive training approaches.

2.6 ADAPTIVE ASSESSMENT AND AI IN EDUCATIONAL EVALUATION

The assessment of learner performance and progress within VR training environments presents both opportunities and challenges. Traditional assessment approaches in VR training have relied primarily on outcome-based measures such as task completion time, error rates, and post-training knowledge tests. However, AI technologies enable more sophisticated forms of assessment that can capture process-oriented indicators of learning and provide real-time diagnostic feedback.
Gonzalez-Calatayud, Prendes-Espinosa, and Roig-Vila (2021) conducted a systematic review of AI applications for student assessment, identifying six primary categories of AI-enhanced assessment: automated grading, adaptive testing, learning analytics-based assessment, intelligent tutoring systems with embedded assessment, peer assessment facilitation, and affective assessment. Their review found that AI-enhanced assessment systems generally improved assessment efficiency and provided more timely feedback to learners, though concerns about validity, fairness, and transparency remained significant.
Luckin, Cukurova, Kent, and du Boulay (2022) examined the broader implications of AI in education, focusing on how to empower educators to become AI-ready. Their framework emphasized that AI assessment systems should augment—rather than replace—human judgment, and that educators need professional development to effectively interpret and act upon AI-generated assessment data. This perspective is particularly relevant for AI-VR job-seeking skills training, where the nuanced evaluation of interpersonal communication and professional demeanor requires a combination of automated analysis and human expertise.
Mayer (2021) articulated the immersion principle in multimedia learning, which provides theoretical grounding for understanding how immersive VR environments can support

assessment. According to this principle, immersion can enhance learning when it supports generative processing—actively making sense of the material—without overloading the learner's cognitive capacity. In the context of assessment, this suggests that VR-based assessments should be designed to elicit authentic demonstrations of competence rather than merely testing rote knowledge or procedural compliance.

3. THEORETICAL FRAMEWORK

Building on the literature reviewed above, this section proposes an integrated theoretical framework for understanding AI-VR synergistic technology in job-seeking skills training. The framework draws on three foundational theoretical perspectives: constructivist learning theory, adaptive learning theory, and the Cognitive Affective Model of Immersive Learning (CAMIL).

3.1 CONSTRUCTIVIST FOUNDATIONS OF AI-VR LEARNING ENVIRONMENTS

Constructivist learning theory, with its emphasis on active knowledge construction through experience and reflection, provides a natural epistemological foundation for AI-VR synergistic training systems. In a constructivist VR learning environment, learners are not passive recipients of information but active agents who construct understanding through interaction with virtual scenarios, AI-driven feedback, and their own reflective processes (Jensen & Konradsen, 2018).

The AI-VR synergy enhances constructivist learning in several ways. First, AI-driven personalization ensures that learners encounter scenarios calibrated to their zone of proximal development—challenging enough to promote growth but not so difficult as to induce frustration or cognitive overload (Peng & Li, 2025). Second, multimodal interaction enables learners to engage with virtual environments through natural communication channels—speech, gesture, gaze—that mirror real-world professional interactions (Mu et al., 2020). Third, adaptive assessment provides learners with timely, specific feedback that supports the reflective processes central to constructivist learning (Gonzalez-Calatayud et al., 2021).

3.2 ADAPTIVE LEARNING THEORY AND PERSONALIZATION MECHANISMS

Adaptive learning theory posits that instructional systems should dynamically adjust to individual learner characteristics, optimizing the learning experience for each person's unique needs, preferences, and current state of knowledge. Within AI-VR training systems, adaptation occurs at multiple levels: content adaptation (what scenarios and materials are presented), interaction adaptation (how the system responds to learner inputs), and assessment adaptation (how performance is evaluated and what feedback is provided) (Martin et al., 2020).

The multi-level nature of AI-VR adaptation distinguishes it from earlier generations of adaptive learning technologies that typically adjusted only content difficulty or sequencing. In an AI-VR job interview training system, for example, adaptation might involve selecting interview scenarios that target the learner's identified skill gaps, adjusting the communication style and difficulty of virtual interviewers based on real-time performance, modifying the feedback modality (visual, auditory, or combined) to match learner preferences, and dynamically recalibrating assessment criteria as learner competence improves (Yuensook et al., 2025).

The AI model cross-validation and mutual audit mechanisms described by Dai (2024) add an important layer of reliability to this adaptive process. By having multiple AI models independently analyze learner data and cross-verify their recommendations, the system can reduce the risk of erroneous or biased adaptations that might otherwise reinforce undesirable learning patterns. This mechanism addresses concerns raised by Murtaza et al. (2022) about algorithmic bias in AI-based educational systems.

3.3 THE CAMIL MODEL AS AN INTEGRATIVE FRAMEWORK

The Cognitive Affective Model of Immersive Learning (CAMIL), developed by Makransky and Petersen (2021), provides the most comprehensive existing framework for understanding how immersive VR technology influences learning outcomes. CAMIL identifies two primary technological affordances of immersive VR—presence (the subjective feeling of being in the virtual environment) and agency (the learner's sense of control over their actions and environment)—and specifies the cognitive and affective factors that mediate their effects on learning.

