Discover Education

https://doi.org/10.1007/s44217-026-01798-y

Article in Press

Engagement responsive pedagogy with the PRISM framework for inclusive and adaptive learning

Received: 18 September 2025

Accepted: 12 June 2026

Published online: 23 June 2026

Cite this article as: Ibitoye A.O., Chishti I., Akinyemi J.D. *et al.* Engagement responsive pedagogy with the PRISM framework for inclusive and adaptive learning. *Discov Educ* (2026). https://doi.org/10.1007/s44217-026-01798-y

Ayodeji Olusegun Ibitoye, Irfan Chishti, Joseph Damilola Akinyemi, Adeyinka Oluwabusayo Abiodun & Olufade F. W. Onifade

Engagement Responsive Pedagogy with the PRISM Framework for Inclusive and Adaptive Learning

Ayodeji Olusegun Ibitoye[1*], Irfan Chishti[2], Joseph Damilola Akinyemi[3], Adeyinka Oluwabusayo Abiodun[4], Olufade F.W Onifade[5]

[1,2] School of Computing and Mathematical Sciences, University of Greenwich, London, UK
[3] Department of Computer Science, University of York, Heslington, York, UK
[4] Africa Centre of Excellence on Technology Enhanced Learning, National Open University of Nigeria, Abuja, Nigeria.
[5]Department of Computer Science, University of Ibadan, Ibadan, Nigeria.

Corresponding Author: Ayodeji Olusegun Ibitoye
Email: a.o.ibitoye@greenwich.ac.uk

Abstract

As online and hybrid learning environments expand, traditional instructional frameworks often struggle to address cognitive overload, fragmented attention, and learner disengagement. This work introduces PRISM, an engagement-responsive pedagogical framework comprising four interdependent pillars: Personalised Adaptive Learning, Responsive Engagement, Immersive Learning Experiences, and Social Collaboration. PRISM is designed to support real-time instructional adaptation based on the evolving states of learners. To computationally operationalise PRISM's core logic, a multimodal transformer model was developed using the publicly available PEEK dataset, containing more than 290,000 learner–video interaction records with behavioural engagement labels. The model achieved better predictive performance (F1 = 82.6%), outperforming text-only and metadata-only baselines. Simulation experiments examined how engagement predictions could trigger micro-level pedagogical interventions within an adaptive instructional loop. Although not tested in live classrooms, ablation studies and temporal attention analyses suggested the additive value of combining textual and behavioural features and highlighted patterns consistent with engagement fluctuations across viewing timelines. A simulated comparison of static versus adaptive conditions indicated that timely interventions may help reduce learner drop-off and improve engagement metrics. These findings suggest the technical feasibility of engagement-aware instruction and provide an initial conceptual foundation for future implementations of PRISM in authentic learning environments. The study concludes with design implications, limitations of simulation-based inference, and directions for future work on deployment, inclusivity, and learner perceptions in adaptive learning systems.

Keywords: PRISM framework, learner engagement, adaptive learning, multimodal modelling, transformer architecture, real-time education, cognitive equity

1. Introduction

The educational landscape is undergoing a profound transformation, driven by technological advancements, evolving learner expectations, and the growing demand for personalised, engaging experiences [1]. Central to this shift is the emergence of Generation Z, learners who have been raised in a digital-first environment shaped by social media, streaming platforms, and interactive content [2]. These students are characterised by shorter attention spans, a strong preference for multimodal content delivery, and an expectation for technology-enhanced, responsive learning environments [3]. Despite the proliferation of innovative pedagogical models, such as flipped classrooms [4] and project-based learning

[5], student disengagement remains a persistent and growing concern. Existing instructional approaches often fall short in aligning with the cognitive and motivational profiles of today's learners. As [6] notes, distractions within the classroom and disinterest in conventional teaching strategies exacerbate disengagement. The prevalence of technostress, the psychological strain induced by overexposure to digital technologies, further hampers academic focus and well-being [7]. This is compounded by cyberloafing, which undermines concentration and academic performance [8].

The COVID-19 pandemic brought these issues into sharper relief, revealing systemic weaknesses in digital learning environments. Studies have shown that online learning during the pandemic often suffered from low levels of sustained engagement and persistence [9]. In response, educational stakeholders are increasingly advocating for learning environments that offer not only technological integration but also meaningful interactivity and learner relevance [10]. Motivational factors further influence academic engagement, with research indicating that aspirations and future orientation significantly affect students' commitment [11]. To address these evolving learner needs, educators are encouraged to adopt strategies such as game-based learning to increase engagement [12] and to leverage peer support systems to boost motivation and reduce anxiety [13]. Moreover, technology acceptance models, such as the Unified Theory of Acceptance and Use of Technology (UTAUT2), provide valuable insights into the determinants of successful digital tool adoption in educational contexts [14]. Although these models and strategies enhance digital learning effectiveness, they remain primarily design-oriented and static. They offer guidance on how to plan instruction but lack mechanisms for real-time pedagogical responsiveness based on fluctuations in learner attention, motivation, or interaction behaviour. As a result, disengagement may go unnoticed until the learning activity is completed, and opportunities for immediate instructional intervention are missed. This gap highlights the need for responsive instructional frameworks that can leverage live engagement signals to adapt learning experiences dynamically.

Addressing this need, the present study introduces PRISM, a pedagogical framework designed to meet the engagement challenges of Generation Z learners.PRISM stands for Personalised Adaptive Learning, Responsive Engagement, Immersive Learning Experiences, and Social Collaboration. These four pillars collectively aim to promote attention, sustainability, learner autonomy, and inclusive instructional practices. PRISM is grounded in established learning theories and operationalises them by integrating technological responsiveness into instructional design. Its core function is to adapt learning activities dynamically based on real-time engagement signals. Unlike frameworks such as Universal Design for Learning (UDL), Technological Pedagogical Content Knowledge (TPACK), or SAMR, which guide instructional planning and technology integration, PRISM offers a real-time operational mechanism for instructional adaptation. It translates learner-state information into actionable pedagogical interventions, allowing instruction to respond to engagement fluctuations during learning rather than solely at the design stage. It emphasises moment-to-moment adaptation during instruction, enabled through intelligent systems. This distinguishes PRISM as a responsive instructional framework rather than a static model of technology integration.

This study articulates the theoretical foundations and instructional design logic of the PRISM framework and presents a pilot computational validation of one of its core mechanisms: real-time responsiveness to learner engagement. Using the PEEK dataset, a large-scale collection of annotated learner-video interactions, we implement a transformer multimodal model to detect shifts in engagement and simulate micro-level instructional adaptations. This empirical component demonstrates the feasibility of PRISM's operational logic within digitally mediated learning environments. This study is guided by the following research questions: (1) How can real-time engagement be detected using multimodal learner data? (2) To what extent can a transformer model simulate the responsive logic of the PRISM framework? and (3) What implications do these results hold for designing adaptive, attention-aware pedagogies in digital learning environments?

By bridging theoretical insights with implementation, this research contributes to the development of next-generation instructional models capable of responding dynamically to learner needs. The PRISM framework has the potential to redefine how engagement is conceptualised and operationalised in educational settings, shifting the focus from static content delivery to adaptive, learner-centred instruction. Its implications extend across curriculum design, educational technology development, and teacher training, offering a scalable approach for enhancing motivation, reducing cognitive overload, and fostering deep learning in digitally mediated environments. Building on these considerations, the present study introduces PRISM as an early-stage computational architecture designed to examine the feasibility of real-time engagement detection and adaptive intervention within asynchronous video-based learning environments. Although not yet validated in live classroom settings, PRISM establishes a conceptual and technical foundation for exploring how responsive, data-informed design may enrich digital learning experiences. The remainder of this paper outlines the theoretical grounding, model design, simulation methods, and preliminary findings that position PRISM as a promising direction for future empirical and applied research.

2. Literature Review

This section surveys four interconnected areas of scholarship that inform the PRISM framework. First, it examines how learner engagement has been conceptualised within digital learning environments. Second, it considers advances in technology-enhanced learning and real-time adaptivity that support continuous monitoring and personalisation. Third, it reviews the theoretical foundations of responsive learning and motivation-driven instructional design, serving as PRISM's theoretical foundation. Then it analyses established pedagogical models to clarify PRISM's positioning, contribution, and areas of theoretical and practical extension.

2.1 The Evolving Nature of Learner Engagement

Learner engagement has emerged as a pivotal focus in educational research, particularly within the context of technology-enhanced learning environments. It is widely conceptualised as a multidimensional construct encompassing behavioural, cognitive, and emotional involvement in academic tasks [15]. As digital technologies continue to reshape pedagogical practices, the nature of engagement has evolved, especially among Generation Z learners, who favour interactive, personalised, and feedback-rich environments [16]. Traditional pedagogical models, such as lecture-driven instruction and static content delivery, increasingly fall short in meeting the expectations and cognitive styles of these digitally native students. Research has shown that short attention spans, heightened exposure to digital distractions, and reduced tolerance for passive consumption significantly impair sustained engagement [17-18]. Consequently, engagement is now being reconceptualised not merely as a psychological state but as a dynamic, observable, and responsive process, one that demands adaptability and cognitive alignment. These evolving definitions of engagement underscore the need for pedagogical models, like PRISM, that not only acknowledge engagement as dynamic but also incorporate real-time adaptability into instructional design

Technological integration has become a cornerstone of this transformation. [19] emphasise that technology now plays a critical role in enhancing interactivity and fostering learner participation in the modern classroom. Effective management of educational environments is equally crucial; [20] underscores the need for systemic strategies that cultivate engaging and academically robust learning contexts. Complementing this, [21] demonstrates that the adoption of mixed-method pedagogies can lead to improved learner commitment and educational outcomes, while [22] explores dynamic, student-centred strategies aimed at amplifying classroom participation. Learning Management Systems (LMSs) further exemplify the structural shift toward digital engagement. [23] argue that LMS platforms serve as vital tools for enhancing learner commitment, particularly in higher education settings. The integration of artificial intelligence (AI) into hybrid education environments has also shown promising potential. As [24] observes, AI-driven systems can personalise instruction and automate feedback, thereby improving student motivation and engagement. Recent research demonstrates the growing impact of immersive technologies, such as Metaverse-style environments and Augmented Reality/Virtual Reality (AR/VR) platforms, in enhancing interactivity, learner presence, and instructional effectiveness, thereby accelerating the shift from static content delivery to dynamic, engagement-driven pedagogies [25 - 26]. These findings reinforce the need for frameworks like PRISM that operationalise technological immersion through real-time adaptivity. Environmental and motivational factors also play significant roles in shaping engagement. [27] explores how the learning environment, both physical and digital, can affect students' academic drive, while [28] presents a conceptual framework that outlines key strategies for fostering sustained online engagement. Their work highlights ongoing challenges and opportunities for both educators and learners navigating digital learning landscapes. These studies underscore a critical consensus: learner engagement is no longer a static attribute but a responsive, context-dependent phenomenon. Emerging pedagogical frameworks, such as PRISM, must therefore be designed to operationalise these insights, integrating adaptive technologies, immersive environments, and socially collaborative tools to sustain attention, motivation, and learning outcomes in increasingly complex educational ecosystems.

2.2 Technology-Enhanced Learning and Real-Time Adaptivity

The integration of technology in education has transformed how learning is delivered, monitored, and personalised, offering unprecedented opportunities for adaptivity and data-informed instructional design [29]. Technology-Enhanced Learning (TEL) environments, powered by learning management systems, AI-driven tutors, and intelligent analytics dashboards, are increasingly capable of capturing, interpreting, and responding to learner behaviours in real time [30]. However, while many of these systems excel at content distribution and performance evaluation, they often struggle to sustain learner engagement or manage attentional shifts over time.

Recent developments in affective computing and learning analytics have begun to address this gap by incorporating behavioural and physiological signals, such as facial expressions, gaze direction, interaction patterns, and pause/play events, to model student engagement dynamically [31, 32]. These multimodal data streams serve as critical inputs for adaptive algorithms that can tailor learning experiences based on live feedback. The PRISM framework builds on this foundation, formalising the logic of micro-responsiveness*,* where instructional interventions are triggered in response to engagement signals. TEL's evolution is characterised by its increasing emphasis on personalisation and adaptivity. [33] highlights the role of emerging computational technologies in enhancing instructional effectiveness across diverse educational settings. Similarly, [34] proposes a comprehensive model for integrating adaptive learning technologies into higher education, underscoring their value in customising instruction and improving learner outcomes.

This argument is extended by [35], demonstrating that adaptive teaching not only aligns content with individual learner needs but also fosters long-term retention and deeper learning. Central to this adaptivity is the use of data analytics, which enables educators to continuously monitor and respond to student performance and behaviour [36]. [37] further argues that technology-mediated environments must go beyond efficiency and embrace transformative learning, where adaptivity is used to empower learners and support self-directed inquiry. A structural understanding of adaptive systems is equally essential. [38] propose taxonomies that map the functional dimensions of TEL environments, providing a systematic basis for evaluating how adaptive mechanisms are implemented. [39] reinforces this perspective by exploring how AI-driven platforms can scaffold cognitive development through personalised learning trajectories.

