Journal of Artificial Intelligence, Virtual Reality, andHuman-Centered Computing ISSN: 3069-2318

PRIME OPEN ACCESS

Volume 2, Issue 2

Research Article

Date of Submission: 16 Apr, 2026
Date of Acceptance: 13 May, 2026
Date of Publication: 22 May, 2026

The Mediating Effect of an Agentic AI Chatbot in Scaffolding Online Engagement: Evidence from a Three-Cycle Design-Based Study

Vincent English*

Longford International College, Ireland; and The Faculty of Education, University of Cambridge, UK

***Corresponding Author:** Vincent English, Longford International College, Ireland; and The Faculty of Education, University of Cambridge, UK.

Citation: English, V. (2026). The Mediating Effect of an Agentic AI Chatbot in Scaffolding Online Engagement: Evidence from a Three-Cycle Design-Based Study. *J AI VR Hum Comput, 2*(2), 01-11.

Abstract
The rapid integration of artificial intelligence (AI) into higher education presents new paradigms for supporting student engagement in online learning environments. This study investigates how an agentic AI chatbot can scaffold dialogue and enhance engagement among postgraduate students enrolled in an online Executive MBA (EMBA) programme. Using a three-cycle design-based research (DBR) methodology, this study analysed the progression of ten design conjectures across a cumulative sample of 145 EMBA students from nine countries. Data were collected from asynchronous forum posts, Moodle learning management system (LMS) metrics, and the Online Student Engagement (OSE) scale. Quantitative analysis employed Kruskal-Wallis H tests, linear regression, and a Naïve Bayes classifier, whilst qualitative analysis utilised the Toolkit for Systematic Educational Dialogue Analysis (Tech-SEDA) to assess dialogue quality and the Structure of Observed Learning Outcome (SOLO) taxonomy to assess cognitive depth. The findings demonstrate a significant and large effect of the interventions on student engagement across all iterations, with Kruskal-Wallis tests yielding large effect sizes (epsilon-squared ranging from 3.78 to 12.08, all $p < .001$). The integration of the AI chatbot in later iterations was associated with a substantial increase in the quality of dialogue and a progressively stronger correlation between academic performance and participation ($R^2 = 0.32$ in Iteration 1, rising to $R^2 = 0.49$ in Iteration 3; Cohen's $f^2 = 0.96$). The Naïve Bayes classifier achieved 73.4% accuracy in predicting SOLO levels from dialogic features, confirming that dialogue quality is a reliable predictor of cognitive depth. The study introduces the ENGAGE (Engage, Navigate, Guide, Articulate, Gather, Evolve) framework as a model for ethically and effectively integrating AI into online pedagogy. The results confirm that a well-designed, AI-enhanced learning environment can act as a powerful mediating tool, scaffolding deeper cognitive engagement and fostering a more robust dialogic space for adult learners.

Keywords: Generative AI, Agentic AI, Online Engagement, Scaffolding, Dialogue, Design-Based Research, EMBA, Adult Learning, Tech-SEDA, SOLO Taxonomy

Introduction
The landscape of higher education has been irrevocably altered by the ascent of online learning, a modality that has become central to the delivery of postgraduate and executive education. For adult learners enrolled in Executive Master of Business Administration (EMBA) programmes, the online environment offers essential flexibility but simultaneously introduces challenges related to isolation, transactional interactions, and sustaining meaningful engagement [1]. Asynchronous discussion forums, a staple of learning management systems such as Moodle, are intended to foster communities of inquiry; however, they frequently devolve into superficial exchanges that fail to promote deep cognitive engagement [2]. Addressing this pedagogical gap is critical for realising the full potential of online education for adult professionals.

The emergence of powerful generative artificial intelligence (AI), particularly large language models (LLMs) such as ChatGPT, presents a transformative opportunity to re-imagine online learning environments. When designed as agentic pedagogical tools, these technologies can function as more than mere information-retrieval systems; they can act as dialogic partners that scaffold learning, personalise feedback, and mediate interactions between learners and knowledge [3]. This study explores the mediating effect of an agentic AI chatbot in scaffolding dialogue and enhancing engagement within an online EMBA programme. Drawing on Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD), this research conceptualises the AI as a 'more knowledgeable other' capable of providing the tailored, responsive support necessary to elevate students' cognitive and dialogic contributions [4].

This paper documents a three-year design-based research (DBR) study that iteratively developed and tested a series of pedagogical interventions across three cohorts of EMBA students. The research systematically evolved from establishing foundational principles of structured dialogue in Iteration 1 to integrating an agentic AI (ChatGPT) as a facilitator and co-participant in Iterations 2 and 3. By evaluating a sequence of ten design conjectures, this study provides empirical evidence of how specific AI-driven strategies influence student engagement, the quality of dialogue, and learning outcomes. The findings culminate in the ENGAGE framework, a practical and ethically grounded model for integrating AI into online learning to foster richer and more effective educational experiences.

The contribution of this paper is threefold. First, it provides a detailed empirical account of how an agentic AI chatbot can scaffold dialogue in an authentic online learning context, with quantified effect sizes across multiple iterations. Second, it demonstrates the utility of combining the Tech-SEDA framework with a Naïve Bayes classifier as a scalable method for assessing dialogue quality and predicting cognitive depth. Third, it proposes the ENGAGE framework as a theoretically grounded and practically applicable model for AI integration in online higher education.