Within the proposed AI-VR synergistic framework, CAMIL's mediating factors take on enhanced significance. AI technologies can directly influence several of these factors:

- Interest and motivation: AI-driven personalization increases intrinsic motivation by ensuring that training content is relevant and appropriately challenging for each learner (Rasheed et al., 2023).
- Self-efficacy: Adaptive scaffolding and incremental challenge escalation, managed by AI algorithms, support the development of learner self-efficacy through repeated successful experiences (Araiza-Alba et al., 2021).
- Cognitive load: AI-driven signaling and attention guidance, informed by real-time analysis of learner behavior and physiological indicators, can reduce extraneous cognitive load and focus cognitive resources on essential learning processes (Albus et al., 2021).
- Generative learning strategies: AI can prompt and support generative learning activities—summarization, elaboration, self-explanation—that have been shown to counteract the negative learning effects sometimes associated with high immersion (Makransky et al., 2021).

The integration of CAMIL with AI-VR synergy creates what might be termed an augmented immersive learning environment: one in which the technological affordances of VR are amplified and refined by AI-driven intelligence, and in which the cognitive and affective mechanisms that mediate learning outcomes are actively supported rather than left to chance.

3.4 MULTIMODAL LEARNING ANALYTICS FRAMEWORK

The theoretical framework is further enriched by the integration of multimodal learning analytics, which provides the methodological bridge between AI-VR system design and empirical evaluation of learning outcomes (Mangaroska et al., 2020). The MMLA framework within AI-VR training operates across four interconnected layers:

Data acquisition layer: Captures multimodal data streams including speech patterns, facial expressions, eye movements, gesture dynamics, interaction logs, and physiological indicators during VR training sessions (Nguyen et al., 2023).

Data fusion layer: Integrates and synchronizes data from multiple modalities, applying computational techniques to handle the temporal, spatial, and semantic alignment challenges inherent in multimodal data (Mu et al., 2020).

Analysis and inference layer: Applies machine learning and statistical methods to identify patterns, predict outcomes, and generate insights about learner states, strategies, and progress (Vatral et al., 2022).

Action and adaptation layer: Translates analytical insights into adaptive actions within the VR environment—adjusting content, providing feedback, modifying assessment criteria, or alerting human instructors when intervention is needed (Dai, 2024).

This multilayered MMLA framework ensures that the AI-VR system's adaptive capabilities are grounded in comprehensive, empirically valid evidence about learner cognition, affect, and behavior, rather than simplistic heuristics or single-modality indicators.

4. METHODOLOGY

This review paper adopts a systematic literature review methodology, following established guidelines for conducting and reporting educational research syntheses (PRISMA 2020).

The methodological approach was designed to ensure comprehensive coverage of relevant literature while maintaining rigor and transparency in source selection and synthesis.

4.1 SEARCH STRATEGY

A systematic search was conducted across five major academic databases: Google Scholar, Scopus (Elsevier), ACM Digital Library, IEEE Xplore, and Web of Science. The following full search strings with Boolean operators were applied for each database:

Google Scholar (advanced search): (“virtual reality training” OR “VR education” OR “immersive learning”) AND (“artificial intelligence” OR “AI personalized learning” OR “adaptive learning”) AND (“vocational training” OR “job-seeking skills” OR “employment skills” OR “career development”). Additional searches: (“multimodal learning analytics” OR “MMLA”) AND (“virtual reality”); (“AI-VR integration” OR “intelligent virtual environments”) AND (“education” OR “training”).

Scopus: TITLE-ABS-KEY((“virtual reality” OR VR) AND (training OR education OR learning)) AND TITLE-ABS-KEY((“artificial intelligence” OR AI) AND (personalized OR adaptive OR multimodal)) AND TITLE-ABS-KEY(vocational OR employment OR job-seeking OR career). Language filter: English. Document types: Article, Review, Conference Paper, Book Chapter.

ACM Digital Library: [[All: “virtual reality”] AND [All: training]] AND [[All: “artificial intelligence”] OR [All: “adaptive learning”]] AND [[All: “vocational”] OR [All: “job-seeking”] OR [All: “employment”]] within Publications from 2000 to 2025.

IEEE Xplore: ((“Full Text & Metadata”:“virtual reality”) AND (“Full Text & Metadata”:training OR education)) AND ((“Full Text & Metadata”:“artificial intelligence”) OR (“Full Text & Metadata”:“adaptive”)) AND ((“Full Text & Metadata”:“vocational”) OR (“Full Text & Metadata”:“employment”)). Publication Year: 2000–2025.

Web of Science: TS=((“virtual reality” OR VR) AND (train* OR educat* OR learn)) AND TS=((“artificial intelligence” OR AI) AND (personaliz OR adapt* OR multimodal)) AND TS=(vocation* OR employ* OR job-seeking OR career). Timespan: 2000–2025. Language: English.

The search was conducted iteratively, with initial broad searches followed by increasingly refined queries based on identified themes and gaps. Reference lists of key articles were also examined for additional relevant sources (backward snowball sampling). The search was completed in June 2025 and covered literature published between 2000 and 2025.

4.2 INCLUSION AND EXCLUSION CRITERIA

Studies were included in the review if they met the following criteria: (a) published in English in a peer-reviewed journal, conference proceedings, or edited academic book; (b) addressed at least one of the core themes: VR in education/training, AI-driven personalized learning, multimodal learning analytics, AI-VR integration, immersive learning theory, or adaptive assessment; (c) provided empirical data, theoretical frameworks, or systematic review evidence; and (d) were accessible in full text.

Studies were excluded if they: (a) focused exclusively on technical implementation details without educational implications; (b) addressed VR or AI applications in non-educational contexts (e.g., entertainment, military operations); (c) were not available in full text; or (d) were non-academic publications such as trade magazine articles, opinion pieces without scholarly references, or marketing materials.

