2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #51539

Beyond One-Off Prompts: Bloom-Aligned Question Chaining in a GPT-based Manufacturing Adviser

Fatemeh Karimi Kenari, University of North Carolina at Charlotte
Mahmoud Dinar, University of North Carolina at Charlotte

Mahmoud Dinar is an assistant professor of mechanical engineering. His main research interest involves integrating AI in multimodal computational frameworks to understand and aid design for manufacturing. His ongoing projects are creating a conversational platform to evaluate and teach manufacturing skills to the future workforce, and solving geometric puzzles to design sustainable and resilient products as a 3D composition of pieces.

Beyond One-Off Prompts: Bloom-Aligned Question Chaining in a GPT-based Manufacturing Adviser

Introduction

Manufacturing education is undergoing transformation as advances in automation, digital manufacturing, and data-driven decision making continue to reshape industrial practice and workforce expectations. As manufacturing systems become more interconnected and digitally mediated, educational programs are expected to prepare learners not only to recognize individual processes and parameters, but also to reason across sequences of decisions and constraints. This shift highlights a mismatch between traditional, static instructional approaches and the dynamic reasoning demands of contemporary manufacturing environments, motivating renewed interest in educational models that can adapt to evolving learner needs [1]. One response to this challenge has been the development of adaptive intelligent tutoring systems designed to accommodate variation in learner background, pace, and goals. Prior work in this area emphasizes tailoring instructional content and progression based on observed learner behavior rather than relying on uniform instructional sequences. Such systems typically incorporate explicit models of learner state and progression, enabling targeted feedback and guidance as learners advance through increasingly complex material [2].

More recently, the emergence of generative AI and large language models has introduced new opportunities for flexible, dialogue-based educational support. At the same time, this work has emphasized that simply deploying generative models is insufficient; educational effectiveness depends on how such models are guided, constrained, and integrated into pedagogical frameworks [3]. Studies exploring LLMs as virtual tutors in engineering education further suggest that these models can support explanation, exploration, and conceptual reinforcement when appropriately framed. However, they also note that effective tutoring behavior does not emerge automatically from isolated prompts. Instead, interaction structure and progression play a central role in shaping responses, motivating principled approaches to question sequencing in LLM-based educational systems [4].

Underlying many adaptive tutoring approaches is the view that learning unfolds over time as a structured progression rather than as a collection of independent responses. Foundational research on procedural knowledge acquisition frames learning as a sequence of cognitive states, where earlier actions and decisions influence subsequent performance. This perspective motivates analyzing educational interactions as ordered sequences and provides a theoretical basis for transition-based models of learning behavior [5].

Advances in natural language processing have led to growing interest in conversational question answering as a mode of educational interaction. Surveys of conversational Q/A systems show that multi-turn interactions differ fundamentally from single-turn question answering, particularly in their reliance on dialogue history and contextual continuity. In these settings, the meaning and relevance of a question often depend on what has been asked previously, underscoring the importance of question sequencing in conversational learning environments [6].

Related work on question generation reinforces this emphasis on sequence and flow. Curiosity-driven question generation studies view questions as mechanisms for guiding exploration, where successive questions are shaped by evolving informational needs rather than posed independently. Although not specific to manufacturing education, these approaches support the

broader idea that structured question progression can be used to scaffold learning in complex domains [7].
This paper presents a step toward structured, adaptive conversational tutoring for manufacturing education by focusing on the role of question sequencing. Building on prior development of a GPT-based Manufacturing Advisor, the work introduces a probabilistic framework for generating fixed-length chains of manufacturing questions WH-type questions aligned with Bloom's taxonomy. Rather than treating questions as independent prompts, the system models how questions evolve across a short interaction for learners at different expertise levels. The study evaluates how this sequencing affects model behavior by comparing responses to identical questions asked in isolation versus within guided chains. Using aggregate textual metrics, the paper investigates whether probabilistic question chaining encourages more elaborate, context-aware explanations. The results are intended as exploratory evidence that structured question progression can move LLM-based educational systems beyond one-off prompting toward more coherent, tutor-like interactions in manufacturing education.

