2026 ASEE ANNUAL CONFERENCE
June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte

Paper ID #53772

WIP: The AI Consultant—A Longitudinal Study of LLM Integration for Team Dynamics, Process Safety, and Ethical Advocacy in Chemical Engineering

Prof. Ryan Toomey, University of South Florida
Michelle L Cardenas, University of South Florida

Work-In-Progress: The AI Consultant - A Longitudinal Study of LLM Integration for Team Dynamics, Process Safety, and Ethical Advocacy in Chemical Engineering

Abstract

This work-in-progress paper presents a three-phase longitudinal study on the impact of Large Language Models (LLMs) on the professional development and learning of chemical engineering students. This paper examines the integration of ChatGPT/Copilot into three assignments across a cohort of senior chemical engineering students ($n \approx 65$) enrolled in Chemical Engineering Lab I, Process Safety and Ethics, and Chemical Engineering Lab II at a public R1 university. Assignments in Fall 2025 positioned ChatGPT as (1) a teamwork-coaching tool to help student teams convert vaguely defined collaboration challenges into actionable improvement plans and (2) a domain scaffold supporting process safety tasks including HAZOP development, identification of relevant engineering standards, construction of a LOPA, and preparation of a neutral "risk-gap" memo addressing tensions between budget constraints and safety recommendations. Students were explicitly reminded that AI output may be incomplete or incorrect and that engineering judgment remained their responsibility. A qualitative synthesis of reflections and submitted work identified themes related to student adoption of AI-suggested practices, strategies for constraining prompts and outputs, evidence of safety and standards literacy within ChatGPT responses, and approaches to ethical reasoning under realistic professional pressures. A follow-up assignment in Spring 2026 (Chemical Engineering Lab II) analyzed AI usage patterns after these instructor-led experiences.
Data from 17 student teams indicate that students transitioned to an independent "human-in-the-loop" model. Results highlight a significant increase in AI literacy: Students shifted from using AI for "answers" to using it for "overhead" tasks (i.e. concept clarification) while acknowledging skepticism towards AI responses. This work-in-progress study suggests that LLMs are most effective in engineering education when students are tasked with assessing the performance of the LLM. Instructors should design assignments that push AI to fail while probing students to assess the strengths and weaknesses of LLMs to enhance learning outcomes.

1. Introduction

Artificial intelligence (AI) and large language models (LLMs) are rapidly reshaping tools, workflows, and expectations across the chemical engineering profession including the chemical engineering curriculum. [1] As programs prepare students for practice in process design, process safety, laboratory operations, and cross-functional engineering environments, educators are exploring how generative AI (GenAI) might support not only technical problem solving but also communication, teamwork, and ethical judgment.[2], [3] Integrating AI into engineering education is often debated as a threat to fundamental learning.[4] Therefore, this paper explores whether instructor-led (or "intervention") AI instruction leads to enhanced adoption and learning outcomes with students recognizing "human-in-the-loop" necessity of discipline specific judgement.

By following a single cohort over a year, we examine how instructor-led AI intervention evolves into adoption of ethical/best use of AI by students. These competencies are central both to ABET student outcomes and to contemporary chemical engineering practice. Within this discipline-specific context, we posed three questions that are particularly salient for instructors integrating GenAI into senior-level curricula: (a) whether GenAI can effectively coach student teams toward actionable and sustainable collaboration practices; (b) whether GenAI can meaningfully scaffold domain-specific mastery such as hazard identification (HAZOP) and Layer of Protection Analysis (LOPA to complement limited classroom instruction); and (c) whether GenAI-supported practices persist longitudinally as students transition across courses and contexts.

