

## **(Board 50/Work in Progress) A Systematic Review of Embedding Large Language Models in Engineering and Computing Education**

**Dr. David Reeping, University of Cincinnati**

Dr. David Reeping is an Assistant Professor in the Department of Engineering and Computing Education at the University of Cincinnati. He earned his Ph.D. in Engineering Education from Virginia Tech and was a National Science Foundation Graduate Research Fellow. He received his B.S. in Engineering Education with a Mathematics minor from Ohio Northern University. His main research interests include transfer student information asymmetries, threshold concepts, curricular complexity, and advancing quantitative and fully integrated mixed methods.

**Aarohi Shah, University of Cincinnati**

# **WIP: A Systematic Review of Embedding Large Language Models in Engineering and Computing Education**

## **Abstract**

This work-in-progress paper explores how students and faculty are employing large language models (LLMs) like ChatGPT in engineering and computing education contexts through a systematic literature review (SLR) with Arxiv. We screened 717 preprint abstracts of emerging literature related to LLMs, ultimately analyzing 63 papers. We extracted the educational applications from each paper, prompts they provided relevant to the application, and details of any study they conducted to explore the efficacy of the proposed application (e.g., research design and subjects). This review aims to highlight promising work using LLMs in engineering and computing education that provides more "out of the box" applications to try in classrooms.

## **Introduction**

As seen in the myriad opinion pieces and articles in the popular press, students are leveraging generative AI models such as ChatGPT to complete their assignments [1], [2], [3] – bringing discussions of academic dishonesty to the forefront, much to the concern of instructors [4], [5], [6]. However, using LLMs like ChatGPT is not entirely fraught with threats to education; work has also emerged about faculty experimenting with incorporating these models into their teaching and evaluation methods. In fact, there is some evidence that ChatGPT is being used more by teachers than students [7]. Current explorations into student perceptions and use cases do not substantiate narratives about rampant cheating [8], [9]; thus, categorizing all student use of LLMs as violating academic integrity is unproductive. Moreover, despite the proliferation of manuscripts offering methods for incorporating LLMs into our teaching practices, much of the advice does not elaborate on practical use cases across disciplines or provides limited data to support the application's efficacy. For example, [10] offers generic examples and sample outputs with little formal evaluation and surveyed student at a high-level about ChatGPT's usefulness.

Previous systematic reviews focusing on using ChatGPT in educational settings have provided several general suggestions for how LLMs can be purposefully integrated into the learning process. For example, Imran and Almusharraf [11] reviewed 30 papers related to how ChatGPT could be used as a writing assistant for instructors and students, but the synthesis does not offer concrete prompts or specific guidance on how one would use such a tool to enhance their work beyond suggestions offered in the reviewed papers (e.g., grammar assistance, text summarization, constructing initial drafts, and brainstorming). Beyond only writing tasks, Montenegro-Rueda et al. [12] explored how ChatGPT was being implemented in educational contexts, including the benefits and the challenges of adopting the technology for classroom use. In their assessment of the final corpus of 12 papers, the authors provide a high-level summary of the findings, including publication location, methods implemented (i.e., quantitative, qualitative, and "theoretical"), and the premise of the papers (e.g., educational supports, educational challenges, teacher training). Moreover, the promise of personalized learning is emphasized – describing ChatGPT as an "easy-to use and accessible tool for teachers and students, allowing for quick integration into the classroom" [12, p. 10]. Similarly, İpek et al. [13] reviewed a broader set of educational applications for ChatGPT – including areas like finance, language, medicine, and law – and catalogued several applications of ChatGPT, including identifying student needs, scaling assessment, personalized tutoring, and generating material. Although the findings from

these reviews – including others such as [14], [15], [16], [17], [18] – can help conceptualize the different possibilities, guidance about how to implement LLM-powered tools like ChatGPT is understandably sparse across these literature reviews.

The best practices for using LLM-powered tools in educational research are developing as well, specifically how we go about unlocking their proclaimed benefits. The key is determining what prompts and practices can be used to generate quality outputs, but so-called "prompt engineering" resources available may be too generic to be useful. Moreover, without disclosing the full prompts used to facilitate the application of the LLM – whether in research or as an educational tool – the transferability of the results and the overall quality of the study is compromised. With the avalanche of literature concerning tools like ChatGPT, scrutinizing articles for evidenced-based applications will become increasingly necessary.

