Learning Scenarios and Tasks for Knowledge Tracing in Virtual Reality: A Systematic Review

Yuying Yang and Geping Liu*

Faculty of Education, Southwest University, Chongqing, China
Email: yyy546685@email.swu.edu.cn (Y.Y.); liugp@swu.edu.cn (G.L.)
*Corresponding author

Manuscript received July 6, 2025; revised July 14, 2025; accepted August 5, 2025; published September 8, 2025

***Abstract*—Virtual Reality (VR) provides immersive, interactive, and personalized learning experiences, yet its integration with Knowledge Tracing (KT) remains underexplored. This systematic literature review synthesizes 27 empirical studies on KT in VR learning environments published between 2015 and 2025. The findings reveal that most studies focus on higher education and vocational training, with medical and engineering disciplines being the most common application areas. Learning scenarios and tasks play a critical role in shaping learner interactions and generating meaningful data for effective KT in VR environments. Based on both manual coding and Latent Dirichlet Allocation (LDA) topic modeling, the learning tasks and educational scenarios identified in the studies involve multimodal knowledge state modeling and prediction, personalized and adaptive learning support systems, instructional design and learning outcome evaluation in VR, and KT applications targeting specific learner groups or contextual needs. Finally, this review identifies key challenges and proposes directions for future research on KT in VR environments.**

***Keywords*—Virtual Reality (VR), Knowledge Tracing (KT), immersive learning**

I. Introduction

The emergence of the metaverse, characterized by immersive, interconnected virtual spaces, has catalyzed renewed attention to Virtual Reality (VR) technologies and their educational applications [1]. VR, as an interactive medium that integrates multidisciplinary technological advances, has been recognized for its potential to reshape traditional learning paradigms by offering advantages in immersion, interactivity, and creative engagement [2, 3]. By providing a multisensory, immersive learning environment, VR facilitates the integration of learners' visual, auditory, and behavioral information within a unified context, thereby supporting multilayered interactions with virtual objects, intelligent agents, avatars, and three-dimensional resources [4]. Such embodied and contextualized interactions have been shown to enhance cognitive engagement and facilitate both the construction and transfer of knowledge [5]. Furthermore, VR systems, with their wearable characteristics and compatibility with various sensing devices, are designed to synchronously capture learners' behavioral trajectories, eye movements, speech, and physiological signals. These features offer distinct advantages for dynamic data acquisition and real-time learning assessment within educational contexts [6].

Knowledge Tracing (KT) is regarded as a core technique for dynamically inferring learners' knowledge states and cognitive trajectories, playing a crucial role in learning analytics and personalized education. With the advancement of deep learning and machine learning techniques in recent years, KT has evolved from traditional probabilistic models, such as Bayesian Knowledge Tracing (BKT), to neural network-based approaches, including Deep Knowledge Tracing (DKT) and Self-Attentive Knowledge Tracing (SAKT) [7]. These models have been widely applied in adaptive learning systems, intelligent tutoring platforms, and online educational interventions, enabling accurate prediction of learners' performance and supporting personalized instructional design. The availability of multimodal data in VR environments presents new opportunities to enhance and expand KT models.

In recent years, an increasing number of researchers have applied KT techniques in VR environments to investigate learners' knowledge acquisition and changes in cognitive states. Some studies have directly adopted KT techniques originally designed for online learning platforms in VR settings. These studies mainly use questionnaires and pre- and post-tests, with limited integration of VR-specific data like behavioral trajectories and multimodal interactions [8, 9]. Other studies have explored more advanced KT methods by incorporating multimodal data, including speech, semantic information, and physiological signals, to capture learners' emotional states and cognitive engagement more comprehensively [10]. In addition, machine learning techniques have been utilized to model and classify these multimodal features, enhancing the prediction of learning outcomes [11].

KT plays a critical role in education by enabling the dynamic modeling of learners' knowledge states, and its integration into immersive VR environments holds great potential for delivering personalized and adaptive learning experiences [12]. However, existing reviews on VR mainly focus on technical advancements or application scenarios across disciplines, whereas research specifically addressing KT within VR environments remains fragmented and limited. Current research is limited by the heterogeneity and under-explored integration of data sources, along with low model generalizability, which constrains the development of a comprehensive understanding in this emerging area [13]. In VR environments, learning scenarios and tasks lay the foundation for implementing KT and play a significant role in determining the performance and practical relevance of KT models. They define how learners interact with virtual content, generate multimodal behavioral data, and engage in cognitive processing [6]. A clear understanding of these elements is essential for designing effective KT systems that align with authentic educational needs. Therefore, this systematic literature review is conducted to synthesize the

doi: 10.18178/ijiet.2025.15.9.2380

current state of KT research in VR environments. While this review emphasizes learning scenarios and tasks, it also identifies broader research limitations that influence the advancement of KT in VR environments as a whole.

II. Literature Review

A. Theoretical Foundations for Applying VR in Education

The application of VR in educational settings is grounded in multiple learning theories. Embodied Cognition Theory emphasizes the positive impact of physical interaction on higher-order cognitive processing [14]. Through the construction of highly interactive three-dimensional environments, VR enables learners to engage directly with virtual worlds through bodily movements, activating the coordinated functioning of sensory, motor, and cognitive systems. This process has been shown to enhance conceptual understanding, spatial reasoning, and problem-solving abilities. Moreover, learners' embodied behaviors, including operational actions, movement trajectories, and gaze patterns, serve as important indicators of cognitive states, providing both a theoretical foundation and a data source for multimodal KT. For example, Walkington *et al.* [10] adopted a theoretical framework based on embodied cognition and employed the multimodal analysis of embodied technology approach to examine learners' gestures, actions, movements, and eye gaze in a VR environment. Their findings demonstrated that students' embodied behaviors indicate cognitive development and can serve as a foundation for understanding knowledge construction in immersive learning contexts.

The Cognitive Affective Theory of Immersive Learning (CATIL) further explains that VR environments enhance learner agency through immersion, interactivity, and a heightened sense of realism [15]. By stimulating psychological factors such as learning interest, motivation, immersion, self-efficacy, and self-regulation, VR fosters increased engagement and deeper cognitive processing. This theory highlights the interaction between cognitive and affective components in immersive learning and suggests that researchers should closely monitor the dynamic evolution of cognitive-affective states during learning. Tracking learner states in immersive environments not only deepens understanding of how learning occurs but also supports personalized feedback and improved instructional design. Makransky *et al.* [16] conducted two between-subjects experiments comparing immersive virtual reality (IVR) with traditional video-based instruction for science learning. The results showed that IVR significantly enhanced learning outcomes in science, suggesting that the sense of immersion stimulates emotional value, which is then transformed into learning gains through cognitive strategies. This study provides empirical support for the CATIL and highlights the importance of tracking learners' cognitive and emotional states in VR environments.