Review of Related Literature

Dialogic Pedagogy in Online Learning

Dialogic education, rooted in the philosophical work of Bakhtin (1981), posits that learning is a social and co-constructive process that unfolds through the interplay of diverse voices in dialogue [5]. Bakhtin's concept of polyphony—the presence of multiple, equally valid voices in a text or discourse—is particularly relevant to online learning environments, where students from varied professional and cultural backgrounds converge in asynchronous forums. Building on this foundation, Wegerif (2013) introduced the concept of the 'dialogic space', a metaphorical and epistemic space where learners engage with alternative perspectives, negotiate meanings, and co-construct understanding [6]. Unlike transmission-based pedagogies, dialogic education is not concerned solely with delivering information but with cultivating the habits of inquiry, reasoning, and collaborative meaning-making.

In the digital era, dialogic pedagogy finds a natural home in asynchronous online forums, which offer students the temporal flexibility to reflect, compose, and respond at their own pace. These environments facilitate sustained engagement, challenge, and responsiveness—core tenets of dialogic theory [7]. However, the asynchronous format also presents challenges, such as the loss of immediacy and spontaneity present in face-to-face interactions. Research consistently shows that without deliberate instructional design, online discussions can lack depth, coherence, and critical inquiry [8]. To address this, frameworks such as the Toolkit for Systematic Educational Dialogue Analysis (Tech-SEDA) have been developed to codify and assess the quality of educational dialogue, moving beyond simplistic metrics like post frequency to analyse the functional and cognitive nature of student interactions [9].

Scaffolding and the Zone of Proximal Development

Vygotsky's (1978) concept of the ZPD—the cognitive space between what a learner can achieve independently and what they can achieve with guidance—is central to the principle of scaffolding [4]. In an educational context, scaffolding refers to the temporary support provided by an instructor, peer, or tool that enables a learner to accomplish a task they would otherwise be unable to complete. As the learner's competence grows, the scaffold is gradually withdrawn. In online learning, particularly with adult learners who possess rich but varied prior knowledge, providing effective, individualised scaffolding is a significant challenge [10]. The asynchronous and often isolated nature of the experience makes it difficult for human instructors to provide the timely, responsive support that is the hallmark of effective scaffolding.

Andragogy, Knowles's (1984) theory of adult learning, further contextualises this challenge [11]. Adult learners are characterised by their self-directedness, their rich reservoir of experience, their readiness to learn in relation to their social roles, and their problem-centred orientation to learning. Effective scaffolding for adult learners must therefore be responsive to these characteristics, offering support that is relevant, respectful of prior experience, and conducive to autonomous inquiry. The challenge for online EMBA programmes is to design environments that honour these principles whilst simultaneously fostering the collaborative dialogue that is central to deep learning.

Generative AI as a Mediating Tool

Generative AI, and LLMs like ChatGPT, represent a new frontier in educational technology, offering the potential to provide scalable, personalised scaffolding. These AI systems can function as mediating tools that facilitate dialogue and support learners within their ZPD [3]. Unlike static resources, generative AI can engage in dynamic, conversational interactions, offering immediate feedback, generating tailored prompts, clarifying complex concepts, and even modelling higher-order thinking through techniques like Chain-of-Thought (CoT) prompting. This capability positions AI not as a replacement for the human educator but as a powerful new kind of dialogic partner.

Research has begun to explore this potential, demonstrating that AI-generated prompts can enhance the quality of student contributions and foster deeper cognitive engagement [9]. The concept of the 'agentic AI' is particularly relevant here; rather than a passive tool that responds to queries, an agentic AI takes initiative, proactively scaffolding the learner's journey by introducing new perspectives, posing challenging questions, and guiding the learner towards higher-order thinking. This aligns with Garrison et al.'s (2000) Community of Inquiry framework, which identifies cognitive presence, social presence, and teaching presence as the three pillars of effective online learning [2]. An agentic AI can contribute to all three, particularly by enhancing cognitive presence through structured inquiry and teaching presence through consistent, responsive facilitation.

However, this integration is not without its challenges. Concerns around the reliability of AI-generated content, the potential for academic dishonesty, the ethical implications of data use, and the risk of diminishing the human element of learning must all be carefully considered [3]. Furthermore, the introduction of AI into a learning community can disrupt existing social dynamics, potentially marginalising less confident students or those with lower digital literacy. A thoughtful, principled approach to AI integration—one that keeps the human educator at the centre—is therefore essential.

Design-Based Research in Educational Technology
Design-based research (DBR) is a methodological approach that is particularly well-suited to studying the complex, context-dependent nature of learning with technology. DBR is characterised by its iterative, pragmatic, and theory-building approach, involving the systematic development and refinement of pedagogical interventions (the conjectures) while simultaneously generating a deeper understanding of the underlying learning processes [12]. Unlike controlled experiments, DBR is conducted in authentic educational settings, acknowledging the messiness of real-world learning environments. This makes it an ideal methodology for studying the integration of AI into online learning, where the interplay between technology, pedagogy, and human factors is complex and dynamic.

Research Questions
This study was guided by three overarching research questions that evolved with the DBR cycles:

- **RQ1:** What strategies can be used to increase attendance and participation in live and asynchronous online MBA sessions?

- **RQ2:** How can online discussion forums be optimised to enhance dialogic engagement in online MBA education?

- **RQ3:** How can ChatGPT and other AI-based educational tools be integrated into online MBA programmes to enhance student engagement and dialogic participation?

Methods
Research Design
This study employed a design-based research (DBR) methodology, guided by Bannan-Ritland's (2003) Integrative Learning Design Framework (ILDF) [13]. The ILDF provides a structured approach to DBR, comprising four phases: Informed Exploration, Enactment, Evaluation/Reflection, and Diffusion. This framework was applied across three primary iterations, with a pilot study preceding them, allowing for the progressive enhancement of the learning environment based on empirical findings from each cycle. Each iteration involved the formulation of specific design conjectures, their implementation in the online course, and a systematic evaluation of their impact.

