Work-In-Progress: Using AI Tools for Supplemental Learning in a Lean Concepts Class

Hugh McManus
Mechanical and Industrial Engineering Department
Northeastern University
Boston, MA, USA
h.mcmanus@northeastern.edu

***Abstract*—In this innovative practice Work-In-Progress paper, Generative AI tools were used in a supplemental assignment in a Lean Concepts class. Students were directed to create a summary of a specific topic covered only partially in the regular class material. Students assigned to the same topic then formed four- or five-student groups and reflected on the effectiveness of the AI tools. The accuracy, consistency and usefulness of the AI-created material was considered. Students also reflected on their own learning from the exercise and their possible future use of AI tools. The results, from 78 students, showed that seven different generative AI tools, with a wide range of prompts, gave similar answers. The answers were judged to be useful and accurate; only two instances of the AI being widely off-topic were seen. The answers were rather vague and general; even when prompted for different topics, the AI results had much text in common. All student groups agreed that human interpretation would be needed before the AI-generated material could be used either for presentation or work. They felt the exercise would make it more likely that they will use generative AI tools in the future. The exercise demonstrates both the usefulness and limitations of AI tools for supplemental learning in technical classes. The exercise was easy to run and could be replicated in any introductory or survey class.**

Keywords—Artificial Intelligence, Industrial Engineering, reflective learning, critical thinking, lifelong learning

I. Introduction

This innovative practice Work-In-Progress paper describes an attempt to include Generative Artificial Intelligence (AI) tools in an independent research project that is part of a course in Lean Process Improvement methods. The course covers many different tools and techniques. It uses a variety of learning modes, including lectures, exercises, simulations, and case studies. This is both best for student learning, and necessary given the experience-based, context-dependent nature of the material. To supplement the course material, students were asked to do a short research project on a particular topic of interest to them. In self-selected teams of 1-5 people, students collected material on a topic or specific application, and presented their results to the class in the last few days of the semester.

This work was supported by discretionary funding provided by the Department of Mechanical and Industrial Engineering, Northeastern University

An experimental exercise included the use of AI tools in the research project. Its original inspiration was student attempts to create a topic abstract (a required step in the project) using AI tools. The results were not acceptable, but they were intriguing. As a supplementary "mini project," teams of students created topic summaries with AI tools and critiqued the results. Individual students created one-page summaries on a lean sub-topic (e.g. Kanban, Design of Experiments, Single Minute Exchange of Dies, etc.), using multiple publicly available AI tools. Students then joined randomly-selected small groups (4-5 students) and compared and contrasted the individual results. They critically assessed how useful the results were. Could they be used as is? Could they be the basis of further efforts? Were there inaccuracies or gaps in the information? The students were graded on the quality of the critique rather than the AI output.

The exercise had three objectives for the students. The first was to give the students an experience in using generative AI tools to provide supplemental information on the course topic. The second was to provide a template for learning more after the class was over, and thus explore a new pathway to lifelong learning. Finally, the students were given an opportunity in teams to think reflectively and critically on their individual work. It was hoped that this would make them more sophisticated customers of AI products.

For the course staff, this was an opportunity to experiment with the use of AI tools as an instructional supplement. The results will be used to inform further work in this area, including an updated version of the exercise for future terms.

II. Background

The work was not derived directly from any previous work, but it was informed by existing knowledge in several broad areas. This paper will provide only a brief and high-level review. The literature on learning *about* AIs and AI related skills is new but fairly well populated. Dori and Magerko, provide an extensive literature review and a framework for AI literacy and its teaching [1]. The *use* of AIs to improve education is even newer, and much of the work is speculative. A recent review of the literature is provided by [2]. Qadir provides an good round-up of the possibilities in [3], which includes a section written by ChatGPT(!). Possible uses for generative AI in university classrooms are identified. These

include language and writing help, personalized learning, learning outside of classroom hours, and enhanced engagement and retention. Pitfalls such as plagiarism, overreliance, and misinformation are identified. Instructors are challenged to teach acceptable use (as opposed to plagiarism), foster creative and critical thinking, and leverage prior knowledge to get the most out of AI tools.