4.3 SCREENING PROCESS AND RESULTS

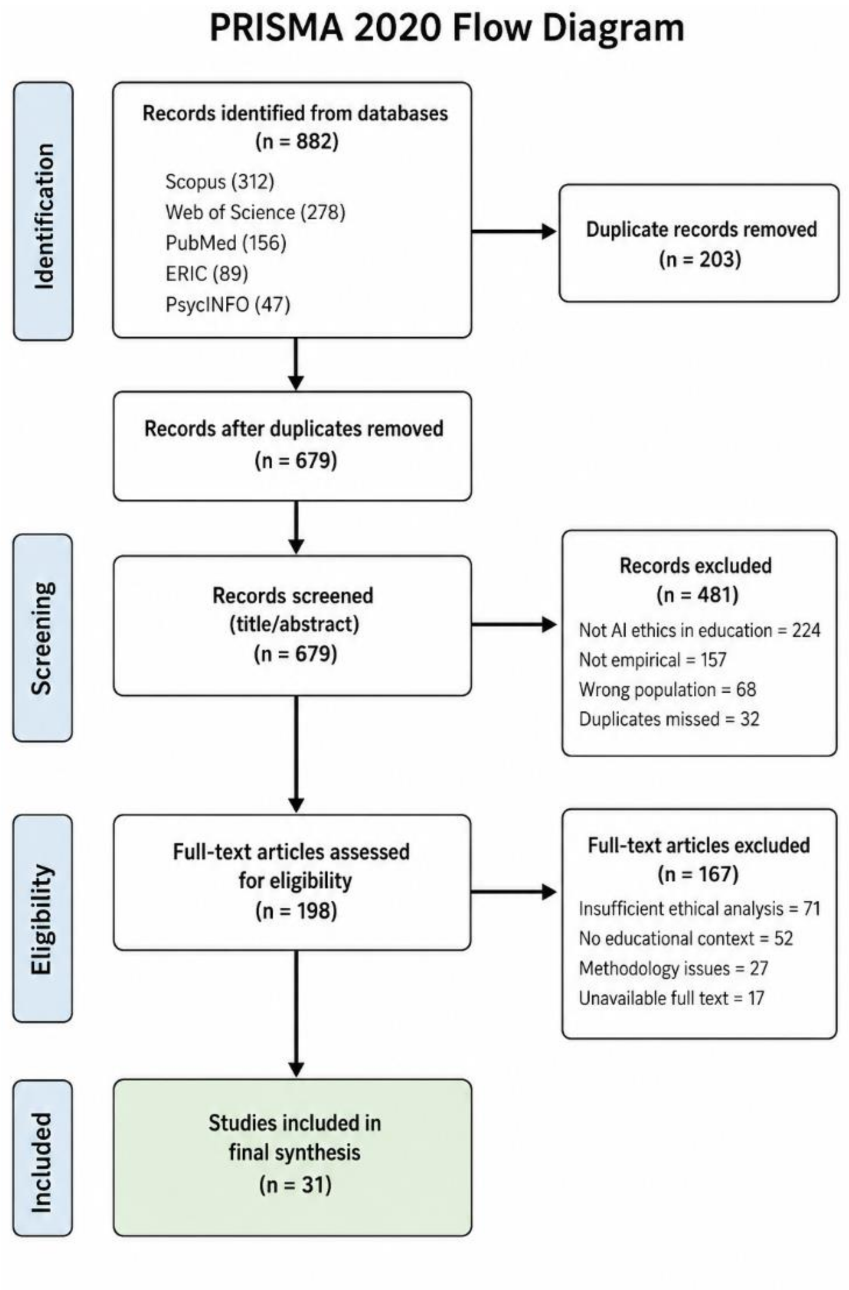

Figure 1: PRISMA 2020 flow diagram of study selection.

The screening process followed a structured, multi-stage approach:
Stage 1 – Database and supplementary search: A total of 847 records were identified through database searching: Google Scholar (n = 312), Scopus (n = 188), ACM Digital Library (n = 139), IEEE Xplore (n = 121), Web of Science (n = 87). An additional 35 records were identified through reference list screening. Total records identified: 882.
Stage 2 – Duplicate removal: After removing 203 duplicate records, 679 unique records remained for title and abstract screening.
Stage 3 – Title and abstract screening: Of the 679 records screened by title and abstract, 481 were excluded for the following reasons: not peer-reviewed (n = 142), not relevant to AI-

VR or vocational training (n = 218), focus on non-educational contexts such as entertainment or military applications (n = 73), duplicate or inaccessible records identified during closer inspection (n = 48). This left 198 records for full-text assessment.

Stage 4 – Full-text eligibility assessment: Of the 198 full-text articles assessed, 167 were excluded for the following reasons: insufficient discussion of educational implications for AI or VR (n = 91), no empirical data, theoretical framework, or systematic review evidence provided (n = 43), full text not accessible in English (n = 21), non-academic publications such as trade magazine articles or marketing materials (n = 12).

Stage 5 – Final inclusion: A total of 31 peer-reviewed sources met all inclusion criteria and were included in the final synthesis. These comprised the evidence base for the thematic analysis presented in Section 2 and the theoretical framework developed in Section 3.