However, the enthusiasm surrounding TEL must be tempered with critical reflection. [40] caution against uncritical reliance on technology that prioritises system efficiency over educational depth. Similarly, [41] and [42] stress the importance of aligning technological interventions with pedagogical goals to preserve educational integrity. These critiques underscore the need for frameworks, like PRISM, that maintain a balanced, learner-centred approach by embedding adaptivity within a coherent pedagogical architecture. The intersection of TEL and adaptivity represents a paradigm shift in instructional design. By leveraging analytics and AI not merely to deliver content but to respond to the cognitive and emotional states of learners, emerging systems can support engagement-sensitive pedagogies. Despite advances in real-time adaptivity, concerns about interpretability and trust in AI-assisted educational systems remain. [43] found that although AI-generated assessments may align with human judgments, both students and instructors still question the contextual depth and adaptability of such systems. PRISM contributes to this shift by integrating multimodal data, responsive feedback, and adaptive pathways, thereby advancing a vision of education that is both technologically advanced and pedagogically grounded.

2.3 Theoretical Foundations for PRISM’s Framework

The PRISM framework is grounded in complementary learning theories that collectively shape its responsive, adaptive, and engagement-focused design. These theories explain how learners construct knowledge, what drives their motivation, and how technology can support meaningful participation. Foundational to this synthesis are constructivist and connectivist perspectives, which frame PRISM’s conceptualisation of learner engagement. Constructivism, rooted in [44], emphasises learner agency, scaffolding, and contextual collaboration, principles mirrored in PRISM’s socially driven, problem-based activities. This view is strengthened by [45], who highlight the value of learning spaces that promote critical thinking and autonomy, and by [46], who demonstrate the adaptability of constructivist approaches for diverse learners. Connectivism [47] extends these ideas into digital contexts by framing learning as networked and distributed across digital environments. As learners increasingly depend on online interactions and knowledge flows, digital platforms become central to learning, a point underscored by [48] and [49], who show how social networks cultivate collaborative, lifelong learning communities. Understanding learning as both constructed and digitally connected naturally foregrounds the cognitive demands placed on learners in such environments.

Cognitive Load Theory (CLT) provides this insight, emphasising the need to balance intrinsic, extraneous, and germane load to prevent overload and optimise working memory [50,51]. In digitally mediated settings, efficient cognitive processing becomes crucial. PRISM responds through microlearning, multimodal delivery, and flexible pacing. The importance of minimising unnecessary mental effort is highlighted by [52], while [53] shows that effectively managed cognitive load enhances knowledge acquisition and retention. However, cognitive efficiency alone does not ensure engagement; motivation remains central. Self-Determination Theory (SDT) [54] explains intrinsic motivation through autonomy, competence, and relatedness, needs especially relevant in self-paced digital learning. While CLT clarifies how learners process information, SDT elucidates why they persist. PRISM addresses these motivational needs through personalised pathways and peer-based social structures, aligning with [55], who emphasise learner agency as a driver of deep engagement and long-term retention. Yet even motivated learners face increasing competition for attention within digital ecosystems.

Educational neuroscience highlights attention as a limited cognitive resource shaped by novelty, relevance, and timely feedback [56]. In an attention economy saturated with competing stimuli, sustaining focus becomes a significant instructional challenge. [57] emphasise that contemporary learning environments require designs capable of capturing and maintaining attention, while [58] and [59] show that relevance, interactivity, and cognitive novelty are essential for reducing distraction and maximising engagement. These insights underpin PRISM's responsive mechanisms, which use multimodal learner data to detect attentional lapses and trigger timely instructional adjustments. These interrelated theories provide a coherent foundation for responsive, technology-mediated pedagogy. Constructivism and connectivism explain how learners build and navigate knowledge; CLT identifies the cognitive conditions that support effective processing; SDT clarifies the motivational factors sustaining engagement; and educational neuroscience and attention-economy perspectives highlight the need to maintain focus in complex digital environments. All this informs PRISM's personalisation, adaptive feedback, immersive design, and collaborative structures, enabling it to meet the cognitive, motivational, and attentional demands of 21st-century learners while anticipating future directions in learner-centred, data-informed practice.

2.4 Related Pedagogical Models

The PRISM framework draws on several influential pedagogical models but advances them through a real-time, engagement-responsive layer of adaptivity. Foundational models such as Universal Design for Learning (UDL) [60], Technological Pedagogical Content Knowledge (TPACK) [61], and the SAMR model [62] have shaped inclusive and technology-enhanced instruction, yet PRISM reorients their principles toward dynamic learner responsiveness. UDL promotes inclusive design through multiple means of engagement, representation, and expression to accommodate learner variability [63]. PRISM extends these principles by allowing learner data to directly shape the instructional pathway as it unfolds. This reflects

[64], who emphasise UDL's importance for meeting diverse learner needs, and [65], who argue that integrating UDL with adaptive technologies mitigates technological determinism by foregrounding learner-centred variability.

TPACK, introduced by [58], provides a framework for integrating technological, pedagogical, and content knowledge. While TPACK largely centres on teacher expertise, PRISM shifts emphasis toward learner behaviour by operationalising instructional adjustments based on live input. This aligns with [66], who highlight the value of TPACK-informed strategies, particularly when coupled with UDL to enhance pedagogical diversity. Building on these models, the SAMR framework conceptualises technology integration across substitution, augmentation, modification, and redefinition [62]. [67] show that progressing through these stages can significantly transform learning design, but PRISM diverges by focusing on how technology responds to fluctuating engagement and attentional states within any stage rather than on the level of innovation itself.

UDL, TPACK, and SAMR provide the pedagogical architecture for inclusive, technology-mediated instruction. Research by [68] demonstrates that their integration can strengthen equity and learner engagement, while [69] highlights the synergistic benefits of applying them simultaneously to enrich both teacher capacity and student experience. Further studies [70–71] show how these combined frameworks enhance digital pedagogies, and [72] reports sustained improvements in learner outcomes when supported by robust implementation structures. The effective use of these models, however, relies heavily on ongoing professional development, [73] underscores the importance of continual training to ensure educators can fully leverage their potential.

PRISM incorporates these insights but extends them by embedding adaptivity and temporal responsiveness into instructional design. Whereas UDL addresses accessibility, TPACK supports teacher decision-making, and SAMR scaffolds technological transformation, none inherently account for moment-to-moment fluctuations in learner engagement. PRISM fills this gap by foregrounding behavioural and attentional dynamics through its data-driven responsiveness. Figure 1 maps how each PRISM pillar aligns with components of the empirical engagement prediction system used in this study, clarifying the theoretical–functional integration that underpins PRISM's pedagogical logic.

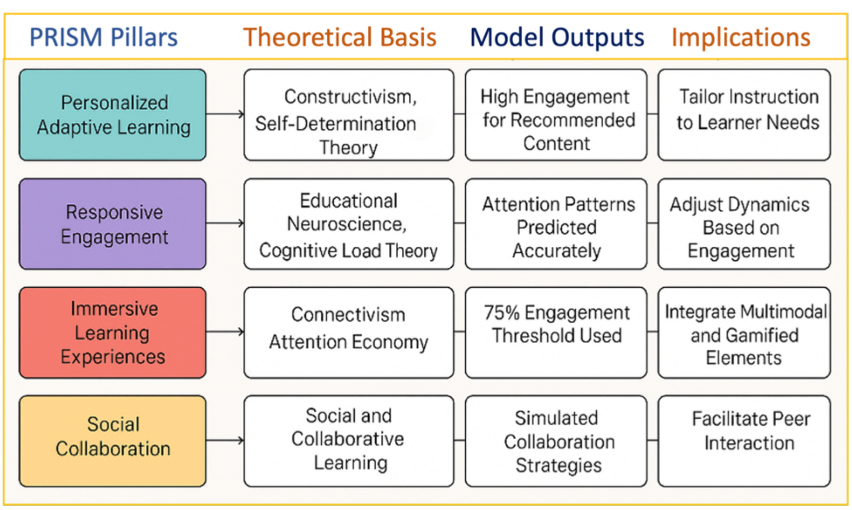

Figure 1: Conceptual Mapping of PRISM Pillars

To bridge the gap between theory, model validation, and instructional practice, Figure 1 offers a conceptual mapping of the PRISM framework. It visually aligns each of the four pillars, Personalised Adaptive Learning, Responsive Engagement, Immersive Learning Experiences, and Social Collaboration, with their respective theoretical foundations, empirical support from our model outputs, and practical implications for classroom application. This mapping reinforces PRISM's positioning not just as a conceptual model, but as a data-informed, pedagogically grounded response to the evolving demands of 21st-century learners

3. The PRISM Framework: A Pedagogical Model for Responsive, Inclusive, and Sustained Engagement

The PRISM framework, short for Personalised Adaptive Learning, Responsive Engagement, Immersive Learning Experiences, and Social Collaboration**,** is proposed as a pedagogical model that translates well-established cognitive and motivational theories into a practical, digitally integrated instructional design. PRISM responds to the evolving needs of Generation Z learners by offering a structured yet flexible approach to sustaining attention, enhancing motivation, and promoting inclusive, personalised learning at scale. Rather than functioning as a linear or prescriptive model, as illustrated in Figure 2, PRISM operates as a dynamic ecosystem of four interdependent pillars. PRISM's four pillars are not intended as independent components, but as mutually reinforcing processes within a dynamic learning ecology. Personalised Adaptive Learning determines the optimal pathway for each learner; Responsive Engagement continuously monitors attentional and behavioural signals to modify that pathway; Immersive Learning Experiences provide the emotionally and cognitively rich environments necessary for sustained motivation; and Social Collaboration ensures communal knowledge-building, which in turn informs adaptive decisions.

These pillars form a cyclical system in which learner data shapes instruction, instruction shapes engagement, and engagement shapes further adaptation. The modularity of this system allows educators to design, monitor, and adjust learning experiences in real time based on engagement feedback, learner profiles, and contextual demands. For example, in an undergraduate Data Science module, students begin with an interactive micro-lecture on regression modelling (Immersive Learning). As learners watch, the system tracks pauses, replay behaviour, and note-taking to estimate engagement levels (Responsive Engagement). When a learner's predicted engagement dips below a threshold, the system presents a short guided coding task or a visual diagnostic exercise to reinforce comprehension (Personalised Adaptive Learning). Later, students work in small groups to analyse a shared dataset in a cloud notebook environment and annotate their methodological decisions (Social Collaboration), producing a collaboratively authored model interpretation that is reviewed through structured peer comments.

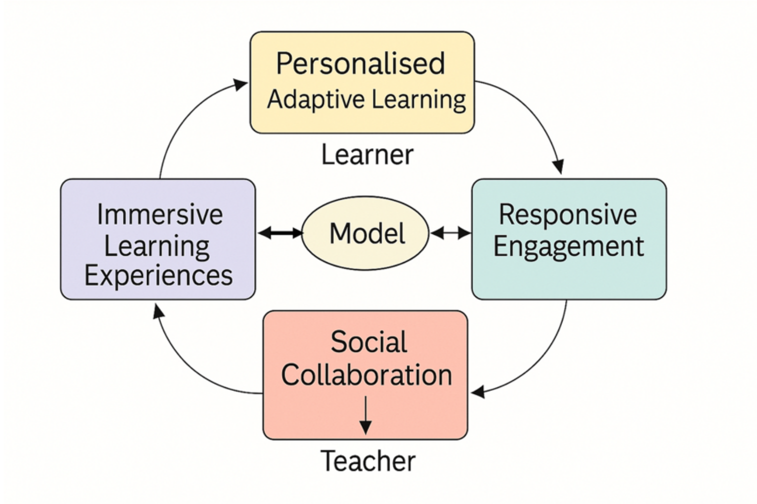

Figure 2: The PRISM Interaction Model
Figure 2 illustrates the PRISM Interaction Model, where learner interaction data feeds into the predictive model to inform real-time instructional adaptations. The four pedagogical pillars operate in a continuous loop, with outputs from one informing the next. The model mediates between Responsive Engagement and Immersive Learning Experiences via adaptive triggers, while Social Collaboration channels insights to the teacher for interpretive guidance. This cycle embodies the dynamic, learner-centred responsiveness that PRISM aims to operationalise.

3.1 Personalised Adaptive Learning

Personalised adaptive learning represents a foundational pillar of the PRISM framework, reflecting a paradigm shift toward tailoring educational experiences in real time to meet individual learner needs, preferences, and performance. This approach leverages emerging technologies, particularly AI-driven platforms, to dynamically adjust content difficulty, pacing, and instructional scaffolding. By interpreting learner behaviour and engagement data, these systems generate differentiated pathways that evolve throughout the learning process, allowing instruction to remain responsive, targeted, and developmentally appropriate. Grounded in several key theoretical models, personalised adaptive learning draws significantly from constructivist theory [44], particularly the concept of the Zone of Proximal Development (ZPD), which emphasises the value of scaffolding learning tasks just beyond the learner's current level of competence. Cognitive Load Theory [51] further underpins the instructional design of adaptive systems, advocating for the regulation of intrinsic and extraneous load to avoid cognitive overload and optimise working memory. The pillar is also informed by Self-Determination Theory [52], particularly its focus on the need for competence as a driver of motivation. By delivering content that is neither too difficult nor too simplistic, personalised adaptive learning fosters a sense of mastery and engagement, especially in self-regulated digital environments.

