

From Automated Writing Evaluation to Generative AI: Technology-Integrated Academic Writing Support in Engineering Education

M. Hepsiba George^{1*}, Dr. Kranthi Priya Oruganti²

¹Assistant professor, Department of English, Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana, India. ORCID : 0009-0004-9587-0494, 19320014@klh.edu.in

²Assistant professor, Department of English, Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana, India. ORCID: 0000-0001-6656-6822, kranthi.priya@klh.edu.in

Abstract

Academic writing is increasingly important in engineering education because students must communicate technical processes, evidence, design decisions, and research findings to academic and professional audiences. At the same time, automated writing evaluation (AWE), intelligent tutoring systems, collaborative platforms, and generative artificial intelligence (GenAI) are expanding computational support for planning, drafting, feedback, and revision. This review synthesizes evidence on technology-integrated academic writing support with engineering education as the principal application context. The analysis is informed by PRISMA 2020 principles and focuses on technology functions, targeted writing dimensions, feedback arrangements, and implementation constraints. Published meta-analyses show that automated feedback can improve writing performance, although effects are heterogeneous and should not be interpreted as technology-independent. Across the literature, lower-level features such as grammar and mechanics are more consistently supported than higher-order dimensions such as argumentation, disciplinary accuracy, and rhetorical effectiveness. The review therefore proposes a human-in-the-loop framework in which student agency, technology support, teacher and peer guidance, feedback literacy, and responsible AI use jointly support independent revision. The contribution is a process-oriented framework for engineering education that treats technology as scaffolding rather than a substitute for authorship, disciplinary judgment, or learning.

Keywords—Automated Writing Evaluation; Generative Artificial Intelligence; Engineering Education; Academic Writing; Feedback Literacy; Human-in-the-Loop.

1. Introduction

Writing is a core engineering competence because engineers must explain technical processes, report results, justify decisions, document designs, and communicate with audiences that differ in expertise. Writing support is therefore not limited to grammar correction. It also involves organization, coherence, evidence use, technical precision, rhetorical awareness, and revision. These requirements create a practical instructional problem: detailed feedback across multiple drafts is time-intensive, particularly in large classes.

Computational writing technologies have expanded from word processors and grammar checkers to automated writing evaluation (AWE), intelligent tutoring systems, online annotation, collaborative writing environments, and large language model (LLM)-based assistants. AWE systems can provide individualized and timely feedback at scale, while GenAI can support idea exploration, explanation, drafting, and interactive revision. These capabilities make technology attractive for engineering education, but technological fluency does not guarantee disciplinary accuracy or learning.

Evidence on AWE is encouraging but heterogeneous. A multi-level meta-analysis of 20 studies (84 effect sizes; $N = 2,828$) reported a medium overall effect of automated feedback on writing performance ($g = 0.55$) [2]. A second meta-analysis of 26 primary studies involving 2,468 participants reported a large positive overall effect on writing quality ($g = 0.861$) [3]. Such results support the potential of computational feedback while also showing that effectiveness depends on learner, intervention, and outcome characteristics.

Generative AI introduces a different form of support. Unlike conventional systems that primarily detect or score selected textual features, LLM-based tools can participate in dialogue, generate alternatives, explain feedback, and support pre-writing, during-writing, and post-writing activities [4]. These capabilities are valuable, but they also raise concerns about inaccurate information, overdependence, authorship, privacy, and assessment validity. The central educational question is therefore not whether technology can improve a finished text, but how technology can be integrated without displacing the cognitive and disciplinary work through which writing competence develops.

Existing studies are distributed across educational technology, computer-assisted language learning, AWE, AI in education, and disciplinary communication. Engineering-specific evidence remains comparatively fragmented, while many influential findings originate in general higher education or language-learning settings. This review addresses that gap by synthesizing technology functions and pedagogical conditions relevant to engineering writing. It asks: (RQ1) which technologies are used to support academic writing; (RQ2) which writing dimensions and outcomes are targeted; (RQ3) how automated, AI-based, teacher, and peer feedback can be combined; and (RQ4) which challenges and research gaps shape implementation.

