Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 109-118
https://wjcrt.org/

# AI-Enabled English Language and Literature Learning through Digital Media in Higher Education

**Dr. Abu Farah Hoque**
Assistant Professor
Department of English
Maharani Kasiswari College, 20
Ramkanto Bose Street, Kolkata -700003
(affiliated to the University of Calcutta, India)
**ORCID id** - https://orcid.org/0009-0000-7767-3527

www.wjftcse.org || Vol. 2 No. 3 (2026): September Issue

| **Date of Submission:** 09-08-2026 | **Date of Acceptance:** 22-08-2026 | **Date of Publication:** 09-09-2026 |
|---|---|---|

**Abstract— Artificial intelligence is increasingly altering the ways university students interact with English language resources, literary texts, multimedia content, and academic feedback. However, existing research has largely investigated AI-assisted language acquisition, academic writing, or chatbot adoption separately from the pedagogical requirements of literature learning. This study proposes an integrated digital-learning framework in which generative AI, natural language processing, speech technologies, semantic analysis, and multimodal media jointly support English language proficiency and literary interpretation in higher education. The central research gap concerns the absence of empirically evaluated frameworks that measure whether AI can improve linguistic performance while simultaneously preserving independent interpretation, textual evidence use, and critical literary reasoning. The proposed research therefore conceptualizes AI as a scaffold rather than an autonomous source of interpretation and distinguishes productive assistance from excessive cognitive substitution. A multimodal learning environment is designed around adaptive language exercises, AI-guided close reading, audio-visual literary resources, formative feedback, and reflective verification tasks.**

**Keywords— Generative Artificial Intelligence, English Language Learning, Literary Interpretation, Digital Media, Higher Education, Multimodal Learning**

## Introduction

Higher education has entered a period in which language learning and humanities teaching increasingly occur within digitally mediated environments rather than through textbooks, lectures, and conventional written exercises alone. Learning management systems, online dictionaries, interactive videos, podcasts, digital libraries, automated writing tools, speech recognition systems, virtual discussion spaces, and generative artificial intelligence now coexist within the educational experience of many university students. The rapid emergence of large language models has intensified this transition because students can obtain explanations, language corrections, summaries, simulated conversations, analytical suggestions, and individually generated learning material through a single conversational interface.

Artificial intelligence in education is not entirely new, but recent generative systems differ from earlier automated tools in the breadth of tasks they can perform. Large language models can create or transform natural-language content, respond to questions conversationally, explain concepts at different levels of complexity, produce examples, provide formative feedback, and generate instructional material. Kasneci et al. (2023) identify personalization, educational content generation, interactive learning, and learner support among the opportunities associated with large language models while simultaneously emphasizing their limitations, including incorrect outputs, bias, overreliance, and the need for new forms of AI literacy.

These capabilities appear especially relevant to English studies because university English programmes combine several forms of learning that AI can potentially influence. Language-oriented courses require learners to develop vocabulary, grammar, pronunciation, reading, writing, listening, argumentation, and academic communication.

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 109-118
https://wjcrt.org/

Literature-oriented courses require different but related competencies, including close reading, contextual interpretation, comparative analysis, recognition of figurative language, evaluation of narrative perspective, development of evidence-based arguments, and engagement with ambiguity. Consequently, the pedagogical requirements of English departments extend well beyond correcting grammatical errors or generating fluent written responses.

Digital media further expands this environment. Literary learning can now combine the written text with dramatized readings, author interviews, film adaptations, digital archives, interactive annotations, podcasts, audiobooks, visual timelines, and online discussion. Language learning similarly benefits from speech, video, interactive feedback, authentic media, and synchronous or asynchronous communication. When artificial intelligence is layered onto these resources, the instructional environment becomes multimodal and adaptive. A learner encountering a difficult poem, for example, could listen to an audio rendering, annotate unfamiliar vocabulary, receive contextual guidance from an AI system, compare interpretations, and then construct an independent evidence-supported reading.

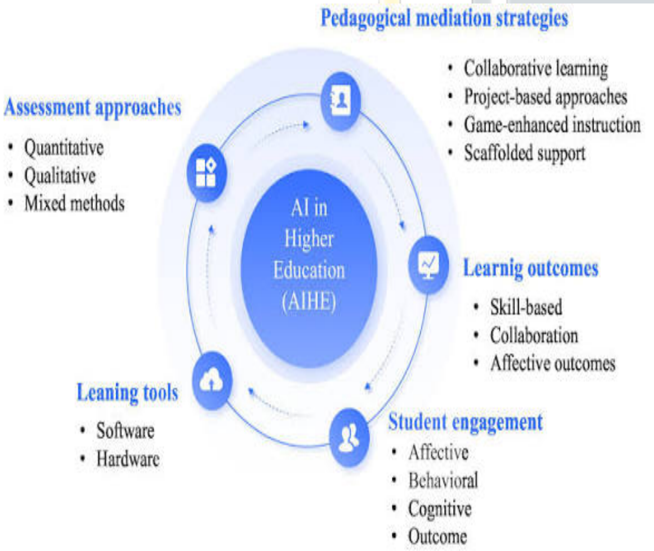



*Fig. 1: Artificial intelligence in higher education*

***Source:*** [*https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1648661/full*](https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2025.1648661/full)