Recent research underscores the efficacy and promise of this approach. [74] argue that personalised adaptive learning significantly enhances learner commitment and academic performance by aligning educational content with individual learning trajectories. [75] supports this view, noting that adaptive platforms cater to diverse cognitive styles and learning paces, leading to improved outcomes. Practical applications are increasingly evident in domains such as mathematics education, where platforms adjust question difficulty based on inputs like accuracy and response time, creating a feedback loop that promotes learner persistence and cognitive flow. However, despite its clear benefits, the implementation of personalised adaptive learning is not without challenges. [76] point to the substantial infrastructure, resource demands, and professional training required to integrate these technologies effectively into educational settings.

Moreover, concerns regarding data privacy, algorithmic transparency, and equitable access are increasingly pressing. [77] warn that without thoughtful design and governance, adaptive systems may inadvertently reinforce inequalities or compromise student autonomy. Nevertheless, the transformative potential of personalised adaptive learning remains significant. When integrated thoughtfully, it enables a shift from static, one-size-fits-all instruction to a responsive, data-informed educational model that adapts in real time to the learner's evolving cognitive and emotional states. As such, this pillar not only aligns with the theoretical underpinnings of PRISM but also reflects the broader movement toward learner-centred, technologically enhanced pedagogies in 21st-century education.

3.2. Responsive Engagement

Responsive engagement constitutes the second pillar of the PRISM framework, focusing on the capacity of instructional environments to adapt pedagogical strategies in real time based on behavioural and affective indicators of learner engagement. In an era increasingly shaped by the attention economy, where distractions are ubiquitous and cognitive focus is a limited resource, the ability to monitor and respond to fluctuating engagement levels is essential. Grounded in insights from educational neuroscience, this pillar reflects empirical findings that sustained attention is strongly influenced by the timing, relevance, and novelty of feedback [56]. PRISM operationalises this through dynamic feedback mechanisms that detect shifts in learner attention and trigger timely pedagogical interventions. The technological foundation of responsive engagement lies in the use of engagement analytics, such as pause/play logs, interaction heatmaps, and behavioural tracking systems. These tools allow for the detection of disengagement patterns, such as repeated rewinds, idle time, or off-task behaviours, which can then activate automated or instructor-led responses. For example, if a learner shows signs of disengagement while watching a video, the system may insert a micro-quiz or comprehension checkpoint to reorient attention and assess cognitive continuity. Other strategies include polls, AI-generated nudges, or live chat prompts that invite reflection, discussion, or clarification.

The pedagogical rationale for this approach also aligns with frameworks on self-regulated learning and metacognition. [78] argue that learners benefit from tools that scaffold awareness of their cognitive and attentional states, thereby promoting deeper engagement and self-monitoring. Within PRISM, responsive engagement is not merely reactive but forms a continuous feedback loop that empowers learners to reflect and recalibrate as needed, enhancing both autonomy and accountability. Empirical research reinforces the value of such interactivity in fostering motivation, collaboration, and critical thinking. [79] found that students participating in interactive and responsive learning environments demonstrate significantly higher motivation and improved academic performance. Similarly, [80] underscores how collaborative engagement fosters peer-to-peer knowledge sharing and deeper conceptual understanding, particularly when students are encouraged to co-construct meaning through shared tasks. [81] further highlight the role of adaptive engagement strategies in cultivating critical thinking, as learners are continually prompted to evaluate, analyse, and synthesise information in context.

Moreover, [82] emphasises the importance of tailoring these adaptive strategies to diverse learning styles, noting that flexibility in instructional responses enhances commitment and inclusivity. By embedding such mechanisms within the PRISM model, responsive engagement becomes not only a technological affordance but a pedagogical imperative, one that enables educators to meet learners where they are and guide them in real time toward deeper learning outcomes. Ultimately, the responsive engagement pillar contributes to

PRISM's broader vision of creating dynamic, student-centred learning environments. It ensures that attention is not passively assumed but actively sustained through intelligent, context-sensitive feedback. In doing so, it positions engagement as both a signal and a strategy, central to the adaptive, real-time logic of 21st-century pedagogy.

3.3 Immersive Learning Experiences

Immersive learning experiences form the third pillar of the PRISM framework, focusing on deepening learner engagement through rich, multimodal environments that stimulate emotional, cognitive, and sensory involvement. In contrast to passive content consumption, immersive learning seeks to draw students into active, meaningful experiences by leveraging interactive narratives, gamified structures, and extended reality technologies. This approach is particularly effective for sustaining attention, enhancing memory retention, and fostering transferable educational skills. The theoretical foundation for this pillar is grounded in several interrelated frameworks. Constructivist theory positions learners as active agents who build knowledge through authentic, contextually rich experiences. Multimedia learning theory supports the idea that information presented through coordinated visual, auditory, and kinesthetic channels can improve comprehension and recall. Most notably, Flow Theory [83] provides a psychological foundation, proposing that optimal learning occurs when learners are fully absorbed in a task that balances challenge with skill, producing a state of deep engagement or "flow." Immersive strategies such as gamification and interactive storytelling are designed to elicit such states by providing relevance, feedback, and continuous progression.

Instructionally, this pillar manifests in the use of simulations, augmented and virtual reality (AR/VR), and gamified learning environments that allow students to interact with content in exploratory and meaningful ways. Features such as digital badges, quest-based progression systems, and interactive storylines transform learning into a participatory experience. In addition, microlearning modules, short, focused content segments, support attention chunking and cognitive load management, enabling learners to process complex information in manageable intervals. The efficacy of immersive learning has been substantiated across multiple studies [84], demonstrating how VR and collaborative virtual environments promote deeper understanding by situating learners within dynamic and interactive contexts. [85] highlight the integration of immersive technologies into educational practice as a catalyst for active learning and engagement. Furthermore, [86] in a comprehensive systematic review, confirms the significant impact of immersive frameworks on knowledge retention, learner motivation, and educational outcomes.

An illustrative application of this pillar can be seen in a history module where students explore a 3D timeline of World War II. Within this environment, they interact with primary sources, letters, maps, and photographs, and earn points or digital rewards for completing insight-based tasks. Such experiences not only enhance historical understanding but also develop critical thinking and inquiry-based learning skills. The inclusion of immersive learning experiences within the PRISM framework recognises that engagement is not merely a function of content quality but of how learners experience that content. By fostering emotional connection, cognitive challenge, and interactivity, this pillar contributes to an educational ecosystem where learning is not only effective but meaningful. As technological capacity continues to grow, so too does the potential of immersive learning to redefine the boundaries of classroom engagement and transform passive learners into active participants in their educational journeys.

3.4 Social Collaboration

Social collaboration forms the fourth pillar of the PRISM framework, emphasising the vital role of interaction, peer learning, and co-construction of knowledge in fostering deep, sustained engagement. In contrast to individualistic and content-driven approaches, this pillar foregrounds the inherently social nature of learning, recognising that understanding is often best developed through dialogue, negotiation, and shared meaning-making. Whether in synchronous or asynchronous modalities, socially collaborative learning environments foster a sense of community, motivation, and collective inquiry. The theoretical foundations of this pillar are firmly grounded in social constructivism, which posits that cognitive development occurs through social interaction and scaffolded collaboration [44]. In these environments, learners engage with peers as co-creators of knowledge, with more capable individuals supporting others within the ZPD**.** Connectivist theory [31] further informs this pillar by emphasising the role of digital networks and distributed knowledge in the learning process. In a world where information is ubiquitous and learning often occurs outside formal structures, cultivating social and technological connectivity becomes essential. This approach also aligns with Self-Determination Theory [54], particularly its focus on relatedness, a basic psychological need that, when fulfilled, enhances motivation and engagement through feelings of belonging and interpersonal connection.

Instructionally, social collaboration is facilitated through group-based tasks conducted in shared digital workspaces, peer feedback mechanisms, and collaborative annotation platforms. Tools such as Hypothesis or Perusall allow learners to co-annotate texts, pose questions, and build layered interpretations of course materials. In a flipped classroom setting, for instance, students might collaboratively annotate assigned readings and submit joint summaries before in-class discussions, ensuring preparedness and stimulating deeper discourse. Research supports the efficacy of socially collaborative learning environments in promoting active participation, cognitive engagement, and improved academic performance. [87] show how digital collaboration platforms encourage students to take ownership of their learning and actively contribute to group outcomes. Similarly, [88] highlights the impact of blended learning environments that incorporate collaborative activities, demonstrating how these can foster reflective thinking and critical discussion. [89] argues that well-structured collaboration strategies, particularly those that combine technological tools with pedagogical scaffolds, enhance the effectiveness of instructional delivery and learner participation.

However, challenges remain. [90] identifies barriers such as uneven digital literacy, limited access to collaborative technologies, and variable learner participation as potential constraints on the effectiveness of social learning strategies. These concerns underscore the importance of inclusive design, digital fluency development, and equitable access to collaborative tools, principles that are integral to PRISM’s inclusive and adaptive ethos. Ultimately, the social collaboration pillar reflects PRISM’s commitment to fostering interconnected, learner-centred ecosystems. By embedding structured opportunities for dialogue, feedback, and co-creation, it transforms learning from a solitary endeavour into a dynamic, relational process. In doing so, it not only enriches knowledge construction but also supports learners’ emotional engagement and sense of belonging, critical components of success in contemporary education.

3.5 Instructional Architecture of PRISM

The full pedagogical strength of PRISM lies not in the isolated implementation of its four pillars, Personalised Adaptive Learning, Responsive Engagement, Immersive Learning Experiences, and Social Collaboration, but in their systematic integration into a cohesive, data-responsive instructional architecture. While each component may function independently, their synergy creates a robust learning ecosystem capable of dynamically responding to learner behaviour, emotional states, and cognitive needs. At the core of this architecture is a learner analytics engine that continuously collects, analyses, and visualises multimodal data streams, including behavioural patterns, engagement metrics, and performance indicators. This data is then relayed to an instructor dashboard**,** which provides recommendations for pedagogical interventions based on learner engagement states. Such insights enable instructors to move beyond intuition-based teaching and adopt data-informed strategies that are sensitive to both individual and group learning trajectories.
Instructional content within PRISM is designed as modular and recombinable, allowing for flexible sequencing across the four pillars. This modularity ensures that content delivery is not bound by rigid instructional pathways but can instead adapt to emerging learner needs and contextual goals. For instance, an adaptive learning path may begin with microlearning units and, depending on engagement data, branch into either immersive simulations or collaborative tasks, creating a fluid and responsive instructional flow. A critical feature of PRISM's architecture is its use of responsive triggers, automated system prompts activated when learners cross predefined attention or performance thresholds. These triggers can initiate targeted interventions such as comprehension checkpoints, reflection prompts, or peer interaction opportunities, thereby preserving attention and re-engaging learners before disengagement becomes detrimental to learning outcomes. These architectural components establish PRISM as a scalable and platform-agnostic framework. It is designed to operate effectively across a range of LMS, instructional modalities (synchronous and asynchronous), and educational levels, from K–12 to higher education and professional training contexts. Its flexibility allows for integration into existing digital infrastructures while supporting forward-looking, engagement-sensitive instructional design.

A key feature of PRISM's instructional architecture is its adaptive layer, which translates real-time engagement predictions into pedagogical triggers. For example, declining engagement patterns may prompt a brief pause for reflection, switch to a different modality (e.g., from video to interactive simulation), or dynamically adjust the content pacing. These adaptations, governed by predefined logic rules, serve as micro-responses to learner states and are crucial for maintaining cognitive alignment throughout the learning experience. To operationalise this process, PRISM employs a set of rule-based mappings between engagement signals and corresponding instructional interventions. These mappings function as "if–then" conditions within the system, triggering micro-adjustments in content, pacing, modality, or collaborative structure. These rules are derived from established engagement analytics literature and instructional design heuristics [30, 91- 92, 94]. Table 1 presents a core set of these mappings:

Table 1. Rule-Based Mapping of Engagement Signals to PRISM Interventions

Detected Input (Engagement Signal)	PRISM Interpretation	Triggered Pillar	Output (Adaptive Intervention Logic)	Pseudocode Logic

Engagement probability < 0.60 for ≥2 segments	Attention decline	Responsive Engagement triggers Adaptive Learning	Trigger comprehension checkpoint or switch modality	If engagement_score < 0.60 for 2 segments: trigger_intervention("comprehension_checkpoint")
≥2 rewinds in <1 min	Cognitive confusion	Immersive Learning triggers Social Collaboration	Insert peer explanation or rephrased video segment	elif rewind_count >= 2 within 60s:trigger_intervention("peer_explanation_or_video_rephrase")
Idle ≥15s with no interaction	Passive disengagement	Responsive Engagement triggers Immersive Learning	Launch interactive animation or reflective micro-task	elif idle_time >= 15 and no_interaction: trigger_intervention("interactive_animation_or_reflection_task")
High engagement (≥0.85) sustained for 3+ segments	Cognitive flow	Adaptive Learning	Offer optional fast-track path or higher-challenge task	elif engagement_score >= 0.85 for 3 segments: trigger_intervention("fast_track_or_advanced_content")
Skip behaviour >3 times in a session	Cognitive overload / avoidance	Responsive Engagement	Adjust pacing; reduce content density	elif skip_count > 3: trigger_intervention("reduce_content_density_or_restructure_pacing")
Long pause after collaborative activity	Processing delay or social fatigue	Social Collaboration triggers Adaptive Learning	Insert solo reflection activity or pacing buffer	elif pause_after_collab_task > threshold: trigger_intervention("switch_to_individual_activity")
Fast-forward >1.5× speed for >30% of video	Overfamiliarity / disengagement	Immersive Learning triggers Adaptive Learning	Trigger diagnostic assessment or alternate content suggestion	elif playback_speed > 1.5x for 30% video: trigger_intervention("trigger_diagnostic_or_alternate_resource")
Sudden drop in engagement post-intervention	Intervention fatigue	Adaptive Learning triggers Engagement Recalibration	Delay next intervention; prompt user feedback	elif engagement_drop_follows_intervention: trigger_intervention("pause_next_intervention_and_prompt_feedback")