2. Methodology

2.1. Question Template Design and Classification

The foundation of the methodology is a structured manufacturing question framework. Questions were organized to reflect four complementary dimensions of manufacturing knowledge: general process understanding, sub-process mechanisms, process parameters, and design for manufacturability. These categories capture increasing levels of specificity and decision-making commonly encountered in manufacturing education, ranging from high-level conceptual understanding to parameter-level and design-oriented reasoning.
Each question was further classified by its WH-form, including *what*, *why*, and *how*. The WH-type serves as a proxy for the learner's cognitive intent, distinguishing between descriptive, explanatory, and procedural modes of inquiry. This dual classification by knowledge category and WH-type enables the system to reason both about the manufacturing aspect being addressed and the form of engagement implied by the question.
To enable scalable and controlled question generation, the questions were constructed using a template-based design approach. Approximately 1,900 manufacturing questions were created by defining reusable WH-type templates across the four knowledge categories. Each template specifies a WH-form, a syntactic structure, and placeholders for manufacturing-relevant concepts, such as processes, materials, or features. Templates were instantiated using a curated domain vocabulary to generate concrete natural language questions.

2.2. Bloom's Taxonomy–Based Question Design Using WH-Types

The design of the question framework and the resulting question chains in this work is grounded in Bloom's revised taxonomy, which provides a structured framework for organizing cognitive processes from lower-order recall to higher-order creation and evaluation [8]. Bloom's taxonomy is suited for manufacturing education, where learners must progressively move from recognizing basic process concepts to reasoning about trade-offs, constraints, and design decisions. Rather than treating questions as isolated prompts, this framework enables questions to be interpreted as indicators of the type and depth of cognitive engagement being elicited. Table 1 summarizes the

Bloom's taxonomy levels used in this work, along with representative cognitive processes, action verbs, and example manufacturing questions.

Table 1: Bloom's taxonomy levels and example WH-type manufacturing questions used to guide question design.

Level	Cognitive Process	Example Action Verb	Example Question
1	Remember	define, list, recall	What is milling?
2	Understand	explain, summarize	Why does surface finish matters is milling?
3	Apply	use, implement	What is the formula for calculating MRR in drilling?
4	Analyze	compare, differentiate	How do casting and forging compare when producing complex shapes?
5	Evaluate	justify, critique	What are the criteria for choice of material for a CNC part
6	Create	design, develop	How would you design a process plan for high-volume turning?

In this study, WH-type questions are used as observable linguistic realizations of Bloom-aligned cognitive processes. Lower levels of the taxonomy, including *Remember* and *Understand*, are primarily associated with *what* and simple *why* questions that focus on defining terms, listing characteristics, or explaining basic concepts. For example, a question such as "What is milling?" reflects recall and conceptual recognition, while "Why surface finish matters in milling?" prompts explanation and comprehension of underlying principles. These forms are essential for establishing foundational knowledge and are therefore prevalent in questions intended for beginner-level learners.

As learners progress to the *Apply* level of Bloom's taxonomy, questions increasingly emphasize procedural usage and calculation. These questions often retain a *what* or *how* structure but shift in intent from description to execution, such as "What is the formula to calculate material removal rate in drilling?" Here, the WH-form signals not only the type of information being requested, but also the expectation that the learner can operationalize previously acquired knowledge in a concrete context. This distinction is critical in manufacturing education, where procedural fluency is a key learning objective.

Higher-order cognitive processes in Bloom's taxonomy, *Analyze*, *Evaluate*, and *Create* are reflected through more complex *why* and *how* questions that require comparison, justification, and synthesis. Questions such as "How do casting and forging compare when producing complex shapes?" align with the *Analyze* level by requiring differentiation across processes, while evaluative prompts like "What are the criteria for choice of material for a CNC part?" require learners to justify decisions based on multiple constraints. At the highest level, *Create*, questions such as "How would you design a process plan for high-volume turning?" require learners to integrate knowledge across processes, parameters, and production goals to generate novel solutions.