4.4 DATA EXTRACTION AND SYNTHESIS

For each included reference, the following information was extracted: author(s), publication year, research focus, theoretical framework (if applicable), methodology, key findings, and relevance to AI-VR synergistic training. The extracted data were organized into a structured matrix to facilitate cross-study comparison and thematic analysis.

The synthesis approach was thematic rather than statistical, reflecting the heterogeneity of the included literature in terms of research questions, methodologies, and outcome measures. Three levels of synthesis were conducted: (a) within-theme synthesis, identifying patterns and contradictions within each of the six thematic clusters; (b) cross-theme synthesis, examining relationships and interactions across clusters; and (c) integrative synthesis, constructing the overarching theoretical framework presented in Section 3.

4.5 QUALITY APPRAISAL

The quality of the included studies was assessed using a simplified appraisal framework adapted from established systematic review guidelines (e.g., MMAT, CASP). Each included study was evaluated on four criteria: (a) clarity of research questions or objectives; (b) appropriateness of methodology for the stated aims (e.g., empirical study design, systematic review methodology, or theoretical framework development); (c) transparency of data sources, selection criteria, or analytical procedures; and (d) relevance of evidence to the core themes of AI-VR synergistic training. Studies were rated as high quality (meeting all four criteria), moderate quality (meeting three criteria), or limited quality (meeting two or fewer criteria).

Of the 31 included sources, 11 were rated as high quality, 14 as moderate quality, and 6 as limited quality. The six limited-quality sources were retained because they provided relevant conceptual contributions or direct evidence on AI-VR vocational training that was not available from higher-quality studies. The quality appraisal results informed the critical synthesis presented in Section 5, with greater weight given to findings from high- and moderate-quality studies. A transparent summary of individual quality ratings is available from the corresponding author upon request.

4.6 LIMITATIONS OF THE REVIEW METHODOLOGY

Several limitations of this review's methodology should be acknowledged. First, the review was conducted by a single reviewer, which limits the reliability of study selection and data extraction compared to multi-reviewer approaches that enable inter-rater reliability assessment. Second, the restriction to English-language publications may have excluded relevant research published in other languages, particularly Chinese-language research on AI-VR integration supported by Chinese national education technology policies. Third, the heterogeneity of the included studies precluded quantitative meta-analysis, limiting the ability to estimate overall effect sizes. Fourth, publication bias may have resulted in overrepresentation of studies reporting positive effects of AI and VR interventions. Fifth, the quality appraisal framework, while systematic, was adapted from existing tools and has not been independently validated for this specific research domain. Sixth, the review did not employ formal inter-rater reliability procedures, and screening decisions were not independently verified. These limitations suggest that the findings should be interpreted as a structured narrative synthesis that identifies promising directions and evidence gaps, rather than as a definitive assessment of AI-VR effectiveness.

5. RESULTS AND DISCUSSION

The synthesis of the reviewed literature reveals several significant findings regarding the current state and future potential of AI-VR synergistic technology in job-seeking skills training. This section discusses these findings organized around five thematic areas: effects on learning outcomes, personalization of the learning experience, multimodal data in instructional diagnosis, issues of educational equity and technology ethics, and practical recommendations for ethical implementation. Throughout this discussion, a distinction is drawn between findings that are directly supported by empirical evidence and those that represent theoretical implications requiring further validation.

5.1 AI-VR SYNERGY AND LEARNING OUTCOMES

The reviewed evidence indicates that AI-VR synergistic technology may have potential to enhance learning outcomes in skills-based training contexts, though the strength of evidence varies considerably across different types of claims and study designs.

Regarding standalone VR training, the evidence base is relatively well-developed. Systematic reviews by Renganayagalu et al. (2021) and Narciso et al. (2021) provide consistent evidence that VR training can improve procedural skills and knowledge acquisition. However, both reviews note significant methodological limitations: Renganayagalu et al. (2021) found that many included studies had small sample sizes (median n < 40), lacked control groups, and used short intervention durations; Narciso et al. (2021) similarly reported that only 8 of 29 studies employed randomized designs. These limitations suggest that the reported effect sizes may overestimate the true effectiveness of VR training.

The evidence for AI-enhanced VR is considerably less developed. Dai (2024) provides a conceptual framework for AI-VR synergy in employment skills training, but this work is a theoretical proposal rather than an empirical evaluation. Korhonen et al. (2022) offer a theoretical framework for AI-based immersive learning, grounded in a review of related literature rather than direct testing of integrated AI-VR systems. These theoretical contributions are valuable for guiding system design but should not be interpreted as evidence of effectiveness.

The strongest available empirical evidence for AI-VR synergy comes from studies in adjacent domains. Moon and Ke (2023) found that adaptive prompts in VR-based social skills training significantly improved outcomes in a controlled experiment (n = 27 children with autism). While the population and context differ from adult vocational learners, this study provides proof-of-concept evidence that AI-driven adaptation within VR environments can produce measurable learning gains. Similarly, the randomized trial by Smith et al. (2022) on VR job interview training (n = 90 adults in correctional settings) provides rigorous empirical support for the effectiveness of VR-based job-seeking skills interventions, though the VR system used did not incorporate AI-driven personalization.

A recurring theme across the literature is the engagement-learning paradox identified by Makransky et al. (2019, 2021). Multiple studies have found that greater immersion increases enjoyment and presence but does not automatically translate into better learning outcomes. This finding is critical for the AI-VR integration agenda: it suggests that adding AI capabilities to VR systems will not automatically improve learning unless those capabilities are designed to activate specific cognitive and metacognitive processes. The generative learning strategies identified by Makransky et al. (2021)—summarizing, self-explaining, elaborating—represent evidence-based mechanisms through which AI could bridge the gap between engagement and learning, but this hypothesis has not been directly tested in integrated AI-VR job-seeking skills training systems.