The literature is growing rapidly. Relevant recent works include attempts to use AI to enhance and/or simulate student reflections; it was found that AIs were useful for enhancing student reflection exercises but did a poor job of actually reflecting on a subject [4]. Another study examined student acceptance of help from AIs. It was found that, surprisingly, students placed less value on information gained from easy-to-use sources [5]. Some training in prompt engineering was found to be very useful for students, both in their skills and their confidence in their interactions with AIs [6].

The direct *application* of AI to the subject of the course, Lean Six-Sigma, is thinly studied to date. Most available material is speculative or promotional in nature. The Harvard Business Review provides a recent short article which refers to (but does not describe) a few application cases, and which calls for Lean Six-Sigma professionals to learn AI tools [7]. The course in which the current work was done already has an emphasis on multiple learning modes [8] and the author is actively investigating reflective learning [9]. In this context the use of generative AI is yet another learning mode to take advantage of.

III. THE EXERCISE

Students were assigned a topic to explore using a generative AI of their choice. The students were randomly assigned to teams of 4-5 people, and each team was given a topic (see Table 1). These topics are deliberately obscure – they contain Japanese words, jargon, and acronyms – but they are all well understood terms of art in Lean Six Sigma. Some of the topics had been covered in the class to varying levels of depth; some were new to the students.

The exercise had two components. The first part of the exercise was for students to individually select a generative AI, and ask it for a one-page summary of the topic. The AI and exact prompt were not specified. The students turned in the results, along with the prompt and AI that they used.

TABLE I. TOPICS

Groups	Topic
1, 9	Kanban
2, 10	Poka-Yoke (mistake Proofing)
3, 11	SMED (Single minute exchange of dies)
4, 12	Heijunka
5, 13	Jidoka
6, 14	Kitting for point of use
7	Flow Cell (U-Cell)
8	Hosin Kanri

The second part of the exercise was done in groups. The students were asked to share their individual responses, and then, as a group, answer these questions:

1. What do you think of the overall quality of the responses?
2. Compare the AI responses. Are there common themes that most or all of the responses hit?
3. Contrast the AI responses. What differs between the responses?
4. Could you use (at work) or present (like for an end-of-term presentation) the results as is, or are there things missing, too vague, or wrong (!) in the responses.
5. How did this exercise affect your likely use of AI in the future?

All but one student (for a total of 78 students) did the individual component; all 14 groups did the group assignment. The students didn't ask any questions about the assignment, and apparently had no difficulty doing the individual assignment, meeting in groups (zoom or teams was suggested), and completing the group discussion.

IV. THE RESULTS

Results are presented in three sections. First, the results of the individual part of the exercise is briefly discussed. The group reflections are discussed in more detail. They are paired with faculty observations on the AI results and the students' reaction to them. Finally, the overall success of the exercise is evaluated.

A. *Individual Exercises*

The students used 7 different AIs to create one page summaries of lean topics, using their own prompt and choice of AI tool. ChatGPT was by far the most commonly used AI. The version was usually unspecified, although students sometimes noted 3.5 or 4.0; these were lumped into a single ChatGPT category in the analyses. The distribution of AIs used can be seen in Fig. 1. The student's choice of prompts were varied but tended to be simple and direct. A typical example was "Give me a one page summery on Poka Yoke." The effects of AI and prompt choice are discussed in the next subsection.

