Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
ASEE
Paper ID #46765

Development of FeedCap: A Tool for Real-Time Writing Feedback in Capstone Design Projects

Zhenlin Yang, University of Florida
Gabriel Castelblanco, University of Florida
Dr. Laura Melissa Cruz Castro, University of Florida

Dr. Laura Cruz is an instructional assistant professor in the engineering education department at University of Florida.

Generating timely, specific, and actionable feedback using chain of thought on a construction capstone assignment

Abstract:

This paper conceptualizes an AI-driven feedback tool enhanced with a chain-of-thought (CoT) mechanism tailored for LEED Narrative assignments in construction capstone projects. Recognizing that timely, specific, and actionable feedback is challenging to deliver at scale, especially in technical writing contexts like LEED documentation, the study explores how leveraging CoT can improve feedback transparency and effectiveness. Grounded in a cyclical learning model aligned with LEED's structured rubric, the framework emphasizes iterative refinement through criterion-referenced, transparent reasoning that connects student deficiencies directly to LEED credit requirements.

Methodologically, the study compares AI-generated feedback using CoT against evaluations provided by experienced teaching assistants (TAs) on anonymized student narratives. The AI system's prompt design guided a systematic, item-by-item review of compliance with LEED credits, detailing its reasoning process for each identified gap. Results indicate that while both AI and human reviewers similarly pinpointed missing technical information, the AI provided more elaborate, rubric-linked explanations. This transparency contrasts with the TAs' more concise remarks, underscoring potential benefits and challenges of deploying AI feedback in educational settings.

Concluding, the paper discusses how CoT-enhanced AI feedback can bridge gaps between student work and LEED standards, proposing future studies in live classroom settings to empirically assess its impact on student learning outcomes and revision processes.

Introduction:

In construction- and engineering-related programs, the capstone project is widely regarded as a pivotal milestone in a student's academic journey, testing not only their mastery of specialized knowledge but also their ability to integrate project management, system-level thinking, and collaborative skills [1], [2]. Given the growing importance of sustainable development in the construction industry, many educational institutions have embedded LEED (Leadership in Energy and Environmental Design) elements into their curricula or capstone projects to ensure students gain a solid understanding of green building principles and sustainable design practices.

For LEED-based projects or coursework, students are often required to produce a LEED Narrative outlining sustainability strategy, energy usage plans, and environmental impact assessments for their projects. LEED provides a comprehensive scoring rubric that covers a wide range of criteria—including site feasibility, energy efficiency, water conservation, materials and resources, and indoor environmental quality—making the evaluation process systematic yet challenging. Compared to more generalized writing tasks, LEED Narratives demand both technical precision and effective communication skills. Because LEED already has clear evaluation frameworks, instructors theoretically have a strong foundation for offering specific and actionable feedback.

Nevertheless, in practice, delivering timely and personalized feedback faces several obstacles. First, many students lack formal training in technical writing, making it difficult to craft a document that balances professional rigor with clarity and coherence [3], [4], [5]. Second, even when these same students transition into faculty or supervisory roles, time constraints and workload often prevent them from offering in-depth, individualized feedback to every student [6]. Realistically, even highly skilled instructors find it impractical to thoroughly review and comment on a large number of LEED Narratives each semester [7]. These limitations in providing robust, iterative feedback can hamper student learning outcomes and undercut the potential impact of the capstone experience [8], [9].

In recent years, artificial intelligence (AI)—particularly large language models (LLMs)—has shown promise in scaling and optimizing feedback processes in educational settings [10]. AI-driven feedback tools can analyze substantial amounts of text in a relatively short time and produce structured comments grounded in a specific domain or rubric. This rapid response capacity can be especially advantageous in large classes, reducing the instructor's workload while supporting students' self-guided learning and revision cycles[11]. A notable development here is the chain-of-thought approach, which explicitly reveals the AI's step-by-step reasoning. This transparency allows both educators and learners to scrutinize how the AI arrives at particular suggestions, thereby helping them gauge the clarity, reliability, and relevance of the feedback.