In summary, the evidence for AI-VR synergy in learning outcomes is emergent rather than established. While there is theoretical coherence across findings from VR training, AI-driven personalization, and immersive learning research, direct empirical evidence on fully integrated AI-VR systems for job-seeking skills is scarce. This represents a significant evidence gap that future research should address.

5.2 PERSONALIZED LEARNING EXPERIENCE: FROM STANDARDIZATION TO INDIVIDUALIZATION

One of the most significant theoretical contributions of AI to VR-based training is the proposed shift from standardized, one-size-fits-all training experiences to genuinely personalized learning pathways. Traditional VR training systems follow predetermined scripts that present the same content, in the same sequence, to all learners regardless of their prior knowledge, skill level, or learning preferences (Xie et al., 2021). AI integration, in principle, fundamentally transforms this model by enabling dynamic adaptation at multiple levels.

The intelligent analysis and multimodal algorithms described by Dai (2024) represent a concrete conceptual instantiation of this adaptive capability. By processing learner behavior data—including speech patterns, response latency, eye movement, and interaction choices—these algorithms could build rich learner models that inform real-time content adaptation. For example, a learner who demonstrates strong verbal communication skills but struggles with nonverbal cues might be presented with interview scenarios that progressively increase the complexity of nonverbal interaction demands, while receiving targeted feedback on body language and eye contact.

The theoretical grounding for this personalization approach is provided by adaptive learning theory (Martin et al., 2020) and the CAMIL framework (Makransky & Petersen, 2021). Together, these frameworks suggest that personalized VR training should address not only what content learners encounter but also how that content is delivered, what support is provided, and how success is defined. This multi-dimensional personalization represents a significant advance over earlier generations of adaptive learning systems that focused narrowly on content sequencing.

However, the reviewed literature also identifies important practical challenges in implementing AI-driven personalization. The “cold-start” problem—the difficulty of providing effective personalization when a learner first begins using the system and little data is available—remains a significant practical concern that has not been adequately addressed in the AI-VR literature (Murtaza et al., 2022). Additionally, the transparency of AI personalization decisions is important for learner trust and agency; learners should understand, at least at a general level, why the system is making particular recommendations or adjustments (Luckin et al., 2022). At present, these challenges remain largely theoretical, as few empirical studies have examined learner experiences with AI-driven personalization in VR vocational training contexts.

5.3 MULTIMODAL DATA IN INSTRUCTIONAL DIAGNOSIS

The integration of multimodal learning analytics into AI-VR training systems opens new possibilities for instructional diagnosis that extend far beyond traditional assessment approaches. Rather than relying solely on outcome measures such as task completion or post-training knowledge tests, MMLA enables continuous, process-oriented assessment that captures the richness and complexity of learning as it unfolds (Mangaroska et al., 2020).

In the context of job-seeking skills training, MMLA can provide insights into dimensions of performance that are difficult to assess through conventional methods. Eye tracking data can reveal patterns of visual attention during virtual interviews, indicating whether a learner maintains appropriate eye contact or becomes distracted. Speech analysis can assess not only the content of verbal responses but also paralinguistic features such as tone, pace, and fluency. Physiological measures can indicate levels of anxiety or stress that might affect interview performance. And gesture analysis can evaluate the appropriateness and effectiveness of nonverbal communication (Chen et al., 2024; Nguyen et al., 2023).

The fusion of these multimodal data streams presents significant analytical challenges, as noted by Mu et al. (2020). Different modalities operate at different timescales, have different noise characteristics, and may provide contradictory signals about learner states. The AI model cross-validation mechanisms described in the AI-VR synergy framework (Dai, 2024) offer one conceptual approach to addressing these challenges, by ensuring that instructional decisions are based on converging evidence from multiple analytical perspectives rather than any single, potentially unreliable data source. However, this approach has not been empirically validated in vocational training contexts, and the practical feasibility of real-time multimodal fusion in VR environments remains uncertain.

The pedagogical value of MMLA-enhanced diagnosis lies not only in its analytical sophistication but in its potential to make assessment more formative and less summative. Rather than simply evaluating whether a learner has achieved a predetermined standard, MMLA can identify specific areas of difficulty, track learning trajectories over time, and provide actionable feedback that supports continuous improvement (Vatral et al., 2022). This formative orientation aligns well with constructivist and adaptive learning theories, which emphasize the importance of timely, specific feedback in supporting knowledge construction and skill development. At present, however, the translation of MMLA insights into actionable pedagogical interventions remains more aspirational than demonstrated.

5.4 EDUCATIONAL EQUITY AND TECHNOLOGY ETHICS

The rapid advancement of AI-VR technology in education raises important questions about equity, access, and ethics that must be addressed to ensure responsible implementation. The reviewed literature, while focused predominantly on technical and pedagogical aspects, provides important insights relevant to these concerns.

The issue of educational equity is particularly acute in the context of AI-VR training systems. High-quality VR hardware remains expensive, and the computational resources required for sophisticated AI processing are substantial. Without deliberate attention to accessibility and affordability, AI-VR training systems risk exacerbating existing educational inequalities by providing enhanced learning opportunities to well-resourced institutions while leaving under-resourced institutions behind (Luckin et al., 2022). The specific application to job-seeking skills training adds another layer of equity concern: if advanced AI-VR interview preparation is available primarily to privileged learners, it could reinforce employment disparities rather than reduce them.