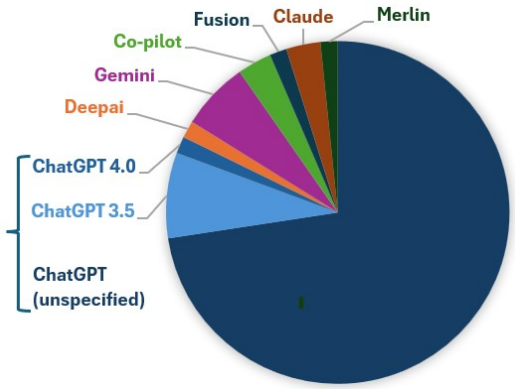

Fig. 1. Distribution of AIs used by students for the individual exercise

The responses to the individual exercises were on topic 97% of the time (76 out of 78 cases). The basic definitions were correct, and did not vary much with AI used or exact prompt. The exceptions included one apparent hallucination and one off-topic response. One ChatGPT response to "Create a one-page summary on the history of Heinjuka (sic) and its application" resulted in "Heinjuka, originating from Japan, merges ancient traditions with contemporary practices to cultivate peace and harmony. Rooted in Zen Buddhism and non-violence principles, it advocates for inner peace as the foundation of societal well-being." This appears to be a full-on hallucination; a web search on the (mis-spelled) term came up with many cases of mis-spellings of Heijunka in lean practice articles, but no Buddhist philosophy. ChatGPT also created a definition of Flow Cell (U-Cell) that involved design of electrochemical battery components rather than the expected definition involving manufacturing line layouts.

B. Group Reflections

The groups all handed in complete answers to the questions. This section will be organized around the exercise questions. A summary of student responses will be followed by a faculty reflection.

1) What do you think of the overall quality of the responses?

The students mostly liked the AI responses, as seen in Fig. 2. Most groups though they were very good, using terms such as thorough, clear, in-depth and solid. Some were less enthusiastic, either saying the responses were good without elaborating, or noting some lack of detail or examples. Two groups had reservations, noting a lack of details, depth or examples in the AI output.

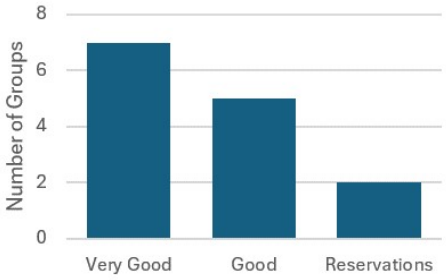

Fig. 2. Student evaluations of the quality of AI topic summaries

The faculty noted some issues that the students seemed to be unaware of. The observation in [3] that AI will be more valuable to those who already have some knowledge seems relevant here. The students did not have deep knowledge of the topics, and therefore missed some issues with the AI responses. The more skeptical students caught a basic weakness of the AI responses – they were generic, without details or specific examples, even when the prompts ask for them. The students noted, but did not call out, the hallucination and off topic response noted in the previous section – they described the responses as providing breadth of application of the lean concepts.

Another weakness of the AI responses noted by the faculty was a strong bias towards topics and interpretations currently under active discussion on the internet. The topic Kanban was meant to include the use of Kanban methods for inventory control; the AI responses were almost exclusively aimed at the currently trending use of Kanban methods for project management and time management for multi-tasking teams and individuals. Some combination of the lack of knowledge on the students' part and the vagueness of the AI summaries caused this to go unnoticed by the students.

2) Compare the AI responses. Are there common themes that most or all of the responses hit?

Ten of the 14 student teams saw very strong similarities in their AI responses. These were often expressed as lists or bullet lists of topics covered. They tended to include not only technical aspects of the topic, but also benefits and implementation strategies. Three teams did note small variations in the responses based on both the AI used and the prompt. One of these teams noted that despite the differences in length and formatting, the underlying information was very consistent. One team was a bit of an outlier, noting only the history of the subject was treated uniformly, but without comment on the technical content.