To investigate these questions, we implemented a set of three coordinated assignments spanning two semesters in a single cohort of senior chemical engineering students. The interventions took place in Chemical Engineering Lab I and Process Safety and Ethics during Fall 2025, followed by Chemical Engineering Lab II in Spring 2026. The Fall 2025 assignments focused on responsible and ethical use of ChatGPT, emphasizing prompt sensitivity, verification, and awareness of potential biases. The Spring 2026 assignment examined whether and how students carried forward AI-supported collaboration and analytical practices. Collectively, these activities were purposefully aligned with ABET outcomes in teamwork, communication, ethics, and professional responsibility [5] while contributing to a rapidly growing body of literature on GenAI in engineering education. [6], [7], [8], [9]

In Assignment 1, teams in Chemical Engineering Lab I used either ChatGPT or Microsoft Copilot to translate high-level team goals into specific, actionable collaborative behaviors. Students then evaluated the AI-generated suggestions by documenting which practices they adopted, which they rejected, and why.

In Assignment 2, students in Process Safety and Ethics conducted a safety analysis of a hydrogenation reactor using ChatGPT or Microsoft Copilot as a structured technical coach and tasked with assessing the AI generated responses. Students had to prompt the LLM with regards to enumeration of HAZOP nodes and guidewords; identification of relevant industry codes and standards (API, ASME, OSHA); development of a LOPA including independent protection layers including calculations; and drafting of a "risk gap memo" that invoked Tenet #1 of the AIChE Code of Ethics when management resisted implementing a required safeguard.

In Assignment 3, students in Chemical Engineering Lab II revisited their earlier practices to assess the longitudinal transfer of GenAI-supported behaviors. Students reflected on their prior experiences, reassessed team goals, documented specific uses of AI in experimental planning, data analysis, and report preparation, and critically evaluated both the strengths and limitations of AI in collaborative technical work.

Taken together, these three assignments offer a multi-semester view of how chemical engineering students adopt, adapt, and use GenAI tools. After instructor-led interventions in Assignments 1 and 2, results from Assignment 3 indicate how students transition to using AI for productivity (e.g., data cleaning and concept distillation) while developing a robust skepticism involving AI for chemical engineering problem solving.

2. Project Approach

This study employed a multi-assignment, multi-semester design to investigate how chemical engineering students use generative AI tools to support teamwork, technical reasoning, and laboratory practice. The same cohort of students ($n \approx 65$), senior-level chemical engineering majors, participated in three sequential assignments across three required courses: two Fall 2025 core courses and a Spring 2026 course. There has been no systematic adoption of GenAI into the curriculum, although students reported that GenAI has been banned (via syllabi instructions) as a form of academic misconduct in a few courses prior to entering the senior year.

2.a Course Context and Participants

Students ($n \approx 65$) were enrolled in three sequential chemical engineering core courses across two semesters in the senior year:

1) Fall 2025: Chemical Engineering Laboratory I and Process Safety and Ethics.
2) Spring 2026: Chemical Engineering Laboratory II, which emphasizes experimental methods, data analysis, and team-based reporting.

Students completed Assignments 1 (Chemical Engineering Laboratory I) and 2 (Safety and Ethics) in Fall 2025 and Assignment 3 (Chemical Engineering Laboratory II) in Spring 2026. Table 1 shows explicitly the three-phase pedagogical framework, role of the LLM, and key deliverables.

Participation in the assignments was integrated into graded course activities as part of the general course participation grade (in which students must physically be in class). The approximate grade of each assignment was 1% of the total course grade.

Table 1. Three Phase Pedagogical Framework for incorporating ChatGPT/Microsoft Copilot

Course	Time Frame	Focus Area	LLM Role	Key Deliverable
I	Fall 2025	Team dynamics	Logistic coach	Team action plans (post-lab huddles, mini deadlines)
II	Fall 2025	Safety & Ethics	Technical/Ethical consultant	HAZOP/LOPA development, verification of calculations, Ethic scenarios and memo writing
III	Spring 2026	Longitudinal adoption	Productivity assistant	Reflection of unprompted AI integration based on lessons learned in Courses I and II

2.b Assignment Design

2.b.1 Assignment 1: Team Dynamics (Fall 2025) in Chemical Engineering Laboratory I

In Assignment 1, student teams created a team contract identifying goals related to timeliness, communication, work distribution, quality assurance, and preparation habits. Student teams were tasked with the following:

1. Prompting the LLM to translate team goals into concrete, actionable collaboration practices.
2. Evaluation of the suggestions from the LLM, noting which recommendations they adopted or rejected and providing rationale.
3. Documenting their decision-making process in a written submission.