Algorithmic bias represents a second critical ethical concern. AI systems trained on data that reflects existing social biases may perpetuate or amplify those biases in their recommendations and assessments (Murtaza et al., 2022). In the context of job interview training, this could manifest as differential feedback or assessment criteria applied to learners from different demographic backgrounds. The sensitive content identification and filtering mechanisms described in the AI-VR synergy framework (Dai, 2024) address some aspects of content safety but do not fully address the deeper challenge of algorithmic fairness in educational AI systems.

Data privacy and learner autonomy represent additional ethical considerations. The rich multimodal data that makes AI-VR personalization possible also creates unprecedented opportunities for surveillance and behavioral prediction. Learners may feel uncomfortable knowing that their eye movements, speech patterns, and physiological responses are being continuously monitored and analyzed, even if this monitoring serves educational purposes. Establishing clear ethical guidelines for data collection, storage, and use in AI-VR educational contexts is essential for maintaining learner trust and protecting fundamental rights (Luckin et al., 2022).

Finally, the role of human educators in AI-VR enhanced training environments requires careful consideration. The reviewed literature consistently emphasizes that AI should augment rather than replace human instruction (Gonzalez-Calatayud et al., 2021). AI-VR training systems are most effective when they function as tools that extend and enhance educators' capabilities—providing rich diagnostic information, automating routine instructional tasks, and enabling more personalized attention—rather than as autonomous systems that operate independently of human judgment and relationship.

5.5 PRACTICAL RECOMMENDATIONS FOR ETHICAL AI-VR IMPLEMENTATION

Based on the reviewed literature and the ethical concerns discussed above, several practical recommendations emerge for institutions and developers implementing AI-VR training systems in vocational education.

First, data privacy must be prioritized through transparent data governance policies. Institutions should clearly communicate to learners what multimodal data are collected (e.g., speech, gaze, physiological indicators), how these data are used for personalization and assessment, and how long they are retained. Data minimization principles should be applied, collecting only what is necessary for the stated educational purposes (Murtaza et al., 2022; Luckin et al., 2022).

Second, algorithmic fairness requires proactive measures. AI models used for learner assessment and content recommendation should be audited for differential performance across demographic groups. Training data should be examined for representativeness, and institutions should establish mechanisms for learners to contest or appeal AI-generated assessments (Gonzalez-Calatayud et al., 2021). The multi-model cross-validation approach described in the AI-VR synergy framework (Dai, 2024) provides one technical mechanism for reducing the risk of biased recommendations, though its effectiveness in practice requires empirical validation.

Third, accessibility and equity should guide system design and deployment. AI-VR training systems should be designed with universal design principles, accommodating learners with disabilities (e.g., alternative input modalities for learners with motor impairments, visual alternatives to gaze-based interaction). Institutions with limited resources should explore cloud-based VR solutions, shared infrastructure models, and partnerships to reduce cost barriers and prevent AI-VR training from exacerbating existing educational inequalities.

Fourth, informed consent and learner autonomy must be maintained. Learners should have the option to opt out of specific data collection modalities without losing access to core training functions. The role of human educators should be preserved as central to the learning process, with AI-VR systems functioning as tools that augment—rather than replace—instructor judgment, mentorship, and relational support (Luckin et al., 2022).

Conclusion:

This review has examined the potential role of AI-VR synergistic technology in virtual simulation training for job-seeking skills. The reviewed literature suggests that virtual reality can provide immersive and repeatable learning environments for practising professional communication, interview behaviour, and employment-related competencies. Artificial intelligence may enhance these environments by enabling personalization, adaptive feedback, multimodal analysis, and more responsive assessment. Together, these functions may address some limitations of conventional virtual training systems, including fixed scenarios, limited learner differentiation, and delayed or general feedback.

However, the educational value of AI-VR integration depends on careful instructional design rather than technology use alone. Immersion may increase engagement, but effective learning requires appropriate scaffolding, cognitive load management, meaningful practice, and reliable assessment. The review also highlights important ethical and practical issues, including data privacy, algorithmic bias, transparency, accessibility, and the continuing role of human educators. At present, the evidence base for fully integrated AI-VR systems in job-seeking skills training remains developing. The majority of supporting evidence comes from studies of VR, AI personalization, or multimodal analytics in isolation; direct empirical evidence on integrated AI-VR systems for employment-related skills is scarce. Future studies should evaluate such systems in authentic vocational education contexts, compare them with traditional and standard VR approaches, and examine long-term transfer to real employment-related performance.

Implications for Higher Vocational Education

The integration of AI and VR technologies in vocational education represents more than a technological upgrade to existing training approaches; it constitutes a fundamental shift in how job-seeking skills can be taught, practiced, and assessed. For higher vocational education institutions—particularly those like the affiliated institution of this research, which have already invested in virtual reality training infrastructure—the AI-VR synergy offers a pathway to substantially enhance the pedagogical value of existing technological resources (Jensen & Konradsen, 2018; Dai, 2024).