### **Research Aim**

In this work-in-progress paper, we delve into the implementation of different approaches to using LLMs like ChatGPT in engineering and computing education by examining how these tools are being leveraged for pedagogical and assessment purposes. The research question guiding this work is: "how are students and faculty using LLMs (i.e., ChatGPT) in engineering and computing education contexts for instruction and assessment?" The primary objective of this work is not just to consolidate the existing disparate strands of emerging research, but wherever possible, also to highlight concrete examples *with prompts* for harnessing LLMs in different classrooms for instructors to readily adapt to their contexts.

### **Methods**

To provide a comprehensive understanding of the current landscape, a systematic literature review (SLR) was conducted, specifically targeting papers relevant to generative AI in engineering and computing education. We first selected a set of "sentinel articles," which are articles selected beforehand that fit within the scope of the review to help develop a set of keywords and eventually form the search string [19]. The sentinel articles were the following: [20], [21], [22]. The search string was a combination of general terms such as "large language model" and specific models "GPT-3.5." These were combined with keywords like "education" to capture a breadth of papers. Our complete search string was:

```
("generative AI" OR "ChatGPT" OR "GPT3" OR "GPT-3" OR "GPT 3" OR "GPT4" OR "GPT-4" OR "GPT 4" OR "Large Language model" OR "GPT3.5" OR "GPT 3.5" OR "GPT-3.5" OR "LLM") AND ("Engineering education" OR "STEM" OR "Education" OR "Engineering Design" OR "Engineering")
```

This search string yielded 717 papers, which were collected in September 2023. We used Arxiv, a popular preprint repository, to retrieve papers because, at the time, it was anticipated that many papers about LLMs in education would not have passed peer review or been published yet. Moreover, the "correctness" of the results was not of concern at this stage; we were interested in which applications were garnering interest to explore through a research lens.

To determine which papers were relevant, we evaluated them against the following two criteria: (1) the paper included reference to an LLM like ChatGPT in an educational setting, and (2) the educational setting was science, technology, engineering, or mathematics (STEM). Following the first round of filtering by examining the abstracts, 51 papers were identified as having a clear

educational purpose in STEM Education, whereas 49 were considered to have a promising application in education but needed additional review. Additionally, 107 papers were filtered out of the sample as having tenuous relevance to education (e.g., brief references to educational applications), and 51 were determined to be out of scope for this study. We have analyzed 63 papers to date.

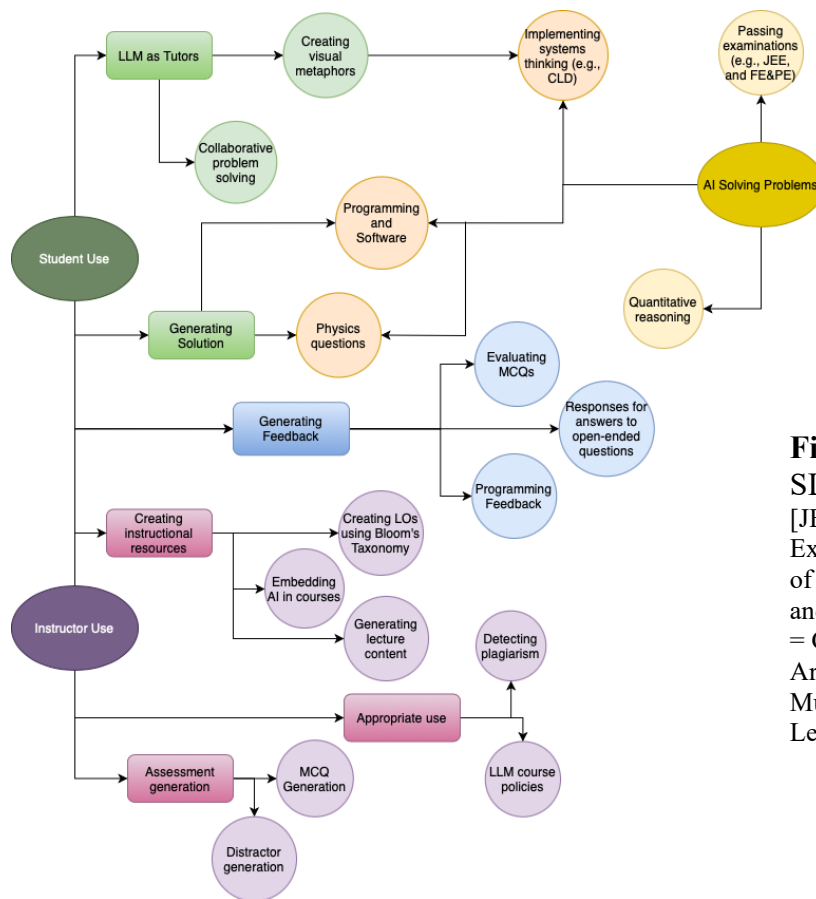
### **Preliminary Results and Tangible Examples of LLM Applications**