This assignment positioned the LLM as a team coach and provided a baseline for analyzing GenAI-supported teamwork behaviors.

2.b.2 Assignment 2: HAZOP/LOPA Safety Analysis (Fall 2025)

Assignment 2 examined how students used GenAI to support structured technical reasoning in process safety. Student teams initially analyzed a hydrogenation reactor scenario developing a HAZOP and a layer of protection analysis. Results of each team's HAZOP was collated and distributed to the entire class. Each student was tasked with using ChatGPT or Copilot to develop a P&ID representing the process and to enumerate process nodes, guidewords, and deviation scenarios in developing an AI led HAZOP. Students also had to task the LLM with the following prompts:

1. Identifying relevant engineering standards (API, ASME, IEC) with AI assistance.
2. Conducting a Layer of Protection Analysis (LOPA) including identification of independent protection layers and appropriate calculations.
3. Drafting a risk-gap memo that invoked the AIChE Code of Ethics Tenet #1 when management resisted implementing a safety measure.

This assignment enabled comparison between human and AI reasoning in safety-critical contexts.

2.b.3 Assignment 3: Integration of AI in Chemical Engineering Laboratory II (Spring 2026)

In Laboratory Chemical Engineering Lab II, students were encouraged to use AI as they saw fit; however there was no instructor-led intervention. In Assignment 3, students revisited their previous AI use reflecting on whether AI-supported strategies from Fall 2025 persisted into Spring 2026. Students were tasked with the following questions:

1. Re-evaluated team goals using the same criteria as in Assignment 1.
2. Provided concrete examples of how AI tools were used during laboratory work (planning, data analysis, report drafting).
3. Assessed limitations, ethical considerations, and modifications to their team functioning with respect to AI.

This assignment served as a longitudinal measure of the transfer of AI-supported practices.

2.c Data Sources

Three types of student responses were collected: (1) Written student/team submissions for each assignment; (2) Reflections and rationales documenting decision-making around AI use; and (c) tables of team goal ratings (Assignment 1 and 3), allowing comparison across semesters. All responses were de-identified prior to analysis.

2.d Data Analysis

All written responses were analyzed using an inductive qualitative coding approach. Words emerged around themes such as:

1. Adoption vs. rejection of AI suggestions
2. Types of tasks delegated to AI
3. Perceived benefits (i.e. efficiency, clarity, brainstorming)

4. Concerns (i.e. inaccuracy, over-reliance, bias)
5. Ethical awareness and verification practices

As part of our data analysis, ChatGPT was used to identify trends based on student submissions.

3. Preliminary Observations and Discussion

3.1 Assignment 1: Team dynamics and logistics coach

Analysis of the team reports (n = 17) revealed several themes that emerged from using ChatGPT or Copilot as a logistics coach, as noted in Table 2.

Table 2. Themes that emerged from the use of ChatGPT/Copilot as a logistics coach

Theme	Typical Action Integrated	Representative Evidence
Internal micro-deadlines	Staggered internal due dates; "finish 1 day early" for review	"Micro-deadlines... meeting as a mini deadline" (T1); "peer-review deadline 24h before" (T2); "finish 1 day early" (T11)
Post-lab huddles	10–15 min role assignment & deadline setting right after lab	"Post-lab 10–15 minute meeting..." (T6); "recurring 10-minute huddle" (T4)
Quality control	Two-person peer review; early merge deadline; one formatter	"Two people review each part" (T1); "merge earlier" (T5); "one formatter" (T12)
Inclusion & etiquette	Hybrid options; comment-only edits; explicit directions/do not assume	"Hybrid for commutes" (T10); "leave comments, not overwrites" (T8); "invite input directly" (T11)
AI literacy (accept/reject)	Modify or reject redundant/not applicable or high overhead advice advice (e.g., rotating coordinator; single-person "quick analysis")	All teams noted at least one or more suggestions that were not applicable