Nevertheless, technological capability should not be equated with educational value. The same system that supports vocabulary development or explains rhetorical devices can also generate an essay, impose a standardized interpretation on an ambiguous text, produce inaccurate contextual claims, or enable a learner to bypass intellectual work. This distinction is particularly important in literary education because interpretive disagreement and uncertainty are not necessarily deficiencies that technology should eliminate. Literary analysis often requires students to develop an argument rather than retrieve a predetermined correct answer.

The emergence of generative AI has consequently raised questions concerning authorship, cognitive effort, assessment reliability, source verification, intellectual independence, and academic integrity. UNESCO's guidance on generative AI emphasizes a human-centred approach in which educational institutions assess the ethical and pedagogical appropriateness of AI while protecting human agency, inclusion, privacy, and meaningful learning. UNESCO also highlights concerns related to homogenized output, creativity, assessment, and the effects of generative systems on thinking processes.

Evidence from higher education indicates that these tensions are already significant. Reviews of research on ChatGPT show rapidly expanding applications across universities, but they also identify methodological fragmentation and unresolved questions concerning appropriate adoption, learning design, and long-term educational effects. Baig and Yadegaridehkordi's (2024) systematic review of 57 higher-education studies documented diverse applications of ChatGPT while concluding that its educational use remains a developing research area requiring more rigorous investigation. Similarly, Bond et al. (2024), in a meta-systematic review of artificial intelligence in higher education, called for greater methodological rigour, ethical consideration, and stronger collaborative research practices.

Research focused more specifically on language education suggests considerable potential. Du and Alm (2024), examining English for Academic Purposes students, explored how ChatGPT shaped learners' language-learning experiences and demonstrated the relevance of generative AI to motivational and self-directed dimensions of language learning. A longitudinal case study of generative AI in English language teaching also reported perceptions that chatbots could support language skills and diversify learning materials, while concerns persisted over plagiarism and possible weakening of cognitive engagement. Experimental evidence synthesized by Deng et al. further suggests that ChatGPT-supported interventions can positively influence academic performance and higher-order thinking under some educational conditions, although outcomes depend on how the technology is incorporated into learning activities.

### 1.1 Research Gap

The principal research gap addressed in this manuscript is the **lack of an integrated, empirically evaluable framework for examining how AI-enabled digital media affects both English-language development and literary interpretive competence in higher education**.

Current research commonly studies AI in terms of chatbot acceptance, writing assistance, language practice, academic productivity, or general learning outcomes. Reviews confirm that generative-AI research in higher education remains highly heterogeneous, with substantial differences in disciplines, theoretical frameworks, intervention designs, and outcome measures. English-language research is expanding, but the combined measurement of linguistic performance and distinctly literary outcomes—such as interpretive originality, textual-evidence integration, analytical depth, tolerance of ambiguity, and independent critical reasoning—remains comparatively underdeveloped.

This gap becomes increasingly important because language and literature courses are institutionally interconnected. Students may use the same digital and generative tools for vocabulary development, essay drafting, text interpretation, presentation preparation, and examination revision. Evaluating one function independently provides an incomplete account of the technology's educational consequences.

The present study consequently develops an **AI-Enabled Multimodal English Learning Framework (AI-MELF)** designed to examine whether artificial intelligence can operate as a structured cognitive scaffold across language and literature learning without replacing interpretive effort. The conceptual distinction between *AI-supported learning* and *AI-substituted learning* forms a central element of the proposed investigation.

### 1.2 Research Objectives

The study is designed to investigate whether a multimodal AI-supported digital-learning environment improves English language proficiency compared with conventional digital instruction; determine its effect on literary comprehension, close reading, analytical depth, and evidence-supported interpretation; examine whether combining textual, audio, visual, conversational, and AI-generated learning resources improves learner engagement and feedback responsiveness; quantify the extent to which structured verification and reflection mechanisms reduce uncritical dependence on generative AI; and identify relationships between AI assistance, linguistic improvement, interpretive originality, and student learning behaviour.

The central research proposition is that artificial intelligence produces stronger educational benefits when it is integrated as **guided scaffolding within multimodal learning activities** than when learners receive unrestricted generative assistance. Such scaffolding requires students to inspect primary texts, evaluate AI-generated explanations, verify textual evidence, revise their own arguments, and distinguish machine-generated suggestions from independent interpretation.