B. KT in VR Environments

KT originated from BKT, a model that dynamically estimates the probability of learners' mastery of individual concepts based on their item response sequences [17]. With the increasing scale and complexity of online learning data, the limitations of BKT in terms of parameter flexibility and sequence length handling have become evident. To address these challenges, Piech *et al.* [18] introduced DKT, which uses recurrent neural networks to model learning sequences and significantly improves prediction accuracy. Subsequently, models such as the Dynamic Key-Value Memory Network (DKVMN) and the Sequential Key-Value Memory Network (SKVMN) have been developed to expand memory capacity, support conceptual relationship modeling, and enhance cross-task transferability, further boosting the predictive power of KT models [19]. As educational data sources have evolved from simple response records to include behavioral logs, physiological signals, speech, and video interactions, extensible DKT models have emerged to integrate heterogeneous data and better capture learning states and cognitive dynamics [7].

Traditional questionnaires and post-hoc tests are unable to capture, in real time, the immersive interaction and highly personalized learning trajectories that characterize VR-based learning [20]. VR systems, by integrating sensors such as eye trackers, motion capture devices, Electroencephalography (EEG), and Electrodermal Activity (EDA) sensors, provide KT with a rich foundation of high-frequency, multimodal data [21]. This enables researchers to infer learners' knowledge states in real time from multiple modalities, including spatial behavior, visual attention, emotional responses, and physiological signals. Consequently, the evolution of KT reflects a shift from probabilistic to deep models, and from unimodal to multimodal data integration, while VR-based KT further extends the observability of the learning process.

However, current research on KT in VR environments remains in its early stages. There is a lack of standardized procedures and technical protocols for multimodal data fusion, and deep models also face limitations in generalizability and interpretability when applied to heterogeneous data sources [22]. For instance, Dubovi [23] combined multimodal objective data, including blink rates, gaze patterns, and facial expressions, with subjective self-reports to predict learning outcomes in VR environments. The study demonstrated that multimodal data outperformed single data sources in predicting learning performance. However, the absence of standardized integration mechanisms and the technical instability of multimodal processing continue to pose significant challenges. Similarly, Walkington *et al.* [10] highlighted that, despite the potential of multimodal analysis in VR-based embodied learning, the absence of coordinated data processing and unified evaluation frameworks continues to limit the application of KT in such settings. These findings underscore the urgent need for improved data fusion algorithms, validation of model generalizability, and standardized evaluation systems to realize the potential of KT in VR.

C. Reviews on VR and KT

Existing reviews on VR primarily focus on technological features and their educational applications in contexts such as K-12, higher education, and vocational training [24–27]. Most of these reviews explore the relationship between technological features, learning contexts, and educational outcomes, while relatively few have examined the dynamic modeling of learning processes. A small number of reviews have explored educational data mining within VR settings. For example, Lampropoulos and Evangelidis [6] reviewed learning analytics and educational data mining in AR, VR, and metaverse contexts, highlighting that methods such as BKT, clustering, traditional machine learning, and deep learning could analyze multimodal learning data. However, their review does not address approaches for constructing student models or predicting knowledge states and skill acquisition. Similarly, Shadiev and Li [28] systematically reviewed eye-tracking research in immersive environments, finding that the studies predominantly involve higher education participants and mainly use questionnaires and performance tests. Although the authors suggest integrating eye-tracking data with physiological measurements (such as EEG, ECG, and fNIRS) to improve cognitive-state recognition, their analysis is restricted to single-modal statistics, without discussing sequential modeling or real-time assessment of learning processes.

KT has typically been reviewed in terms of model development, algorithmic comparisons, data adaptation, and evaluation metrics, with recent reviews increasingly highlighting the integration of multimodal data and deep learning approaches [29, 30]. Empirical KT studies in educational contexts have largely relied on online platforms or intelligent tutoring systems [31, 32]. Yan *et al*. [33] collected learner behavior logs from learning management systems (such as MOOCs) and applied various methods, including BKT, IRT and DKT, to predict learners' knowledge states. They comprehensively compared the predictive accuracy of multiple models and discussed connections to self-directed learning design. Similarly, Trifa *et al*. [34] integrated KT models into intelligent tutoring systems to analyze data from online practices and interactions, dynamically predicting students' knowledge mastery and enhancing personalized learning experiences through intelligent feedback mechanisms.

Currently, few KT studies focus on immersive and interactive VR learning environments. The limited research in this area often employs multimodal analysis to model learners' cognitive changes and learning trajectories. These approaches aim to capture fluctuations in knowledge states throughout immersive learning processes.

In summary, although the educational advantages of VR and the diagnostic potential of KT are well recognized, systematic integration of the two remains underexplored. Therefore, this study presents a systematic review of existing KT research conducted in VR environments, providing an integrated analysis of data collection methods, modeling approaches, and evaluation strategies. This paper aims to clarify the research landscape, identify typical methodological approaches, and highlight promising avenues for future research.

Specifically, this study addresses the following research questions:

RQ1: What research contexts and design features characterize KT studies conducted in VR environments?

RQ2: What types of learning tasks and educational scenarios have been targeted by KT applications in VR?

RQ3: What are the main limitations identified in existing research, and what future directions can be proposed for advancing KT within VR-based education?

III. Methods

A. Identification

This study followed the PRISMA 2020 guidelines to systematically identify and review prior research [35]. Studies were included if they met the following criteria:

1) Eligibility criteria

Studies were conducted on learners within VR environments specifically designed for educational or training purposes, with empirical analyses of knowledge mastery or cognitive states. Two types of studies were considered eligible. First, classic KT research explicitly aimed at modeling learners' knowledge states. Second, we also included studies that indirectly addressed knowledge mastery or cognitive states through intermediate variables such as attention, cognitive load, or motivation, even if they did not explicitly employ KT models. Additionally, studies that indirectly inferred cognitive changes or trends in knowledge acquisition through statistical analysis of learner behavior or performance data were also included. Such studies provide evidence relevant to understanding learning processes and informing future modeling efforts.