Several specific implications emerge from this review. First, vocational education institutions should prioritize the development of AI-enhanced adaptive capabilities in their VR training systems, moving beyond static simulations to dynamic, responsive learning environments that personalize the training experience (Peng & Li, 2025). Second, investment in multimodal data collection and analysis infrastructure is necessary to realize the full potential of AI-VR synergy for instructional diagnosis and formative assessment (Mu et al., 2020). Third, faculty development programs should prepare educators to effectively integrate AI-VR training tools into their pedagogical practice, developing the skills to interpret AI-generated learner data and to balance automated and human-mediated instruction (Luckin et al., 2022).

The practical implementation pathway suggested by this review aligns with the phased approach described in the underlying AI-VR synergy research project: beginning with needs analysis and technology prototyping, proceeding through iterative testing and refinement, and culminating in deployment within authentic educational contexts. This phased approach allows for the gradual development of institutional capacity while generating evidence to inform continuous improvement.

Future Research Directions

Several research directions emerge from this review. First, rigorous comparative studies are needed to directly evaluate fully integrated AI-VR training systems against traditional VR-only and conventional classroom approaches in authentic vocational education settings. Such studies should employ randomized or quasi-experimental designs, report effect sizes, and include longitudinal follow-up to assess long-term transfer to employment-related outcomes (Renganayagalu et al., 2021; Smith et al., 2022). Second, future reviews should adopt multi-reviewer protocols, transparent quality appraisal frameworks such as GRADE or MMAT, and formal PRISMA flow diagrams to strengthen methodological rigor and reproducibility. Third, research is needed on the specific mechanisms through which AI-driven personalization and multimodal feedback influence learning processes in VR environments, potentially using process-tracing methods such as think-aloud protocols, interaction log analysis, and physiological measures (Makransky & Petersen, 2021). Fourth, implementation studies in diverse institutional contexts—including under-resourced settings—are essential for understanding the scalability, cost-effectiveness, and equity implications of AI-VR training systems. Finally, research on ethical frameworks and governance models for AI-VR in education should move beyond identifying concerns to developing and testing practical solutions, such as algorithmic auditing protocols, learner data rights frameworks, and inclusive design guidelines.

Limitations

This review has several limitations that should be considered when interpreting its findings. First, the analysis is based on selected published literature rather than a complete meta-analysis, and the reviewed studies differ considerably in research design, educational context, technology configuration, and outcome measures. This heterogeneity limits the extent to which firm conclusions can be drawn about the overall effectiveness of AI-VR integration. Second, much of the available evidence concerns virtual reality, artificial intelligence, adaptive learning, or multimodal analytics separately, rather than fully integrated AI-VR systems for job-seeking skills training. Therefore, some conclusions are based on theoretical synthesis rather than direct empirical confirmation, and the distinction between evidence-based findings and theoretical implications should be carefully maintained. Third, the review was conducted by a single reviewer, which limits the reliability of study selection, data extraction, and quality appraisal compared to multi-reviewer approaches with established inter-rater reliability procedures. Fourth, the restriction to English-language sources may introduce language bias, excluding relevant research published in Chinese, Japanese, Korean, and other languages where AI-VR educational research is active. Fifth, publication bias may have resulted in overrepresentation of studies reporting positive or statistically significant effects. Sixth, the quality appraisal framework, while systematic, was adapted from existing tools rather than using a validated instrument. These limitations collectively suggest that further research adopting multi-reviewer protocols, formal quality assessment instruments, and transparent screening procedures (e.g., PRISMA flow diagrams) is needed to strengthen the evidence base for AI-VR integration in vocational education.

Declaration of AI Use

This manuscript was prepared through the combined contributions of all author(s), including contributions to the study design, data, content development, results, interpretation, and related scholarly work. The author(s) acknowledge the use of Grammarly and ChatGPT to assist with grammar checking, language refinement, and reference formatting. These AI-assisted tools were not used as authors and did not replace the intellectual contributions or scholarly judgment of the author(s). All AI-assisted outputs, including content, references, and interpretations, were carefully reviewed, revised, verified, and approved by the author(s). The author(s) accept full responsibility for the accuracy, integrity, and final content of the manuscript.

REFERENCES

Albus, P., Vogt, A., & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. Computers & Education, 166, 104154. https://doi.org/10.1016/j.compedu.2021.104154

Araiza-Alba, P., Keane, T., Chen, W. S., & Kaufman, J. (2021). Immersive virtual reality as a tool to learn problem-solving skills. Computers & Education, 164, 104121. https://doi.org/10.1016/j.compedu.2020.104121

Baceviciute, S., Mottelson, A., Terkildsen, T., & Makransky, G. (2020). Investigating representation of text and audio in educational VR using learning outcomes and EEG. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–9). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376872

Checa, D., & Bustillo, A. (2020). A review of immersive virtual reality serious games to enhance learning and training. Multimedia Tools and Applications, 79(9), 5501–5527. https://doi.org/10.1007/s11042-019-08348-9

Chen, J., Shen, Z., Bian, S., Zhao, H., Liu, X., & Xiong, G. (2024). Multimodal behavior analysis for virtual reality learning environment. In 2023 China Automation Congress (CAC) (pp. 1–6). IEEE. https://doi.org/10.1109/CAC59555.2023.10451050

Dai, W. (2024). Application of AI and VR synergistic enhancement in employment skills training: A virtual simulation approach. Academic Journal of Management and Social Sciences, 6(1), 38–46. https://doi.org/10.54097/7pdb3510

González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. Applied Sciences, 11(12), 5467. https://doi.org/10.3390/app11125467