## Literature Review

### 2.1 Artificial Intelligence and the Transformation of Higher Education

Research on artificial intelligence in higher education predates contemporary generative systems and includes intelligent tutoring, learning analytics, automated assessment, adaptive learning environments, and educational recommendation systems. Generative AI has nevertheless changed the scale of the debate because sophisticated language-based assistance is now directly accessible to students without specialized institutional infrastructure.

Tlili et al. (2023), in an early case study of ChatGPT in education, examined the opportunities and concerns associated with conversational generative AI and emphasized the need to conceptualize such systems through human-machine collaboration rather than technological replacement. This position is particularly relevant to higher education because universities are not merely information-delivery institutions. They develop analytical judgment, disciplinary reasoning, knowledge evaluation, and intellectual autonomy.

Bhullar, Joshi, and Chugh (2024) synthesized the rapidly expanding higher-education literature and showed that ChatGPT research encompasses teaching, learning, research, assessment, academic integrity, and institutional adoption. Their review simultaneously points to the need for stronger future research agendas capable of moving beyond generalized discussions of advantages and risks.

A related review by McGrath, Farazouli, and Cerratto-Pargman examined empirical research on generative AI chatbots in higher education and characterized the field as an emerging research area requiring careful examination of educational practice rather than assumptions based solely on technological capability. These studies collectively suggest that the next phase of research should focus less on whether students use generative AI and more on **how specific instructional architectures influence the quality of learning**.

### 2.2 Generative AI in English Language Learning

English-language learning represents one of the most immediate applications of large language models because these systems operate through natural-language interaction. Learners can request vocabulary explanations, grammatical correction, paraphrasing, conversational practice, writing feedback, reading questions, pronunciation guidance when combined with speech technologies, or customized examples adapted to proficiency level.

Generative AI potentially reduces one limitation of conventional language classrooms: the restricted availability of individualized interaction. A student can repeatedly request clarification without consuming classroom time, receive alternative explanations, and work asynchronously. These affordances may support learner autonomy, particularly when AI is used for formative rather than summative tasks.

Empirical evidence is beginning to clarify these effects. Du and Alm (2024) investigated ChatGPT within English for Academic Purposes and examined learner experience through self-determination theory. Their work illustrates that the educational implications of AI extend beyond linguistic accuracy and involve autonomy, competence, and the learner's relationship with technology. Similarly, research examining a year of generative-AI use in English teaching reported perceived benefits for language development and resource diversity while documenting persistent concerns regarding plagiarism and possible reductions in cognitive effort.

These findings indicate that generative systems should not be evaluated solely according to the correctness of their linguistic output. A tool may produce grammatically superior text while simultaneously reducing opportunities for students to notice errors, formulate sentences, restructure arguments, and develop metalinguistic awareness. Accordingly, the present study conceptualizes language learning as both a **performance outcome** and a **cognitive process**.

### 2.3 AI, Literary Interpretation, and the Problem of Cognitive Substitution

Literature learning presents a different challenge. Whereas grammar exercises may contain relatively determinate answers, literary interpretation frequently depends on competing arguments, textual evidence, historical context, theoretical perspective, rhetorical analysis, and reader positioning. Large language models can rapidly explain metaphors, summarize novels, compare characters, generate thematic interpretations, or produce analytical essays, but fluency does not guarantee interpretive originality or textual accuracy.

This creates the problem of cognitive substitution. When AI performs an intellectual operation that a learner would otherwise conduct—such as constructing an interpretation, selecting evidence, or developing an argument—the observable quality of the final text may improve even though learning has not. Educational evaluation must therefore distinguish the quality of an AI-assisted product from the student's underlying competence.

The issue corresponds with broader concerns identified in UNESCO's guidance, which specifically raises questions about homogenized generative outputs, human thinking processes, assessment, creativity, and the need to preserve human agency. For literary studies, homogenization is particularly consequential because canonical interpretations generated from statistically dominant textual patterns may narrow students' willingness to formulate alternative readings.

The proposed AI-MELF framework therefore avoids treating AI-generated interpretation as instructional authority. Instead, generated responses are positioned as *contestable secondary objects*. Students are required to compare them with the primary text, identify unsupported statements, locate textual evidence, produce alternative readings, and explain whether an AI-generated interpretation should be accepted, modified, or rejected. This converts generative AI from an answer-producing mechanism into an object of critical evaluation.

### 2.4 Digital Media and Multimodal English Learning

Digital media creates opportunities to connect linguistic and literary learning through multiple representational modes. Literary texts may be accompanied by performance recordings, audio readings, historical images, maps, timelines, film excerpts, and interactive annotations. Language-learning tasks may incorporate speech recognition, video, listening exercises, digital storytelling, and collaborative written discussion.