The study makes two contributions. First, it consolidates evidence across AWE, intelligent writing support, collaborative technologies, and GenAI into a concise process-oriented synthesis. Second, it proposes a human-in-the-loop framework for engineering education that assigns complementary roles to students, technologies, teachers, and peers. The framework is intended to support independent writing development rather than technology-dependent text production.

II. RELATED WORK AND RESEARCH GAP

From Automated Evaluation to Generative AI

AWE and intelligent tutoring systems demonstrate the computational potential of writing support. Writing Pal combines strategy instruction, game-based practice, essay writing, and automated formative feedback, illustrating that technology can be embedded within a broader instructional design rather than used as a stand-alone corrector [5]. Earlier genre-oriented work also showed that automated evaluation can be tailored to academic writing tasks when computational analysis is connected to rhetorical and disciplinary expectations [6]. Student perceptions of automated grammar feedback also show that usefulness is mediated by how feedback is interpreted and incorporated into revision [7].

Recent evidence nevertheless indicates an imbalance in writing technologies: support for micro-level revision is more established than support for macro-level writing strategies, rhetorical moves, self-monitoring, and discipline-specific research writing. This limitation is particularly relevant to engineering, where technically correct communication depends on more than linguistic form. A grammatically improved sentence may still be technically inaccurate, poorly justified, or unsuitable for its intended audience.

GenAI has intensified this issue. Su et al. [4] describe ChatGPT as a potential evaluator and virtual peer that can be integrated across pre-writing, during-writing, and post-writing stages. However, its flexibility also increases the need for critical evaluation. A useful engineering-writing design must therefore combine computational assistance with disciplinary expertise and explicit responsibility for verification.

B. Research Gap

Three gaps motivate the present review. First, technology-supported writing research is often organized by tool type, making it difficult to understand how multiple technologies can support the connected stages of writing. Second, results from general higher education are frequently applied to engineering without sufficient attention to technical genres, disciplinary vocabulary, data interpretation, and evidence-based explanation. Third, recent AI discussions often emphasize either opportunity or risk, while providing limited guidance on how student agency and human feedback should be retained in instructional design.

Accordingly, the present synthesis uses the writing process as the organizing principle. Technologies are considered according to the functions they perform, the writing dimensions they target, the feedback relationships they enable, and the conditions required for responsible implementation.

III. REVIEW METHOD

This review uses a structured evidence-synthesis approach informed by PRISMA 2020 reporting principles [1]. The topic spans educational technology, automated feedback, artificial intelligence in education, and engineering communication; consequently, the review prioritizes transparent eligibility criteria and thematic synthesis over a single-tool comparison.

Search concepts combined three domains: (1) academic writing and writing development; (2) technology support, including AWE, automated feedback, AI-assisted writing, generative AI, digital feedback, and collaborative writing; and (3) engineering, STEM, and higher education. Representative Boolean structures combined these concepts using OR within domains and AND across domains.

Studies were considered when they reported peer-reviewed evidence relevant to technology-supported writing, feedback, revision, or writing instruction in engineering, STEM, or higher education. Studies were excluded when they focused only on algorithm development without an educational writing outcome, general digital literacy without a writing component, or non-peer-reviewed opinion material. Evidence was interpreted by context: engineering-specific studies were treated as direct evidence; STEM studies as related evidence; and broader higher-education studies as supporting evidence.

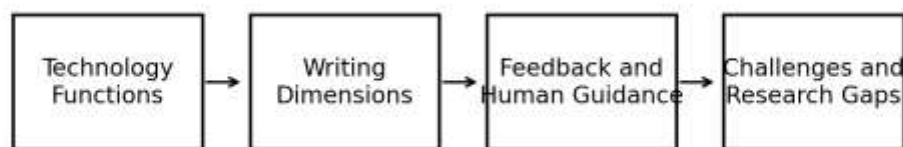
For synthesis, each source was coded by technology function, educational context, targeted writing dimension, feedback source, reported outcome, and implementation challenge. Because the available evidence is heterogeneous

and the manuscript does not reproduce database-specific search logs, no PRISMA numerical counts are fabricated. Quantitative values are reported only where they are explicitly provided by published meta-analyses.