Against this backdrop, the present work sets out to design and conceptualize an AI-driven feedback tool tailored to LEED Narrative assignments. Specifically, we focus on a central research question:

How can leveraging a CoT mechanism in an AI feedback tool improve the specificity, actionability, and transparency of feedback on LEED Narrative assignments, thereby enhancing students' understanding of sustainable design requirements and their ability to make informed revisions?

While we have not yet conducted an extensive experiment with student participants, this paper presents the conceptual framework, design strategies (including prompt engineering and chain-of-thought reasoning), and potential benefits and challenges of integrating AI-driven feedback in the context of LEED Narrative assignments. The goal is to lay the groundwork for future studies that will empirically validate the effectiveness of such a system in real-world educational settings.

Theoretical framework (feedback cycle):

The provision of feedback in a LEED-based capstone assignment can be understood through a cyclical model of learning, wherein students iteratively refine their work based on formative insights [12], [13]. At the core of this framework is the premise that feedback should be both criterion-referenced—anchored in LEED's structured scoring system—and tailored to the student's evolving understanding of sustainable design [14]. Traditional theories of formative assessment emphasize the role of clear benchmarks in guiding student progress [15], and LEED's credit categories offer precisely such benchmarks, defining the thresholds and evidence needed to claim points [16]. By aligning feedback with these well-defined criteria, instructors (or automated systems) can help students identify discrepancies between the project's proposed strategies and LEED's requirements, thus directing revisions toward measurable improvements.

A defining feature of this cycle is the recognition that student narratives are not static statements but living documents subject to successive refinement. Foundational feedback research often underscores the importance of "closing the gap" between learners' current performance and desired learning outcomes [12], [15]. In the context of a LEED Narrative, this means pinpointing precisely which aspects of site selection, energy use, water efficiency, or materials sourcing remain incomplete or unsupported by adequate evidence. The cyclical nature arises when students, upon receiving feedback indicating a shortfall—such as insufficient calculations for water reduction or absent documentation for recycled materials—revise their narrative to strengthen, clarify, or correct those areas. Subsequently, the revised submissions can undergo another round of scrutiny, either by the same or a complementary feedback provider, sustaining an iterative loop of continuous improvement [17].

Central to this theoretical framework is the idea of transparent, step-by-step reasoning when conveying how well a given piece of writing aligns with LEED standards. Chain-of-thought approaches in AI are particularly relevant here [18], although such a mechanism is grounded in the broader literature on explainable feedback [19]. By exposing the logical process behind the review—whether human or AI—students gain insight into why their current writing may or may not fulfill a credit's prerequisites and what precise additions or alterations are necessary to move closer to the formal requirement. This form of transparent explanation fosters metacognitive awareness: learners see not only what they must change, but also the rationale underlying those suggestions, which is crucial for long-term skill development in both technical writing and sustainable design [20].

Another dimension to this theoretical model is the balance between timeliness and depth of feedback. In large classes, instructors may struggle to provide frequent, individualized commentary that delves into each credit's nuances [17]. Existing scholarship on feedback emphasizes the powerful effect of near-immediate responses, which allow students to stay engaged and swiftly correct misconceptions [15]. When a cycle is constrained by infrequent feedback, students risk continuing down an incorrect path or misunderstanding specific LEED requirements. In contrast, AI-supported feedback cycles have the potential to accelerate this exchange, offering immediate formative input that is mapped to the recognized LEED rubric. Even though ultimate validation must often come from a human assessor or a formal LEED evaluation, the faster iterative cycle can ensure students revise their narratives before major gaps or conceptual errors become entrenched.