Our initial results suggest that the papers are converging around a set of common use cases for ChatGPT and similar tools. The current mapping of papers can be seen in Figure 1, which is divided into student uses, instructor uses, and where LLMs have been tested with established assessments (i.e., "AI Solving Problems"). Many papers focused on how well ChatGPT can perform on established assessments such as the Force Concept Inventory [23], the Fundamentals of Engineering Exam, and the Principles and Practice of Engineering Exam [24] in addition to quantitative reasoning questions [25] and introductory programming problems [26]. In terms of student use, implementing LLMs as a tutor was the prevailing theme (especially for programming), in addition to specific tasks like generating visual metaphors for learning STEM concepts [27] and determining how well these tools can help students solve physics problems [20] as more niche topics. Regarding instructor use cases, the papers predominantly discussed generating assessment questions, particularly multiple-choice questions (MCQs) [28], [29], [30], as well as creating general course content such as learning outcomes and lecture outlines [31], [32]. Moreover, another subset of studies evaluated student-facing feedback generated by LLMs in educational contexts [33], [34]. The areas highlighted in Figure 1 show the emerging use cases that have been explored in the literature.

Notably, *nearly half of the papers contained no prompts*. Others only contained partial prompts, such as our next example. We expect this will be a growing pain for this literature, as replicating results will be difficult or impossible without contacting the authors for verification.

#### **Instructor Use – Constructing Learning Outcomes**

One tangible example of a more ready-to-use educational application of LLMs was in the context of constructing learning outcomes. In particular, Sridhar et al. [35] crafted a prompt that considers aspects of quality LOs, such as their sensibility in terms of grammar and relevance, measurability, and alignment with levels of Bloom's Taxonomy [36]. They provided GPT-4 with a two-stage input by individually submitting the inputs to OpenAI's GPT-4 API using Python; the first being a system prompt (which is the initial text given to the API to set the model's role, persona, or behavior), followed by the user input. The system prompt consists of what we want the model to act like, which in this case is a curricular development expert, along with instructions about what makes LOs well-constructed in terms of three main parts: Behaviour (the observable action of the student), Conditions (in what context the student will perform the action), and Degree (how well the behavior must be performed). Moreover, the prompt contains examples of what conceptual LOs (which they define as Remembering and Understanding in Bloom's Taxonomy) and project LOs (based on the remaining levels of Bloom's taxonomy) are to give the model a basis for what output to produce. The authors include the criteria that need to be satisfied for LOs to be effective; readers are strongly encouraged to borrow the checklist from Kennedy [37] to expand on their list. Lastly, they list what the user input would consist of. On the user side, the input states the course name, course goals, module name, and LO types desired.



**Figure 1: Mapping of the SLR Results to Date**  
 [JEE = Joint Entrance Examination, FE = Fundamentals of Engineering, PE = Principles and Practice of Engineering, CLD = Causal Loop Diagram, AI = Artificial Intelligence, MCQ = Multiple Choice Questions, LO = Learning Outcomes]

Sridhar et al. [36] do not provide the entire prompt in their paper, so we provide an example of how the prompt as given might be used in Appendix A in the context a module on brainstorming in an introduction to engineering course. The free version of ChatGPT (i.e., GPT 3.5) was used because we expect this is the option instructors will be drawn toward. Although we do not find the Conditions and Degree element particularly illuminating, the learning outcomes are either suitable as-is or could be ready with tweaking despite being generated with little effort.

After evaluating the outputs given their prompt structure, [35] concluded that their LOs were highly sensible, but action verbs were not as consistent as desired. This behavior can be seen as well in our example, where action verbs did not correspond to the requested levels of Bloom's taxonomy. Even when the original authors gave GPT-4 specific words to choose from, it generated unmentioned verbs while leaving most of the provided action verbs unused. As for identifying whether it used the correct levels of Bloom's Taxonomy, it was found that generated LOs primarily operated on the expected levels. We should note that Dickey and Bejarano [31] also explored how instructors can create LOs, which we demo in Appendix A. Although neither prompt is necessarily perfect, there is evidence that using a GPT model to create draft learning outcomes could become a handy use case for instructors.