Participants and Setting
The study was conducted in a fully online EMBA programme delivered via the Moodle LMS. The programme was designed for working professionals and was delivered asynchronously, with optional synchronous sessions. The total sample comprised 145 participants from nine countries, reflecting the diverse, international nature of executive education. The three main iterations had the following sample sizes: Iteration 1 (n = 37), Iteration 2 (n = 38), and Iteration 3 (n = 14). These sample sizes, whilst modest from a purely statistical standpoint, are representative of typical EMBA cohorts and are consistent with the established norms of design-based research, which prioritises depth of analysis over breadth of sample.

Data Collection
A mixed-methods approach was employed. Quantitative data were extracted from the Moodle LMS, including forum participation metrics (post frequency, word count, response latency), assignment grades, and examination results. The Online Student Engagement (OSE) scale a validated 19-item Likert-type instrument, was administered to all students at the beginning and end of each iteration to measure self-reported engagement characteristics [1]. Qualitative data consisted of the full text of 1,273 forum posts, which were systematically analysed using the Tech-SEDA framework. Semi-structured interviews and a focus group were also conducted in Iteration 3 to gather in-depth qualitative perspectives on students' experiences with AI integration.

Data Analysis
- **Tech-SEDA Coding:** Forum posts were coded using the Tech-SEDA framework, which identifies specific dialogic moves including *Challenge* (CH), *Elaboration* (EL), *Elaboration with Invitation* (ELI), *Refer to Wider Context* (RW), *Reasoning-Based comment* (RB), *Genuine Response* (GR), *Reinforcement* (RE), and *Humour*

Contribution (HC1, HC2). To ensure reliability, two independent human coders applied the framework, and Cohen's Kappa (κ) was calculated to measure inter-rater agreement. In Iteration 3, ChatGPT was introduced as a third coder, and weighted Kappa ($\kappa\omega$) was calculated to account for the additional complexity.

• **SOLO Taxonomy:** The cognitive complexity of student contributions was assessed using the SOLO taxonomy, which categorises learning outcomes from prestructural (no understanding) to extended abstract (generalisation and transfer) [14]. This provided a measure of the depth of learning evident in forum posts.

• **Statistical Analysis:** Kruskal-Wallis H tests were used to compare OSE scores across groups, given the ordinal nature of the data and the small sample sizes. Epsilon-squared (ε^2) was calculated as the effect size measure, using the formula $\varepsilon^2 = H / (n - 1)$, where H is the Kruskal-Wallis statistic and n is the total sample size. Linear regression analysis was conducted to examine the relationship between participation scores and academic performance, with Cohen's f^2 (calculated as $f^2 = R^2 / (1 - R^2)$) used to determine the effect size of the model. A Naïve Bayes classifier was trained in SPSS to predict SOLO levels from Tech-SEDA codes, sentiment scores, and lexical diversity metrics. Pseudo-BIC values and average log-likelihood scores were used to assess the predictive influence of each feature.

Results

The findings are presented across the three DBR iterations, charting the evolution of the interventions and their measured impact on student engagement and dialogue quality. A comprehensive comparative analysis synthesises these results to evaluate the overall effectiveness of integrating an agentic AI.

Iteration 1: Establishing a Baseline with Structure, Personalisation, and Incentives

Iteration 1 focused on three conjectures designed to establish a foundation for engagement through structural clarity, personalisation, and motivational incentives.

• **Conjecture 1: Clear Guidelines will Increase the Frequency and Quality of Forum Dialogue:** The introduction of explicit guidelines for forum participation, including clear expectations for the number and nature of posts, led to a 133% increase in weekly forum posts (from 3 to 7 per week) and a 30% improvement in assessed discussion quality. This confirmed the foundational importance of structure in online learning environments, consistent with the literature on instructional design [2]. However, whilst participation quantity increased, the dialogue often remained at a surface level, with students listing ideas in isolation rather than engaging in substantive exchange. Tech-SEDA analysis confirmed this, showing that *Elaboration* (EL) was the most frequent code, but *Challenge* (CH) and *Refer to Wider Context* (RW)—the codes most strongly associated with deeper learning—were less prevalent.

• **Conjecture 2: Personalised Interactive Elements will Reduce Impersonality and Improve Engagement:** The introduction of personalised elements, including SCORM-based interactive modules and recorded lectures, was associated with a 25% higher course completion rate. Students valued the flexibility afforded by recorded materials. However, the SCORM modules were not consistently revisited, and drop-in support sessions were poorly attended, suggesting that personalisation alone was insufficient to sustain deep engagement. The asynchronous revisit rate for course materials showed only a weak positive trend over the course ($y = 0.01x + 0.74$, $R^2 = 0.14$), indicating that session progression explained only 13.6% of the variance in revisit rates.

• **Conjecture 3: Incentivising Participation with Rewards will Motivate More Dialogue:** The introduction of a participation points system (contributing 20% to the final grade) initially boosted forum activity. However, a paired-samples t-test found no statistically significant difference in engagement between the first and second halves of the course ($t(3) = 1.65$, $p = .198$), indicating that the incentive effect was not sustained. External factors, including the professional commitments of students (several of whom were professional athletes with international commitments), significantly disrupted participation patterns. This finding aligns with the literature on the limitations of extrinsic motivation, suggesting that incentives alone cannot sustain meaningful engagement [15].

• The regression analysis for Iteration 1 revealed a moderate positive correlation between assignment grades and participation ($R^2 = 0.32$, slope = 8.34), yielding a large effect size (Cohen's $f^2 = 0.47$). The OSE post-course Kruskal-Wallis test yielded a highly significant result ($\chi^2 = 177.61$, $df = 9$, $p < .001$, $\varepsilon^2 = 4.93$), indicating a large effect of the course on student engagement.