Within this framework, prompt design is pivotal for aligning the feedback process with the rubric. The theoretical perspective here borrows from studies in criterion-referenced testing and specialized writing assessment [12], [14], suggesting that the language and structure of any prompting (human or AI-based) must mirror the precise vocabulary and logical flow of LEED credit requirements [16]. This conceptual alignment ensures a consistent bridge between the defined learning outcome (achieving specific LEED credits) and the mechanism of feedback delivery (identifying deficiencies, citing relevant credit thresholds, and proposing adjustments). Consequently, prompt design underlies the ability of the feedback cycle to function effectively, determining whether the resulting comments are diagnostic and actionable or merely generic statements that fail to foster measurable improvements.

Taken together, this theoretical framework positions LEED's standardized criteria as the anchor for a dynamic feedback loop, emphasizes transparent chains of reasoning, and highlights the importance of well-structured prompts. Under these principles, each iteration of review and revision becomes an opportunity for students to incrementally adjust their narrative to meet quantifiable sustainability targets, ultimately reinforcing both technical proficiency and deeper engagement with the principles of green building.

Methods:

This study involved the development of an AI-driven feedback tool for LEED Narrative assignments and a comparative review of its outputs alongside feedback from two TAs. Although the project did not involve live student submissions, the methodological framework was designed to reflect typical scenarios in which LEED Narratives are graded. The procedures described here establish how the data were gathered and analyzed, forming a basis for subsequent studies that will incorporate real classroom implementations.

Human Feedback Providers

Two TAs with multiple semesters of experience grading LEED-related coursework served as human reviewers. they will follow official LEED rubric for evaluating student submissions, making them well-suited to provide comparative feedback for this study. Each TA was responsible for reading and commenting on a set of LEED narratives that have been anonymized in accordance with the standards and conventions they would normally use in a real grading setting. Their commentary thus served as a baseline for human-generated feedback against which the AI's feedback could be compared.

AI Feedback Configuration and Prompt Design

An AI-based system was configured to generate feedback specifically for LEED Narratives, referencing the same rubric used by the TAs. A COT feature was incorporated into the system to allow it to articulate the logical steps behind each observation. The prompt design, which guided how the AI would analyze the text and structure its commentary, emphasized four main areas. First, the system verifies input sufficiency by checking whether a narrative provides enough relevant information—such as design strategies, calculations, or baseline data—to support meaningful feedback. Next, it limits its scope to LEED-related content, ensuring that all generated remarks align with the specific credit requirements. The AI then adopts an item-by-item approach, examining each claimed credit (for example, Water Efficiency or Energy and Atmosphere) against the corresponding thresholds or documentation standards. Finally, it upholds reasoning transparency by explaining any missing details or inconsistencies, referencing pertinent rubrics or metrics to demonstrate the logical route used to draw its conclusions.

These elements of prompt design aim to replicate the systematic grading approach commonly employed by TAs, wherein instructors progress through the LEED criteria to confirm that each credit's requirements have been adequately addressed. By detailing how the system arrives at its findings, the AI tool aspires to offer feedback that is clear, rigorous, and closely aligned with recognized standards in sustainable construction.

Sample Narrative Collection

Nine student assignments from Construction Management Capstone Class in a Southeast Institution were randomly selected for this experiment, we anonymized them to make sure all of the assignments will be evaluated fairly. The samples naturally varied in scope and depth—some included detailed discussions on topics such as energy usage or water conservation, while others were shorter, reflecting common variations in student work. Each unmodified sample was provided separately to both the AI system and the teaching assistants, allowing us to compare the tool's feedback with that of the TAs.

4. Comparative Review of Feedback

The AI-generated feedback and the TAs' written feedback were collected separately for each narrative. Both data sources were then compiled for comparison. Points of interest included the types of deficiencies each reviewer identified, the specificity of any corrective suggestions, and the clarity or brevity of remarks. To maintain consistency, the TAs worked independently of the

AI feedback, ensuring that they produced commentary based on their usual grading style rather than being influenced by the automated outputs.

5. Data Management and Analysis

The study team organized all feedback artifacts—AI outputs and TA comments—into a qualitative dataset for subsequent examination. The aim at this stage was to document the nature and scope of each reviewer's remarks rather than to measure student outcomes. Because no real students were involved, the analysis centered on whether the AI's feedback was comparable to, or distinct from, standard TA feedback in referencing the LEED rubric and identifying missing information. Although the TAs were not asked to evaluate the AI's chain-of-thought mechanism, their feedback provided a reference point for whether the AI's final statements would align with typical human assessments of LEED Narratives.