### Student Use – Generating Feedback and Collaborative Problem Solving

With tools like ChatGPT being heralded as the next step in personalized learning, our sample of papers detailed how LLMs could provide immediate instructor-quality feedback for students. For

example, Phung et al. [37] show practical applications of how LLMs can be integrated into programming education. The authors curated five specific categories of buggy programming problems that students might come across (e.g., finding the greatest common divisor of two positive integers and checking if a string is a palindrome) and tasked ChatGPT (i.e., GPT 3.5) and GPT-4 with six different scenarios (program repair, hint generation, grading feedback, pair programming, contextualized explanation, task synthesis). For each scenario, they compared feedback from ChatGPT, GPT-4, and human tutors – highlighting that these LLMs have varied performance on programming feedback tasks, sometimes performing as well as human tutors. The prompts provided in their appendix showcase the details and outputs generated. For example, they provide the following prompt for the program repair task: *"I'm working on a Python programming problem. The current program below is not working well. Can you help in fixing this program with as few changes as possible? Below I first provide the problem description and then the current buggy program. {problem\_description} Buggy Program: `` {buggy\_program} `` Can you fix the above buggy program?"* Instructors may find these prompts useful to share with students to model using LLMs responsibly.

Moving away from programming, Arndt [38] delves into the application of LLMs in explaining concepts from system thinking and system dynamics, in addition to creating visualizations such as causal loop diagrams (a model showing causal relationships between variables with +'s and -'s to denote the direction of the relationship). Leveraging the ability of tools like ChatGPT to write scripts in Python (and other languages), it was found that creating such visualizations was possible by running the output outside of the LLM's interface – albeit with expected hallucinations, as usual. Given causal loop diagrams can be created in this fashion, we encourage instructors to adapt this approach to prepare other diagrams that represent relationships between ideas or variables, such as concept maps [39]. Concept maps, in particular, can be useful as assessments [40] and as curricular design tools; for example, students could use the output and build upon a concept map by critiquing its formulation and adding missing connections.

### **Conclusion and Future Work**

As we refine our pedagogical practices to adjust to the democratization of generative AI, there is not necessarily a clear answer for addressing the wicked problems and fruitful capabilities inherent to ChatGPT and its competitors. When interviewing students and instructors about generative AI's benefits, concerns, and future prospects, there is remarkable convergence but notable disagreements [41]. For example, instructors discuss using generative AI to reduce effort while writing code, increasing the focus on design and problem solving, and providing alternative perspectives on demand; on the other hand, students iterated on a similar thought, mentioning how generative AI helps in understanding code, aids in creative performance, and provides a convenient way of learning. Both groups promote generative AI in the classroom instead of banning the tool, yet how students are assessed remains contentious.

With these considerations in mind, many ideas are being advanced in the literature; the efficacy of these approaches will no doubt be tested in the coming years. This systematic review highlights these early-tested concepts and represents a preliminary phase of our ongoing processes. We anticipate uncovering further insights and tested applications as we continue to analyze papers. Through this, we further aim to bridge the gap between the existing literature on LLMs and their practical applications.