By mapping WH-type questions to Bloom's taxonomy in this manner, the methodology provides a principled mechanism for controlling the cognitive demands of generated questions without explicitly labeling users by expertise. Instead, learner level is reflected implicitly through the distribution and progression of Bloom-aligned WH-types within question sequences. Beginner-oriented question sets emphasize lower and intermediate Bloom levels to support scaffolding and concept formation, while advanced question sets concentrate on analysis, evaluation, and creation to reflect expert reasoning patterns. Representative examples of Bloom-aligned WH-type questions used in this study for a non-expert user are shown in table 1.

Table 1: Example of a generated question chain for non-expert user

Step	Question	Bloom's Level	Cognitive Process
Q1	What is bulk deformation?	Remember	Recall the definition of a key concept.
Q2	What are the main steps involved in bulk deformation?	Understand	Summarize or describe a process.
Q3	Why is rolling important in bulk deformation?	Analyze	Explain the role of a sub-process (rolling) within a broader context.
Q4	How is temperature controlled during bulk deformation?	Apply	Show understanding by explaining real-world control mechanisms.
Q5	Why does changing temperature affect the outcome of deformation?	Evaluate	Justify and reason about process sensitivity and outcomes.

To evaluate the effect of question sequencing, two prompting conditions were defined. In the isolated condition, each question was posed to the language model in a separate interaction, with no prior dialogue history retained between queries. Practically, this corresponds to initiating a new session for each question. In contrast in the chained condition, questions were presented sequentially within a single interaction, allowing earlier questions and responses to influence subsequent responses.

Results

The effect of question sequencing on model responses was examined through both qualitative comparison and aggregate quantitative analysis. The goal of this evaluation was not to assess factual correctness, but to understand how asking questions in isolation versus within a guided sequence alters the structure, depth, and framing of the generated explanations.

To illustrate how contextual sequencing influences response quality, we compare two answers generated for the question *"How is speed controlled during bulk deformation?"* under isolated and chained prompting conditions. In the isolated condition, the response presents speed control as a set of parallel mechanisms organized through enumeration. For example, it states that "speed is controlled primarily through the deformation equipment and process parameters," and then proceeds to list distinct controls such as "ram or punch velocity," "roll speed," and "drawing / extrusion speed," each discussed in separate sections. While the answer correctly notes that "speed control aims to regulate the strain rate, which influences flow stress, work hardening, and dynamic recovery," this fundamental concept appears midway through the response and functions more as an additional factor than as an organizing principle. As a result, the explanation emphasizes completeness and technical coverage, but the relationships among machine settings, strain rate, and material behavior remain largely implicit.

In contrast, the chained response reflects the influence of prior questions by embedding the explanation of speed control within an established conceptual context. Having previously defined bulk deformation and its process steps, the answer introduces speed control by stating that it is managed "to ensure proper material flow, dimensional accuracy, and defect prevention," immediately situating speed within process outcomes. The response repeatedly anchors details to strain rate, noting that "deformation speed directly affects the strain rate" and that "higher speeds increase flow stress and forming loads; lower speeds improve ductility." Machine-specific controls, such as "hydraulic presses" allowing "precise, variable, and slow speeds," are presented as mechanisms for achieving these strain-rate and temperature objectives rather than as independent parameters. Additionally, the chained answer explicitly links speed to prior context by discussing hot versus cold working and by concluding with why speed control matters, emphasizing defect prevention and uniform deformation.

The contrast between these answers highlights a shift from a reference-style explanation to a more instructional, causally connected narrative. The isolated response prioritizes breadth and technical listing, while the chained response leverages conversational context to foreground underlying principles and reinforce cause–effect relationships across sentences. This qualitative difference supports the broader claim that question chaining shapes not only the amount of information provided, but also how explanations are structured, integrated, and aligned with tutoring-oriented reasoning.

While this qualitative example highlights how sequencing changes the framing of responses, aggregate quantitative metrics were used to examine whether such differences appear consistently across a larger set of questions. At this step we used 100 questions to do the comparison.