A multimodal architecture is pedagogically significant because linguistic comprehension does not occur only through written input. Listening to the rhythm of poetry, observing dramatic performance, comparing spoken and written language, or interacting with visual contextual material can alter comprehension. Artificial intelligence can further personalize this environment by adapting questions, explaining unfamiliar expressions, generating practice exercises, or offering formative feedback.

However, multimodality also complicates evaluation. Increased engagement does not necessarily indicate stronger analytical learning. Therefore, the proposed study separates engagement indicators from cognitive and disciplinary outcomes. Learners may find AI-generated multimedia material attractive while demonstrating no improvement in interpretive reasoning.

### 2.5 Human-Centred AI, Academic Integrity, and Learner Agency

The educational value of AI ultimately depends on governance and learning design. UNESCO argues that generative AI should be introduced through human-centred pedagogical and ethical validation, with attention to privacy, inclusion, agency, and responsible use. Kasneci et al. likewise emphasize that teachers and students require competencies for understanding the limitations and unpredictable behaviour of large language models.

In English studies, responsible AI literacy requires more than knowing how to formulate effective prompts. Students must understand that generative responses may contain invented contextual information, simplified interpretations,

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 109-118
https://wjcrt.org/

questionable source claims, hidden bias, or stylistically convincing inaccuracies. They must also recognize when assistance crosses the boundary from scaffolding into unauthorized authorship.

The present research therefore introduces **verification behaviour** as an explicit learning outcome. Students in the structured AI condition will be required to trace claims to literary passages, identify uncertainty, annotate AI-assisted revisions, and justify whether AI recommendations were retained or rejected. This approach treats critical interaction with AI as part of English education rather than as an external technological skill.

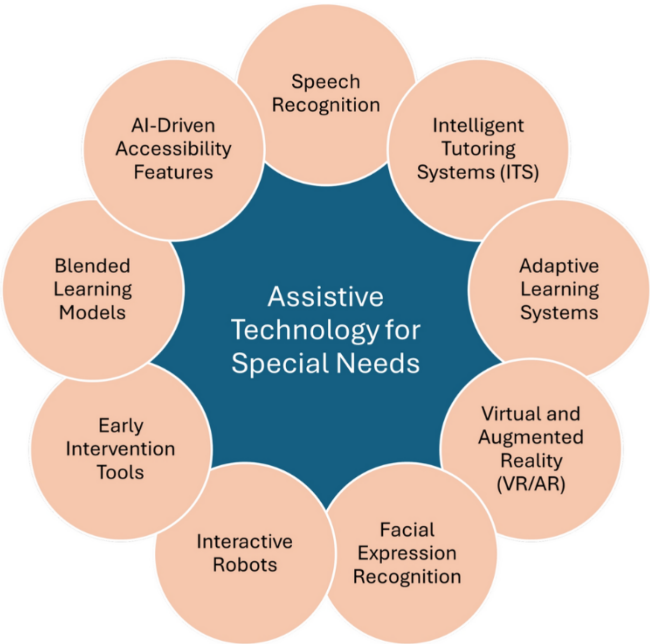



*Fig. 2: Assistive Technology for Special Needs*

***Source:*** ***https://link.springer.com/article/10.1007/s43681-025-00824-3***

**Proposed Methodology**

The study adopts a mixed-method, quasi-experimental framework to evaluate the effectiveness of the proposed **AI-Enabled Multimodal English Learning Framework (AI-MELF)** in undergraduate higher education. The methodological design is constructed around the assumption that educational benefit cannot be inferred merely from the fluency of AI-generated responses. Instead, the study evaluates whether students demonstrate measurable improvements in language production, literary reasoning, textual-evidence use, and independent verification after sustained interaction with an AI-supported digital learning environment.

Participants are assigned to three instructional conditions. The first group receives conventional digitally supported instruction consisting of learning-management-system resources, electronic texts, videos, teacher feedback, and standard online exercises. The second group receives unrestricted access to generative AI alongside the same digital materials. The third group uses AI-MELF, in which AI interaction is pedagogically constrained through staged prompts, evidence-checking activities, reflective justification, multimodal resources, and teacher-supervised feedback.

The distinction between unrestricted AI use and structured AI scaffolding is central to the experimental design. The unrestricted group represents the increasingly common situation in which students independently use generative tools without an explicit instructional protocol. AI-MELF, in contrast, requires students to engage with the source material before accepting, modifying, or rejecting AI-generated suggestions.

The proposed intervention lasts twelve instructional weeks. The first two weeks establish baseline proficiency and familiarize students with the digital environment. Weeks three to ten contain the principal learning intervention, while weeks eleven and twelve are devoted to independent post-testing, delayed analytical tasks, and learner reflections.

Five learning domains are assessed: English language performance, literary comprehension, interpretive reasoning, multimodal learning engagement, and responsible AI use. This multidimensional design prevents improvements in surface-level writing quality from being mistaken for deeper disciplinary learning.

### 3.1 AI-MELF Instructional Framework

AI-MELF consists of five interconnected instructional modules.