Results and disscussion:

This study compared feedback generated by an AI tool, which utilized a COT mechanism, with feedback written by two TAs. Because the TAs did not provide suggestions or comments on the process of tool to generate feedback, the comparison was limited to the final statements each source produced for a series of simulated LEED Narratives. The analysis focused on three key aspects: the kinds of missing information identified, the level of specificity in suggesting improvements, and which source aligned with the LEED rubric.

Similarity in Identifying missing information

Both the AI and the TAs show similar gaps in the sample narratives. When a text failed to include calculations supporting a claim of reduced water consumption, both the AI output and the TAs' feedback noted this omission. Likewise, for narratives containing incomplete references to materials or recycling content, both sources highlighted the need for additional documentation. In these instances, the AI's COT process typically produced an explanation that it had looked for certain keywords or baseline data but did not find them in student's writing; the TAs simply mentioned, in concise statements, that more detailed information was required.

Differences in Feedback Specificity

A more noticeable divergence arose in how elaborately each source explained its suggestions. The TAs, drawing on their routine grading practices, often wrote brief remarks such as "How do you achieve this?" or "Detailed information is needed," without necessarily indicating precisely which calculations or documentation were missing. The AI's feedback, by contrast, was generally more explicit about connecting a missing piece of evidence to a specific LEED credit threshold. For instance, if the example text claims to reduce water usage by 20% but does not mention the type of fixtures, the AI will not only notice the lack of these details, but will also point out the need for a base rate and suggest solutions based on that, such as installing water-saving fixtures like low-flow showerheads and toilets. In doing so, it offered a more structured rationale for why further data were needed. Nonetheless, the TAs' concise approach frequently arrived at the same fundamental conclusion—namely, that the student's justification was insufficient to earn the credit.

Chain-of-Thought Transparency

The AI's COT mechanism enabled it to outline which elements of the rubric it searched for and how it determined the student's compliance. However, the TAs did not evaluate the appropriateness of that chain-of-thought or comment on whether the AI's logical sequence was correct. Instead, the TAs generated their own feedback as if grading the narratives independently, paying minimal attention to the AI's underlying reasoning. Because of this, no data exist on how acceptable the TAs found the AI's internal logic; the comparison thus centered exclusively on final statements of feedback. In the instances where the AI spelled out a thorough rationale, the TAs neither confirmed nor disputed that reasoning. They simply wrote a shorter, more direct version of the same observation or recommendation.

Alignment with the LEED Rubric

In most cases, the AI's commentary and the TAs' remarks both corresponded to the official LEED rubric's criteria. Whether dealing with Energy and Atmosphere or Water Efficiency, both sources singled out omissions like missing calculations, contradictory baseline assumptions, or unsubstantiated claims of performance improvement. When a student narrative appeared only partially compliant with a credit requirement, the AI typically suggested providing measurements or references to applicable LEED standards; the TAs similarly acknowledged that the narrative needed elaboration, though their feedback was usually confined to a sentence or two. Despite these stylistic differences, both sets of feedback converged on the same determination about the credit's validity.

References:

[1] G. Castelblanco, L. Cruz-Castro, and Z. Yang, "Performance of a Large-Language Model in scoring construction management capstone design projects," *Computer Applications in Engineering Education*, vol. n/a, no. n/a, p. e22796, doi: 10.1002/cae.22796.

[2] L. Debs, B. Benhart, and F. R. Moraes, "A Review and Comparison of Associated Schools of Construction (ASC) Capstone Course Content," presented at the ASC2022. 58th Annual Associated Schools of Construction International Conference, pp. 560–550. doi: 10.29007/vp46.

[3] J. Yoritomo *et al.*, *Examining engineering writing instruction at a large research university through the lens of writing studies*. 2018. doi: 10.18260/1-2--30467.