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. Education and Information Technologies, 23(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Khan, H., Lo Presti, A., & Briscoe, J. P. (2025). Virtual reality in vocational interventions: A systematic literature review. International Journal for Educational and Vocational Guidance. Advance online publication. https://doi.org/10.1007/s10775-025-09785-9

Korhonen, T., Lindqvist, T., Laine, J., & Hakkarainen, K. (2022). Training hard skills in virtual reality: Developing a theoretical framework for AI-based immersive learning. In H. Niemi, R. D. Pea, & Y. Lu (Eds.), AI in learning: Designing the future (pp. 195–213). Springer. https://doi.org/10.1007/978-3-031-09687-7_12

Luck, M., & Aylett, R. (2000). Applying artificial intelligence to virtual reality: Intelligent virtual environments. Applied Artificial Intelligence, 14(1), 3–32. https://doi.org/10.1080/088395100117142

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. Computers and Education: Artificial Intelligence, 3, 100076. https://doi.org/10.1016/j.caeai.2022.100076
Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. Journal of Educational Psychology, 113(4), 719–735. https://doi.org/10.1037/edu0000473
Makransky, G., & Mayer, R. E. (2022). Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. Educational Psychology Review, 34(3), 1771–1798. https://doi.org/10.1007/s10648-022-09675-4
Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. Educational Psychology Review, 33, 937–958. https://doi.org/10.1007/s10648-020-09586-2
Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and Instruction, 60, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007
Mangaroska, K., Sharma, K., Gašević, D., & Giannakos, M. (2020). Multimodal learning analytics to inform learning design: Lessons learned from computing education. Journal of Learning Analytics, 7(3), 79–97. https://doi.org/10.18608/jla.2020.73.7
Martin, F., Chen, Y., Moore, R. L., & Westine, C. D. (2020). Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. Educational Technology Research and Development, 68(4), 1903–1929. https://doi.org/10.1007/s11423-020-09793-2
Mayer, R. E. (2021). The immersion principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning (3rd ed., pp. 296–303). Cambridge University Press. https://doi.org/10.1017/9781108894333.031
Moon, J., & Ke, F. (2023). Effects of adaptive prompts in virtual reality-based social skills training for children with autism. Journal of Autism and Developmental Disorders, 54, 2218–2232. https://doi.org/10.1007/s10803-023-06021-7
Mu, S., Cui, M., & Huang, X. (2020). Multimodal data fusion in learning analytics: A systematic review. Sensors, 20(23), 6856. https://doi.org/10.3390/s20236856
Murtaza, M., Ahmed, Y., Shamsi, J. A., Sherwani, F., & Usman, M. (2022). AI-based personalized e-learning systems: Issues, challenges, and solutions. IEEE Access, 10, 81323–81342. https://doi.org/10.1109/ACCESS.2022.3193938
Narciso, D., Melo, M., Rodrigues, S., Cunha, J. P. S., & Bessa, M. (2021). A systematic review on the use of immersive virtual reality to train professionals. Multimedia Tools and Applications, 80(9), 13195–13214. https://doi.org/10.1007/s11042-020-10454-y
Nguyen, A., Järvelä, S., Rosé, C., Järvenoja, H., & Malmberg, J. (2023). Examining socially shared regulation and shared physiological arousal events with multimodal learning analytics. British Journal of Educational Technology, 54(1), 293–312. https://doi.org/10.1111/bjet.13280
Peng, J., & Li, Y. (2025). Frontiers of artificial intelligence for personalized learning in higher education: A systematic review of leading articles. Applied Sciences, 15(18), 10096. https://doi.org/10.3390/app151810096
Rasheed, Z., Ghwanmeh, S., & Abualkishik, A. Z. (2023). Harnessing artificial intelligence for personalized learning: A systematic review. Data and Metadata, 2, 146. https://doi.org/10.56294/dm2023146
Renganayagalu, S., Mallam, S. C., & Nazir, S. (2021). Effectiveness of VR head mounted displays in professional training: A systematic review. Technology, Knowledge and Learning, 26, 999–1041. https://doi.org/10.1007/s10758-020-09489-9
Smith, M. J., Parham, B., Mitchell, J., Blajeski, S., Harrington, M., Ross, B., Johnson, J., Brydon, D. M., Johnson, J. E., Cuddeback, G. S., Smith, J. D., Bell, M. D., McGeorge, R., Kaminski, K., Suganuma, A., & Kubiak, S. (2022). Virtual reality job interview training for

adults receiving prison-based employment services: A randomized controlled feasibility and initial effectiveness trial. Criminal Justice and Behavior, 49(9), 1346–1367. https://doi.org/10.1177/00938548221081447
Vatral, C., Biswas, G., Cohn, C., Davalos, E., & Mohammed, N. (2022). Using the DiCoT framework for integrated multimodal analysis in mixed-reality training environments. Frontiers in Artificial Intelligence, 5, 941825. https://doi.org/10.3389/frai.2022.941825
Xie, B., Liu, H., Alghofaili, R., Zhang, Y., Jiang, Y., Lobo, F. D., Li, C., Li, W., Huang, H., Akdere, M., Mousas, C., & Yu, L.-F. (2021). A review on virtual reality skill training applications. Frontiers in Virtual Reality, 2, 645153. https://doi.org/10.3389/frvir.2021.645153
Yuensook, T., Jantakoon, T., & Limpinan, P. (2025). AI-driven adaptive learning systems in higher education: A systematic review. Journal of Education and Learning, 15(2), 117–131. https://doi.org/10.5539/jel.v15n2p117.