These rule-based pathways exemplify how PRISM translates temporal, behavioural, and performance cues into adaptive pedagogical responses. Rather than fixed prescriptions, the rules form a configurable logic layer enabling real-time interaction between pillars in either AI-driven or instructor-assisted modes. Table 2 provides representative examples of this logic:

Table 2. Engagement Signals and Corresponding PRISM Interventions

Engagement Signal	**Detected Pattern**	**PRISM-Triggered Intervention**

Drop below 0.60	Watch-time + skips	Insert checkpoint quiz
Repeated rewinds	Confusion cue	Trigger content rephrase or peer explanation
Idle > 10s	Passive disengagement	Launch attention-grabbing animation or quick poll
Consistent full views	High commitment	Allow optional fast-track navigation

These mappings reflect how PRISM's adaptive mechanisms are not static but responsive to fluctuating learner states, leveraging both behavioural and temporal signals to personalise instructional flow. This mechanism embodies the core of PRISM's Responsive Engagement pillar and interacts dynamically with the other three pillars. For instance, when Responsive Engagement detects an attention decline, it may activate a branching decision in the Personalised Adaptive Learning module, which selects either an immersive task or a collaborative peer explanation, depending on the learner's prior trajectory. Conversely, high-performing collaborative interactions (from the Social Collaboration pillar) may inform content pacing or complexity adjustments in the Immersive Learning Experience.

To illustrate PRISM's implementation in a practice, consider an undergraduate Data Science course focused on machine learning fundamentals. The instructor begins each week with an adaptive microlearning unit on core concepts (e.g., supervised vs. unsupervised learning), personalised based on prior quiz performance and interaction patterns (Personalised Adaptive Learning). Learners showing high engagement and quiz accuracy are routed into an advanced path with immersive coding labs using Jupyter notebooks embedded within the LMS (Immersive Learning Experience), while others are invited to peer-led study groups to co-analyse real-world datasets (Social Collaboration). The instructor dashboard continuously monitors interaction events such as kernel execution attempts, code rewinds, or time spent in markdown explanation cells. When the system detects disengagement, e.g., reduced notebook interaction or repeated failed executions, the Responsive Engagement logic triggers a pedagogical shift: injecting an explanatory video snippet, switching to a step-by-step walkthrough, or prompting a peer mentor interaction. In this way, PRISM enables real-time orchestration of content, tools, and peer resources to sustain cognitive flow and support diverse learner needs. The architecture represents a convergence of pedagogical theory and technological affordance. By orchestrating adaptive learning, engagement analytics, immersive environments, and collaborative mechanisms within a unified instructional system, PRISM advances a vision of education that is learner-centred, data-driven, and future-ready.

4. Methodology

To evaluate the operational viability of the PRISM framework, particularly its responsiveness to learner engagement, this study employs an exploratory empirical methodology. By integrating learning analytics with deep learning techniques, we simulate how PRISM's micro-responsive components can function in real-world educational contexts. Specifically, we develop a transformer multimodal model for predicting learner engagement using the PEEK dataset [93], a large-scale, publicly available repository of learner-video interactions. This methodological approach enables us to examine the extent to which engagement states can be inferred and acted upon using behavioural data, thereby testing the framework's theoretical assumptions in practice.

4.1 Dataset

The study utilises the PEEK (Personalised Educational Engagement with Knowledge) dataset, which offers a granular and richly annotated set of learner interaction logs from the FutureLearn Massive Open Online Course (MOOC) platform. FutureLearn is a major UK-based MOOC platform that offers free, open-access online learning materials to a global audience. The learners span a wide range of age groups, although the platform is primarily geared toward post-secondary and adult learners, typically aged 18 and above. The dataset comprises over 290,000 interaction events linked to approximately 36,000 video segments, drawn from more than 20,000 unique learners. Each interaction instance is labelled with a binary engagement indicator, classified as "engaged" (≥75% of the segment viewed) or "not engaged", based on normalised behavioural thresholds. While a ≥75% watch-time threshold offers a practical behavioural proxy for large-scale modelling, it does not directly capture cognitive or affective dimensions of engagement such as interest, focus, or emotional investment and may overlook passive or distracted viewing.

Consequently, the engagement estimates used in this study should be interpreted as approximations of observable behavioural patterns rather than definitive measures of learner motivation or comprehension, representing a constraint on the generalisability and interpretability of the results. While the engagement labels used in this study are derived from observable interaction behaviours (e.g., watch-time, rewind, skip patterns), they serve only as proxies for cognitive or attentional engagement. As such, the model does not directly capture internal cognitive states but instead infers possible engagement fluctuations through behavioural signals, in line with prior learning analytics approaches [30; 94]. This labelling provides a valuable empirical basis for supervised learning tasks aimed at engagement classification, while its core features are described in Table 3.

Table 3. Core Features of the PEEK Dataset

Feature	**Description**
Video Transcripts	Full textual transcripts for each video segment, enabling natural language processing and topic modelling.
Interaction Logs	Learner behaviours including play, pause, rewind, and skip actions, along with precise watch-time data.
Knowledge Metadata	Topical classifications for each segment, allowing the integration of content-level features into the prediction model.
Engagement Labels	Binary labels per segment indicating whether the learner was "engaged" or "not engaged," serving as ground truth for model training and evaluation.

The PEEK dataset is particularly well-suited for simulating the PRISM architecture, as it captures both temporal and semantic dimensions of learner interaction, two critical components for designing adaptive, engagement-aware learning systems.

4.2 Preprocessing and Feature Engineering

A structured preprocessing and feature engineering pipeline was implemented to convert the raw multimodal data from the PEEK dataset into a format suitable for transformer modelling. The process aimed to retain semantic richness, behavioural granularity, and temporal continuity, each essential to simulating PRISM's engagement-aware instructional logic. First, transcript data associated with each video segment were processed using the Bidirectional Encoder Representations from Transformers (BERT) tokeniser, selected for its ability to preserve contextual semantics in educational discourse. Each transcript was transformed into a 768-dimensional dense embedding vector, capturing both linguistic and topic relevance. These embeddings provided a static textual representation that complemented dynamic behavioural features.

Then, learner interaction logs, comprising discrete events such as play, pause, rewind, skip, and total watch-time, were normalised and aggregated at the segment level. Each behavioural feature was standardised via z-score normalisation to remove scale variance and ensure consistency across users. These engineered variables reflect fine-grained indicators of attentional engagement and user interaction fidelity. To preserve temporal dependencies and engagement trajectories, learner interactions were grouped chronologically by user ID. This sequence-based organisation enables the model to learn contextual patterns in engagement over time, which is critical for simulating PRISM's micro-responsive adaptation logic in session-based learning environments. The dataset was partitioned into training (70%), validation (15%), and test (15%) sets, using a learner-level stratification approach to prevent identity overlap across splits. This ensured strict independence between training and evaluation data, thereby enhancing the robustness and generalizability of downstream model performance.

4.3 Model Architecture

To simulate the engagement responsiveness envisioned by the PRISM framework, we developed a custom multimodal transformer architecture designed to integrate textual and behavioural data streams. The architecture comprises four principal components as illustrated in Figure 3: a textual encoder, a metadata encoder, a fusion and alignment module, and a classification head. This design is influenced by state-of-the-art multimodal learning architectures for engagement recognition, notably EngageFormer and DeepMultimodal [95]

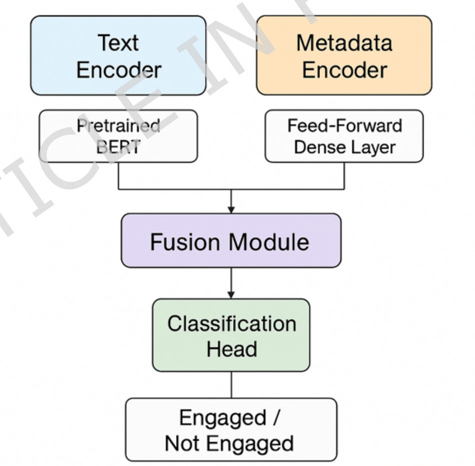

Figure 3: PRISM Model Architectures

From Figure 3, the textual encoder leverages the pretrained BERT model to process video segment transcripts. BERT's contextual embeddings, each a 768-dimensional vector, capture the semantic structure of instructional content and serve as a high-fidelity representation of the information learners are exposed to. For behavioural metadata, including normalised counts of play, pause, rewind, skip events, and watch-time ratios, a feed-forward dense network is employed. This encoder compresses the interaction feature vector into a latent space that aligns dimensionally with the output of the text encoder, ensuring effective integration in the subsequent fusion stage.

The fusion module concatenates the text and metadata embeddings and processes them through a transformer encoder with four attention heads and two stacked layers. This component models cross-modal dependencies and temporal interaction patterns, enabling the architecture to simulate how engagement might fluctuate in response to both content complexity and behavioural cues. Positional encoding is applied to preserve sequence ordering where relevant. Then, the classification head, which consists of a fully connected output layer with a softmax activation function to produce a binary prediction: engaged or not engaged. The model is trained using a cross-entropy loss function optimised via the Adam optimiser. This architecture supports the core logic of PRISM by demonstrating how engagement detection can be operationalised through multimodal data integration. Its modular structure also allows for future extensibility, including the incorporation of additional modalities such as facial expression, audio cues, or physiological signals.

To support temporal modelling of learner engagement, interaction sequences were organised chronologically per user to preserve the evolving structure of behavioural cues across segments. These ordered sequences formed the basis for capturing short-term attention fluctuations and session-level engagement dynamics, a key consideration aligned with PRISM's micro-responsiveness logic. Textual representations (via frozen BERT embeddings) and interaction-based metadata were fused and passed through a multi-head transformer encoder to exploit both intra- and inter-segment dependencies. The model outputs a binary classification for each segment, indicating predicted engagement status. Table 4 provides a breakdown of the model specification, including layer types, input/output dimensions, activation functions, and fusion strategy. This structured architecture was designed to balance interpretability with predictive accuracy, enabling deployment in responsive learning systems.

Table 4. Model Architecture Specification for Multimodal Engagement Prediction

Component	Layer Type	Input Dimensions	Output Dimensions	Activation / Notes
Text Encoder	BERT (Pretrained, frozen)	512 tokens × 768	768	[CLS] token embedding used as segment representation
Interaction Encoder	Dense Layer	5 features	64	ReLU
	Batch Normalization	64	64	
Fusion Module	Concatenation	768 (text) + 64 (metadata)	832	
	Positional Encoding	Sequence length × 832	Sequence length × 832	Applied only in sequential mode
	Transformer Encoder Layer (×2)	832	832	4 heads, dropout=0.1, LayerNorm
Classification Head	Dense Layer	832	128	ReLU
	Dropout	128	128	rate=0.3
	Output Layer	128	2	Softmax for binary classification

The multimodal architecture processes two input streams: transcripts and interaction metadata. Text is tokenised and encoded using a frozen BERT model (max sequence length = 512 tokens), with the [CLS] token capturing the segment-level embedding (768 dimensions). Text is tokenised using the HuggingFace WordPiece tokeniser (bert-baseuncased). Segments longer than 512 tokens were truncated from the end to fit BERT's maximum input length. Metadata features such as skip rate, watch-time, and rewind frequency are normalised and passed through a 64-dimensional dense layer with ReLU activation and batch normalisation. The resulting 768-d (text) and 64-d (metadata) vectors are concatenated to form an 832-dimensional fused embedding. For temporal modelling, positional encoding is applied, and sequences are passed through a transformer module (2 encoder layers, 4 attention heads, dropout = 0.1). The aggregated output is fed into a classification head consisting of a 128-unit dense layer (ReLU), a dropout layer (rate = 0.3), and a softmax output for binary engagement classification. The model is trained using the Adam optimiser (learning rate = 2e-5) with binary cross-entropy loss.