• The Naïve Bayes analysis of Iteration 1 data identified *Challenge* (CH) as the most influential predictor of learning outcomes (Rank 1, pseudo-BIC = 0.431, average log-likelihood = −0.132), followed by *Elaboration with Invitation* (ELI) (Rank 2, pseudo-BIC = 0.512) and *Refer to Wider Context* (RW) (Rank 3, pseudo-BIC = 0.602). Crucially, Elaboration (EL)—the most frequent code—was only the fourth-ranked predictor (pseudo-BIC = 0.688, average log-likelihood = −0.539), demonstrating that frequency of dialogue is not equivalent to quality or impact on learning. Inter-rater reliability for Tech-SEDA coding was substantial, with Cohen's Kappa values ranging from 0.76 to 0.88 across six weeks (average $\kappa = 0.80$, $p < .050$).

Iteration 2: Integrating an Agentic AI for Deeper Dialogue

Building on the findings of Iteration 1, Iteration 2 introduced ChatGPT as an agentic AI tool to scaffold discussions. Four conjectures guided this iteration.

- **Conjecture 4: Integrating ChatGPT will Improve the Quality of Student Dialogue:** The introduction of ChatGPT as a dialogic scaffold produced marked qualitative improvements in discussion quality. Tech-SEDA analysis revealed significant increases in the frequency of *Elaboration* (EL), *Reasoning-Based comments* (RB), and *Reference to Wider Context* (RW) codes, particularly in Week 4, which coincided with the formal integration of ChatGPT into the forum activity. The Tech-SEDA graph showed that EL codes increased from an average of approximately 50–60 occurrences per week to 166 in Week 4 alone, a 177–232% increase. Students began using dialogue codes associated with higher-order thinking, suggesting that the AI was successfully scaffolding more sophisticated cognitive engagement. Spearman correlation analysis of Tech-SEDA codes revealed a strong negative correlation between *Challenge* (CH) and *Humour Contribution without context* (HC1) ($\rho = -0.82$, $p = .044$), and between *Elaboration* (EL) and HC1 ($\rho = -0.85$, $p = .034$), indicating that contextually irrelevant contributions were associated with reduced depth of dialogue.

- **Conjecture 5: ChatGPT Personalisation will Heighten Student Engagement:** Qualitative data from student feedback indicated that ChatGPT's ability to provide personalised, immediate responses was highly valued. Students reported that the AI reduced the cognitive load associated with navigating complex business concepts, allowing them to focus their intellectual energy on higher-order analysis. However, some students noted that the AI lacked the emotional nuance of a human instructor, and a minority reported feeling that interactions with the AI were less personally meaningful than those with peers or the educator. This finding underscores the importance of maintaining a human presence in AI-enhanced learning environments.

- **Conjecture 6: Incentivising Participation with ChatGPT's Guidance will Elevate Engagement:** The regression analysis for Iteration 2 revealed a weaker correlation between exam results and participation than in Iteration 1 ($R^2 = 0.19$, slope = 12.29), yielding a medium effect size (Cohen's $f^2 = 0.23$). Whilst this might initially appear to be a regression, it is interpreted as evidence that the AI was providing alternative pathways to engagement, reducing the dominance of exam performance as the primary driver of participation. The slope of 12.29 represents a 47.4% increase from Iteration 1 (slope = 8.34), indicating that each percentage point increase in exam results was associated with a larger increase in participation points when AI support was available.

- **Conjecture 7: The Synergy Between ChatGPT and CoT Prompts will Promote Deeper Critical Thinking:** The introduction of Chain-of-Thought (CoT) prompts within the ChatGPT interface produced notable qualitative improvements in the depth of student reasoning. Students were guided to articulate their thinking step by step, which fostered more structured and reflective contributions. Dendrogram analysis of forum posts using Jaccard correlations revealed distinct dialogue clusters, with the highest thematic similarity observed between students who engaged most actively with the CoT-prompted discussions (e.g., Cluster A: $p = 0.58$; Cluster B: $p = 0.50$ in Session 2). The OSE post-course Kruskal-Wallis test for Iteration 2 yielded $\chi^2 = 140.03$, df = 8, $p < .001$, $\varepsilon^2 = 3.78$, confirming a large effect of the interventions on engagement.

Iteration 3: Synthesising Structure and AI Facilitation

Iteration 3 represented the synthesis of the previous cycles, combining structured discussion boards with AI facilitation to create a highly scaffolded learning environment. Three conjectures guided this final iteration.

- **Conjecture 8: Well-Structured Discussion Boards Enhanced by AI will Foster Collaboration and Improve Learning Outcomes:** The results from this iteration were the most compelling of the study. The forum on Apple's Strategic Use of Blockchain Technology was designed with a structured format that required students to post an initial contribution, respond to peers, and then engage in a reflective synthesis. ChatGPT was integrated as a dialogic partner, available to provide prompts, challenge assumptions, and introduce divergent perspectives. The regression analysis revealed the strongest correlation between assignment grades and participation observed in the study ($R^2 = 0.49$, slope = 19.83), with a correspondingly large effect size (Cohen's $f^2 = 0.96$). This indicates that nearly half the variance in participation scores could be explained by students' academic performance in this highly structured, AI-enhanced environment. The regression slope of 19.83 represents a 61.3% increase from Iteration 2 and a 137.8% increase from the baseline in Iteration 1.

- **Conjecture 9: ChatGPT as a Facilitator will Enhance Collaboration and Knowledge Exchange in Group Learning:** Qualitative analysis of group learning activities, including asynchronous breakout discussions, revealed that ChatGPT effectively catalysed collaborative exchanges. Students reported that the AI helped to balance participation, giving less confident voices a platform to contribute. The AI's ability to introduce new perspectives and synthesise diverse viewpoints was particularly valued. Weighted Kappa analysis for Iteration 3, which included ChatGPT as a third coder alongside two human coders, showed substantial agreement between human coders ($\kappa\omega$ up to 0.774 in Week 5, $p < .05$), though the AI coder showed a systematic anomaly in Week 3 ($\kappa\omega = -0.50$), highlighting the limitations of AI-based coding and the importance of human oversight.