2) Exclusion criteria

Studies focusing solely on VR hardware or learning environment development without empirical assessment of learners' cognitive or knowledge states were excluded.

3) Publication types

Only peer-reviewed journal articles or conference papers published in English with accessible full-text versions were included.

4) Publication dates

Considering the rapid advancements in consumer-grade VR technology since 2015, especially with the emergence of devices such as Oculus Rift and HTC Vive, the review specifically focused on literature published within the past decade.

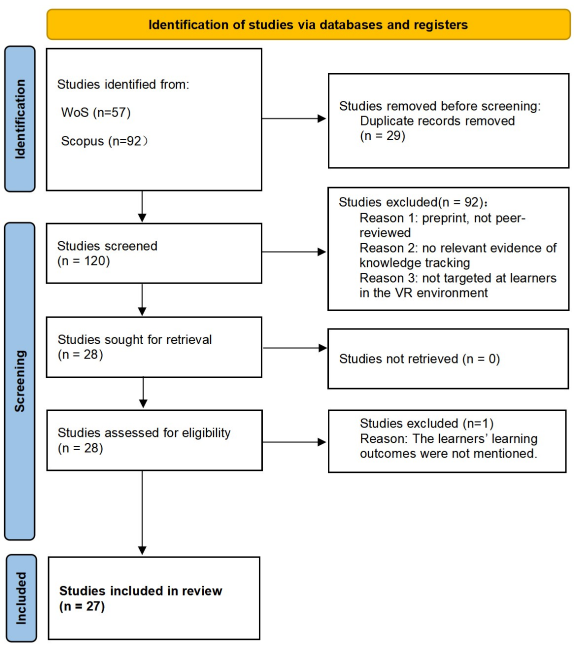

Fig. 1. PRISMA literature selection process.

A comprehensive literature search was conducted in the Web of Science Core Collection and Scopus databases using the following combination of keywords: ("knowledge tracking" OR "knowledge tracing") AND ("virtual reality" OR "VR") AND (education). This search yielded a total of 149 studies. After removing duplicates, 120 unique articles remained. Titles and abstracts were independently screened by two researchers, with disagreements resolved by discussion, resulting in the inclusion of 28 studies. Following a detailed full-text assessment conducted independently by both researchers, one study was excluded, leaving a final selection of 27 articles. The detailed selection process is illustrated in Fig. 1.

B. Coding

A coding scheme was developed based on the research questions to extract six categories of information from each selected study (see Table 1).

Table 1. Coding

No.	Category	Items (examples)
1	Basic publication information	Year, publication source, country, publication type (journal/conference)
2	Educational contexts	K-12, higher education, vocational education, medical training, language learning, STEM education
3	Research objectives	Evaluating VR course effectiveness, proposing theoretical frameworks

C. Analyses

The bibliographic details of each study (title, authors, abstract, and related metadata) were first compiled in Microsoft Excel. Full-text articles identified for in-depth coding were then imported into NVivo for qualitative analysis. In parallel, Latent Dirichlet Allocation (LDA) topic modeling was applied to the corpus to extract latent thematic structures and triangulate the qualitative findings. The combined qualitative coding and LDA results were used to address the research questions.

IV. Results

A. To Answer RQ 1: Research Context and Study Design

This section reports the general characteristics of the included studies, including their educational contexts and research design features.

In terms of publication types, 17 studies (62.96%) were published in peer-reviewed journals, while 10 studies (37.04%) appeared in conference proceedings. The journal and conference sources are highly diverse, with the 27 studies distributed across 27 publication outlets.

Based on the first author's affiliation, the 27 studies were produced by researchers from 13 countries and regions. The United States accounts for the largest share, with 10 papers (37.03%), followed by China with 3 papers (11.11%). Germany, Israel, and Japan each contributed 2 papers (7.41%). Additional contributions also originated from countries across Europe, Asia, North America, and South America, reflecting the global reach of VR-based KT research. In terms of annual publication volume, the peak occurred in 2014, with a total of 9 studies, followed by 2023, which recorded 5 studies. Overall, the number of publications has shown a generally increasing trend over time (see Fig. 2).

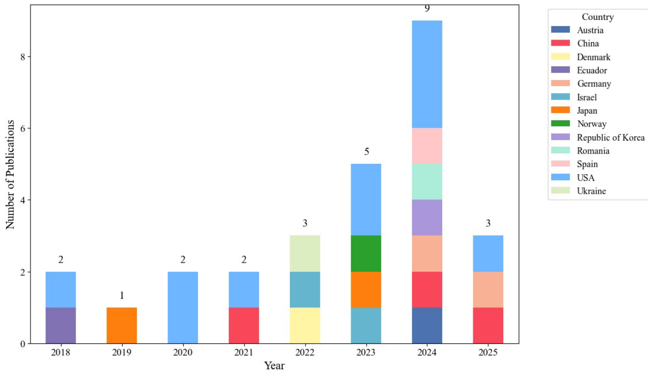

Fig. 2. Publications by country and year.

The educational levels targeted in the 27 reviewed studies are summarized in Fig. 3 Since some studies involved multiple educational levels, the total count across all levels exceeds 27.

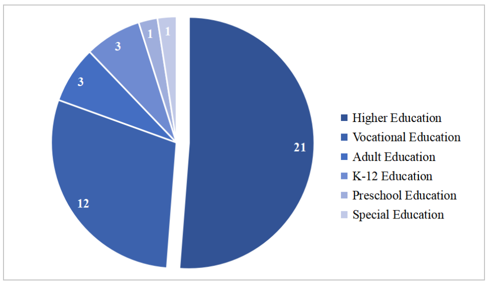

Fig. 3. Educational stages.

The disciplinary focus of the studies is presented in Table 2, reflecting the diverse academic fields in which VR-based KT has been applied.