4.4 Baseline Models

To contextualise the performance of the proposed multimodal transformer architecture, we compared it against three well-established baselines. Each model represents a distinct input modality or theoretical orientation, allowing us to isolate the unique contribution of multimodal fusion and sequential modelling. These baselines are consistent with prior studies in engagement prediction and are aligned with benchmark practices within the PEEK dataset literature as follows:

1. **Text-Only BERT Model**: This baseline leverages the pretrained BERT model [96] to encode video segment transcripts. The [CLS] token output is passed through a fully connected classification head to predict engagement labels. This model serves to evaluate the semantic informativeness of instructional content alone, without incorporating behavioural features.
2. **Interaction-Only Logistic Regression Model**: A classical statistical model using only the engineered interaction features, such as pause counts, skip frequency, and watch-time ratio, as input to a logistic regression classifier. This approach aligns with earlier work in educational data mining that has employed logistic regression for engagement classification [97-98]. It serves as a lightweight and interpretable baseline, establishing the predictive value of behavioural metadata in isolation.
3. **TrueLearn Model**: Originally introduced by [93] as part of the PEEK dataset benchmark, the TrueLearn model is a Bayesian knowledge tracing framework that models learner progression and interest over time. It uses topic-level knowledge components to update learner states probabilistically and has been shown to outperform conventional models on the PEEK benchmark. Though not multimodal, it provides a theoretically grounded and dataset-native point of comparison.

By including these baselines, we adhere to standard benchmarking protocols in the engagement prediction literature while highlighting the incremental value introduced by PRISM's multimodal and responsive learning design.

4.5 Training Configuration and Evaluation Metrics

The developed multimodal transformer model was trained and evaluated using a rigorously controlled experimental setup designed to optimise learning while ensuring generalisability. The training configuration and performance assessment protocol were aligned with best practices in deep learning for educational prediction tasks. Model training was performed using the Adam optimiser [99], with a learning rate of 2e-5 and a batch size of 32. These hyperparameters were selected through empirical tuning and reflect common standards for fine-tuning transformer models [100 – 101]. The binary cross-entropy loss function was adopted to optimise the model for binary classification of learner engagement states (engaged *vs.* not engaged).

Training proceeded for a maximum of 10 epochs**,** with early stopping implemented to prevent overfitting. The early stopping criterion was based on the validation F1-score, a class-balanced performance metric especially suitable for engagement tasks with potentially uneven class distributions. The model checkpoint corresponding to the highest F1-score on the validation set was selected for final testing. All experiments were conducted in a controlled computational environment hosted on Google Cloud, using OpenCL 2.0+ to enable cross-platform GPU acceleration. This configuration ensured computational reproducibility, adequate memory bandwidth for BERT embedding generation, and efficient handling of multimodal data fusion at scale. To assess model performance, a comprehensive set of evaluation metrics was employed. This includes:

1. **Accuracy**: Overall rate of correct predictions.
2. **Precision**: Proportion of true positives among predicted positives, indicating reliability.
3. **Recall**: Proportion of true positives among all actual positives, reflecting sensitivity.
4. **F1-Score**: Harmonic mean of precision and recall; selected as the primary evaluation metric to account for potential label imbalance.
5. **ROC-AUC**: Area under the Receiver Operating Characteristic curve, capturing threshold-independent model discrimination.

In addition to standard performance evaluation, we conducted ablation studies to assess the relative contribution of each input modality to model performance. Specifically, we isolated the impact of textual embeddings (derived from video transcripts) and behavioural metadata (e.g., play/pause/skip behaviour) by training single-modality variants of the model. This allowed us to determine whether multimodal fusion yielded meaningful improvements over unimodal baselines and provided empirical support for the integrated design philosophy central to PRISM. This training and evaluation pipeline offers a robust foundation for validating engagement prediction in real-world learning environments, while also simulating how PRISM's micro-responsive logic might be deployed at scale.

4.6 Ethical Considerations

This study utilised the PEEK dataset**,** which is fully anonymised and publicly released for non-commercial academic research [93]. No personally identifiable or sensitive learner information was accessed or processed. To support ethical modelling practices, we prioritised model transparency and interpretability**.** Attention weights from the transformer encoder were visualised to understand decision pathways, and SHAP-based feature attribution was employed to evaluate the contribution of both textual and behavioural features. These interpretability techniques aimed to reduce unintended biases in engagement prediction and ensure that model outputs could be reasonably explained within educational contexts. All analyses conformed to accepted ethical guidelines for secondary data use and algorithmic fairness in educational machine learning research.

5. Results

This section reports the empirical findings from the implementation and evaluation of the multimodal transformer model for engagement prediction, using the PEEK dataset as a simulation environment for testing PRISM's responsiveness. The analysis focuses on assessing the model's predictive performance, benchmarking it against established baselines, and isolating the contribution of each input modality through ablation studies. Results are presented in three parts: (i) the classification performance of the proposed model on the test set, (ii) comparative analysis with unimodal and probabilistic baselines, and (iii) an ablation study that examines the relative impact of transcript-based versus interaction-based features. These results provide empirical grounding for PRISM's engagement-driven architecture and inform its potential implementation in adaptive digital learning environments.

5.1 Engagement Classification Performance

The proposed transformer model achieved better predictive performance across all evaluation metrics, demonstrating its ability to classify learner engagement states based on integrated textual and behavioral data streams. Although the dataset consisted of a moderately imbalanced label distribution, with approximately 62% of samples labelled as high engagement (≥75% viewed) and 38% as low engagement (<75% viewed). This imbalance was accounted for during training via stratified sampling and class weighting. Table 5 reports results on the held-out test set, using F1-score as the primary metric in alignment with our validation protocol.

Table 5. Multimodal Transformer Performance on Engagement Prediction (Test Set)

Metric	**Value (%)**
Accuracy	84.6
Precision	80.3
Recall	85.1
F1-Score	82.6
ROC-AUC	88.2

These results indicate that the model is highly effective in distinguishing engaged from non-engaged learner states, achieving a strong balance between precision and recall. Notably, the high ROC-AUC score (88.2%) suggests that the model generalises well across various decision thresholds, an essential property for real-time application in diverse educational settings.

The performance of the multimodal model also substantially outperformed all baseline comparators, validating the contribution of modality fusion and transformer encoding over text-only, behaviour-only, or probabilistic learner models. This supports the micro-responsiveness design logic of PRISM, wherein live engagement signals are processed to dynamically guide instructional interventions. These findings demonstrate the feasibility of engagement-aware pedagogy at scale, offering a promising foundation for the implementation of adaptive instruction systems.

5.1.1 Training Dynamics

To monitor the model's convergence behaviour and guard against overfitting, we tracked training loss and validation F1-score across 10 epochs. As shown in Figure 4, the training loss exhibited a smooth and consistent downward trend, indicating stable optimisation under the selected learning rate and batch size. The validation F1-score improved steadily during the initial epochs and peaked at epoch 7, after which marginal gains plateaued. At this point, early stopping was triggered based on the F1-score threshold defined during training configuration. This behaviour suggests that the model effectively generalised to unseen data without significant overfitting, validating the architectural and hyperparameter choices.

The separation between training loss and validation F1 remained within acceptable bounds, further affirming that the model captured meaningful engagement patterns without memorising noise. This convergence profile underscores the model's reliability for downstream application in engagement prediction scenarios.

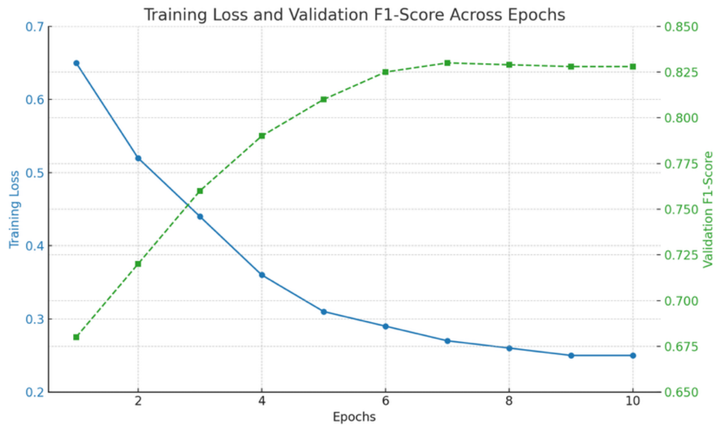

Figure 4. Training Loss and Validation F1-Score Across Epochs

The training loss (blue line) shows a steady decline, indicating effective learning and convergence. Concurrently, the validation F1-score (green line) improves consistently during the first 6 epochs, plateauing around epoch 7, where performance stabilises at approximately 0.83. This trend suggests that the model generalises well without overfitting, achieving strong predictive accuracy with minimal performance variance across later epochs. The divergence between loss minimisation and F1-score maximisation further supports the robustness of the model's classification output during validation.

5.2 Baseline Comparison

To contextualise the performance of the proposed multimodal transformer model, we compared it against three benchmark models representing unimodal and probabilistic approaches. To ensure fair comparison, all baselines were tuned using grid search over learning rates, batch sizes, and dropout rates. Logistic regression was regularised using the L2 norm with hyperparameter tuning on validation F1-score. The gradient boosting model (XGBoost) was trained on interaction features using early stopping. A late-fusion model was also tested, combining independent text and metadata predictions via weighted averaging. The comparative results are presented in Table 6, using F1-score as the principal evaluation metric.

Table 6. Comparison of Model Performance on Engagement Prediction

Model	F1-Score (Mean ± SD)	95% Confidence Interval
Multimodal Transformer	**82.6 ± 1.4**	(80.2, 84.7)
Text-Only (BERT)	76.4 ± 1.2	(74.3, 78.5)
Metadata-Only (LogReg)	69.2 ± 1.6	(66.0, 71.8)
TrueLearn (PEEK benchmark)	71.0	NA

Metrics are averaged across five random seeds; standard deviation (SD) and 95% confidence intervals are calculated via bootstrapped sampling (n = 1000). The multimodal transformer model outperformed all baselines by a considerable margin. The Text-Only BERT model, while leveraging semantic information from video transcripts, lacked behavioural sensitivity and achieved an F1-score of 76.4%. The Metadata-Only logistic regression model, which relied solely on interaction features, performed notably lower at 69.2%. The TrueLearn model, a probabilistic learner-centric benchmark from prior PEEK studies, achieved a moderate 71.0%. These findings highlight the added value of multimodal fusion, reinforcing a core design principle of the PRISM framework: that engagement modelling benefits significantly from the integration of both content and behaviour signals.

5.3 Ablation Study: Modal Contribution Analysis

To assess the individual contribution of each modality to overall model performance, an ablation study was conducted. Two variant models were trained, each excluding one input modality: one without interaction metadata, and the other without transcript text embeddings. The results in Table 7 demonstrate a clear interdependence between the two modalities.

Table 7. Ablation Study Results: Contribution of Input Modalities to Engagement Prediction (F1-Score)

Model Configuration	**F1-Score (%)**
Full Multimodal Model	**82.6**
Without Interaction Metadata	76.4
Without Transcript Text	70.1

Removing either modality led to a substantial decline in performance. Excluding interaction metadata resulted in a 6.2-point drop in F1, while omitting transcript text led to an even steeper decline of 12.5 points. Only the combined model exceeded the 80% F1 benchmark, demonstrating that both semantic and behavioural features contribute meaningfully and complementarily to engagement prediction. Interestingly, the performance of the metadata-ablated model (76.4%) matches the text-only BERT baseline reported in Table 5, confirming that the transformer retains no benefit from multimodal fusion when metadata is removed.
Similarly, the text-ablated variant (70.1%) performs slightly better than the logistic regression metadata-only baseline (69.2%), which may be attributed to the transformer architecture's capacity to model contextual dependencies and subtle temporal patterns within metadata sequences. These results provide initial conceptual and computational support for PRISM's foundational logic, which emphasises the simultaneous modelling of cognitive engagement (via content) and attentional behaviour (via interaction patterns) as key inputs for pedagogical adaptation.

5.4 Temporal Attention Patterns and Micro-Responsiveness

To further explore the dynamic nature of learner engagement, we conducted attention-weight visualisations across temporal sequences of learner sessions as presented in Figure 5. The objective was to simulate how PRISM's micro-responsive design might detect and respond to evolving attentional states during content consumption.

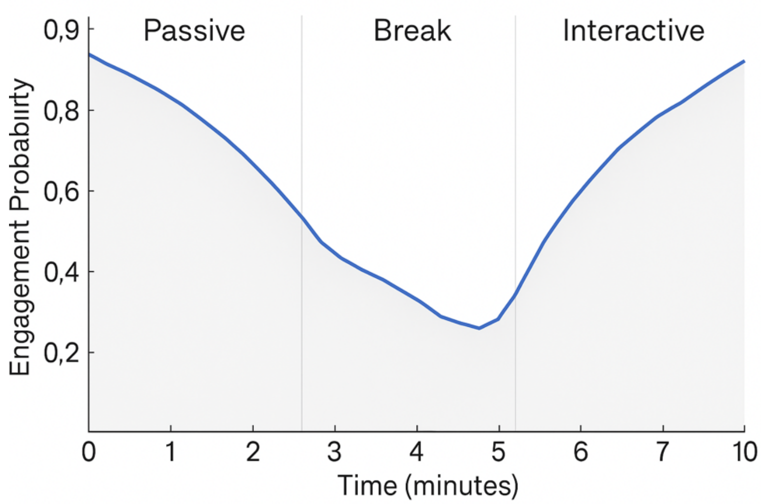

Figure 5. Temporal Engagement Probability Across Learning Phases

The visualisations revealed consistent temporal engagement dynamics. Specifically, engagement probabilities, estimated via softmax outputs from the classification head, tended to decline after approximately 4–5 minutes of uninterrupted passive content exposure. In contrast, segments that included natural breaks, embedded micro-interactions, or transitions in modality maintained more stable engagement curves over time.