## References

- [1] J. Purtil, "ChatGPT was tipped to cause widespread cheating. Here's what students say happened," *ABC News*, Nov. 21, 2023. Accessed: Feb. 07, 2024. [Online]. Available: <https://www.abc.net.au/news/science/2023-11-22/how-high-school-students-used-chatgpt-2023-education-cheating/103108620>
- [2] H. Duffy and S. Weil, "ChatGPT, Cheating, and the Future of Education | Magazine | The Harvard Crimson," *The Harvard Crimson*. Accessed: Feb. 07, 2024. [Online]. Available: <https://www.thecrimson.com/article/2023/2/23/chatgpt-scrut/>
- [3] O. Terry, "Opinion | I'm a Student. You Have No Idea How Much We're Using ChatGPT.," *The Chronicle of Higher Education*. Accessed: Feb. 07, 2024. [Online]. Available: <https://www.chronicle.com/article/im-a-student-you-have-no-idea-how-much-were-using-chatgpt>
- [4] M. Sullivan, A. Kelly, and P. McLaughlan, "ChatGPT in higher education: Considerations for academic integrity and student learning," *J. Appl. Learn. Teach.*, Jan. 2023, doi: 10.37074/jalt.2023.6.1.17.
- [5] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innov. Educ. Teach. Int.*, vol. 0, no. 0, pp. 1–12, 2023, doi: 10.1080/14703297.2023.2190148.
- [6] S. A. Bin-Nashwan, M. Sadallah, and M. Bouteraa, "Use of ChatGPT in academia: Academic integrity hangs in the balance," *Technol. Soc.*, vol. 75, p. 102370, Nov. 2023, doi: 10.1016/j.techsoc.2023.102370.
- [7] The Walton Foundation, "ChatGPT Used by Teachers More Than Students, New Survey from Walton Family Foundation Finds," Walton Family Foundation. Accessed: Feb. 07, 2024. [Online]. Available: <https://www.waltonfamilyfoundation.org/chatgpt-used-by-teachers-more-than-students-new-survey-from-walton-family-foundation-finds>
- [8] T. T. A. Ngo, "The Perception by University Students of the Use of ChatGPT in Education," *Int. J. Emerg. Technol. Learn. Online*, vol. 18, no. 17, pp. 4–19, 2023, doi: 10.3991/ijet.v18i17.39019.
- [9] A. Shoufan, "Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-Up Survey," *IEEE Access*, vol. 11, pp. 38805–38818, 2023, doi: 10.1109/ACCESS.2023.3268224.
- [10] M. M. Rahman and Y. Watanobe, "ChatGPT for Education and Research: Opportunities, Threats, and Strategies," *Appl. Sci.*, vol. 13, no. 9, Art. no. 9, Jan. 2023, doi: 10.3390/app13095783.
- [11] M. Imran and N. Almusharraf, "Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature," *Contemp. Educ. Technol.*, vol. 15, no. 4, p. ep464, Oct. 2023, doi: 10.30935/cedtech/13605.
- [12] M. Montenegro-Rueda, J. Fernández-Cerero, J. M. Fernández-Batanero, and E. López-Meneses, "Impact of the Implementation of ChatGPT in Education: A Systematic Review," *Computers*, vol. 12, no. 8, p. 153, Jul. 2023, doi: 10.3390/computers12080153.
- [13] Z. H. İpek, A. İ. C. Gözümlü, S. Papadakis, and M. Kallogiannakis, "Educational Applications of the ChatGPT AI System: A Systematic Review Research," *Educ. Process Int. J.*, vol. 12, no. 3, 2023, doi: 10.22521/edupij.2023.123.2.
- [14] A. W. Prananta, R. R. P. M. S. N. Susanto, and J. H. Raule, "Transforming Education and Learning through Chat GPT: A Systematic Literature Review," *J. Penelit. Pendidik. IPA*, vol. 9, no. 11, Art. no. 11, Nov. 2023, doi: 10.29303/jppipa.v9i11.5468.
- [15] M. Pradana, H. P. Elisa, and S. Syarifuddin, "Discussing ChatGPT in education: A literature review and bibliometric analysis," *Cogent Educ.*, vol. 10, no. 2, p. 2243134, Dec. 2023, doi: 10.1080/2331186X.2023.2243134.
- [16] I. Purnama, F. Edi, R. R. P. M. S. N. Wijanarko, "ChatGPT for Teachers and Students in Science Learning: A Systematic Literature Review," *J. Penelit. Pendidik. IPA*, vol. 9, no. 10, Art. no. 10, Oct. 2023, doi: 10.29303/jppipa.v9i10.5259.
- [17] Y. AlBadarin, M. Tukiainen, M. Saqr, and N. Pope, "A Systematic Literature Review of Empirical Research on ChatGPT in Education." Rochester, NY, Sep. 06, 2023. doi: 10.2139/ssrn.4562771.