• **Conjecture 10: ChatGPT in Synchronous Breakout Rooms will Promote Equitable Participation:** The integration of ChatGPT into synchronous breakout rooms was found to promote more equitable participation in real-time discussions. Observational data indicated that the AI's presence as a neutral facilitator reduced the dominance of more vocal students and encouraged contributions from those who might otherwise have remained passive. Students reported that having access to instant, AI-generated insights during discussions helped them to engage more confidently with complex topics. The OSE post-course Kruskal-Wallis test for Iteration 3 yielded $\chi^2 = 157.02$, df = 9, $p < .001$, $\varepsilon^2 = 12.08$, the largest effect size of any iteration, confirming a large and significant effect of the synthesised interventions on student engagement.

Comparative Analysis Across Iterations

The progression across the three iterations reveals a clear and statistically significant trend of improvement. Table 1 summarises the OSE Kruskal-Wallis results, and Table 2 presents the comparative regression and effect size analysis.

Iteration	χ^2	df	n	p	Epsilon-Squared (ε^2)	Effect Size Interpretation
IT1 Post-course	177.61	9	37	<.001	4.93	Large
IT2 Post-course	140.03	8	38	<.001	3.78	Large
IT3 Post-course	157.02	9	14	<.001	12.08	Large

Table 1: Kruskal-Wallis H Test Results for Post-Course OSE Scale Scores

Note. Epsilon-squared calculated as $\varepsilon^2 = H / (n - 1)$. Effect size interpretation follows Cohen (1988): small ≥ 0.01, medium ≥ 0.06, large ≥ 0.14 [16].

Iteration	Predictor	R^2	Cohen's f^2	Effect Size	Slope	% Change in Slope
IT1	Assignment grades	0.32	0.47	Large	8.34	Baseline
IT2	Exam results	0.19	0.23	Medium	12.29	+47.4%
IT3	Assignment grades	0.49	0.96	Large	19.83	+61.3%

Table 2: Comparative Regression and Effect Size Analysis: Participation vs. Academic Performance

Note. Cohen's f^2 calculated as $f^2 = R^2 / (1 - R^2)$. Effect size interpretation: small ≥ 0.02, medium ≥ 0.15, large ≥ 0.35 [16].

The Naïve Bayes classifier, trained to predict the cognitive depth of forum posts (SOLO levels) based on their dialogic features, achieved an overall accuracy of 73.4% across 1,273 forum posts. The model's performance was strongest at SOLO Levels 3 and 4 (Relational and Extended Abstract), confirming that the dialogic features captured by Tech-SEDA are most reliably associated with the higher levels of cognitive engagement. Table 3 summarises the evolution of the top Naïve Bayes predictors across iterations.

Iteration	Top Predictor	Rank	Pseudo-BIC	Avg. Log-Likelihood	Interpretation
IT1	Challenge (CH)	1	0.431	−0.132	Critical dialogue most influential
IT1	Refer to Wider Context (RW)	3	0.602	−0.005	Contextual relevance important
IT1	Elaboration (EL)	4	0.688	−0.539	Frequency ≠ impact
IT2	Refer to Wider Context (RW)	1	—	—	Contextual relevance dominant
IT3	Elaboration (EL)	1	—	—	Elaborative dialogue dominant

Table 3: Evolution of Top Naïve Bayes Predictors of Learning Outcomes Across Iterations

Note. Lower pseudo-BIC indicates better model fit; lower average log-likelihood indicates stronger predictive influence.

The evolution of the top predictor—from *Challenge* in Iteration 1, to *Refer to Wider Context* in Iteration 2, to *Elaboration* in Iteration 3—reflects a meaningful pedagogical progression. In the early stages of the course, challenging assumptions was the most powerful driver of learning. As students developed confidence and contextual knowledge, connecting ideas to broader real-world scenarios became paramount. In the final iteration, with a fully scaffolded AI environment, the ability to elaborate and extend ideas became the hallmark of deep learning.

Table 4 provides a comprehensive summary of all ten conjectures, their outcomes, and the actions taken.

Conjecture	Iteration	Focus	Key Outcome	Statistical Evidence	Action
C1	IT1	Clear guidelines	133% increase in posts; 30% quality improvement	Qualitative (Tech-SEDA)	Retain
C2	IT1	Personalised elements	25% higher completion rate; weak revisit trend	$R^2 = 0.14$ (revisit rate)	Retain
C3	IT1	Incentivised participation	Initial boost; not sustained	$t(3) = 1.65$, $p = .198$	Retain
C4	IT2	ChatGPT integration	Marked gains in EL, RB, RW codes	EL codes: 50–60 → 166 (Week 4)	Retain
C5	IT2	AI personalisation	Increased ownership; lacked emotional nuance	Qualitative (interviews)	Retain
C6	IT2	AI-incentivised participation	Moderate correlation; AI as alternative pathway	$R^2 = 0.19$, $f^2 = 0.23$	Retain
C7	IT2	CoT prompts + AI	Deeper critical thinking; richer discussions	Jaccard clusters (p = 0.50–0.58)	Retain
C8	IT3	Structured boards + AI	Strongest participation-performance correlation	$R^2 = 0.49$, $f^2 = 0.96$	Retain
C9	IT3	AI as group facilitator	Enhanced collaboration; balanced participation	$\kappa\omega$ up to 0.774	Retain
C10	IT3	AI in synchronous rooms	Equitable participation; sustained dialogue	$\chi^2 = 157.02$, $\varepsilon^2 = 12.08$	Retain

Table 4: Summary of All Ten Design Conjectures: Outcomes and Actions

The SOLO taxonomy analysis further illustrates the progression of cognitive engagement across iterations. In Iteration 1, student contributions were predominantly at the Multistructural level (SOLO Level 3), characterised by the listing of relevant ideas without integration. By Iteration 2, following the introduction of ChatGPT, a significant shift towards Relational and Extended Abstract thinking was observed, with students making connections across concepts and generalising to new contexts. In Iteration 3, the structured forum design combined with AI facilitation consistently produced contributions at the Extended Abstract level, characterised by the generation of new hypotheses and the application of learning to novel professional contexts.