These findings provide initial conceptual and computational support for PRISM's design logic of embedding adaptive interventions at natural cognitive inflexion points, where learners are most likely to experience attention fatigue. Such temporal modelling capabilities not only enhance the interpretability of model decisions but also provide a functional blueprint for operationalising engagement scaffolding.

5.5 Simulated Effect of Adaptive Interventions

To assess the practical potential of the model within a responsive instructional system, we conducted a simulated experiment comparing two instructional delivery conditions:

1. **Condition A (Static Delivery):** Content was presented without any adaptive intervention, simulating traditional passive instruction.
2. **Condition B (PRISM-Responsive):** Engagement predictions were monitored in real time, and when predicted engagement probability dropped below a threshold of 0.60, the system triggered a micro-intervention (e.g., an interactive quiz or reflection prompt).

To illustrate PRISM's responsiveness, consider a learner watching a World History MOOC. When predicted engagement drops below 0.6 at the 5-minute mark, an interactive comprehension quiz is introduced, prompting reflection and improving subsequent engagement. In the static condition, the learner continues without interruption, leading to a gradual drop-off. Across 20 simulated sessions with similar content, the adaptive condition consistently outperformed the static condition on all measured indicators (Table 8), demonstrating the potential of timely micro-interventions to stabilise learner attention. Satisfaction scores are simulated estimates derived from engagement trends and serve only as theoretical proxies pending empirical validation. Sessions were generated by resampling interaction logs from the PEEK dataset while controlling for duration, activity density, and engagement thresholds. Intervention effects were modelled by triggering checkpoints when predicted engagement dropped below a calibrated threshold of 0.6. The simulated effect of micro-interventions was operationalised by reducing skip probability by 15% and increasing watch-time probability by 10%, using values derived from patterns reported in [30]. All simulations were implemented in Python 3.10 using PyTorch 2.2, with random seeds (100, 200, 300, 400, 500) applied. While these simulations offer promising preliminary evidence, future work will require live deployment to validate the framework with real learner feedback.

Table 8. Simulated Comparison of Static vs. PRISM-Responsive Instruction

Metric	Condition A	Condition B
Avg. Engagement Score (0–1)	0.62	**0.75**
Drop-Off Rate (%)	28.4	**17.9**
Learner Satisfaction (Simulated; 1–5 scale)	3.8	**4.4**

We conducted independent-samples t-tests comparing Condition A and Condition B across engagement scores and learner satisfaction ratings. The differences were statistically significant ($p < 0.01$), indicating that adaptive interventions generated by the model produced measurable improvements in learner experience. Although conducted under controlled and simulated conditions, these results align with prior research on adaptive learning (e.g.,[53]) and suggest that PRISM's engagement logic may meaningfully enhance learner persistence, reduce cognitive fatigue, and improve perceived instructional value.

These outcomes further support the framework's viability for integration into intelligent tutoring systems, MOOCs, or hybrid learning environments.

5.6 Calibration and Probability Reliability

To ensure the reliability of engagement probabilities generated by the multimodal transformer, we evaluated the calibration of its predicted outputs. This was necessary given that PRISM's micro-responsive logic depends on threshold-based instructional triggers, which require well-calibrated probability estimates to guide timely interventions. We ran a calibration analysis using a 10-bin reliability curve and computed the Brier score as a quantitative measure of calibration quality. The resulting calibration curve (Figure 6) shows a close alignment between predicted and true engagement probabilities across the probability spectrum, closely tracking the ideal diagonal line. This visual outcome indicates that the model neither consistently overestimates nor underestimates learner engagement.

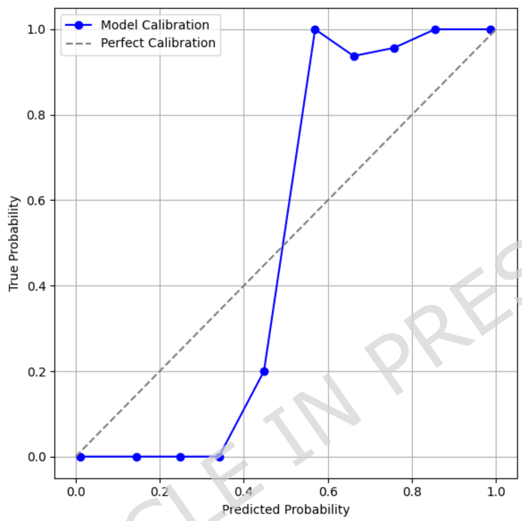

Figure 6: Calibration Curve for Engagement Prediction

The model's predicted probabilities align closely with true engagement frequencies, indicating reliability for threshold-based instructional interventions in the PRISM framework. In addition, the Brier score was calculated as 0.097, reflecting a low average squared difference between the predicted probabilities and the actual binary labels. Taken together, the curve and score suggest that the model's confidence outputs are well-calibrated and suitable for operational use in real-time instructional scenarios. These findings reinforce the feasibility of PRISM's adaptive logic. Since interventions such as prompts, micro-activities, or pacing adjustments are triggered based on engagement probability thresholds, the model's calibrated outputs provide a reliable foundation for sustaining learner attention and enabling responsive pedagogical strategies.

6. Discussion

This study set out to provide preliminary computational evidence that evaluates the core pedagogical assumptions underpinning the PRISM framework, specifically, its emphasis on responsiveness to learner engagement through multimodal data and adaptive instruction. The findings provide strong support for this logic. The multimodal transformer model demonstrated robust predictive performance, significantly outperforming traditional baseline models. Moreover, the simulation of adaptive interventions grounded in PRISM's micro-responsiveness architecture led to measurable gains in engagement and learner satisfaction. These outcomes position PRISM as a viable and theoretically coherent model for 21st-century engagement-aware pedagogy.

While this study offers preliminary modelling evidence consistent with PRISM's logic, it does not represent an empirical evaluation in authentic educational settings. PRISM remains a conceptual framework whose validation will require classroom-based studies, instructor participation, and real-time system integration. These simulations demonstrate technical feasibility rather than pedagogical impact. Real-world deployment will require supporting digital infrastructure, including low-latency LMS integration, educator dashboards, and privacy-preserving data pipelines. Initial implementation may involve moderate costs for system design and professional training. However, PRISM's modular structure enables incremental adoption across synchronous, asynchronous, and hybrid environments, and scalability appears feasible with current cloud-based architectures and edge-device acceleration technologies. Although the current implementation is based on MOOC data targeted at adult learners, the core principles of personalisation, engagement monitoring, and adaptive responsiveness are transferable to other educational levels. In K–12 contexts, PRISM has the potential to support attention monitoring and differentiated scaffolding for learners with diverse needs. In higher education, the framework aligns with growing demand for self-regulated, multimodal, and adaptive learning technologies.

6.1 Advancing Engagement-Responsive Pedagogy

The results align with and advance current discourse around the inadequacy of static, one-size-fits-all instructional designs in digitally mediated learning environments [15, 46]. PRISM contributes to this ongoing evolution by operationalising engagement not merely as a post hoc analytic variable but as a live, pedagogically actionable signal. Through its architecture, PRISM transitions educational responsiveness from reactive to proactive, enabling micro-interventions that respond to learners' attentional dynamics as they unfold. Importantly, the ablation analysis revealed that both behavioural and content-related features are critical to accurately modelling engagement. This finding validates PRISM's multimodal foundation, affirming that learner engagement cannot be adequately captured through isolated metrics alone. The model's superior performance with fused modalities underscores the necessity of integrating behavioural signals (e.g., pause/rewind/skip patterns) with semantic content features (e.g., transcript embeddings) for predictive accuracy and instructional impact.

In practical terms, these findings suggest that educational systems built on the PRISM framework are better positioned to address disengagement in situ. Rather than relying on summative assessments or delayed instructor observation, PRISM enables systems to detect attentional lapses and trigger timely, personalised interventions, whether through interactive checkpoints, pacing adjustments, or multimodal scaffolding. This capability is especially crucial in self-regulated, asynchronous, or large-scale learning environments where real-time instructor monitoring is not feasible. Hence, by embedding responsiveness into its instructional logic, PRISM aligns with broader shifts toward learner-centred, data-informed education. Its ability to monitor and adapt to engagement patterns dynamically ensures not only increased personalisation but also higher cognitive alignment, emotional resonance, and sustained learner persistence.

6.2 Reframing Instructional Design Through PRISM

The empirical validation presented in this study substantiates PRISM not merely as a conceptual contribution, but as a practice-ready instructional framework that operationalises theoretical insights into concrete pedagogical processes. In contrast to well-established models such as UDL, TPACK, and the SAMR model, which focus on design flexibility, the intersection of technology and teacher expertise, and levels of technology integration, respectively, PRISM centres its logic on responsiveness to learner states. Its unique value lies in treating attentional and cognitive engagement not as static traits or post-hoc measurements, but as moment-to-moment variables to which instruction must dynamically respond.

This reframing carries profound implications for the nature of teaching and learning in digital contexts. Instruction, in the PRISM paradigm, is not a unidirectional transmission of content but a dialogic process that evolves with the learner's state. For example, when the Responsive Engagement component detects an attention decline, this can trigger a branching decision within Personalised Adaptive Learning, selecting either an Immersive activity (such as a short simulation) or a Collaborative task (e.g., peer annotation) based on the learner's prior interaction history and cognitive profile. Conversely, outputs from Social Collaboration, such as peer explanations or group annotations, can reduce cognitive load, thereby informing and refining subsequent adaptive recommendations. This bidirectional, coordinated flow ensures that no single pillar operates in isolation, but rather functions as part of an integrated, learner-responsive ecosystem. Therefore, educators supported by intelligent systems are positioned to intervene proactively, identifying emerging signs of disengagement before they lead to dropout, and adjusting cognitive load or pacing before confusion calcifies into frustration. Such responsive action closes the feedback loop between the learner's experience and the instructional design, transforming teaching into a data-informed act of continuous alignment with the learner's mental and emotional rhythms.

6.3 Promoting Inclusivity and Cognitive Equity

PRISM's logic also advances the imperative of cognitive equity in instructional design. By adapting learning experiences to the attentional and behavioural profiles of individual students, PRISM ensures that instruction is not only personalised but equitable, providing learners with the type and timing of support that aligns with their needs in real time. This extends the work of UDL by embedding responsiveness directly into the instructional sequence, allowing for not just multiple representations of content but multiple trajectories of support based on learner behaviour.

Crucially, PRISM's multimodal foundation supports inclusivity at both a structural and experiential level. It facilitates assistive compatibility for learners with sensory or cognitive differences, while also accommodating diverse engagement styles through behavioural analytics and semantic adaptation. In this way, it reduces the systemic biases that arise when instruction is built around normative learner assumptions. However, translating this potential into practice will require equitable access to digital tools, as well as the cultivation of digital literacy among both educators and students. Moreover, ethical oversight is necessary to prevent the use of engagement tracking in ways that reinforce surveillance culture, particularly in under-resourced or vulnerable learning contexts.

6.4 Implementation Challenges and Systemic Considerations

While the study provides promising evidence in support of PRISM's theoretical architecture and simulated effectiveness, significant challenges remain for real-world implementation. One core issue concerns the fidelity of the data used to model engagement. Behavioural proxies such as watch-time percentages, rewind frequency, or skip events offer important signals but may not fully capture the cognitive or emotional engagement. The interpretability of such signals, particularly when automated systems use them to trigger pedagogical changes, must be carefully calibrated to avoid misalignment between system action and learner intent. In addition, the practical readiness of educators to engage with real-time analytics remains a concern. Effective use of PRISM-aligned systems requires teachers to possess a new set of professional competencies, namely, the ability to interpret engagement indicators and make informed instructional decisions on the fly.

This demands investment in both professional development and the co-design of usable, intuitive interfaces that bridge data with pedagogical action. Technical infrastructure also presents a constraint: real-time responsiveness depends on low-latency systems, seamless data integration, and scalable platforms, features that are not yet universal across educational settings. Consequently, the ethical dimensions of PRISM's deployment cannot be overlooked. Behavioural and biometric engagement tracking raises important questions around data ownership, learner consent, and the interpretability of algorithmic decisions.

Transparency in how models are trained and deployed, along with mechanisms for learner opt-in and feedback, will be critical for building trust and ensuring ethical alignment. These concerns do not undermine the framework's validity, but rather highlight the importance of carefully managed implementation studies and participatory design with all stakeholders involved.

Similarly, PRISM's implementation requires not only technological integration but pedagogical shifts. Educators need targeted professional development to interpret real-time engagement analytics and design branching content flows responsive to learner states. This includes training on dashboard use, multimodal data interpretation, and co-authoring logic rules aligned with curriculum outcomes. At the curriculum design level, PRISM encourages modular, flexible content development that supports recombination across its four pillars. Institutions may benefit from embedding PRISM-aligned design patterns into learning design templates, supporting a more scalable transition toward data-informed, engagement-responsive instruction.