From the perspective of the faculty, the AI responses regardless of topic tended to discuss generic benefits of lean improvement such as improved efficiency, quality, responsiveness and flexibility. These benefits were unsubstantiated and often stated in promotional language, perhaps reflecting the source material on which the AIs were trained. The responses also tended to make generic implementation recommendations such as continuous improvement, culture change, and teamwork regardless of the actual topic. The students were only looking at a single topic so did not get this perspective.

3) Contrast the AI responses. What differs between the responses?

The students noticed some differences in the AI answers, summarized in Table II. It is in this section that they pointed out the off-topic answers, as "different examples." Most of the differences were small, although one group noted (as a positive) that the responses they received gave differing perspectives on the topic.

TABLE II. DIFFERENCES NOTED BETWEEN AI RESPONSES ON THE SAME TOPIC

Number of Groups	Differences Noted
6	Structure – bullets list, paragraphs, etc.
6	Specific examples used
5	Focus – implementation, benefits, tools, etc
4	Level of detail, length of answer
3	Words used, writing style

4) Could you use (at work) or present (like for an end-of-term presentation) the results as is, or are there things missing, too vague, or wrong (!) in the responses.

The students were unanimous in answering that the AI generated material was not sufficient to be presented or used. They all noted that more specific information would be needed. Two groups gave long bullet lists of things that would need to

be added before the AI product could be used professionally. One team was particularly cautious about using AI results in their present state. They noted:

> *"We think… the AI gives good structure for the topic, but it clearly lacks depth and specificity. It … uses fancy language and plenty of jargon. [The AI responses] not only lack detail, but clearly overuse generic information and use lengthy and wordy sentences. They are quite overoptimistic and sometimes reluctant to explicitly mention any … weakness in the matter it was asked about [10]."*

It appears at least some of the students are skeptical, and all are cautious, about using AI results for work or presentation. It is not certain that this is a learning from this exercise, as students may have prior experience with AIs. They may have heard this caution expressed by faculty, which might create an expectation that this is the "right" answer. They may also have had low expectations of the AI results based on the ease of the exercise [5]. We explicitly asked them not to use AIs to write the summary, so hopefully this is not a factor, although no explicit controls or checks were put in place. In any case, it is encouraging to see a thoughtful approach to using AI expressed by the students.

5) How did this exercise affect your likely use of AI in the future?

The students were unanimous in saying they would use AI more in the future based on their experience. Only two groups suggested that the exercise was entirely successful and they would do similar things in the future. Twelve groups out the 14 noted that they would be careful to use AI not as a sole source for information. How this would work varied from a cynical group that postulated AI would ultimately prove just a better search engine, to groups that thought they would use AIs for brainstorming and quick summaries. Typically, the students suggested that AI queries would be a good place to *start* an exploration that would also require human judgment. One group summed up their experience nicely: "[The exercise] showcases AI's capability in synthesizing information swiftly but also highlights the need for human oversight [11]."

C. Faculty Reflection

As a supplement to the course, the AI summaries were of limited use in an absolute sense. The summaries provided were usually correct. Some topics, notably the Single Minute Exchange of Die (SMED) method, were defined quite well. The steps needed to implement the method were also laid out correctly. However, even in this case, the wording was generic and vague. Examples and illustrations of the sort necessary for a deep enough understanding to actually use the method were lacking. On the positive side, these summaries were very easy to create and provided a starting point for further explorations. Finally, the exercise probably engaged the students more deeply than a short lecture would have.

As a way of introducing the students to AI applications in Lean Six Sigma in particular, and engineering topics more generally, the exercise was quite successful. The students had no trouble using the tools, and the answers to the reflective exercises indicated that they learned something about the tools' use and capabilities.

V. Summary and Conclusions

The exercise was extremely successful as an experiment. The results provide guidance for further work in supplemental learning of Lean Six Sigma topics, teaching of generative AI tools, and use of these tools in the engineering profession.