- [18] C. K. Lo, "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature," *Educ. Sci.*, vol. 13, no. 4, Art. no. 4, Apr. 2023, doi: 10.3390/educsci13040410.
- [19] C. M. L. Phillips, J. S. London, W. C. Lee, A. S. Van Epps, and B. A. Watford, "Reflections on the messiness of initiating a systematic literature review on broadening participation in engineering and computer science," in *2017 IEEE Frontiers in Education Conference (FIE)*, Oct. 2017, pp. 1–8. doi: 10.1109/FIE.2017.8190482.
- [20] L. Krupp *et al.*, "Unreflected Acceptance -- Investigating the Negative Consequences of ChatGPT-Assisted Problem Solving in Physics Education." arXiv, Aug. 21, 2023. doi: 10.48550/arXiv.2309.03087.
- [21] B. Ünlütürk, Z. C. C. Ozdil, and E. S. Cirit, "Exploring the Use of ChatGPT as a Learning and Teaching Tool in Material Science and Nanotechnology Engineering Education." ChemRxiv, Sep. 08, 2023. doi: 10.26434/chemrxiv-2023-4d99t.
- [22] S. Onal and D. Kulavuz-Onal, "A Cross-Disciplinary Examination of the Instructional Uses of ChatGPT in Higher Education," *J. Educ. Technol. Syst.*, p. 00472395231196532, Sep. 2023, doi: 10.1177/00472395231196532.
- [23] C. G. West, "AI and the FCI: Can ChatGPT Project an Understanding of Introductory Physics?" arXiv, Mar. 26, 2023. doi: 10.48550/arXiv.2303.01067.
- [24] M. Z. Naser *et al.*, "Can AI Chatbots Pass the Fundamentals of Engineering (FE) and Principles and Practice of Engineering (PE) Structural Exams?" arXiv, Apr. 02, 2023. doi: 10.48550/arXiv.2303.18149.
- [25] A. Lewkowycz *et al.*, "Solving Quantitative Reasoning Problems with Language Models," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 3843–3857, Dec. 2022.
- [26] C. Geng, Y. Zhang, B. Pientka, and X. Si, "Can ChatGPT Pass An Introductory Level Functional Language Programming Course?" arXiv, May 03, 2023. doi: 10.48550/arXiv.2305.02230.
- [27] C. Cao, Z. Ding, G.-G. Lee, J. Jiao, J. Lin, and X. Zhai, "Elucidating STEM Concepts through Generative AI: A Multi-modal Exploration of Analogical Reasoning." arXiv, Aug. 21, 2023. doi: 10.48550/arXiv.2308.10454.
- [28] S. Elsayed, "Towards Mitigating ChatGPT's Negative Impact on Education: Optimizing Question Design through Bloom's Taxonomy." arXiv, Mar. 30, 2023. doi: 10.48550/arXiv.2304.08176.
- [29] S. K. Bitew, J. Deleu, C. Develder, and T. Demeester, "Distractor generation for multiple-choice questions with predictive prompting and large language models." arXiv, Jul. 30, 2023. doi: 10.48550/arXiv.2307.16338.
- [30] H. McNichols *et al.*, "Automated Distractor and Feedback Generation for Math Multiple-choice Questions via In-context Learning." arXiv, Jan. 11, 2024. doi: 10.48550/arXiv.2308.03234.
- [31] E. Dickey and A. Bejarano, "A Model for Integrating Generative AI into Course Content Development." arXiv, Aug. 24, 2023. doi: 10.48550/arXiv.2308.12276.
- [32] T. Wang, D. Vargas-Díaz, C. Brown, and Y. Chen, "Exploring the Role of AI Assistants in Computer Science Education: Methods, Implications, and Instructor Perspectives," in *2023 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, Oct. 2023, pp. 92–102. doi: 10.1109/VL-HCC57772.2023.00018.
- [33] J. K. Matelsky, F. Parodi, T. Liu, R. D. Lange, and K. P. Kording, "A large language model-assisted education tool to provide feedback on open-ended responses." arXiv, Jul. 25, 2023. doi: 10.48550/arXiv.2308.02439.
- [34] N. Kiesler, D. Lohr, and H. Keuning, "Exploring the Potential of Large Language Models to Generate Formative Programming Feedback," in *2023 IEEE Frontiers in Education Conference (FIE)*, Oct. 2023, pp. 1–5. doi: 10.1109/FIE58773.2023.10343457.
- [35] P. Sridhar, A. Doyle, A. Agarwal, C. Bogart, J. Savelka, and M. Sakr, "Harnessing LLMs in Curricular Design: Using GPT-4 to Support Authoring of Learning Objectives." arXiv, Jun. 30, 2023. Accessed: Dec. 13, 2023. [Online]. Available: <http://arxiv.org/abs/2306.17459>
- [36] D. R. Krathwohl, "A Revision of Bloom's Taxonomy: An Overview," *Theory Pract.*, vol. 41, no. 4, pp. 212–218, Nov. 2002, doi: 10.1207/s15430421tip4104\_2.



- [37] T. Phung *et al.*, “Generative AI for Programming Education: Benchmarking ChatGPT, GPT-4, and Human Tutors,” in *Proceedings of the 2023 ACM Conference on International Computing Education Research - Volume 2*, Chicago IL USA: ACM, Aug. 2023, pp. 41–42. doi: 10.1145/3568812.3603476.
- [38] H. Arndt, “AI and Education: An Investigation into the Use of ChatGPT for Systems Thinking.” arXiv, Jul. 26, 2023. doi: 10.48550/arXiv.2307.14206.
- [39] T. Sherborne, “Mapping the Curriculum: How Concept Maps can Improve the Effectiveness of Course Development,” in *Knowledge Cartography: Software Tools and Mapping Techniques*, A. Okada, S. B. Shum, and T. Sherborne, Eds., in *Advanced Information and Knowledge Processing*, London: Springer, 2008, pp. 183–198. doi: 10.1007/978-1-84800-149-7\_9.
- [40] A. Jackson, E. Barrella, and C. Bodnar, “Application of concept maps as an assessment tool in engineering education: Systematic literature review,” *J. Eng. Educ.*, vol. n/a, no. n/a, doi: 10.1002/jee.20548.
- [41] C. Zastudil, M. Rogalska, C. Kapp, J. Vaughn, and S. MacNeil, “Generative AI in Computing Education: Perspectives of Students and Instructors.” arXiv, Aug. 08, 2023. Accessed: Feb. 07, 2024. [Online]. Available: <http://arxiv.org/abs/2308.04309>