Table 2. Disciplinary domains of application

Disciplinary Domain	Specific Field	k	%
Medicine	Nursing	4	14.815
	Clinical Medicine	4	14.815
	Neuroanatomy	1	3.704
	Physiology	1	3.704
Engineering	Architecture and Construction	2	7.406
	Additive Manufacturing (3D Printing)	1	3.704
	Surveying and Mapping	1	3.704
	Power Engineering	1	3.704
	Maritime Technology	1	3.704
Biology		2	7.406
Physics		1	3.704
Energy Science		1	3.704
Marine Science		1	3.704
STEM		2	7.406
Linguistics (Japanese)		1	3.704
Arts		1	3.704
Basic Education (Handwriting and Shape Recognition)		1	3.704
Safety Education		1	3.704
Total		**27**	**100**

B. To Answer RQ 2: Learning Tasks and Educational Scenarios

Drawing on the thematic focus of KT research in VR environments and the relationship between data sources and research objectives, the existing literature can be grouped into four core themes (see Table 3).

1) Knowledge state modeling and multimodal learning process analysis

Studies in this category focus on the real-time, dynamic modeling of learners' knowledge states and cognitive processes. To capture key indicators of the learning process, these studies typically utilize multimodal data sources including eye-tracking, EEG, EDA, and behavioral trajectories. Sequence modeling, pattern recognition, and dynamic inference methods are employed to reveal attention distribution, information processing patterns, and cognitive load fluctuations within VR environments.

2) Adaptive learning and personalized instructional optimization

Research in this category explores how individual learner differences, such as prior knowledge, motivation, and learning styles, influence learning pathways and outcomes. Data-driven instructional design and environment adaptation strategies are applied to enhance personalized learning support.

3) Learning outcome evaluation and instructional intervention validation

Research in this category focuses on the systematic and empirical evaluation of learning outcomes, emphasizing testing the effects of VR-based instructional programs, tools, or interventions on knowledge acquisition, skill development, cognitive progress, and learning motivation. These studies commonly adopt between-group comparisons, longitudinal tracking, and multi-indicator integrated analysis to assess the effectiveness, advantages, and limitations of VR-enhanced instruction, providing scientific evidence to guide teaching practice and the application of educational technologies.

4) Development of instructional systems and methodological frameworks

This category focuses on the innovation and optimization of technical tools, system platforms, and analytical methods to support KT in VR environments. Key topics include the development of data collection systems, algorithmic architectures, multimodal data integration mechanisms, and user interface designs.

Table 3. Research Classification

Category	Reference
Knowledge state modeling and multimodal learning process analysis	[10, 11, 23, 36–43]
Adaptive learning and personalized instructional optimization	[44–47]
Learning outcome evaluation and instructional intervention validation.	[8, 9, 21, 48–52]
Development of instructional systems and methodological frameworks	[53–56]

To gain deeper insights into the learning tasks and educational scenarios (RQ 2), we applied Latent Dirichlet Allocation (LDA) topic modeling to uncover underlying thematic structures. LDA is an unsupervised probabilistic model that assumes each document is a mixture of latent topics, and each topic is characterized by a distribution over words [57]. Specifically, the titles and abstracts of the 27 included studies were pre-processed through standard data cleaning steps, including lowercasing, stop-word removal, and lemmatization. The cleaned corpus was then used to train multiple LDA models with varying topic numbers. Among these, the model with five topics achieved the highest topic coherence score, indicating the best balance between topic distinctiveness and interpretability. These five topics served as the basis for subsequent qualitative interpretation and categorization of learning scenarios and tasks (see Fig. 4).

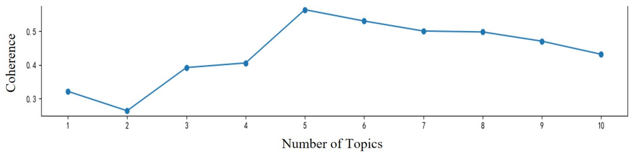

Fig. 4. Topics number and coherence.

The LDA topic analysis revealed five research themes regarding the learning tasks and educational scenarios within VR-based KT research (see Table 4).

Table 4. Topic and keywords

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
vr 0.055	learning 0.052	learning 0.076	virtual 0.031	vr 0.048
learning 0.034	based 0.027	vr 0.053	risk 0.031	data 0.018
immersive 0.023	virtual 0.027	training 0.039	children 0.031	personalized 0.018
eye 0.023	design 0.027	immersive 0.024	learning 0.031	environment 0.018
knowledge 0.023	outcomes 0.015	outcomes 0.016	effects 0.017	system 0.018
understand 0.023	evidence 0.015	tracking 0.016	environment 0.017	effect 0.018
utilizing 0.023	achievement 0.015	mixed 0.016	development 0.017	tracking 0.018
environment 0.013	evaluate 0.015	multimodal 0.016	innovative 0.016666	training 0.018
data 0.013	skill 0.015	predict 0.016	protocol 0.017	construction 0.018
monitor 0.013	ability 0.015	attention 0.009	primary 0.017	support 0.018

[1] The number below each keyword represents its occurrence probability within the topic.

Theme 1: Eye-tracking and cognitive monitoring in IVR learning is characterized by keywords such as VR, learning, immersive, eye, knowledge, understand, utilizing, environment, data, monitor. Studies under this theme primarily focus on capturing learners' attention, comprehension levels, and knowledge states in real time by utilizing eye-tracking and other perceptual data within immersive VR settings.

Theme 2: Evidence-based VR instructional design and learning outcome evaluation is associated with keywords including learning, based, virtual, design, outcomes, evidence, achievement, evaluate, skill, ability. Research within this theme emphasizes the instructional design of VR courses or training programs, experimental validation, and the evaluation of learning outcomes.

Theme 3: Multimodal prediction and mixed-method approach for VR-based KT is reflected by keywords such as learning, VR, training, immersive, outcomes, tracking, mixed, multimodal, predict, attention. This theme centers on integrating multimodal data with conventional test results to develop multimodal KT models and predictive models for learning performance.

Theme 4: VR applications for children and risk education with developmental impacts includes keywords like virtual, risk, children, learning, effects, environment, development, innovative, protocol, primary. Studies in this category explore the use of VR for safety education, risk perception training, and skill development, particularly among children and school-age learners.

Theme 5: Data-driven personalized and adaptive VR learning systems is defined by keywords such as VR, data, personalized, environment, system, tracking, effect, training, construction, support. Research within this theme focuses on leveraging real-time data collection to develop learner profiles and adaptive learning environments, aiming to support personalized learning pathways and tailored instructional support.

V. Discussion

The application of KT in VR settings is undergoing rapid development, characterized by emerging trends such as multimodal data-driven modeling of learning states, personalized instruction supported by diverse evaluation mechanisms, and the systematic integration of data fusion techniques with intelligent algorithms.