Further distillation of the data suggests the following 3 major outcomes of this activity with student acceptance/rejection of AI:

1. **Actionable Specificity**: AI coaching helped teams move beyond ambiguous framings of group dynamics such as "trying harder" to implementing actionable team practices such "mini-deadlines."
2. **Resistance to high-overhead activities**: Several teams rejected AI suggestions for shared calendars or formal acknowledgement protocols, citing that these added "excess work" to an already heavy engineering workload.
3. **Validation and Morale**: Teams that were already performing well used ChatGPT to validate their current informal strengths, which served to boost group cohesion.

A significant observation from self-reflections of student teams was the ability to recognize the LLM as a consultant rather than an authority. For example, one team noted: "ChatGPT does not like to be brief... we confirmed the idea that whatever ChatGPT outputs needs some degree of interpretation and cannot be blindly agreed with." This suggests that the exercise not only improved team dynamics but also fostered AI literacy.

Students generally recognized that the LLM aids as a process coach but is not an authority. Students were agreeable to such suggestions as "internal mini deadlines, post lab huddles for immediate role assignment, and two-person peer review with early merge points." However, students selectively rejected advice perceived as inequitable or high overhead. Several teams adopted inclusion-oriented practices (hybrid options, comment etiquette, explicit invitations to quieter members). Overall, the AI generated responses translated vague teamwork goals into actionable team practices. After submission of the individual assignments, the instructor uploaded all submissions directly to ChatGPT to create an instantaneous heat map for the entire class to view. ChatGPT zeroed in on that the number one obstacle for students was insufficient communication in writing lab reports and suggested mini deadlines where all team members convened for short period of times to address difficulties or other challenges before the submission deadline.

3.2 Assignment 2: Safety and Ethics – Technical and Ethical Consultant

Students performed a manual HAZOP on a hydrogenation reactor and then used ChatGPT to generate a comparative P&ID, HAZOP table, and LOPA with calculations. The exercise culminated in ethical role-play where students pushed ChatGTP or Copilot to respond to a manager who prioritized the budget over safety, forcing students to defend the AIChE Code of Ethics.

Students reported that ChatGPT was particularly useful for: 1) enumerating nodes and guidewords they might otherwise overlook (e.g., reverse flow, vacuum on cooldown) in a HAZOP; 2) Identifying appropriate standards; and 3) writing memos in neutral tones absent of emotionally charged language. However, students explicitly noted that AI-generated HAZOPS were often generic or over inclusive. Students improved quality by selectively modifying the prompt to match the provided P&ID and by discarding irrelevant or redundant information. This process reinforced the role of the LLM to generate a large range of outputs but not necessarily an arbiter of the information provided. Table 3 provides a general comparison of student-led vs AI led technical and ethics performance.

Table 3. Comparison of Human vs. AI Performance in Safety Analysis

Task	Human Intervention	LLM (ChatGPT/Copilot) Performance
P&ID	Accurate P&ID logic.	**Fail:** "Hallucinated" loops and combined streams.
HAZOP/LOPA	Specific to P&ID; identified ~6 nodes.	Broad/Generic; identified 12+ nodes with several guidewords that were not applicable
Ethical Script	Emotionally charged/fear of job loss.	Professional; centered on AIChE First Tenet.

In terms of assessing the performance of the ChatGPT or CoPilot in their reflection statements, students centered on the following observations:

1. **Technical Benchmarking:** Students observed that LLMs can be "confidently generic." For instance, it could list 15-20 ASME/API standards in seconds but often missed the specific nuances of the provided problem.
2. **The hallucination check**: Students explicitly noted that the weakest aspect of ChatGPT and Copilot was the generation of P&ID with outputs looping cooling water into product streams or creating non-functional loops. This served as a high-value lesson in AI skepticism.
3. **Ethical Scaffolding**: Students found the role of ChatGPT and Copilot in manager conflict to be very helpful. Students reported that that both provided a "diplomatic script." It moved students from a position of fear/emotion ("I'm scared to lose my job") to a position of professional authority ("Based on ASME/API standards, this risk is intolerable").
4. **The "Greater Good" Fallacy**: Students successfully used ChatGPT or Copilot to dismantle the argument that green products (biodiesel) justify safety shortcuts. The LLMs refused to back down on the AIChE First Tenet acted as a moral anchor for the students.