## Appendix A: Example Applications of using GPT 3.5 to Author Learning Outcomes

Conversation Date: 3/31/2024

Model: GPT 3.5 using ChatGPT platform

Prompt: Sridhar et al. [35]

User:	<p>You are a curricular development expert system focused on authoring LOs. Learning outcomes are brief, clear statements that describe the desired learning outcomes of instruction. LOs should use action verbs. LOs should be measurable.</p> <p>A well-constructed learning objective contains three parts</p> <ol style="list-style-type: none"> <li>1. BEHAVIOR The behavior is the real work to be accomplished by the student specified by an action verb that connotes observable and measurable behaviors.</li> <li>2. CONDITIONS This is a statement that describes the exact conditions under which the defined behavior is to be performed.</li> <li>3. DEGREE This is a statement that specifies how well the student must perform the behavior</li> </ol> <p>Conceptual LOs are focused on students' knowledge and understanding (i.e., the first two levels of Bloom's taxonomy). Project LOs are focused on students' skills and behaviors (i.e., the higher levels of Bloom's taxonomy).</p> <p>Here are some criteria to satisfy in order to create an effective learning outcome:</p> <ol style="list-style-type: none"> <li>1. LOs should be student-centered.</li> <li>2. LOs should focus on specific cognitive processes.</li> <li>3. LOs should use action verbs.</li> <li>4. LOs should be measurable.</li> </ol> <p>The user will provide you with the name of the course, brief description of the course goals, the name of the module, and the type of the LOs to be developed. Based on these you respond with a list of well-designed effective LOs (5-10 items).</p>
GPT 3.5:	<p><i>[generates learning outcomes for an "Introduction to Programming" course unprompted...]</i></p>
User:	<p>Name of the course: Introduction to Engineering Description: The course introduces methods of generating and exploring creative ideas and alternatives for designing engineering systems. The course also introduces concepts in creativity, innovation, engineering fundamentals, and problem-solving methodologies. Students learn, through experience, the process of design and analysis in engineering including how to work effectively on a team. Finally, they develop skills in project management, sustainability, written, oral and graphical communication, logical thinking, and modern engineering tools (e.g., Excel, Python, LabVIEW, MATLAB, Visual Basic, CAD, Rapid Prototyping). Module Name: Brainstorming methods Type of learning outcomes: Conceptual LOs</p>
GPT 3.5:	<ol style="list-style-type: none"> <li>1. Behavior: Describe the principles of brainstorming as a method for generating creative ideas in engineering design. Conditions: Given a theoretical overview of brainstorming techniques. Degree: Demonstrate understanding by articulating the key components and objectives of brainstorming sessions in engineering contexts.</li> </ol>