Discussion

The findings of this three-year design-based study demonstrate a clear and powerful pathway for enhancing student engagement in online learning through the strategic integration of an agentic AI chatbot. The progression across the iterations illustrates an evolution from basic structural interventions to a sophisticated, AI-mediated dialogic environment. The large effect sizes observed in both the OSE scale results and the regression models confirm that these interventions had a substantial and meaningful impact on student engagement and its relationship with academic performance.

The Scaffolding Trajectory: From Structure to Agency

The three-iteration trajectory of this study can be understood as a progressive scaffolding of the learning environment itself, mirroring the developmental arc that Vygotsky (1978) described at the level of the individual learner [4]. In Iteration 1, the scaffolding was primarily structural and external—clear guidelines, participation incentives, and personalised resources provided the necessary conditions for engagement but did not, in themselves, generate the internal motivation or cognitive challenge needed for deep learning. This is consistent with the literature on instructional design, which distinguishes between the conditions for learning and the conditions that actually produce learning [2]. The paired-samples t-test result ($t(3) = 1.65$, $p = .198$) from Iteration 1 is particularly instructive: it indicates that the structural interventions produced no statistically significant sustained change in engagement across the course, suggesting that the initial boost in participation was driven by novelty and extrinsic motivation rather than genuine cognitive engagement.

The introduction of the AI chatbot in Iteration 2 marked a qualitative shift in the nature of the scaffolding. Rather than providing a static framework within which students were expected to engage, the AI offered a dynamic, responsive scaffold that adapted to the content and direction of each student's contributions. This is the essence of Vygotskian scaffolding: not a fixed support structure but a responsive, interactive process that meets the learner where they are and guides them towards where they need to be. The increase in the regression slope from 8.34 to 12.29 between Iterations 1 and 2 suggests that the AI was beginning to mediate the relationship between effort and outcome, making engagement a more sensitive indicator of academic performance.

Iteration 3 synthesised these structural and cognitive scaffolds into a coherent, integrated learning environment. The dramatic increase in the R^2 value to 0.49 and the regression slope to 19.83 indicates that the AI, when combined with a well-designed pedagogical structure, created a powerful feedback loop where high-quality engagement was strongly predictive of, and likely conducive to, better learning outcomes. The shift in the top Naïve Bayes predictor to Elaboration (IT3) suggests that students had, by this stage, internalised the skills of challenge and contextualisation, and were now ready to engage in the deeper, more collaborative work of elaborating and extending ideas. This is consistent with Vygotsky's (1978) notion of the gradual internalisation of scaffolded skills—what begins as an external, AI-mediated process eventually becomes part of the learner's own cognitive repertoire [4]. This scaffolding trajectory has important implications for the design of online learning environments: the integration of AI should not be conceived as a single intervention but as a progressive, iterative process that evolves in response to the developing needs of the learner community.

Dialogue Quality as a Predictor of Learning Depth

One of the most significant contributions of this study is the demonstration that the quality of dialogue—as captured by the Tech-SEDA framework—is a reliable predictor of the depth of learning, as assessed by the SOLO taxonomy. The Naïve Bayes classifier achieved an overall accuracy of 73.4% in predicting SOLO levels from dialogic features, with the strongest performance at SOLO Levels 3 and 4 (Relational and Extended Abstract). This finding has profound implications for the assessment and facilitation of online learning. It suggests that by attending to the quality of dialogue—specifically, the presence of *Challenge, Elaboration*, and *Reference to Wider Context*—educators can gain reliable insights into the depth of learning that is occurring, without the need for formal assessments.

The finding that Elaboration (EL), the most frequent Tech-SEDA code, was only the fourth-ranked predictor of learning outcomes in Iteration 1 (pseudo-BIC = 0.688) is particularly noteworthy. It demonstrates that frequency of dialogue is not equivalent to quality or impact on learning—a finding that challenges the common practice of using post frequency as a proxy for engagement in online learning. The most influential predictors were *Challenge* (CH) and *Refer to Wider Context* (RW), which are associated with higher-order thinking and the integration of knowledge across contexts. This suggests that educators should design forum activities that specifically elicit these types of contributions, rather than simply requiring students to post a certain number of times.

The evolution of the top predictor across iterations—from *Challenge* (IT1) to *Refer to Wider Context* (IT2) to *Elaboration* (IT3)—also provides a developmental map of the learning process in this context. It suggests that as students progress through a well-scaffolded learning environment, the nature of the dialogue that is most strongly associated with deep learning changes, reflecting the shifting cognitive needs of the learner community. This has important implications for the design of AI-enhanced learning activities: the AI's scaffolding strategies should evolve in response to the developing needs of the learner, shifting from challenge-based prompting in the early stages to contextualisation and elaboration in the later stages.

The Agentic AI as a Dialogic Partner

A central finding of this study is that the most significant improvements in engagement and dialogue quality occurred when the AI was positioned not as a passive resource but as an active, agentic participant in the learning process. This is consistent with the concept of the 'more knowledgeable other' in Vygotsky's (1978) ZPD framework [4]. The AI's ability to introduce new perspectives, challenge assumptions, and guide students through complex reasoning processes—particularly through CoT prompting—was a key driver of the improvements observed in Iterations 2 and 3.