TABLE I. ELIGIBILITY CRITERIA AND SYNTHESIS VARIABLES

Criterion	Included	Excluded
Publication	Peer-reviewed studies and authoritative reviews/meta-analyses	Editorials, opinion pieces, non-scholarly sources
Context	Engineering, STEM, or higher education	Contexts without relevance to higher-education writing
Technology	AWE, automated feedback, intelligent tutoring, collaborative/digital writing, AI or GenAI	Algorithm-only studies without educational writing outcomes
Outcome	Writing quality, revision, feedback use, engagement, self-regulation, or implementation	Studies unrelated to writing development
Synthesis	Technology function, writing dimension, feedback source, outcome, challenge	Insufficient methodological or findings information

Evidence synthesis logic used in the review



Technology functions are interpreted together with targeted skills, feedback arrangements, and implementation conditions.

Fig. 1. Evidence synthesis logic used to organize technology-integrated writing research.

IV. RESULTS AND SYNTHESIS

A. Technology Functions and Targeted Writing Dimensions

The literature supports a functional classification rather than a single category of “writing technology.” Digital environments facilitate drafting and versioning; AWE and language tools primarily support grammar, mechanics, vocabulary, and selected coherence features; intelligent tutoring systems combine feedback with explicit strategy instruction; collaborative platforms enable peer interaction; and GenAI extends support to brainstorming, outlining, explanation, and interactive revision.

The strongest evidence concerns lower-level or readily observable dimensions of writing. Automated systems can repeatedly identify patterns in grammar, spelling, sentence construction, and lexical use. The two published meta-analyses summarized above provide quantitative support for positive effects on writing performance [2], [3]. However, higher-order outcomes are less straightforward because argument quality, disciplinary relevance, originality, and technical accuracy depend strongly on context and human judgment.

For engineering education, this distinction is critical. Technical writing requires precision in describing methods, interpreting results, and linking claims to evidence. Consequently, the same technology may be useful for linguistic revision while remaining insufficient for judging whether a technical explanation is correct. The evidence therefore supports matching technology to the writing problem rather than adopting a tool as a general-purpose solution.

B. Feedback, Revision, and Learner Agency

Feedback is the mechanism through which many writing technologies influence revision. Automated systems provide speed and repeatability; GenAI can provide conversational explanations and alternatives; teachers contribute pedagogical and disciplinary judgment; and peers contribute reader-oriented perspectives. These

sources are complementary rather than interchangeable.

The educational benefit of feedback depends on what students do after receiving it. A correction can improve a draft without improving the student's ability to recognize the same problem later. For this reason, feedback literacy is central: students need to understand a suggestion, evaluate its relevance, compare competing advice, and decide whether and how to revise. In AI-supported environments, these skills also include checking factual accuracy and preserving intended meaning.

Evidence therefore favors iterative workflows rather than one-shot correction. Students should move through planning, drafting, feedback, evaluation, revision, and reflection, with opportunities to return to earlier stages. Technology can make this cycle faster and more visible, but the student remains the decision-maker.

TABLE II. SYNTHESIS OF MAJOR EVIDENCE AND PEDAGOGICAL IMPLICATIONS

Evidence area	Reported value or finding	Implication for engineering writing
Automated feedback meta-analysis [2]	20 studies; 84 effect sizes; $N = 2,828$; $g = 0.55$	Useful evidence for scalable feedback, but effects are heterogeneous.
AWE meta-analysis [3]	26 studies; $N = 2,468$; $g = 0.861$	Positive effects on writing quality; transfer to engineering genres requires contextual testing.
Intelligent tutoring [5]	Combines strategy instruction, practice, and automated formative feedback	Technology is stronger when embedded in instruction rather than used only for correction.
ChatGPT-supported writing [4]	Can support pre-writing, during-writing, and post-writing activities	Interactive support expands possibilities but requires verification and authorship safeguards.
Engineering application	Direct evidence remains fragmented relative to broader higher education	Teacher/disciplinary guidance remains essential for technical accuracy and genre expectations.

C. Challenges and Research Gaps

The principal challenges are not purely technical. Automated and AI-generated feedback may be inaccurate, overly general, or insensitive to disciplinary meaning. Extensive assistance can create learner dependence when students accept corrections without understanding them. GenAI further complicates authorship, academic integrity, privacy, and assessment validity. Unequal access and differences in digital literacy can also produce unequal benefits.