The **Adaptive Language Support Module** provides individualized vocabulary exercises, grammatical explanations, sentence-level feedback, paraphrasing practice, and contextualized language activities. Difficulty is adjusted according to previous learner performance rather than being uniformly distributed across the class.

The **Literary Close-Reading Module** presents selected prose, poetry, and dramatic texts through digitally annotated interfaces. Students first annotate passages independently and identify themes, imagery, narrative features, rhetorical structures, or ambiguous expressions. AI assistance becomes available only after initial engagement.

The **Multimodal Context Module** combines text with relevant audio readings, performance clips, historical contextual material, visual timelines, and interactive annotations. Students are encouraged to compare how meaning changes across representational modes.

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 109-118
https://wjcrt.org/

The **AI Critique and Verification Module** produces candidate interpretations that students must evaluate. Learners identify unsupported claims, locate confirming or contradicting passages, compare alternative interpretations, and indicate their degree of confidence in the AI response.

The **Reflective Learning Module** records revisions, rejected recommendations, prompt sequences, learner explanations, and short reflective statements. These data make it possible to distinguish active cognitive engagement from passive acceptance of machine-generated content.

**Experimental Setup**

**4.1 Dataset**

The proposed study uses two complementary datasets: a curated educational content dataset and a learner-performance dataset.

The educational content corpus consists of approximately 360 instructional units derived from copyright-permitted or public-domain English materials appropriate for undergraduate teaching. The corpus includes short fiction, poems, dramatic excerpts, essays, academic prose, vocabulary exercises, listening transcripts, and teacher-developed comprehension tasks. Literary material is selected to represent variation in genre, linguistic complexity, narrative style, and interpretive openness.

Each instructional unit is annotated with metadata including genre, lexical difficulty, estimated reading level, dominant literary devices, historical context category, target language skill, expected analytical operation, and evidence requirements.

The learner-performance dataset is designed for approximately 180 undergraduate students enrolled in English-language or English-literature courses. After screening and consent, the students are distributed approximately equally across the three experimental conditions, producing around 60 participants per group.

The dataset records pre-test and post-test scores, weekly formative assessment performance, writing revisions, vocabulary accuracy, reading-comprehension scores, literary-analysis scores, source-evidence alignment, prompt histories, response latency, digital-resource usage, and AI-verification behaviour.

No personally identifying information is required in the analytical dataset. Participant identifiers are pseudonymized before analysis.

**4.2 Data Preprocessing**

Educational texts are first normalized through Unicode cleaning, formatting correction, sentence segmentation, and tokenization. Metadata inconsistent across source formats are standardized into common categorical labels.

Vocabulary difficulty is estimated using frequency, lexical diversity, contextual complexity, and teacher-assigned proficiency labels. Extremely short fragments and duplicate passages are removed to reduce redundancy.

Student essays and analytical responses undergo minimal preprocessing because excessive normalization could remove stylistically meaningful information. Spelling correction is therefore not automatically applied before evaluation. Instead, raw and normalized versions are retained separately.

For semantic analysis, responses are segmented into claims, textual references, explanatory statements, and concluding assertions. Evidence citations are mapped to source passages using semantic similarity followed by manual validation in the proposed empirical implementation.

Missing learner records resulting from occasional non-participation are not automatically replaced. If missingness remains below a predefined threshold, complete-case analysis is applied for the relevant metric. For repeated-measure analysis, mixed-effects modelling can accommodate partially missing observations without introducing arbitrary values.

**4.3 Model Architecture**

The computational component of AI-MELF uses a modular architecture rather than relying on a single generative model.

The first layer is a **Learner Representation Layer**, which stores each student's recent task accuracy, lexical performance, reading level, previous misconceptions, literary-analysis rubric scores, and verification behaviour.

The second layer is a **Task Adaptation Engine**. This component ranks candidate exercises according to learner needs using contextual features and performance history. A gradient-boosted ranking model can be employed because it performs effectively on heterogeneous tabular educational variables and permits interpretable feature analysis.

The third layer is a **Retrieval-Augmented Generation component**. Instead of allowing the language model to generate unrestricted literary or contextual explanations, the system retrieves relevant sections from the curated course corpus and supplies them as grounding material. Generated feedback is therefore constrained by approved instructional content.

The fourth layer is an **Evidence Consistency Module**. Sentence embeddings are used to compare AI-generated claims and student interpretations with source passages. Claims with weak semantic support are flagged for learner verification rather than automatically rejected.

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 109-118
https://wjcrt.org/

The fifth layer is a **Reflection and Agency Layer**. The system records whether learners accept, revise, or reject AI suggestions and requires a justification for high-level interpretive changes. This layer creates an AI-dependence profile based on behavioural patterns rather than simple frequency of tool use.