In particular, the ethical prompts produced some of the strongest learning evidence. When asked how to respond to a manager refusing a safeguard due to budget constraints, the LLM helped frame professional, standards-based arguments. Many reported that the exercise increased their confidence to advocate for safety, even when organizational incentives conflicted with risk reduction.

3.3 Assignment 3: Longitudinal Adoption

Across Assignments 1 and 2, students were led through instructional use of ChatGPT/Copilot forcing them to assess both the strengths and weaknesses of LLMs. Students articulated a consistent theme: LLMs can be powerful tool for brainstorming but that its output should not be accepted uncritically. Students learned to refine prompts, to request justification, and to verify outputs against physical reasoning.

In Assignment 3, students revisited their previous AI use reflecting on whether AI-supported strategies from Fall 2025 persisted into Spring 2026. Table 4 shows the main themes found with respect to six application areas.

Table 4. Longitudinal Adoption/Integration of LLM usage

Application Area	Initial Intervention (Fall 2025)	Longitudinal Integration (Spring 2026)	Key Student Reflection
Team Management	**Forced Use:** AI suggested "Mini-deadlines" and "Post-lab huddles."	**Selective Autonomy:** Most teams abandoned AI for logistics once their social "flow" was established.	*"Suggestions lacked significance... we were already doing these as a team or they were impractical."*
Ethical Advocacy	**Simulated Conflict:** AI used to script a defense of the AIChE First Tenet against a simulated manager.	**Conflict Framing:** AI used unprompted to "neutralize" language in emails to TAs or resolve minor peer friction.	*"AI removes the emotional aspect... it helps with phrasing without 'challenging' the authority in a disrespectful way."*
Technical Research	**Information Gathering:** AI listed 20+ API codes for a hydrogenation reactor.	**Information Distillation:** AI used to "skim and scan" 30-page lab guides to identify key operating steps.	*"It acts like an additional office hour... helps distill long documents to save hours of preliminary study."*
Data Analysis & Math	**Technical Generation:** AI attempted to generate HAZOP tables and LOPA calculations.	**Skeptical Verification:** AI used to "sanity check" human calculations (e.g., Tau/Transfer Functions).	*"We don't use AI for final data analysis because it is inaccurate... it is a 'confusing peer' that must be fact-checked."*
Workplace Productivity	**General Writing**: AI used for basic report drafting.	**Tedious Task Automation:** AI used to clean/format un-delimited CSV data exports from lab equipment.	*"Reduced time spent on tedious organization by automating the separation of data points into columns."*
Tool Literacy	**Exploration:** Students "pushed" the AI to see its limits.	**Principled Opt-Out:** Some teams deliberately prohibited AI in their Spring contracts to preserve learning.	*"We prefer to consult professors/TAs... we chose not to use generative AI in our team contract."*

3.4 Discussion

A clear developmental arc emerged between the Fall and Spring semesters. In the Fall semester, students were tasked with implementing ChatGPT or Copilot as a coach for both soft and hard skills. While not initially assessed as to their skepticism of LLMs, the same cohort clearly demonstrated skepticism of both LLMs in Spring 2026. Students shifted their usage away from generating answers towards automating overhead (such as writing e-mails) and distilling complexity (summarizing lab manuals). Most importantly, the Spring semester reflections reveal that students recognize that the ability of ChatGPT or Copilot to solve engineering problems is full of hallucinations and false confidence; however, there is utility to using AI as a comparative benchmark for their own derived calculations. Along such lines, while students reported that ChatGPT is “confidently incorrect” or “unable to perform the math,” this failure could also be used to facilitate learning. By recognizing such limitations, students were forced to return to primary sources (Geankoplis, instructors) to verify their results.