A. Research Context and Study Design

Analysis of the 27 studies shows that KT research in VR environments primarily focuses on higher and vocational education, with 77.78% conducted in higher education and 44.44% in vocational settings. While the disciplinary coverage is broad, the research exhibits a clear concentration in specific professional domains. Specifically, medical-related fields, such as nursing, clinical medicine, neuroanatomy, and physiology, constitute 37% of the studies, followed by engineering and technical disciplines, including architecture, additive manufacturing, and surveying. This concentration may be attributed to the fact that both medicine and engineering involve high-stakes, skill-intensive training that poses challenges for traditional instruction. In medical education, procedures such as surgery and rehabilitation require ethical caution and a low tolerance for error, which makes VR's capacity for high-fidelity, risk-free simulation especially valuable [37, 38, 50]. Learners can repeatedly practice critical tasks, while KT enables detailed monitoring of progress. In engineering fields, training often demands expensive equipment, technical precision, and large-scale infrastructure, making hands-on learning costly and resource-intensive [41, 48, 52]. VR offers an efficient solution, and KT allows for tracking learners' spatial and procedural knowledge development in these virtual environments.

These patterns suggest that VR is particularly effective in educational settings that require a high degree of specialization and practical operation, especially for skills training, professional practice, and the construction of complex knowledge structures. Current KT research in VR commonly focuses on dynamically tracking learners' mastery of professional knowledge, operational skill development, and spatial ability acquisition. This aligns with the principles of embodied cognition theory, which emphasizes that cognitive processes are deeply embedded within the interaction between individuals and their environment, rather than occurring in isolation [14]. In VR settings, highly interactive three-dimensional environments enable learners to engage with virtual spaces through physical movements. Embodied behaviors such as operational actions and motion trajectories both facilitate knowledge construction and provide valuable data for tracking learners' knowledge states. Consequently, educational domains that emphasize operational tasks, spatial complexity, or collaborative processes, such as clinical nursing, construction engineering, and neuroanatomy, are more likely to become concentrated areas for KT research in VR environments.

B. Learning Tasks and Educational Scenarios

Based on content analysis and LDA topic modeling, we identified several key research themes related to learning tasks and educational scenarios in the context of KT in VR environments. These themes encompass knowledge state modeling and performance prediction, multimodal learning process analysis and cognitive mechanism exploration, learning outcome evaluation and instructional design optimization, adaptive learning driven by individual differences, the development of instructional systems and methodological frameworks, as well as learning outcome evaluation and comparative studies. Notably, the themes derived from LDA closely align with the manually coded categories, providing strong triangulation that highlights current research priorities and representative application scenarios.

Consistent with existing review studies, which have pointed out that VR research often focuses on "technology–context–outcome" perspectives while paying relatively little attention to dynamic modeling of learning processes [6, 28], the present study further refines the understanding of dynamic learning process analysis. The findings reveal a growing trend in VR-based KT research, shifting from static outcome evaluation towards more process-oriented, multimodal data integration approaches. In particular, within the first and third thematic clusters, existing studies have combined eye-tracking, EEG, physiological signals, and behavioral trajectory data to analyze learners' attention, cognitive, and emotional state changes, reflecting a clear tendency toward cross-modal data fusion. The observed shift from static outcome assessment to process-oriented, multimodal analysis of learning embodies the core principles of the CATIL, which posits that immersive experiences foster emotional engagement, thereby enhancing cognitive processing and improving learning outcomes [15, 16]. Studies that incorporate eye-tracking, EEG, and behavioral data to monitor attention and engagement are consistent with CATIL's emphasis on the dynamic interplay between cognitive and affective factors in immersive learning environments.

The second and fifth thematic clusters further demonstrate that KT not only supports the evaluation of VR instructional design but also provides the technical foundation for real-time personalized feedback and adaptive learning. Such research trends, emphasizing multimodality, process orientation, and personalization, offer valuable empirical evidence and methodological insights for advancing KT model development in VR environments. The results of topic modeling reveal that systematic research on constructing robust knowledge state prediction and tracing models within VR environments is still underdeveloped. Most existing studies are still at an early stage, focusing on data description or basic predictive tasks, with few establishing unified technical standards or modeling paradigms. In addition, studies in the fourth thematic cluster have explored the application of VR-based KT in child safety education, risk awareness training, and skill development. These studies indicate that the application of KT in VR for children and primary education has begun to emerge, but the overall volume of research in this area remains limited [55, 56].

C. Current Challenges and Future Directions of Reviewed Studies

Although the integration of VR and KT has made notable progress in recent years, the field remains in an early exploratory stage, with several limitations that require further breakthroughs and refinement.

First, the distribution of research samples and educational contexts remains uneven. As noted earlier, existing studies are predominantly concentrated in higher education and vocational training settings, with a clear emphasis on skill-oriented disciplines such as medicine and engineering [42, 56]. In contrast, research at the foundational education levels, including preschool education and special education, remains scarce [42, 56]. This limits the generalizability of VR-based KT findings across diverse educational scenarios. Future research should expand to more authentic classroom contexts, with particular attention to learners of different ages, ability levels, and cultural backgrounds. Stratified and cross-context sampling strategies are recommended to select representative courses across K-12, higher education, and vocational training. Aligning VR tasks with national curriculum standards or professional competency frameworks will enhance the comparability and educational relevance of experimental designs. Furthermore, establishing multi-institutional collaborative platforms to create cross-school and cross-regional data networks can facilitate the collection of longitudinal, multi-cycle datasets, laying the groundwork for transfer learning and generalization studies.

Second, research on KT in VR environments remains exploratory, with no unified theoretical framework or standardized modeling paradigm yet established. Although various studies have applied KT techniques to analyze learners' cognitive states and learning outcomes from different perspectives, significant inconsistencies persist in the research design, data handling, and result reporting [58]. These discrepancies limit the comparability and reproducibility of findings across studies. Furthermore, the immersive and highly interactive nature of VR presents new challenges in ensuring the adaptability and robustness of KT models within complex learning environments. Future research should prioritize the development of systematic, theory-driven KT frameworks tailored to the unique characteristics of VR learning. In particular, greater attention is needed to promote model generalizability across diverse learning tasks, educational contexts, and learner groups. Enhancing model interpretability and aligning KT research more closely with instructional needs will be critical for supporting real-time learning diagnostics and personalized teaching interventions within intelligent learning environments.