Figure 1 compares the total word count of responses generated under isolated and chained prompting conditions. At the aggregate level, isolated responses exhibit a slightly higher total word count than chained responses. This result could suggest that, when questions are asked independently, the language model tends to produce more self-contained explanations that restate background information and enumerate multiple considerations within a single response. In contrast, chained responses distribute explanatory content across multiple turns, allowing later answers to build on previously established context rather than repeating it. As a result, overall verbosity does not necessarily increase under chaining, despite qualitative evidence of greater contextual coherence. Figure 2 presents aggregate sentence count and unique/technical word count for the two conditions. While the total sentence count is comparable across conditions, chained responses show a modest increase relative to isolated responses. This pattern indicates that chaining encourages slightly more elaborated sentence-level structure, even when overall word count remains similar or lower. More notably, chained responses exhibit a higher count of unique and technical words. This increase suggests that contextual sequencing promotes greater lexical diversity and technical specificity, likely because chained interactions enable the model to introduce new terminology progressively rather than reusing the same definitional language in isolation.

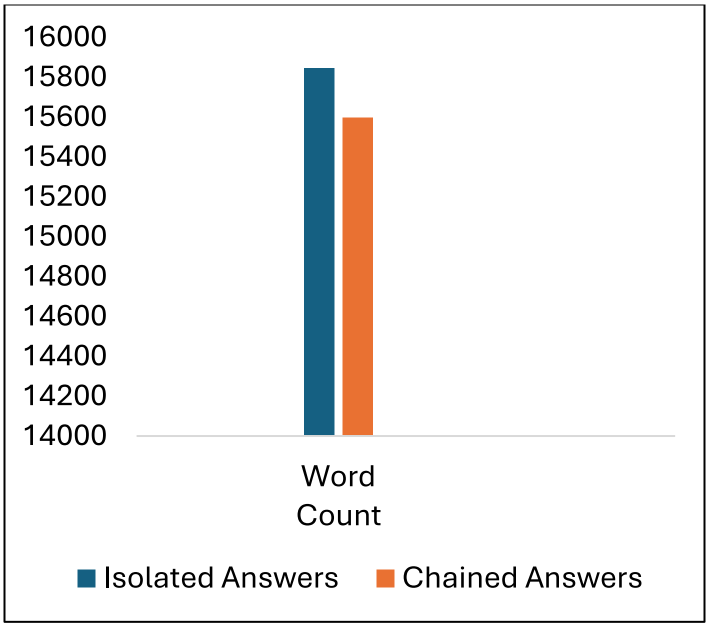

Figure 1: Aggregate word count comparison of model responses under isolated and chained prompting conditions across 100 questions

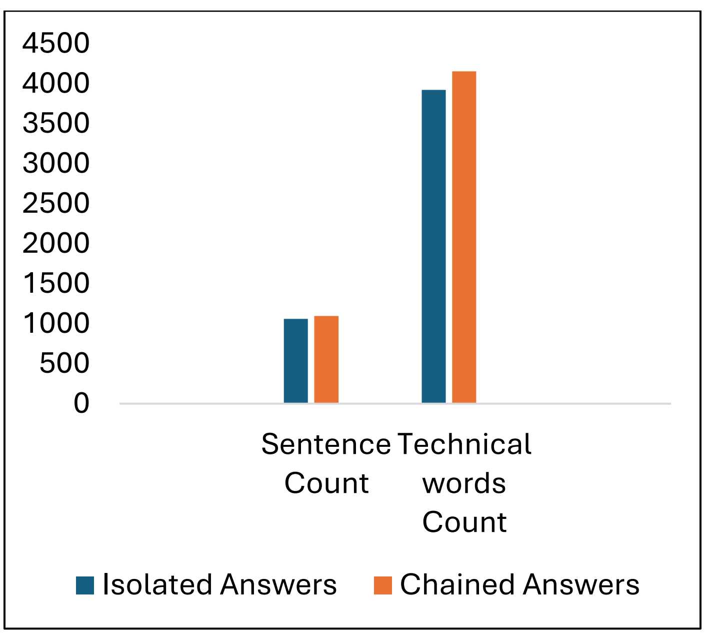

Figure 2: Aggregate sentence count and unique/technical word count of model responses under isolated and chained prompting conditions across 100 questions