2. Behavior: Explain the importance of divergent and convergent thinking in the brainstorming process. Conditions: Presented with examples of divergent and convergent thinking activities. Degree: Differentiate between divergent and convergent thinking strategies and their respective roles in idea generation and refinement.
3. Behavior: *Analyze* the effectiveness of various brainstorming techniques in different engineering design scenarios. Conditions: Presented with case studies or scenarios involving engineering design challenges. Degree: Evaluate the suitability of different brainstorming methods based on factors such as problem complexity, team dynamics, and project constraints. *[This LO is not within the first two levels of Bloom's taxonomy as asked.]*
4. Behavior: *Critique and improve* upon brainstorming sessions to enhance creativity and productivity. Conditions: Participating in or observing a brainstorming session. Degree: Identify strengths and weaknesses of the brainstorming process, proposing modifications or enhancements to optimize outcomes. *[LOs should not have more than one action verb; these verbs also do not correspond to the first two levels of Bloom's taxonomy.]*
5. Behavior: Relate the concept of ideation to the iterative nature of the engineering design process. Conditions: Given an overview of the engineering design process and its iterative nature. Degree: Recognize how ideation, particularly through brainstorming, contributes to the iterative refinement of design solutions in engineering projects.
6. Behavior: Discuss the role of interdisciplinary collaboration in fostering innovative ideas during brainstorming sessions. Conditions: Presented with examples of interdisciplinary engineering projects. Degree: Articulate how diverse perspectives and expertise from multiple disciplines contribute to the richness of ideas generated during brainstorming sessions.
7. Behavior: Identify potential barriers to effective brainstorming and propose strategies to overcome them. Conditions: Given examples of common challenges encountered during brainstorming sessions. Degree: Demonstrate awareness of factors such as groupthink, lack of participation, and dominance dynamics, offering solutions to mitigate these obstacles.
8. Behavior: *Evaluate* the ethical implications of design decisions resulting from brainstorming sessions. Conditions: Presented with ethical dilemmas related to engineering design choices. Degree: Assess the ethical considerations inherent in the ideation and selection of design solutions, considering factors such as safety, sustainability, and societal impact. *[This LO is not within the first two levels of Bloom's taxonomy as asked.]*
9. Behavior: *Reflect* on personal contributions to and experiences in brainstorming sessions, *identifying* areas for improvement. Conditions: Participating in multiple brainstorming activities throughout the module. Degree: Engage in self-assessment and reflection to recognize strengths and weaknesses in individual ideation skills and collaborative practices, setting goals for personal development. *[LOs should*

	<p><i>not have more than one action verb and should be rewritten to foreground identifying areas for improvement.]</i></p> <p>10. Behavior: <i>Synthesize</i> concepts from multiple disciplines to propose innovative solutions during brainstorming exercises. Conditions: Presented with interdisciplinary design challenges requiring creative problem-solving. Degree: Integrate knowledge and techniques from various engineering disciplines and related fields to generate novel ideas and approaches, demonstrating interdisciplinary thinking prowess. <i>[This LO is not within the first two levels of Bloom's taxonomy as asked.]</i></p>
--	--

Conversation Date: 3/31/2024  
 Model: GPT 3.5 using ChatGPT platform  
 Prompt: Dickey and Bejarano [31]

User:	<p>Act as an <u>expert engineering curriculum developer</u>. The topic we are covering is brainstorming methods in the context of engineering design. Students know <u>the general premise of brainstorming, general science principles from physics, biology, and chemistry from high school, and mathematics up to the first semester of Calculus</u>. Students are in their <u>first year of studies in engineering</u>. Give me <u>10</u> learning outcomes for the current topic. Use Bloom's revised taxonomy of verbs for the outcomes. <i>[underlined elements are inserted context from base prompt]</i></p>
GPT 3.5:	<p>Certainly! Here are 10 learning outcomes for the topic of brainstorming methods in the context of engineering design, aligned with Bloom's Revised Taxonomy:</p> <ol style="list-style-type: none"> <li>1. Identify various brainstorming techniques used in engineering design.       <ul style="list-style-type: none"> <li>• Verb: Identify</li> </ul> </li> <li>2. Explain the importance of brainstorming in the engineering design process.       <ul style="list-style-type: none"> <li>• Verb: Explain</li> </ul> </li> <li>3. Apply different brainstorming methods to generate creative solutions for engineering problems.       <ul style="list-style-type: none"> <li>• Verb: Apply</li> </ul> </li> <li>4. Analyze the effectiveness of different brainstorming techniques in generating innovative ideas.       <ul style="list-style-type: none"> <li>• Verb: Analyze</li> </ul> </li> <li>5. Evaluate the advantages and limitations of specific brainstorming methods in different engineering design contexts.       <ul style="list-style-type: none"> <li>• Verb: Evaluate</li> </ul> </li> <li>6. Create a structured plan for conducting a successful brainstorming session in a team setting.       <ul style="list-style-type: none"> <li>• Verb: Create</li> </ul> </li> <li>7. Compare and contrast traditional brainstorming with electronic or virtual brainstorming tools.       <ul style="list-style-type: none"> <li>• Verb: Compare and contrast</li> </ul> </li> <li>8. Synthesize multiple ideas generated during a brainstorming session into cohesive design concepts.       <ul style="list-style-type: none"> <li>• Verb: Synthesize</li> </ul> </li> <li>9. <i>Design and implement</i> variations of brainstorming techniques tailored to specific engineering design challenges. <i>[LOs should not have more than one action verb.]</i> <ul style="list-style-type: none"> <li>• Verb: Design and implement</li> </ul> </li> </ol>

	<p>10. Propose improvements or adaptations to existing brainstorming methods to enhance creativity and efficiency in engineering design processes.</p>
--	--

- Verb: Propose