This finding has important implications for how AI is designed and deployed in educational contexts. An AI that merely answers questions or provides information is unlikely to produce the same scaffolding effect as one that is designed to engage in genuine dialogue, to ask probing questions, and to model the kind of thinking that leads to deep learning. The concept of the 'agentic AI' therefore represents a significant advance over earlier conceptions of educational technology, which tended to position the computer as a repository of knowledge rather than a participant in the learning process.

The Bakhtinian concept of polyphony is also relevant here [5]. The introduction of the AI into the forum created a new kind of polyphonic space, in which the voices of students, the educator, and the AI interacted to produce meanings that none could have generated alone. The AI's contributions were not merely informational but dialogic—they opened up new lines of inquiry, introduced divergent perspectives, and invited students to engage in the kind of reflective, critical thinking that is the hallmark of deep learning. This is consistent with Wegerif's (2013) concept of the dialogic space, and suggests that AI can play a powerful role in expanding and enriching this space [6].

Balancing AI and Human Presence

A recurring theme across the three iterations was the tension between the efficiency and scalability of AI-mediated scaffolding and the irreplaceable value of human presence in the learning community. Students consistently reported that whilst the AI was highly effective at providing information, generating prompts, and scaffolding cognitive processes, it lacked the emotional intelligence, empathetic responsiveness, and contextual sensitivity of a human instructor. This finding is consistent with the literature on the Community of Inquiry framework, which identifies social presence—the ability of participants to project their personal characteristics into the community—as a critical component of effective online learning [2].

The weighted Kappa anomaly observed in Iteration 3, where the AI coder showed a systematic disagreement with human coders in Week 3 (κω = −0.50), is a concrete illustration of this limitation. Whilst the AI was generally reliable as a coding tool, it was susceptible to systematic errors in contexts that required nuanced, contextually sensitive judgement. This underscores the importance of maintaining human oversight in AI-enhanced learning environments and of designing AI systems that are transparent about their limitations. The ENGAGE framework's emphasis on the 'human-in-the-middle' approach reflects this principle, advocating for a model in which the educator remains the primary orchestrator of the learning experience, with the AI serving as a powerful but ultimately subordinate tool.

The findings also suggest that the appropriate balance between AI and human presence may vary depending on the stage of the learning process and the specific needs of the learner community. In the early stages of a course, when students are still developing their confidence and familiarity with the learning environment, a higher degree of human presence may be needed to establish the social and emotional foundations of the learning community. As students become more confident and self-directed, the AI can take on a more prominent role in scaffolding cognitive processes, whilst the educator focuses on facilitating the social and collaborative dimensions of learning.

Implications for Practice: The ENGAGE Framework

Based on the cumulative findings of this research, the ENGAGE framework is proposed as a model for practitioners seeking to integrate AI into online learning. The framework consists of six principles, each of which addresses a critical dimension of effective AI integration:

• **Engage with Relevant Themes and Equitable Access:** AI integration must begin with a commitment to relevance and equity, ensuring that the themes and tasks presented are meaningful to all learners and that all students have equal access to the AI tools.

• **Navigate to the Instructions with Ethical Clarity:** The use of AI in education raises important ethical questions, including issues of data privacy, academic integrity, and bias. Practitioners must navigate these issues with clarity and transparency, communicating openly with students about how AI is being used and why.

• **Guide with Clarity and Inclusivity:** The AI should be designed to guide students through the learning process in a clear, structured, and inclusive manner, providing support that is tailored to the needs of individual learners.

• **Articulate Guidelines with Reflective Responsibility:** Practitioners must articulate clear guidelines for the use of AI, reflecting carefully on the potential risks and benefits and taking responsibility for the outcomes.

• **Gather Feedback Through Inclusive and Iterative Interaction:** Effective AI integration requires ongoing feedback from students, which should be gathered through inclusive and iterative processes that allow for continuous improvement.

• **Evolve and Iterate:** The integration of AI into education is not a one-time event but an ongoing process of evolution and iteration, in which practitioners continuously refine their approach based on evidence and feedback.

The ENGAGE framework advocates for a human-in-the-middle approach, where the educator orchestrates the learning environment, leveraging AI as a tool to enhance, not replace, their pedagogical role. This is consistent with the research of Garrison et al. (2000), who argue that teaching presence—the design, facilitation, and direction of cognitive and social processes—is a critical component of effective online learning [2]. The educator's role in an AI-enhanced environment is not diminished but transformed, requiring new skills in the design of AI-mediated learning experiences and the critical evaluation of AI-generated content.

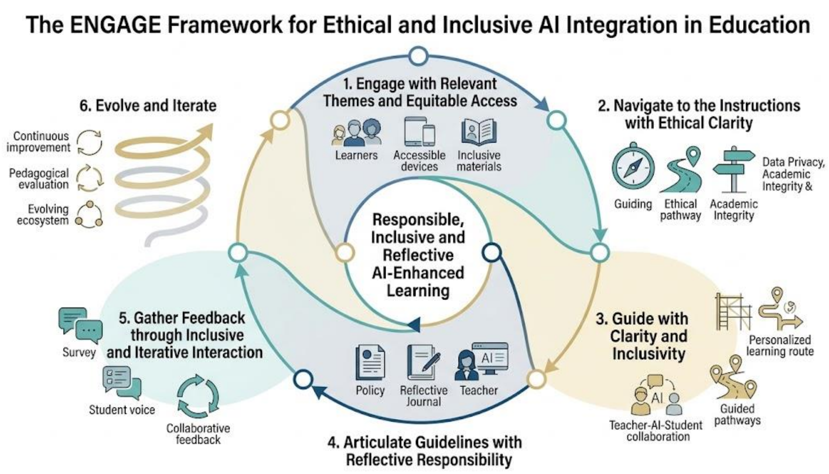

Figure 1

Limitations

This study has several limitations that should be acknowledged. The sample sizes, whilst typical for EMBA programmes, are small, which limits the generalisability of the statistical findings. The epsilon-squared values for the Kruskal-Wallis tests, whilst consistently large, should be interpreted with caution given the small sample sizes, as they may be inflated in small samples. The research was conducted within a single institutional context, and the findings may not be directly transferable to other disciplines, student populations, or technological contexts.