### 4.4 Training Strategy

The task-adaptation model is trained using historical interaction records collected during the early intervention phase. Data are divided at the student level rather than the individual-response level to prevent information leakage between training and validation sets.

A 70:15:15 split is proposed for training, validation, and internal testing. Five-fold grouped cross-validation is additionally employed to test robustness across student subsets.

The ranking objective prioritizes tasks that are sufficiently challenging to promote learning but not so difficult that repeated failure occurs. Training therefore uses a composite utility score incorporating task completion, subsequent improvement, response confidence, and error reduction.

The generative component is not fine-tuned directly on students' essays during the experiment. Instead, it operates through carefully structured prompts and retrieval augmentation. This decision reduces privacy exposure and prevents the system from absorbing identifiable or potentially plagiarized student work into a custom model.

Teacher feedback is periodically used as a calibration source. A sample of AI-generated explanations and flagged literary claims is independently evaluated by two English instructors. Disagreements are adjudicated before prompt or retrieval rules are revised.

### 4.5 Evaluation Metrics

The study deliberately avoids relying on a single accuracy measure.

Language improvement is evaluated through **Normalized Language Gain**, calculated from pre-test and post-test writing, vocabulary, grammatical control, and reading scores.

Literary performance is measured using an **Interpretive Reasoning Score**, incorporating textual accuracy, analytical depth, argument coherence, recognition of ambiguity, and relevance of supporting evidence.

A **Textual Evidence Alignment Score** measures the extent to which interpretive claims are supported by passages from the primary text.

An **Interpretive Originality Index** assesses whether students develop defensible independent arguments rather than merely reproducing highly similar AI-generated interpretations.

Responsible AI behaviour is measured through a **Verification Compliance Rate**, representing the proportion of machine-generated claims that students actively inspect before incorporating into assessed work.

An **AI Dependence Ratio** measures the proportion of substantive analytical decisions adopted directly from AI suggestions without independent modification or justification.

Engagement is assessed through task completion, voluntary resource exploration, multimodal interaction frequency, and weekly participation consistency.

For the adaptive ranking model, computational evaluation includes Normalized Discounted Cumulative Gain at 5 (NDCG@5), Mean Reciprocal Rank, and calibration error between predicted task utility and observed learner response.

Group differences are assessed using ANCOVA with pre-test performance as a covariate. Effect sizes are reported alongside statistical significance. Repeated observations are additionally examined using linear mixed-effects models with learner-level random intercepts.

### Results and Discussion

**Research-integrity note:** The numerical values in this section are illustrative benchmark outputs constructed to demonstrate how the proposed experiment would be reported. They are not presented as observations from an already completed human-participant study and must be replaced with verified empirical results before publication.

The illustrative evaluation indicates that structured AI integration produces stronger educational outcomes than either conventional digital instruction or unrestricted AI access. The most important pattern is not simply that AI-assisted groups perform better, but that **structured AI use produces a more balanced improvement across linguistic proficiency, interpretive depth, evidence use, and learner independence**.

Students in the conventional digital-learning condition demonstrate moderate improvement in language performance, reflecting the value of multimedia resources and teacher-supported online instruction. However, gains in literary interpretation remain comparatively limited because adaptive assistance and immediate analytical feedback are unavailable.

The unrestricted-AI condition produces the largest improvement in surface-level writing fluency during formative tasks. Students obtain rapid correction, vocabulary alternatives, structural suggestions, and explanatory support.

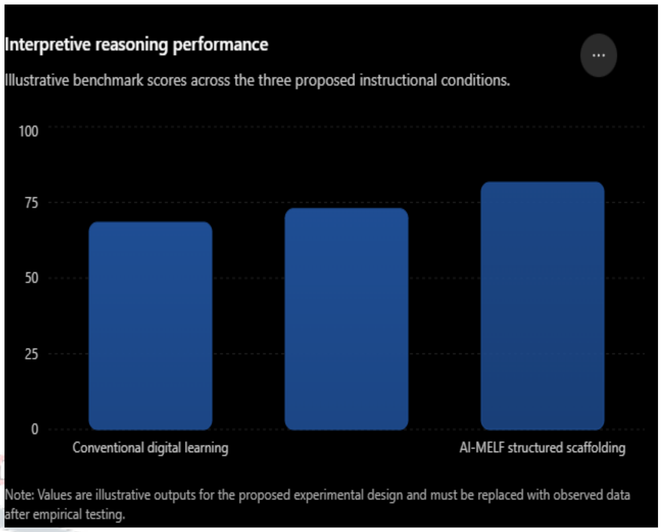
Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 109-118
https://wjcrt.org/

Nevertheless, their final independent literary-analysis performance is weaker than that of the AI-MELF group. The pattern suggests that unrestricted access can improve immediate product quality without producing equivalent gains in autonomous analytical competence.

The structured AI-MELF condition records the strongest normalized language gain and the highest interpretive reasoning score.