Third, the current evaluation systems in VR-based KT research remain limited, with insufficient emphasis on explaining the underlying cognitive mechanisms. Although existing studies have introduced various evaluation dimensions, most remain at the technical validation stage, focusing primarily on short-term outcomes rather than long-term, theory-driven assessments [8, 9, 21, 23, 40, 45–49, 51]. Longitudinal studies, sustained instructional interventions, and evaluations based on real-world teaching feedback are still lacking. Furthermore, many studies prioritize the question of whether learning states can be predicted, but pay limited attention to why such predictions are effective from a cognitive perspective. To address these gaps, future research should strengthen theoretical modeling of cognitive processing, emotional regulation, and knowledge-construction dynamics. Advancing KT from a purely performance-oriented approach toward interpretable, mechanism-driven modeling will be essential for enhancing its educational relevance and practical value.

MOREOVER, with the rapid advancement of Artificial intelligence (AI) technologies, recent studies have applied AI to assessment and feedback in language education. For example, Obaidoon and Wei [59] conducted a comparative analysis of four mainstream AI tools used for writing feedback. Their findings indicate that while these tools are relatively effective at addressing vocabulary and grammar issues, they fall short in offering higher-level writing guidance compared to human teachers. Furthermore, Han and Li [60] employed ChatGPT to assist teachers in delivering writing feedback and found that this approach significantly improved both instructional efficiency and student engagement. These studies provide important insights for KT in VR environments. From a technological and system development perspective, integrating AI with KT methods can promote the realization of interpretable modeling, enabling systems to identify learners' knowledge states in real time and generate initial feedback. From an instructional implementation perspective, teachers should continue to play a leading role in intervention and guidance by supplementing and adjusting AI-generated feedback based on their professional judgment, thereby establishing a collaborative instructional model that combines AI-based prediction with human intervention.

In summary, while the application of KT in VR environments shows promising potential, significant challenges remain in terms of educational applicability, theoretical development, and cognitive interpretability. Addressing these limitations will require systematic efforts to diversify research contexts, standardize modeling approaches, and deepen the understanding of underlying learning mechanisms. Such advancements are essential to fully leverage VR and KT integration for supporting adaptive, data-driven, and cognitively grounded learning in future educational environments.

VI. Conclusion

This study conducted a systematic literature review to examine the development of KT in VR environments for education over the past decade. The review focused on the following key dimensions: (1) research context and study design, (2) learning tasks and educational scenarios, and (3) current challenges and future directions.

The findings indicate that current VR-based KT research is primarily concentrated in higher education and vocational training, with learning tasks covering a wide range of objectives, including knowledge state modeling, multimodal cognitive mechanism exploration, personalized learning path support, and instructional outcome evaluation. Triangulation of the LDA topic modeling results with manually coded research objectives reveals a clear shift in VR-KT research from purely assessment-driven approaches toward a greater focus on understanding learning mechanisms and optimizing instructional design.

Nonetheless, several challenges persist. Current research shows an uneven distribution across educational stages, a lack of standardized modeling mechanisms, limited longitudinal evaluation, and insufficient emphasis on cognitive interpretability. These factors present opportunities for further exploration and refinement, particularly in advancing the practical implementation of KT in VR-based learning environments. Future research should expand to diverse learner groups and educational contexts, develop cross-modal and transferable KT model frameworks, and promote the integration of theoretical modeling with data-driven, explainable mechanisms. Such efforts are essential to enhancing the practical utility and long-term sustainability of KT in VR learning environments.

This review provides a structured synthesis of VR-based KT research, but it includes only peer-reviewed journal articles and conference papers published in English. Although this criterion ensured consistency in screening and analysis, it may have excluded relevant studies published in other languages. Future reviews are encouraged to broaden the scope by incorporating multilingual databases to achieve a more comprehensive global perspective. Moreover, while the present review primarily focused on learning scenarios and tasks, core enabling technologies and specific learning outcomes received comparatively less attention. Future research could explore these dimensions to offer a more nuanced and comprehensive understanding of KT in VR environments.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Yuying Yang conducted the research, analyzed the data and wrote the paper; Geping Liu analyzed the data and wrote the paper; all authors had approved the final version.

Funding

This research was funded by the National Natural Science Foundation of China (Grant No. 62277044) and the Chongqing Municipal Graduate Research and Innovation Project (Grant No. CYS240102).

References

[1] J. N. Han, G. P. Liu, and S. Y. Wang, "Characteristics and development directions of social interactive learning in the metaverse," *Mod. Dist. Educ. Res.*, vol. 35, no. 2, pp. 57–66, 2023.

[2] F. Hekele, J. Spilski, S. Bender, and T. Lachmann, "Remote vocational learning opportunities—A comparative eye-tracking investigation of educational 2D videos versus 360° videos for car mechanics," *Br. J. Educ. Technol.*, vol. 53, no. 2, pp. 248–268, 2022.

[3] W. S. Wang, C. J. Lin, H. Y. Lee, Y. M. Huang, and T. T. Wu, "Integrating feedback mechanisms and ChatGPT for VR-based experiential learning: Impacts on reflective thinking and Aiot physical hands-on tasks," *Interact. Learn. Environ.*, vol. 33, no. 2, pp. 1770–1787, 2025.

[4] M. Singh, D. Sun, and Z. Zheng, "Enhancing university students' learning performance in a metaverse-enabled immersive learning environment for STEM education: A community of inquiry approach," *Futur. Educ. Res.*, vol. 2, no. 3, pp. 288–309, 2024.

[5] L. Jensen and F. Konradsen, "A review of the use of virtual reality head-mounted displays in education and training," *Educ. Inf. Technol.*, vol. 23, no. 4, pp. 1515–1529, 2018.

[6] G. Lampropoulos and G. Evangelidis, "Learning analytics and educational data mining in augmented reality, virtual reality, and the metaverse: A systematic literature review, content analysis, and bibliometric analysis," *Appl. Sci.*, vol. 15, no. 2, p. 971, 2025.

[7] L. He, X. Li, J. Tang, and T. Wang, "EDKT: An extensible deep knowledge tracing model for multiple learning factors," in *Proc. International Conference on Database Systems for Advanced Applications*, Springer, 2021, pp. 340–355.

[8] S. Y. Yang and Y. H. Oh, "Development of neonatal apgar scoring training program utilizing contactless hand tracking in immersive virtual reality," *Nurse Educ. Today*, vol. 140, 106294, 2024.