Care must be taken not to overstate the value of the exercise. Just as the subject matter learning from the AI responses was somewhat superficial, the learning about the AIs was not deep. Based on the AI education framework in [1], the exercise covers only three of the 17 Competencies (Recognizing AI, AI's Strengths and Weaknesses, and Imagine Future AI) and one of 15 Design Elements (Critical Thinking). The students were not given any guidance on creating the prompts, which may have negatively impacted both the quality of the results and the students' confidence in them [6]. That said, the exercise does at least introduce future Lean Six Sigma professionals to AI concepts, as called for in [7].

The exercise could be improved. The students clearly lacked the knowledge necessary to see where the AI responses may have been incorrect or weak. They also did not get to see how the results were generic across different topics. Doing the exercise earlier in the term would allow one more round of reflection, at the class level. Students could share findings from their teams with the whole class, and faculty could provide some perspective on what was good, bad, or ugly about the results. We plan on doing so in future iterations of the class, and will report the results in the conference presentation.

The exercise is extremely simple to run, and could easily be implemented in other courses. Anyone reading this is welcome to do so. The specific topics and questions would probably change, but the overall approach should work.

Acknowledgment

This work was supported by discretionary funding provided by the Department of Mechanical and Industrial Engineering, Northeastern University.

References

[1] D. Long and B. Magerko, "What is AI literacy? competencies and design considerations," *Proc. of the 2020 CHI Conf. on Human Factors in Computing Systems*, CHI '20, April 25–30, 2020, Honolulu, HI, USA, doi: 10.1145/3313831.3376727

[2] S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Rittberger, and H. Drachsler, "Are we there yet? - A systematic literature review on chatbots in education," *Frontiers in artificial intelligence*, vol. 4, 2021, doi: 10.3389/frai.2021.654924

[3] J. Qadir, "Engineering education in the era of ChatGPT: promise and pitfalls of generative AI for education," *2023 IEEE Global Engineering Education Conference (EDUCON),* Kuwait City, Kuwait, 2023, pp. 1-9, doi: 10.1109/EDUCON54358.2023.10125121

[4] K. Kanont et al. 2024. "Generative-AI, a Learning Assistant? Factors Influencing Higher-Ed Students' Technology Acceptance," *Electronic Journal of e-Learning*, vol. 22, no.6, pp. 18-33, 2024, doi: 10.34190/ejel.22.6.3196

[5] M. Håkansson Lindqvist and C. Arvidsson, "Exploring Student and AI Generated Texts: Reflections on Reflection Texts," *Electronic Journal of e-Learning*, vol. 22 no. 6, pp. 52-59, 2024, doi: 10.34190/ejel.22.6.3473

[6] D. J. Woo, D. Wang, T. Yung and G. Kai, "Effects of a Prompt Engineering Intervention on Undergraduate Students' AI Self-Efficacy, AI Knowledge, and Prompt Engineering Ability: A Mixed Methods Study," Researchgate Preprint, July 2024, doi: 10.13140/RG.2.2.32577.36968

[7] M. Holweg, T. Davenport, and K. Sneyder, "How AI fits into lean six sigma," *Harvard Business Review*, Nov. 9, 2023.

[8] E. Murman, H. McManus, and A. Weigel, "The LAI lean academy experience: introductory lean curriculum," *Journal of Enterprise Transformation,* vol. 4, no. 3, pp. 205-225, 2014.

[9] H. McManus and B. Jaeger-Helton, "Reflective exercises to bridge the gap in a two-term industrial engineering capstone sequence," *Capstone Design Conference 2024*, June 3-4, 2024, Knoxville TN.

[10] A. Almuhanna, A. Baddipudi, G. Eeraboina, D. Mehta, S. Nikalje and S. Patil, "Lean Concepts and Applications AI-Special Extra Credit Assignment," Spring 2024 (unpublished).

[11] S. Antony, N, Budh, R. Butani, H. Dalal, H. Jain and Lavanya, "Lean Concepts and Applications AI-Special Extra Credit Assignment," Spring 2024 (unpublished).