Teacher readiness is another implementation condition. Technology can reduce repetitive feedback tasks, but it creates new responsibilities: setting boundaries for AI use, designing process-based assessment, teaching students to evaluate feedback, and deciding when human disciplinary judgment is required. The resulting instructional shift is from simple tool adoption to feedback orchestration.

The literature also reveals several research gaps. More engineering-specific studies are needed; higher-order and discipline-specific outcomes require stronger measurement; longitudinal work is necessary to distinguish temporary text improvement from sustained competence; and integrated designs using multiple feedback sources require empirical evaluation. These gaps motivate the framework proposed in the next section.

V. DISCUSSION AND PROPOSED HUMAN-IN-THE-LOOP FRAMEWORK

The synthesis indicates that technology should be evaluated as part of a writing system rather than as an isolated intervention. AWE, intelligent tutoring, collaborative tools, and GenAI perform different functions. Their educational value depends on the stage of writing, the nature of the task, the quality of feedback, and the degree of student engagement. This explains why positive average effects in meta-analyses coexist with substantial heterogeneity [2], [3].

The proposed framework therefore places student agency at the center. Technology provides support for idea exploration, language revision, feedback, and reflection. Teacher and peer guidance provide contextual, disciplinary, and reader-oriented perspectives. Feedback literacy connects these sources by requiring students to interpret, verify, compare, and apply suggestions. Responsible practice establishes boundaries through authorship, transparency, privacy, and information verification.

The model is particularly suitable for engineering education because technical accuracy cannot be inferred from linguistic fluency alone. Automated tools may flag surface problems efficiently, while teachers can concentrate on higher-order issues such as evidence interpretation, design justification, technical conventions, and audience appropriateness. GenAI may be used for explanation or alternative formulations, but students must verify

generated claims and retain responsibility for the submitted text.

Instructionally, the framework supports a recursive process of Planning → Drafting → Feedback → Critical Evaluation → Revision → Reflection. A student may return from feedback to planning or from revision to further consultation. The objective is not maximum automation; it is progressively greater capacity for independent writing. This also suggests a practical assessment implication: when AI assistance is permitted, assessment should capture process evidence—draft history, feedback evaluation, revision rationale, and verification—rather than relying exclusively on the final text.

Technology selection should consequently begin with pedagogical questions: Which writing difficulty requires support? At which stage is support needed? What kind of feedback is appropriate? Which decisions require disciplinary expertise? How will student independence be protected? And what privacy or integrity constraints apply? These questions align computational capability with educational purpose.