[9] M. V. Miyusov, L. L. Nikolaieva, and V. V. Smolets, "The effectiveness of immersive learning in maritime education and training," *Trans. Maritime Sci.*, vol. 11, no. 2, 2022.

[10] C. Walkington, M. J. Nathan, W. Huang, J. Hunnicutt, and J. Washington, "Multimodal analysis of interaction data from embodied education technologies," *Educ. Technol. Res. Dev.*, vol. 72, no. 5, pp. 2565–2584, 2024.

[11] A. G. Moore, R. P. McMahan, H. Dong, and N. Ruozzi, "Extracting velocity-based user-tracking features to predict learning gains in a virtual reality training application," in *Proc. IEEE Int. Symp. Mixed Augment. Reality (ISMAR)*, 2020, pp. 694–703.

[12] G. P. Liu, W. Y. Ran, Y. Y. Yang, and H. L. Hu, "A review of key technologies for cognitive tracing in virtual reality environments," *Artif. Intell. Sci. Eng.*, vol. 2, pp. 1–17, 2024.

[13] A. Ismayilzada, A. Karimov, and M. Saarela, "Serious games analytics in VR environments: A two-stage systematic literature review," *J. Interact. Learn. Res.*, vol.36, no. 1, pp. 57–69, 2025.

[14] A. Clark, *Being There: Putting Brain, Body, and World Together Again*, Cambridge, MA: MIT Press, 1997.

[15] G. Makransky and L. Lilleholt, "A structural equation modeling investigation of the emotional value of immersive virtual reality in education," *Educ. Technol. Res. Dev.*, vol. 66, no. 5, pp. 1141–1164, 2018.

[16] G. Makransky, N. K. Andreasen, S. Baceviciute, and R. E. Mayer, "Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality," *J. Educ. Psychol.*, vol. 113, no. 4, pp. 719–735, 2021.

[17] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," *User Model. User-Adapt. Interact.*, vol. 4, pp. 253–278, 1994.

[18] C. Piech, J. Bassen, J. Huang, *et al*., "Deep knowledge tracing," *Adv. Neural Inf. Process. Syst.*, vol. 28, 2015.

[19] J. Zhang, X. Shi, I. King *et al*., "Dynamic key-value memory networks for knowledge tracing," in *Proc. Int. Conf. World Wide Web*, 2017, pp. 765–774.

[20] S. Safikhani, L. Nacke, and J. Pirker, "A literature review and taxonomy of in-VR questionnaire user interfaces," *Int. Conf. Immersive Learn.*, 2024, pp. 95–111.

[21] M. Lui, R. McEwen, and M. Mullally, "Immersive virtual reality for supporting complex scientific knowledge: Augmenting our understanding with physiological monitoring," *Br. J. Educ. Technol.*, vol. 51, no. 6, pp. 2181–2199, 2020.

[22] F. Zhao, C. Zhang, and B. Geng, "Deep multimodal data fusion," *ACM Comput. Surv.*, vol. 56, no. 9, pp. 1–36, 2024.

[23] I. Dubovi, "Cognitive and emotional engagement while learning with VR: The perspective of multimodal methodology," *Comput. Educ.*, vol. 183, 104495, 2022.

[24] L. Muñoz-Saavedra, L. Miró-Amarante, and M. Domínguez-Morales, "Augmented and virtual reality evolution and future tendency," *Appl. Sci.*, vol. 10, no. 1, p. 322, 2020.

[25] H. Luo, G. Li, Q. Feng, Y. Yang, and M. Zuo, "Virtual reality in K-12 and higher education: A systematic review of the literature from 2000 to 2019," *J. Comput. Assist. Learn.*, vol. 37, no. 3, pp. 887–901, 2021.

[26] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Comput. Educ.*, vol. 147, 103778, 2020.

[27] K. H. Tai, J. C. Hong, C. R. Tsai, C. Z. Lin, and Y. H. Hung, "Virtual reality for car-detailing skill development: Learning outcomes of

procedural accuracy and performance quality predicted by VR self-efficacy, VR using anxiety, VR learning interest and flow experience," *Comput. Educ.*, vol. 182, 104458, 2022.

[28] R. Shadiev and D. Li, "A review study on eye-tracking technology usage in immersive virtual reality learning environments," *Comput. Educ.*, vol. 196, 104681, 2023.

[29] M. Dai, J. L. Hung, X. Du, H. Tang, and H. Li, "Knowledge tracing: A review of available technologies," *J. Educ. Technol. Dev. Exch.*, vol. 14, no. 2, pp. 1–20, 2021.

[30] A. Zanellati, D. Di Mitri, M. Gabbrielli, and O. Levrini, "Hybrid models for knowledge tracing: A systematic literature review," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 1021–1036, 2024.

[31] W. L. Chan and D. Y. Yeung, "Clickstream knowledge tracing: Modeling how students answer interactive online questions," in *Proc. LAK21: 11th Int. Learn. Anal. Knowl. Conf.*, 2021, pp. 99–109.

[32] K. Takami and B. Flanagan, "Toward educational explainable recommender system: explanation generation based on Bayesian knowledge tracing parameters," *Int. Conf. Comput. Educ.*, 2021.

[33] H. Yan, F. Lin, and Kinshuk, "Including learning analytics in the loop of self-paced online course learning design," *Int. J. Artif. Intell. Educ.*, vol. 31, no. 4, pp. 878–895, 2021.

[34] A. Trifa, A. Hedhili, and W. L. Chaari, "Knowledge tracing with an intelligent agent, in an e-learning platform," *Educ. Inf. Technol.*, vol. 24, pp. 711–741, 2019.

[35] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow *et al*., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *BMJ*, vol. 372, no. 71, 2021.

[36] S. M. Asish, A. K. Kulshreshth, and C. W. Borst, "Detecting distracted students in an educational VR environment utilizing machine learning on EEG and eye-gaze data," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces Workshops (VRW)*, 2023, pp. 703–704.

[37] G. Nguyen, S. Ishikawa, M. Kobayashi, and M. Ito, "Evaluating observation skill in nursing education through gaze-based objective assessment in immersive simulation," in *Proc. IEEE Int. Symp. Mixed Augment. Reality Adjunct (ISMAR-Adjunct)*, 2023, pp. 271–275.