[4] L. Cruz Castro, G. Castelblanco, and P. Antonenko, *LLM-based System for Technical Writing Real-time Review in Urban Construction and Technology*. 2024. doi: 10.29007/d9j3.

[5] M. A. Cantera, M.-J. Arevalo, V. García-Marina, and M. Alves-Castro, "A Rubric to Assess and Improve Technical Writing in Undergraduate Engineering Courses," *Education Sciences*, vol. 11, no. 4, Art. no. 4, Apr. 2021, doi: 10.3390/educsci11040146.

[6] E. Fife, "Making the Case for Technical Communication Courses in Ph.D. Engineering Curricula," in *2019 ASEE Annual Conference & Exposition Proceedings*, Tampa, Florida: ASEE Conferences, Jun. 2019, p. 33079. doi: 10.18260/1-2--33079.

[7] D. Carless, "Differing perceptions in the feedback process," *Studies in Higher Education*, vol. 31, no. 2, pp. 219–233, Apr. 2006, doi: 10.1080/03075070600572132.

[8] K. M. Getchell and D. A. Pachamanova, "Writing to Learn: A Framework for Structuring Writing Assignments to Support Analytics Course Learning Goals," *INFORMS Transactions on Education*, vol. 22, no. 2, pp. 103–120, Jan. 2022, doi: 10.1287/ited.2021.0249.

[9] M. Younis, P. Imdad, and A. A. U. Rahman, "EFFECTS OF CONSTRUCTIVE AND TIMELY FEEDBACK ON ACADEMIC PERFORMANCE OF STUDENTS," *Pakistan Journal of Educational Research*, vol. 4, no. 4, Art. no. 4, Dec. 2021, doi: 10.52337/pjer.v4i4.464.

[10] J. Shin, L. Cruz-Castro, Z. Yang, G. Castelblanco, A. Aggarwal, and B. F. Carroll, "UNDERSTANDING OPTIMAL INTERACTIONS BETWEEN STUDENTS AND A CHATBOT DURING A PROGRAMMING TASKLATEX".

[11] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100033, 2021, doi: 10.1016/j.caeai.2021.100033.

[12] D. J. Nicol and D. Macfarlane-Dick, "Formative assessment and self-regulated learning: a model and seven principles of good feedback practice," *Studies in Higher Education*, vol. 31, no. 2, pp. 199–218, Apr. 2006, doi: 10.1080/03075070600572090.

[13] D. R. Sadler, "Formative assessment and the design of instructional systems," *Instr Sci*, vol. 18, no. 2, pp. 119–144, Jun. 1989, doi: 10.1007/BF00117714.

[14] S. M. Brookhart, *Formative Assessment Strategies for Every Classroom: An ASCD Action Tool*. ASCD, 2010.

[15] J. Hattie and H. Timperley, "The Power of Feedback," *Review of Educational Research*, vol. 77, no. 1, pp. 81–112, Mar. 2007, doi: 10.3102/003465430298487.

[16] U. S. G. B. Council (USGBC), *LEED reference guide for green building design and construction*. U.S. Green Building Council, 2009. Accessed: Dec. 02, 2024. [Online]. Available: https://thuvienso.hoasen.edu.vn/handle/123456789/9868

[17] D. Wiliam, *Embedded Formative Assessment*. Solution Tree Press, 2011.

[18] J. Wei *et al.*, “Chain-of-Thought Prompting Elicits Reasoning in Large Language Models”.

[19] F. Doshi-Velez and B. Kim, “Towards A Rigorous Science of Interpretable Machine Learning,” Mar. 02, 2017, *arXiv*: arXiv:1702.08608. doi: 10.48550/arXiv.1702.08608.

[20] B. J. Zimmerman, “Becoming a Self-Regulated Learner: An Overview,” *Theory Into Practice*, vol. 41, no. 2, pp. 64–70, May 2002, doi: 10.1207/s15430421tip4102_2.