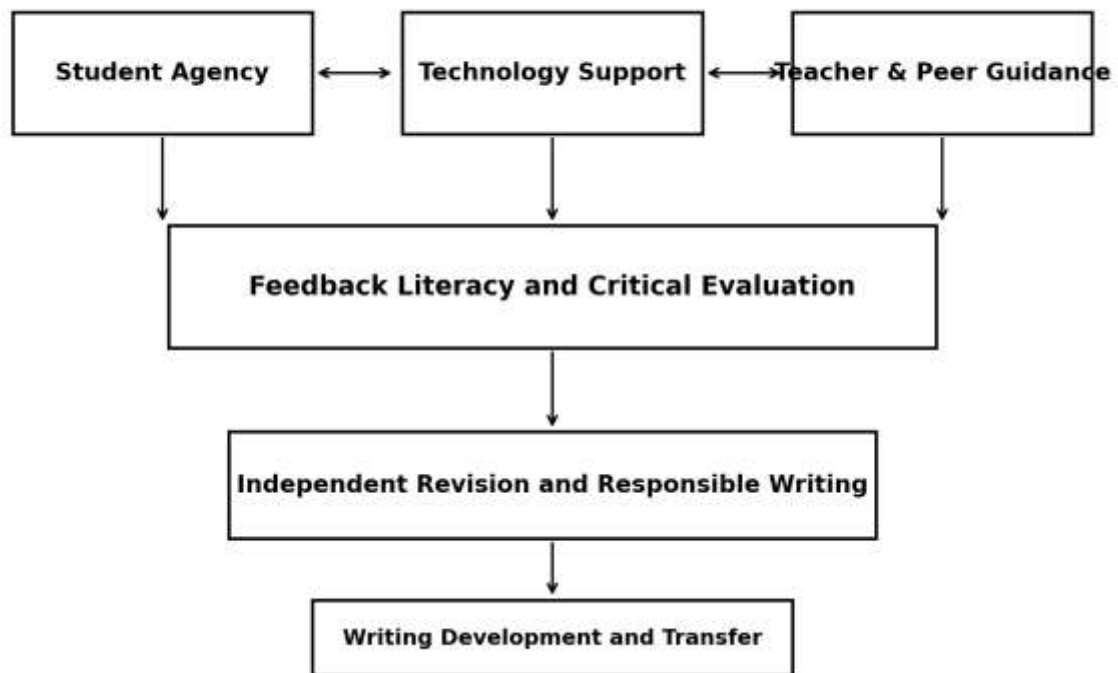


Fig. 2. Proposed human-in-the-loop framework for technology-integrated academic writing development in engineering education.

VI. CONCLUSION

Technology-integrated writing support offers clear opportunities for engineering education, particularly by expanding access to practice, timely feedback, and iterative revision. Published meta-analyses provide quantitative evidence that automated feedback and AWE can improve writing performance and quality [2], [3]. However, the evidence also shows that technology effects are not uniform and that lower-level language features are more readily supported than higher-order and discipline-specific writing.

The principal conclusion is that technology should function as scaffolding rather than as a replacement for authorship, disciplinary judgment, or instruction. The proposed human-in-the-loop framework combines student agency, technology support, teacher and peer guidance, feedback literacy, and responsible AI use across the full writing process. For engineering education, this means using computational tools to extend support while retaining human oversight where technical accuracy, argument quality, and disciplinary conventions are central.

This review is limited by the heterogeneity of the literature and by the comparatively fragmented engineering-specific evidence base. Future work should test the proposed framework in engineering courses, measure long-term transfer to new technical genres, and examine how process-based assessment can distinguish learning from technology-dependent text production. Such research can move the field from tool-centered adoption toward evidence-based integration.

DECLARATION ON GENERATIVE AI

Generative AI tools were used during manuscript development for drafting and language refinement. The authors

are responsible for reviewing, editing, verifying, and taking full responsibility for the final content, citations, interpretations, and conclusions. This declaration follows the journal's responsible AI-use requirement [8].

REFERENCES

- [1] M. J. Page et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, n71, 2021, doi: 10.1136/bmj.n71.
- [2] J. Fleckenstein, L. W. Liebenow, and J. Meyer, "Automated feedback and writing: a multi-level meta-analysis of effects on students' performance," *Frontiers in Artificial Intelligence*, vol. 6, 1162454, 2023, doi: 10.3389/frai.2023.1162454.
- [3] N. Zhai and X. Ma, "The effectiveness of automated writing evaluation on writing quality: A meta-analysis," *Journal of Educational Computing Research*, vol. 61, no. 4, pp. 907–937, 2023, doi: 10.1177/07356331221127300.
- [4] Y. Su, Y. Lin, and C. Lai, "Collaborating with ChatGPT in argumentative writing classrooms," *Assessing Writing*, vol. 57, 100752, 2023, doi: 10.1016/j.asw.2023.100752.
- [5] R. D. Roscoe, L. K. Allen, J. L. Weston, S. A. Crossley, and D. S. McNamara, "The Writing Pal intelligent tutoring system: Usability testing and development," *Computers and Composition*, vol. 34, pp. 39–59, 2014, doi: 10.1016/j.compcom.2014.09.002.
- [6] E. Cotos, *Genre-Based Automated Writing Evaluation for L2 Research Writing: From Design to Evaluation and Enhancement*. New York, NY, USA: Palgrave Macmillan, 2014, doi: 10.1057/9781137317370.
- [7] R. O'Neill and A. Russell, "Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly," *Australasian Journal of Educational Technology*, vol. 35, no. 3, pp. 42–56, 2019, doi: 10.14742/ajet.3795.
- [8] Science and Information Organization, "Author Guidelines," *International Journal of Advanced Computer Science and Applications*, accessed Aug. 21, 2026.