Furthermore, the technology itself (ChatGPT) evolved during the study period, and the specific capabilities of the AI model used may have influenced the results. The study also relied on self-reported engagement data from the OSE scale, which is subject to the limitations of self-report measures, including social desirability bias. Future research should seek to replicate these findings with larger, more diverse samples, across different institutional and technological contexts, and using more objective measures of engagement and learning.

Conclusions and Recommendations

This study provides compelling evidence that an agentic AI chatbot can serve as a powerful mediating tool to scaffold dialogue and enhance engagement in online adult learning. The design-based research approach allowed for the iterative development of a pedagogical model that progressed from simple structural enhancements to a sophisticated, AI-driven learning environment. The statistical analysis, including the large effect sizes observed across multiple metrics, confirms the significant positive impact of these interventions. The strong correlation between participation and academic performance in the final iteration ($R^2 = 0.49$, Cohen's $f^2 = 0.96$) suggests that a well-orchestrated, AI-enhanced environment can successfully align the processes of engagement with the outcomes of learning.

The evolution of the Naïve Bayes predictor rankings—from *Challenge to Refer to Wider Context* to *Elaboration*—provides a nuanced picture of how the nature of effective dialogue changes as students develop within a scaffolded learning environment. This finding has important implications for the design of AI-enhanced learning activities, suggesting that the type of scaffolding provided by the AI should evolve in response to the developing needs of the learner.

The ENGAGE framework offers a research-grounded model for practitioners to guide the integration of AI into online learning, emphasising the importance of equity, ethics, clarity, responsibility, feedback, and continuous improvement. As higher education continues to navigate the opportunities and challenges of the digital age, the thoughtful integration of AI will be paramount. By leveraging AI to create more responsive, personalised, and dialogically rich learning spaces, we can move closer to realising the full potential of online education for all learners.

Future research should focus on three key areas. First, longitudinal studies with larger samples are needed to confirm and extend the findings of this study. Second, research should explore the differential impact of AI integration on students with different learning styles, prior knowledge, and cultural backgrounds, to ensure that AI-enhanced environments are genuinely inclusive. Third, the development of more sophisticated AI tools that can adapt their scaffolding strategies in real time, based on ongoing analysis of student dialogue, represents a promising direction for future innovation.

Acknowledgements

The author wishes to thank the students who participated in this research and the colleagues who provided feedback on earlier drafts of this manuscript.

Ethics Board Statement

This research was conducted in accordance with the ethical guidelines of the participating institution. All participants provided informed consent, and all data were anonymised prior to analysis.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

AI Use Statement

The authors affirm that this manuscript was written entirely by the human author. Generative AI was the subject of the research but was not used in the drafting, editing, or production of this manuscript, in accordance with the Online Learning Journal's policy on the use of generative AI.

Conflict of Interest Statement

The author declares no conflict of interest.

References

1. Dixson, M. D. (2015). Measuring student engagement in the online course: The Online Student Engagement scale (OSE). *Online Learning, 19*(4), n4.
2. Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2–3), 87–105.
3. English, V. (2025). The mediating effect of generative AI in scaffolding dialogue in adult online learning. *Advances*

in Online Education: A Peer-Reviewed Journal, 3(4), 385–400.
4. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.
5. Bakhtin, M. M. (1981). *The dialogic imagination: Four essays* (C. Emerson & M. Holquist, Trans.). University of Texas Press.
6. Wegerif, R. (2013). *Dialogic: Education for the internet age.* Routledge.
7. Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach.* Routledge.
8. Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O., & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9,* 16–44.
9. English, V., & Jagha, N. (2025). Predicting learning depth through dialogue: A design-based approach to SOLO classification using Naive Bayes in online MBA education. *Advances in Online Education: A Peer-Reviewed Journal, 3*(4), 385-400.
10. Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education, 104,* 18–33.
11. Lippitt, G. L., Knowles, M. S., & Knowles, M. S. (1984). Andragogy in action: applying modern principles of adult learning.
12. Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.
13. Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher, 32*(1), 21–24.
14. Biggs, J., & Collis, K. (1982). Evaluating the quality of learning: the solo taxonomy new york: Academic pres.
15. Deci, E. L., & Ryan, R. M. (2013). *Intrinsic motivation and self-determination in human behavior.* Springer Science & Business Media.
16. Cohen, J. (2013). *Statistical power analysis for the behavioral sciences.* routledge.
17. Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences, 11*(1), 105–121.
18. Kali, Y., & Hoadley, C. (2020). Design-based research as a methodology for studying learning in context. In N. Rushby & D. Surry (Eds.), *The Wiley handbook of learning technology* (pp. 1–22). Wiley.
19. Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education, 35*(5), 501–517.
20. Rovai, A. P. (2007). Facilitating online discussions effectively. *The Internet and Higher Education, 10*(1), 77–88.
21. Swan, K., & Shih, L. F. (2005). On the nature and development of social presence in online course discussions. *Journal of Asynchronous learning networks, 9*(3), 115-136.
22. Tatar, D. (2007). The design tensions framework. *Human–Computer Interaction, 22*(4), 413–451.