Perhaps the most striking aspect of the longitudinal data is how students transitioned to critical filtering of AI. Students in Spring 2026 reported: “We plan to only use ChatGPT to understand the question better... we agreed not to use it for data analysis because it is inaccurate.” This indicates that the forced encounters with technical failures of AI in the Fall 2025 successfully built the skepticism required for responsible practice.

4. Conclusions

This work-in-progress study suggests that LLMs are most effective in engineering education when positioned as an assistant coach (through instructor led intervention), wherein students are tasked with assessing the performance of the LLM. Instructors should design assignments that push AI to fail (e.g., complex calculations) while incentivizing students to assess the strengths and weaknesses of LLMs to enhance learning outcomes.

Moving forward, the application areas of LLM use provided in Table 4: Longitudinal Adoption/Integration of LLMs (i.e. team management, ethical advocacy, technical research) will help frame the types of interventions to be provided next year in both Chemical Engineering Lab 1 and Safety and Ethics. One element lacking from this year’s cohort is that we did not assess or survey AI literacy before the interventions. Next year, we will gather such information as well as follow LLM usage into capstone design.

5. References

[1] T. L. Khoo, T. S. Lee, S.-T. Bee, C. Ma, and Y.-Y. Zhang, “A Comparative Review of Large Language Models in Engineering with Emphasis on Chemical Engineering Applications,” *Processes*, vol. 13, no. 9, Aug. 2025, doi: 10.3390/pr13092680.

[2] Y. Albadarin, M. Saqr, N. Pope, and M. Tukiainen, “A systematic literature review of empirical research on ChatGPT in education,” *Discov. Educ.*, vol. 3, no. 1, p. 60, May 2024, doi: 10.1007/s44217-024-00138-2.

[3] Z. Chu *et al.*, “LLM Agents for Education: Advances and Applications,” Mar. 14, 2025, *arXiv*: arXiv:2503.11733. doi: 10.48550/arXiv.2503.11733.

[4] V. Okulich-Kazarin, A. Artyukhov, Ł. Skowron, N. Artyukhova, and T. Wołowiec, “Will AI Become a Threat to Higher Education Sustainability? A Study of Students’ Views,” *Sustainability*, vol. 16, no. 11, p. 4596, Jan. 2024, doi: 10.3390/su16114596.

[5] R. M. Felder and R. Brent, “Designing and Teaching Courses to Satisfy the ABET Engineering Criteria,” *J. Eng. Educ.*, vol. 92, no. 1, pp. 7–25, 2003, doi: 10.1002/j.2168-9830.2003.tb00734.x.

[6] I. Adeshola and A. P. Adepoju, “The opportunities and challenges of ChatGPT in education,” *Interact. Learn. Environ.*, vol. 32, no. 10, pp. 6159–6172, Nov. 2024, doi: 10.1080/10494820.2023.2253858.

[7] M.-L. Tsai, C. W. Ong, and C.-L. Chen, “Exploring the use of large language models (LLMs) in chemical engineering education: Building core course problem models with Chat-GPT,” *Educ. Chem. Eng.*, vol. 44, pp. 71–95, Jul. 2023, doi: 10.1016/j.ece.2023.05.001.

[8] E. E. Danahy, M. K. Vora, Y. M. Xu, and W. Church, “Full Paper: A Generative AI Approach to Better Teamwork in First-Year Engineering,” presented at the 15th Annual First-Year Engineering Experience Conference (FYEE), Jul. 2024. Accessed: Jan. 21, 2026. [Online]. Available: https://peer.asee.org/full-paper-a-generative-ai-approach-to-better-teamwork-in-first-year-engineering
[9] J. A. Ovi, G. Fierro, and C. E. Smith, “Assessing Student Adoption of Generative Artificial Intelligence across Engineering Education from 2023 to 2024,” Mar. 06, 2025, *arXiv*: arXiv:2503.04696. doi: 10.48550/arXiv.2503.04696.