6.5 Toward a Research-Informed Evolution of PRISM

The results of this study offer an encouraging foundation for the further development of PRISM, yet they also point toward a series of unresolved research questions that should guide the next phase of inquiry. First, it remains essential to investigate the longitudinal impact of PRISM-based interventions. While this study demonstrated immediate gains in engagement and satisfaction, further research is needed to determine whether such gains translate into improved retention, deeper learning outcomes, and long-term academic persistence. Second, the adaptability of PRISM across varied instructional contexts warrants systematic evaluation. While the framework was tested using MOOC-based interaction data, its application in disciplines such as the arts, laboratory sciences, or K–12 education may require domain-specific modifications. Similarly, different delivery modes, including hybrid, in-person, or mobile-first instruction, pose distinct challenges for real-time engagement detection and responsiveness.

Third, questions remain about the optimal orchestration between human educators and AI-driven systems. Determining the balance between system-initiated micro-interventions and teacher discretion is critical to avoiding cognitive overload, preserving pedagogical authenticity, and ensuring instructional transparency. As engagement-responsive learning becomes more widespread, the division of labour between machine and human actors will need to be intentionally designed and empirically refined. Subsequently, learner perceptions of engagement-based adaptation must be better understood. While some students may benefit from and appreciate timely interventions, others may perceive them as intrusive or mechanistic. Understanding how responsiveness affects learner agency, motivation, and trust in digital systems is vital for ensuring that adaptation serves the learner, rather than merely optimising for system-defined engagement metrics. These questions define a roadmap for the evolution of PRISM from a theoretically grounded and computationally promising model into a more fully developed and ethically aligned pedagogical paradigm.

7. Conclusion and Future Directions

This study presented PRISM as a pedagogical response to the urgent need for engagement-aware, data-informed instruction in contemporary digital learning environments. Designed around four interrelated pillars- Personalised Adaptive Learning, Responsive Engagement, Immersive Learning Experiences, and Social Collaboration, PRISM is underpinned by a robust theoretical foundation and operationalised through interaction data and multimodal modelling. The simulated pilot using the PEEK dataset and a transformer model demonstrated that learner engagement can be effectively predicted and used to inform simulated micro-interventions. These findings affirm that pedagogical responsiveness is not only conceptually sound but technically achievable, with direct implications for improving attentional continuity, personalisation, and instructional relevance. PRISM thus offers a viable design model for education systems seeking to move beyond static content delivery toward learner-state-informed adaptation.

Looking ahead, the framework opens critical lines of inquiry for advancing both research and practice in engagement-driven education. While the findings provide promising support for PRISM's logic, they are constrained by reliance on behavioural proxy labels and simulations rather than live interventions. Future research must evaluate PRISM in authentic classroom environments to assess pedagogical impact, behavioural persistence, and scalability. The present study provides only computational support for one component of the framework, engagement prediction, rather than full empirical validation. Understanding how learners interpret and respond to adaptive mechanisms will further shape the framework's ethical and motivational grounding. In addition, as intelligent systems become more integrated into teaching workflows, research must clarify the evolving roles of educators and algorithms in shared instructional decision-making. Finally, testing PRISM across diverse cultural, technological, and institutional contexts will be crucial to ensure its scalability and equity. This includes developing governance structures that uphold data privacy, algorithmic transparency, and learner agency. PRISM offers not just a model for designing next-generation learning experiences but a research agenda for ensuring that such experiences remain inclusive, ethical, and pedagogically sound.

Statements and Declarations

Funding

This research received no specific grant or funding.

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Clinical Trial Number
Not Applicable
Consent to Publish declaration
Not applicable
Consent to Participate declaration
Not applicable

Data Availability

The dataset used in this study (PEEK) is publicly available and can be accessed via: https://github.com/sahanbull/PEEK-Dataset

Ethics Approval

Not applicable. The study used publicly available, anonymised data that does not involve human subjects research requiring institutional review.

References

1. Prain, V., Cox, P., Deed, C., Dorman, J., Edwards, D., Farrelly, C., ... & Yager, Z. (2013). Personalised learning: Lessons to be learnt. British Educational Research Journal, 39(4), 654-676. https://doi.org/10.1080/01411926.2012.669747
2. Langenfeld, T., Burstein, J., & von Davier, A. A. (2022, May). Digital-first learning and assessment systems for the 21st century. In Frontiers in Education (Vol. 7, p. 857604). Frontiers Media SA. https://doi.org/10.3389/feduc.2022.857604
3. Sher, V. (2019). Technology-enhanced learning: Using learning analytics to unveil students' use of multiple devices in blended learning. https://summit.sfu.ca/_flysystem/fedora/sfu_migrate/19864/etd20636.pdf
4. Tucker, B. (2012). The flipped classroom. Education next, 12(1), 82-83. https://www.educationnext.org/wp-content/uploads/2023/04/ednext_XII_1_what_next.pdf
5. Gary, K. (2015). Project-based learning. Computer, 48(9), 98-100. doi: 10.1109/MC.2015.268
6. Stout, S. K. (2023). Distracted, bored, and disengaged: How to reengage Gen Z learners in the secondary school classroom (Doctoral dissertation, The University of North Dakota). https://search.proquest.com/openview/714a01db138ddf8c852f20fcd848cb27/1?pq-origsite=gscholar&cbl=18750&diss=y
7. Sharma, P., & Singh, D. (2025). Digitally Drained: Reviving Gen Z Student Engagement in the Face of Technostress. Journal of Engineering Education Transformations, 529-538. https://doi.org/10.16920/jeet/2025/v38is2/25065
8. Mihelič, K. K., Lim, V. K. G., & Culiberg, B. (2023). Cyberloafing among Gen Z students: the role of norms, moral disengagement, multitasking self-efficacy, and psychological outcomes. European Journal of Psychology of Education, 38(2), 567-585. https://doi.org/10.1007/s10212-022-00617-w
9. Adeshola, I., & Agoyi, M. (2023). Examining factors influencing e-learning engagement among university students during covid-19 pandemic: A mediating role of "learning persistence". Interactive Learning Environments, 31(10), 6195-6222. https://doi.org/10.1080/10494820.2022.2029493
10. Yalçin-Incik, E., & Incik, T. (2022). Generation Z Students' Views on Technology in Education: What They Want What They Get. Malaysian Online Journal of Educational Technology, 10(2), 109-124. https://doi.org/10.52380/mojet.2022.10.2.275
11. Holzer, B. M., Ramuz, O., Minder, C. E., & Zimmerli, L. (2022). Motivation and personality factors of Generation Z high school students aspiring to study human medicine. BMC medical education, 22(1), 31. https://doi.org/10.1186/s12909-021-03099-4
12. Balakrishna, C. (2023). The impact of in-classroom non-digital game-based learning activities on students transitioning to higher education. Education sciences, 13(4), 328. https://doi.org/10.3390/educsci13040328
13. Hou, R., Jin, L., He, J., & Wang, J. L. (2025). Peer support and learning outcomes in "Study With Me" among Generation Z college students: mediating roles of motivation, test anxiety, and self-efficacy. Frontiers in Psychology, 16, 1582857. doi: 10.3389/fpsyg.2025.1582857

14. Mihelič, K. K., Lim, V. K. G., & Culiberg, B. (2023). Cyberloafing among Gen Z students: the role of norms, moral disengagement, multitasking self-efficacy, and psychological outcomes. European Journal of Psychology of Education, 38(2), 567-585. https://doi.org/10.1007/s10212-022-00617-w
15. Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74(1), 59–109. https://doi.org/10.3102/00346543074001059
16. Borotić, G., & Jagušt, T. (2022, October). Enhancing student engagement with personalized gamification and adaptive learning strategies. In 2022 IEEE Frontiers in Education Conference (FIE) (pp. 1-5). IEEE.. doi: 10.1109/FIE56618.2022.9962647.
17. Mark, G., Gudith, D., & Klocke, U. (2008, April). The cost of interrupted work: more speed and stress. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 107-110). https://doi.org/10.1145/1357054.1357072
18. Rosen, L. D., Carrier, L. M., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task-switching while studying. Computers in human behavior, 29(3), 948-958. https://doi.org/10.1016/j.chb.2012.12.001
19. Shabidinova, Z., & Mamadiyeva, M. (2025). THE ROLE OF TECHNOLOGY IN ENHANCING STUDENT ENGAGEMENT IN MODERN CLASSROOMS. Journal of Applied Science and Social Science, 1(4), 209-215. https://inlibrary.uz/index.php/jasss/article/view/109138
20. Clinciu, R. A. (2023). Optimizing Educational Management: Strategies for Effective Learning Environments and Academic Excellence. LOGOS, UNIVERSALITY, MENTALITY, EDUCATION, NOVELTY. Section Social Sciences, 12(1), 77-89. https://www.ceeol.com/search/article-detail?id=1258559
21. Vitus, O. C. (2025). INNOVATIVE METHODOLOGIES IN EDUCATIONAL SETTINGS: ENHANCING STUDENT ENGAGEMENT AND LEARNING OUTCOMES THROUGH MIXED-METHODS RESEARCH. World Journal of Educational Studies, 10. : https://doi.org/10.61784/wjes3026
22. Nazarova, Y. (2025). STRATEGIES FOR ENHANCING STUDENT ENGAGEMENT IN A DYNAMIC LEARNING ENVIRONMENT. International Journal of Artificial Intelligence, 1(1), 903-907. https://inlibrary.uz/index.php/ijai/article/view/71033
23. Veluvali, P., & Surisetti, J. (2022). Learning management system for greater learner engagement in higher education—A review. Higher Education for the Future, 9(1), 107-121. https://doi.org/10.1177/23476311211049855
24. Almusaed, A., Almssad, A., Yitmen, I., & Homod, R. Z. (2023). Enhancing student engagement: Harnessing “AIED”’s power in hybrid education—A review analysis. Education Sciences, 13(7), 632. https://doi.org/10.3390/educsci13070632
25. Buragohain, D., Meng, Y., & Chaudhary, S. (2025). Metaverse Impact and Trends in Sustainable Integration of Immersive Technologies in Education. *Computer Applications in Engineering Education*, *33*(3), e70024. https://doi.org/10.1002/cae.70024
26. Buragohain, D., Deng, C., Sharma, A., & Chaudhary, S. (2024). The impact of immersive learning on teacher effectiveness: a systematic study. *IEEE Access*, *12*, 35924-35933. doi: 10.1109/ACCESS.2024.3373541.
27. Cayubit, R. F. O. (2022). Why learning environment matters? An analysis on how the learning environment influences the academic motivation, learning strategies and engagement of college students. Learning Environments Research, 25(2), 581-599. https://doi.org/10.1007/s10984-021-09382-x

28. Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International journal of educational technology in higher education*, *17*(1), 2. https://doi.org/10.1186/s41239-019-0176-8
29. Rogers, D. L. (2000). A paradigm shift: Technology integration for higher education in the new millennium. AACE Review (Formerly AACE Journal), 1(13), 19-33.
30. Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher-AI complementarity. Grantee Submission. https://files.eric.ed.gov/fulltext/ED618924.pdf
31. Arciniegas, et al. (2023). Attention monitoring of students during online classes using XGBoost classifier. Computers and Education: Artificial Intelligence, 5, 100191. https://doi.org/10.1016/j.caeai.2023.100191
32. Smith, A. J., Van Wyk, B. J., & Du, S. (2019, October). CNNs and transfer learning for lecture venue occupancy and student attention monitoring. In International Symposium on Visual Computing (pp. 383-393). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-33723-0_31
33. Isaeva, R., Karasartova, N., Dznunusnalieva, K., Mirzoeva, K., & Mokliuk, M. (2025). Enhancing learning effectiveness through adaptive learning platforms and emerging computer technologies in education. Jurnal Ilmiah Ilmu Terapan Universitas Jambi, 9(1), 144-160. https://doi.org/10.22437/jiituj.v9i1.37967
34. Isaias, P. (2018). Model for the enhancement of learning in higher education through the deployment of emerging technologies. Journal of Information, Communication and Ethics in Society, 16(4), 401-412. https://doi.org/10.1108/JICES-04-2018-0036
35. Sibley, L., Fabian, A., Plicht, C., Pagano, L., Ehrhardt, N., Wellert, L., ... & Lachner, A. (2025). Adaptive teaching with technology enhances lasting learning. Learning and instruction, 99. https://doi.org/10.1016/j.learninstruc.2025.102141
36. Sohan, M. A. K., & Orpa, S. A. (2025). A Comprehensive Review of Technology Enhanced Teaching and Learning Experiences. Journal of Applied Linguistics, 5(1), 148-161. https://doi.org/10.52622/joal.v5i1.450
37. Noroozi, O. (2025). Towards Technology-Enhanced Transformative Learning Environments. International Journal of Technology in Education, 8(2), 541-556. https://doi.org/10.46328/ijte.1209
38. Bockshecker, A., Ebner, K., & Smolnik, S. (2022, January). Technology-Enhanced Learning Environments and Adaptive Learning Systems-Development of Functionality Taxonomies. In HICSS (pp. 1-10). https://www.academia.edu/download/93033664/0006.pdf
39. Kumaresan, V., Jeyakarthic, M., Mohan, E., SIMATS, C., & Surendra, K. R. (2025, March). Personalized AI-Driven Online Education Technologies for Adaptive Learning and Cognitive Skill Enhancement. In 2025 7th International Conference on Signal Processing, Computing and Control (ISPCC) (pp. 1045-1050). IEEE
40. Baneres, D., Whitelock, D., Ras, E., Karadeniz, A., Guerrero-Roldán, A. E., & Rodríguez, M. E. (2019). Technology enhanced learning or learning driven by technology. International journal of educational technology in higher education, 16(5), 26-40. https://educationaltechnologyjournal.springeropen.com/articles/collections/telletedit orial
41. Kyza, E. A. (2023). Technology-enhanced learning: A learning sciences perspective. In Learning, Design, and Technology: An International Compendium of Theory,