[38] A. G. Moore, R. P. McMahan, and N. Ruozzi, "Exploration of feature representations for predicting learning and retention outcomes in a VR training scenario," *Big Data Cogn. Comput.*, vol. 5, no. 3, p. 29, 2021.

[39] J. Lee, C. C. Chen, and A. Basu, "From novelty to knowledge: A longitudinal investigation of the novelty effect on learning outcomes in virtual reality," *IEEE Trans. Vis. Comput. Graph.*, vol. 31, no. 5, pp. 3204–3212, 2025.

[40] L. Liberman and I. Dubovi, "The effect of the modality principle to support learning with virtual reality: An eye-tracking and electrodermal activity study," *J. Comput. Assist. Learn.*, vol. 39, no. 2, pp. 547–557, 2023.

[41] M. Tomori, O. Ogunseiju, M. Tummalapudi, and S. Bangaru, "Towards personalized learning environments: Using machine learning to predict students' learning preferences in a mixed reality environment," in *Proc. IEEE Front. Educ. Conf. (FIE)*, 2024, pp. 1–9.

[42] Z. Sokolikj, F. Ke, S. Chakraborty, and J. Moon, "Using deep learning to track representational flexibility development of children with autism in a virtual world," in *Proc. Int. Conf. Inf. Educ. Technol. (ICIET)*, 2023, pp. 51–55.

[43] J. Orlosky, B. Huynh, and T. Hollerer, "Using eye tracked virtual reality to classify understanding of vocabulary in recall tasks," in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, 2019, pp. 66–667.

[44] M. Holly, S. Brettschuh, A. S. Tiwari, K. K. Bhagat, and J. Pirker, "Game-based motivation: Enhancing learning with achievements in a customizable virtual reality environment," in *Proc. ACM Symp. Virtual Reality Softw. Technol.*, 2024, pp. 1–11.

[45] C. C. Wang, J. C. Hung, and H. C. Chen, "How prior knowledge affects visual attention of Japanese mimicry and onomatopoeia and learning outcomes: Evidence from virtual reality eye tracking," *Sustainability*, vol. 13, no. 19, 11058, 2021.

[46] X. Wu, X. Chen, J. Zhao, and Y. Xie, "Influences of design and knowledge type of interactive virtual museums on learning outcomes: An eye-tracking evidence-based study," *Educ. Inf. Technol.*, vol. 29, no. 6, pp. 7223–7258, 2024.

[47] X. X. Feng, S. Powers, L. Eberman, Z. Liu, X. Xin, and Y. R. Zhang, "The effects of a three-dimensional virtual learning medium on spatial ability and learning achievement in the anatomical sciences," *Think. Skills Creat.*, vol. 58, 101886, 2025.

[48] R. R. Rafa, T. Rahman, M. H. Kobir, Y. Yang, and S. Deb, "Enhancing experiential learning through virtual reality: System design and a case study in additive manufacturing," *Hum. Factors Ergon. Manuf. Serv. Ind.*, vol. 34, no. 6, pp. 649–666, 2024.

[49] D. M. Markowitz, R. Laha, B. P. Perone, R. D. Pea, and J. N. Bailenson, "Immersive virtual reality field trips facilitate learning about climate change," *Front. Psychol.*, vol. 9, 2364, 2018.

[50] A. S. Baetzner, Y. Hill, B. Roszipal, S. Gerwann, M. Beutel, T. Birrenbach, M. Karlseder, S. Mohr, G. A. Salg, H. Schrom-Feiertag, M. O. Frenkel, and C. Wrzus, "Mass casualty incident training in immersive virtual reality: Quasi-experimental evaluation of multimethod performance indicators," *J. Med. Internet Res.*, vol. 27, e63241, 2025.

[51] E. Pérez-Martín, S. L. C. Medina, T. R. Herrero-Tejedor, A. Ezquerra-Canalejo, and D. López-Fernández, "Using virtual reality in the learning of geomatic engineering education," *IEEE Comput. Graph. Appl.*, vol. 44, no. 6, pp. 77–88, 2024.

[52] E. L. Jacobsen, A. Solberg, O. Golovina, and J. Teizer, "Active personalized construction safety training using run-time data collection in physical and virtual reality work environments," *Constr. Innov.*, vol. 22, no. 3, pp. 531–553, 2022.

[53] V. Häfner, T. Li, F. Michels, P. Häfner, H. Yu, and J. Ovtcharova, "A framework for intelligent virtual reality tutoring system using semantic web technology," in *Proc. Int. Conf. Comput. Supported Educ.*, 2024, pp. 141–152.

[54] E. E. Opait, D. Silion, A. Iftene, C. Luca, and C. Corciova, "Mixed realities tools used in biomedical education and training," in *Proc. Int. Conf. INnov. Intell. Syst. Appl. (INISTA)*, 2024, pp. 1–6.

[55] E. B. H. Sandseter, O. J. Sando, H. Lorås, R. Kleppe, L. Storli, M. Brussoni, A. Bundy, D. C. Schwebel, D. J. Ball, M. Haga, and H. Little, "Virtual risk management—exploring effects of childhood risk experiences through innovative methods (ViRMa) for primary school children in Norway: Study protocol for the ViRMa project," *JMIR Res. Protoc.*, vol. 12, e45857, 2023.

[56] M. Pilatásig, E. Tobar, L. Paredes, F. M. Silva, A. Acurio, E. Pruna, I. Escobar, and Z. Sánchez, "Virtual system for teaching-learning of initial education using a haptic device," in *Augmented Reality, Virtual Reality, and Computer Graphics*, L. T. De Paolis and P. Bourdot, Eds., vol. 10850, Springer, 2018, pp. 118–132.

[57] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Jan. 2003.

[58] N. Gaw, S. Yousefi, and M. R. Gahrooei, "Multimodal data fusion for systems improvement: A review," *Handbook of Scholarly Publications from the Air Force Institute of Technology (AFIT)*, vol. 1, pp. 101–136.

[59] S. Obaidoon and H. Wei, "ChatGPT, Bard, Bing Chat, and Claude generate feedback for Chinese as foreign language writing: A comparative case study," *Futur. Educ. Res.*, vol. 2, no. 3, pp. 184–204, Jun. 2024.

[60] J. N. Han and M. M. Li, "Exploring ChatGPT-supported teacher feedback in the EFL context," *System*, vol. 126, 103502, 2024.