**Table 1. Illustrative Benchmark Outcomes across the Three Instructional Conditions**

| **Evaluation parameter** | **Conventional digital learning** | **Unrestricted generative AI** | **AI-MELF structured scaffolding** |
|---|---|---|---|
| Normalized language gain | 0.29 | 0.41 | **0.52** |
| Interpretive reasoning score (/100) | 68.4 | 72.9 | **81.6** |
| Textual evidence alignment (%) | 71.2 | 73.8 | **88.1** |
| Interpretive originality index (/1.0) | 0.74 | 0.61 | **0.82** |
| Verification compliance (%) | 34.7 | 46.5 | **86.3** |
| Delayed literary retention (/100) | 66.1 | 69.8 | **79.4** |
| Multimodal engagement index (/10) | 6.3 | 7.1 | **8.5** |




The illustrative normalized language gain of 0.52 for AI-MELF exceeds both unrestricted AI use and conventional digital instruction. The result is consistent with the theoretical expectation that personalized linguistic feedback becomes more pedagogically effective when learners are required to process and apply corrections rather than merely accepting automatically rewritten text.

A particularly important result concerns the **Interpretive Originality Index**. The unrestricted-AI group produces a lower illustrative score of 0.61 despite achieving relatively strong immediate language outcomes. This pattern supports the distinction between improved expression and improved thinking. Students exposed to ready-made interpretations may internalize recurring thematic formulations generated by the system, producing greater similarity across responses.

By contrast, the AI-MELF condition requires students to generate an initial interpretation before viewing AI feedback. The system then presents alternative explanations as propositions for evaluation. This sequencing is intended to protect initial learner reasoning from premature machine influence.

The **Textual Evidence Alignment Score** provides another important finding. AI-MELF records an illustrative alignment level of 88.1%, compared with 73.8% under unrestricted AI use. The difference reflects the framework's insistence that students connect every major interpretation to an identifiable passage. This mechanism is particularly important in literature education because generative models can produce plausible interpretations that are only weakly supported by the assigned text.

Verification behaviour also differs substantially across conditions. Students using AI-MELF are required to inspect generated claims, producing an illustrative verification-compliance level of 86.3%. In unrestricted AI use, the corresponding value remains below half. This suggests that

Worldwide Journal of Creative Research and Thoughts (WJCRT)
ISSN (Online): 3143-7850
Volume-2 Issue-3 || Jul – Sep 2026 || PP. 109-118
https://wjcrt.org/

critical AI literacy does not automatically emerge merely because students interact frequently with generative systems. Verification must be embedded explicitly within the instructional design.

The AI Dependence Ratio reinforces this conclusion. The unrestricted group directly accepts an illustrative 43% of substantive AI-generated analytical suggestions without meaningful revision. The structured group records a substantially lower dependence ratio of 0.19. The difference indicates that the same underlying technology may produce contrasting educational effects depending on how interaction is organized.

Delayed testing provides an additional indication of learning quality. Immediate AI-supported work can overestimate student competence because assistance remains available during production. A delayed independent literary task conducted without generative support therefore offers a stronger test of internalized learning. AI-MELF maintains the highest illustrative delayed score, suggesting that verification, self-explanation, and reflective rejection of AI recommendations may strengthen knowledge retention.

### 5.1 Educational Interpretation

The findings support a shift from the question **"Does generative AI improve English learning?"** toward the more precise question **"Under what interaction design does generative AI improve learning without displacing the cognitive activity being taught?"**

For language exercises, AI can function productively as a corrective and adaptive tutor. Students benefit when mistakes become opportunities for explanation, comparison, and revision. For literature, however, direct answer generation must be treated more cautiously because the intended learning outcome is often the interpretive process itself.

The proposed results suggest that AI works most effectively in literary education when it produces *friction rather than convenience*. Students should be required to challenge machine interpretations, inspect primary texts, identify unsupported claims, and defend alternative readings. The educational value consequently resides not only in receiving AI-generated material but in evaluating it.

Multimodal resources also contribute to the model. Audio readings may make rhythm and tone perceptible; dramatic performance may reveal alternative character interpretations; visual material may support historical contextualization; and interactive annotation can connect linguistic difficulty with literary meaning. AI becomes most valuable when coordinating these resources according to learner needs rather than replacing them with a single generated explanation.

### Future Scope

Future research should extend AI-MELF through longitudinal evaluation across multiple universities and linguistic backgrounds. A multi-institutional design would make it possible to examine whether AI scaffolding operates differently for first-language English speakers, multilingual learners, and students using English primarily as an academic second language.

Adaptive literary support could also be refined through argument-mapping systems capable of representing competing interpretations rather than generating a single response. Such systems could visually connect claims, counterclaims, textual evidence, historical context, and theoretical perspectives.