Research, Practice, and Policy (pp. 221-244). https://doi.org/10.1007/978-3-319-17461-7_56
42. Sukhera, J., & Poleksic, J. (2021). Adapting compassion education through technology-enhanced learning: an exploratory study. Academic Medicine, 96(7), 1013-1020. DOI: 10.1097/ACM.0000000000003915
43. Lo, N., & Chan, S. (2025). Examining the role of generative AI in academic writing assessment: a mixed-methods study in higher education. *ESP Review*, *7*(2), 7-45. DOI : 10.23191/espkor.2025.7.2.7
44. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
45. Elmer, J., & Dingli, A. (2024). ENHANCING STUDENT LEARNING. In INTED2024 Proceedings (pp. 7771-7779). IATED doi: 10.21125/inted.2024.2076
46. Yang, C. (2025). Adapting teaching methods to accommodate diverse learning styles in education. Journal of Higher Education Research. https://doi. org/10.32629/jher. v5i6, 3382
47. Siemens, G. (2005). Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning, 2(1), 3–10.
48. Alam, M. A. (2024). Connectivism and Traditional Learning Theories: Implications for Contemporary Educational and Pedagogical Practices. Bhartiyam Int. J. Educ. Res., 14, 1-24
49. ṢAHİN, M., & TEBEANU, A. V. (2024). Approaches to learning: From traditional theories to connnectivism
50. Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285.
51. Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), The psychology of learning and motivation: Cognition in education (Vol. 55, pp. 37–76). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8
52. Bui, H. P., Ulla, M. B., Tarrayo, V. N., & Pham, C. T. (2023). The roles of social media in education: affective, behavioral, and cognitive dimensions. Frontiers in Psychology, 14, 1287728. https://doi.org/10.3389/fpsyg.2023.1287728
53. Singh, N. The Impact of Technology on Student Engagement and Learning Outcomes: A Critical Evaluation. https://www.echetana.com/wp-content/uploads/2025/05/30.-R-E-Dr.-Nahar-Singh-1.pdf
54. Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York, NY: Plenum.
55. Chapman, J. (2019). Blended learning and the motivation to learn the English language: A mixed methods study based on first-year university students in Japan. The University of Liverpool (United Kingdom).
56. Posner, M. I., & Rothbart, M. K. (2007). Educating the human brain. American Psychological Association.
57. Stalbert, M. (2022). Attention-Driven Design: How Instructional Designers Design to Capture the Learner's Attention. Capella University.
58. Martín-Monje, E., & Bárcena, E. (Eds.). (2015). Language MOOCs: Providing learning, transcending boundaries. Walter de Gruyter GmbH & Co KG.
59. Toboula, Z., & Martinien, C. (2023, February). Exploring the impact of AI-powered collaborative and interactive NLP apps on EFL teaching in the post-COVID-19 era. In IJCI Conference Proceedings-International Conference on Education in Post

Pandemic (EDUPAN 2023)-Volume (Vol. 12). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4398817
60. Rose, D. (2001). Universal design for learning. *Journal of Special Education Technology*, *16*(4), 64-67. https://doi.org/10.1177/016264340101600411
61. Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers college record*, *108*(6), 1017-1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x
62. Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use. TechTrends, 60(5), 433-441. DOI: 10.1007/s11528-016-0091-y
63. Almeqdad, Q. I., Alodat, A. M., Alquraan, M. F., Mohaidat, M. A., & Al-Makhzoomy, A. K. (2023). The effectiveness of universal design for learning: A systematic review of the literature and meta-analysis. *Cogent Education*, *10*(1), 2218191. https://doi.org/10.1080/2331186X.2023.2218191
64. Ghosh, N. (2019). Adopting technology and UDL in enhancing effectiveness of teaching learning process. International Journal of Science and Research (IJSR), 8(2), 957-964. https://www.academia.edu/download/98341125/ART20195122.pdf
65. Basham, J. D., Blackorby, J., & Marino, M. T. (2020). Opportunity in crisis: The role of universal design for learning in educational redesign. *Learning Disabilities: A Contemporary Journal*, *18*(1), 71-91.
66. Harris, L., & Yearta, L. (2020, April). The UDL, TPACK, and TIM frameworks to evaluate pre-service teacher technology integration. In Society for Information Technology & Teacher Education International Conference (pp. 2010-2016). Association for the Advancement of Computing in Education (AACE). https://www.learntechlib.org/d/216062/
67. Bauder, D. K., Cooper, K. M., & Simmons, T. J. (2019). SAMR strategies for the integration of technology through UDL. In Universal access through inclusive instructional design (pp. 141-152). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780429435515-18/samr-strategies-integration-technology-udl-debra-bauder-katherine-cooper-thomas-simmons
68. Bajracharya, J. R. (2021). Technology integration models and frameworks in teaching and training. Journal of Training and Development, 6(01), 3-11. https://doi.org/10.3126/jtd.v6i01.41674
69. Kimmons, R. & Hall, C. (2016). Toward a Broader Understanding of Teacher Technology Integration Beliefs and Values. *Journal of Technology and Teacher Education, 24*(3), 309-335. Waynesville, NC USA: Society for Information Technology & Teacher Education. Retrieved September 18, 2025 from https://www.learntechlib.org/primary/p/172627/.
70. Pilgrim, J., Vasinda, S., & Lisenbee, P. S. (2020). Integrating Technology in Literacy Instruction: Models and Frameworks for All Learners. https://doi.org/10.4324/9780429340154
71. Lisenbee, P. S., Pilgrim, J., & Vasinda, S. (2020). Integrating technology in literacy instruction: Models and frameworks for all learners. Routledge. https://books.google.com/books?hl=en&lr=&id=M2bnDwAAQBAJ&oi=fnd&pg=PT13&dq=UDL+TPACK+SAMR+frameworks+comparison+in+education+technology+integration+strategies&ots=xQAz2LDACG&sig=_Yukq-V7wJqLFDnrk1PjsK1NzY8

72. Adeshina, A. E. (2024). The transformative role of digital resources in teaching and learning. *Open Journal of Educational Development (ISSN: 2734-2050)*, *5*(1), 1-9. https://doi.org/10.52417/ojed.v5i1.520
73. van Thiel, L. (2018). Professional learning design framework: supporting technology integration in Alberta. *Research in Learning Technology*, *26*. https://doi.org/10.25304/rlt.v26.1989
74. Taylor, D. L., Yeung, M., & Bashet, A. Z. (2021). Personalized and adaptive learning. In Innovative learning environments in STEM higher education: Opportunities, Challenges, and Looking Forward (pp. 17-34). https://doi.org/10.1007/978-3-030-58948-6
75. Mirari, K. (2022). The effectiveness of adaptive learning systems in personalized education. Journal of Education Review Provision, 2(3), 107-115. https://doi.org/10.55885/jerp.v2i3.194
76. Du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. Heliyon, 10(21). DOI: 10.1016/j.heliyon.2024.e39630
77. Okonji, N. N., & Igwe, E. A. (2025). Prospects and Challenges of Personalized Learning through Artificial IntelligenceDriven Adaptive Learning Systems in N. Journal of Education in Developing Areas, 33(1), 139-147. https://journals.journalsplace.org/index.php/JEDA/article/viewFile/665/568.
78. Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. Computers in human behavior, 89, 98-110. https://doi.org/10.1016/j.chb.2018.07.027
79. Al Hussaini, M. H., Kousar, S., Kousar, N., Batool, U., & Munawar, N. (2024). Exploring Effective Strategies for Student Engagement and Learning Outcomes in Elementary Education. International Journal of Learning Reformation in Elementary Education, 3(02), 90-101. https://doi.org/10.56741/ijlree.v3i02.573
80. Fasco, P. S., Asiimwe, S., Ssekabira, G., & Tushabe, J. A. (2024). Enhancing student engagement and learning outcomes: Effective strategies in institutional pedagogy. International Journal of Multidisciplinary Research and Growth Evaluation, 5(04), 758-767.
81. Tahiru, F. (2021). AI in education: A systematic literature review. *Journal of Cases on Information Technology (JCIT)*, *23*(1), 1-20. DOI: 10.4018/JCIT.2021010101
82. El-Sabagh, H. A. (2021). Adaptive e-learning environment based on learning styles and its impact on development students' engagement. International Journal of Educational Technology in Higher Education, 18(1), 53. https://doi.org/10.1186/s41239-021-00289-4
83. Csikszentmihalyi, M., & Csikzentmihaly, M. (1990). *Flow: The psychology of optimal experience* (Vol. 1990, p. 1). New York: Harper & Row.
84. Fitrianto, I., & Saif, A. (2024). The role of virtual reality in enhancing Experiential Learning: a comparative study of traditional and immersive learning environments. International Journal of Post Axial: Futuristic Teaching and Learning, 97-110. https://doi.org/10.59944/postaxial.v2i2.300
85. Beck, D., Morgado, L., & O'Shea, P. (2023). Educational practices and strategies with immersive learning environments: Mapping of reviews for using the metaverse. IEEE Transactions on Learning Technologies, 17, 319-341. doi: 10.1109/TLT.2023.3243946

86. Fernandes, F. A., Rodrigues, C. S. C., Teixeira, E. N., & Werner, C. M. (2023). Immersive learning frameworks: A systematic literature review. IEEE Transactions on Learning Technologies, 16(5), 736-747. doi: 10.1109/TLT.2023.3242553
87. Gopinathan, S., Kaur, A. H., Veeraya, S., & Raman, M. (2022). The role of digital collaboration in student engagement towards enhancing student participation during COVID-19. Sustainability, 14(11), 6844. https://doi.org/10.3390/su14116844
88. Pramesworo, I. S., Fathurrochman, I., Sembing, D., Bangkara, B. A., & Sudrajat, D. (2023). Relevance between Blended Learning and Students' Independent Learning Curriculum: An Overview of Digital Age Education, Student and Teacher Engagement, Technological Resources. Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran, 9(3), 858-869.
89. Zitha, I., Mokganya, G., & Sinthumule, O. (2023). Innovative strategies for fostering student engagement and collaborative learning among extended curriculum programme students. Education Sciences, 13(12), 1196 https://doi.org/10.3390/educsci13121196
90. Han, S. (2022). Flipped classroom: Challenges and benefits of using social media in English language teaching and learning. Frontiers in Psychology, 13, 996294. https://doi.org/10.3389/fpsyg.2022.996294
91. Toro-Troconis, M., Alexander, J., & Frutos-Perez, M. (2019). Assessing student engagement in online programmes: Using learning design and learning analytics. International Journal of Higher Education, 8(6). https://doi.org/10.5430/ijhe.v8n6p171
92. Fritz, J., Hawken, M., & Shin, S. (2021). Using learning analytics and instructional design to inform, find, and scale quality online learning. In Online Learning Analytics (pp. 95-114). Auerbach Publications. https://doi.org/10.1201/9781003194620
93. Bulathwela, S., Perez-Ortiz, M., Yilmaz, E., & Shawe-Taylor, J. (2020, April). Truelearn: A family of bayesian algorithms to match lifelong learners to open educational resources. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 01, pp. 565-573). https://doi.org/10.1609/aaai.v34i01.5395
94. Banihashem, S. K., Gašević, D., & Noroozi, O. (2025). A Critical Review of Using Learning Analytics for Formative Assessment: Progress, Pitfalls and Path Forward. *Journal of Computer Assisted Learning*, *41*(3), e70056. https://doi.org/10.1111/jcal.70056
95. Qiu, Y., Djemili, K., Elezi, D., Shalman, A., Pérez-Ortiz, M., Yilmaz, E., & Shawe-Taylor, J. (2023). A toolbox for modelling engagement with educational videos. *arXiv*. https://doi.org/10.48550/arXiv.2401.05424
96. Devlin, Jacob, et al. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 conference of the North American chapter of the association for computational Linguistics: human language technologies, volume 1 (long and short papers). https://doi.org/10.18653/v1/N19-1423
97. Wen, M., Yang, D., & Rose, C. (2014, July). Sentiment Analysis in MOOC Discussion Forums: What does it tell us?. In *Educational data mining 2014*.
98. Whitehill, J., Serpell, Z., Lin, Y. C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student engagementfrom facial expressions. *IEEE Transactions on Affective Computing*, *5*(1), 86-98. doi: 10.1109/TAFFC.2014.2316163.

99. Adam, K. D. B. J. (2014). A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, *1412*(6). https://doi.org/10.48550/arXiv.1412.6980

100. Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2020, October). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations* (pp. 38-45). https://doi.org/10.18653/v1/2020.emnlp-demos.6

101. Liu, X., & Wang, C. (2021). An empirical study on hyperparameter optimization for fine-tuning pre-trained language models. *https://doi.org/10.48550/arXiv.2106.09204*