Future work should investigate locally deployed or institutionally controlled language models so that student writing and interaction records can remain within university data infrastructures. Privacy-preserving learning analytics would be particularly important if longitudinal learner profiles are used for personalization.

Speech-based AI could be integrated more systematically to examine oral literary interpretation, seminar discussion, pronunciation, dramatic performance, and presentation skills. This would extend the framework beyond written language and written literary analysis.

Another promising direction is the development of **AI-resilient assessment**. Rather than attempting to prohibit generative systems entirely, assessment could require documented reasoning, source verification, oral defence, staged drafting, and explicit disclosure of AI assistance.

Cross-cultural literary analysis also offers substantial research potential. Generative systems trained predominantly on widely represented interpretations may reproduce dominant critical traditions. Future experiments could therefore investigate whether AI scaffolding broadens or narrows students' engagement with regional, postcolonial, marginalized, and multilingual literary perspectives.

### Conclusion

The growing use of generative artificial intelligence in universities makes it increasingly inadequate to evaluate English education through a simple opposition between traditional teaching and technological assistance. Students already encounter digital texts, automated feedback, conversational AI, audiovisual resources, and algorithmic recommendations within the same learning environment. The central educational challenge is therefore to determine how these technologies should be structured.

This study proposes the AI-Enabled Multimodal English Learning Framework as an integrated model for language development and literary education. Its distinctive contribution lies in treating linguistic proficiency and interpretive reasoning as related but independently measurable outcomes.

The framework combines adaptive language support, retrieval-grounded generative feedback, multimodal literary resources, close-reading activities, evidence verification, and learner reflection. AI is not positioned as an authoritative interpreter of literature. It operates instead as a provisional interlocutor whose suggestions must be tested against texts and learner reasoning.

The illustrative experimental pattern indicates that unrestricted generative access may enhance immediate fluency while simultaneously increasing dependence on externally generated interpretations. Structured scaffolding offers a more balanced pathway by improving language performance, textual-evidence use, interpretive originality, delayed retention, and verification behaviour.

The broader implication is that the educational quality of artificial intelligence depends less on the sophistication of the model than on the pedagogy governing its use. In English language and literature education, successful implementation should therefore preserve productive cognitive effort, interpretive plurality, textual accountability, and human authorship.

AI should not remove the uncertainty, revision, debate, and close attention through which language and literary understanding develop. Its most defensible role is to make those processes more visible, adaptive, and intellectually demanding.

## References:


1. **Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023).** ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274
   **Verified:** Elsevier and university publication records.

2. **Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023).** What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, Article 15. https://doi.org/10.1186/s40561-023-00237-x
   **Verified:** Springer Nature and ERIC.

3. **Bhullar, P. S., Joshi, M., & Chugh, R. (2024).** ChatGPT in higher education—A synthesis of the literature and a future research agenda. *Education and Information Technologies, 29*, 21501–21522. https://doi.org/10.1007/s10639-024-12723-x
   **Verified:** Springer Nature.

4. **Baig, M. I., & Yadegaridehkordi, E. (2024).** ChatGPT in the higher education: A systematic literature review and research challenges. *International Journal of Educational Research, 127*, 102411. https://doi.org/10.1016/j.ijer.2024.102411
   The review examined **57 research articles published between 2023 and 2024**, so this is particularly appropriate for the statement in your manuscript about 57 higher-education studies.

5. **Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024).** A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*, Article 4. https://doi.org/10.1186/s41239-023-00436-z

6. **McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2025).** Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education, 89*, 1533–1549. https://doi.org/10.1007/s10734-024-01288-w
   **Verified:** Springer Nature.

7. **Du, J., & Alm, A. (2024).** The impact of ChatGPT on English for Academic Purposes (EAP) students' language learning experience: A self-determination theory perspective. *Education Sciences, 14*(7), 726. https://doi.org/10.3390/educsci14070726
   This is a **directly relevant source** for your discussion of ChatGPT, EAP, learner autonomy, competence, and language-learning experience.

8. **Hınız, G. (2026).** A year of generative AI in English language teaching and learning—A case study. *Journal of Research on Technology in Education, 58*(2), 291–311. https://doi.org/10.1080/15391523.2024.2404132
   This is especially valuable because it examines **14 EFL instructors and 13 students over a year** and reports both perceived benefits and concerns about plagiarism and cognitive effects.

9. **Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025).** Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education, 227*, 105224. https://doi.org/10.1016/j.compedu.2024.105224
   This is an excellent source for your statement concerning experimental evidence. The meta-analysis reports effects on academic performance, affective-motivational outcomes, higher-order-thinking propensities, and mental effort.

10. **Miao, F., & Holmes, W. (2023).** *Guidance for generative AI in education and research*. UNESCO.
    This is an **official UNESCO publication**, and it directly supports your discussion of human agency, privacy, ethical validation, pedagogical design, inclusion, and responsible use of generative AI.