Taking together, these quantitative trends align with the qualitative comparison discussed earlier. Although isolated responses may be longer in aggregate, chained responses appear to use language more efficiently, with a greater proportion of distinct and technical terms and a comparable number of sentences. These findings support the claim that question sequencing influences not only response length, but also how explanatory content is distributed and specialized across an interaction. Importantly, the overlap and modest magnitude of differences underscore the exploratory nature of this analysis; the results indicate systematic tendencies rather than categorical shifts in model behavior.

Conclusion and Future Work

This paper examined how structured question sequencing influences the behavior of a GPT-based Manufacturing Advisor when responding to manufacturing education queries. By modeling question progression as a structured process over WH-type questions aligned with Bloom's taxonomy, the study moved beyond one-off prompting and introduced guided five-question chains tailored to different expertise levels. Rather than evaluating correctness or learning gains, the analysis focused on how response characteristics change when identical questions are asked in isolation versus within a contextualized sequence. Comparisons indicate that chained questioning leads to more structured, and more lexically diverse explanations, suggesting that question sequencing plays a meaningful role in shaping how language models frame and elaborate manufacturing concepts.

The results highlight that the educational value of conversational AI systems is not determined solely by the underlying language model, but also by how interactions are structured. Isolated questions tend to elicit self-contained explanations, whereas chained interactions encourage continuity, contextual grounding, and multi-sentence reasoning that more closely resembles tutoring dialogue. While these findings are exploratory and based on simulated interactions, they provide initial evidence that Bloom-aligned question chaining can serve as a lightweight mechanism for guiding LLM behavior in educational settings without requiring explicit control over internal reasoning processes.

A few directions for future work emerge from this study. First, the current evaluation relies on surface-level textual metrics and qualitative inspection; future studies could incorporate more nuanced measures of relevance, consistency, and conceptual alignment, potentially combining deterministic text analysis with targeted human annotation. Second, while this work focused on simulated interactions, future evaluations should involve real learners to examine how question sequencing affects engagement, comprehension, and perceived instructional value. Finally, integrating the question chaining framework with expertise estimation and curriculum-level planning could enable the Manufacturing Advisor to support longer-term learning trajectories, positioning it as a scalable, adaptive tutoring tool for manufacturing education.

Acknowledgment

This work is supported by the National Science Foundation under Grant No. 2229260. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] A. Sethupathy, P. Thakkar Singh, G. Gooding, J. Reich, D. Soukup, M. Lamarre, and J. S. Warrick, *Industry 4.0 and Modernizing Manufacturing Education*, New York, NY, USA: American Society of Mechanical Engineers (ASME), 2024.

[2]W. Villegas-Ch, D. Buenano-Fernandez, A. M. Navarro, and A. Mera-Navarrete, "Adaptive intelligent tutoring systems for STEM education: analysis of the learning impact and effectiveness of personalized feedback," *Smart Learn. Environ.*, vol. 12, no. 1, p. 41, Jun. 2025, doi: 10.1186/s40561-025-00389-y.

[3] M. Giannakos *et al.*, "The promise and challenges of generative AI in education," *Behav. Inf. Technol.*, vol. 44, no. 11, pp. 2518–2544, Jul. 2025, doi: 10.1080/0144929X.2024.2394886.

[4] F. Caccavale, C. L. Gargalo, K. V. Gernaey, and U. Krühne, "Towards Education 4.0: The role of Large Language Models as virtual tutors in chemical engineering," *Educ. Chem. Eng.*, vol. 49, pp. 1–11, Oct. 2024, doi: 10.1016/j.ece.2024.07.002.

[5] A. T. Corbett and J. R. Anderson, "Knowledge tracing: Modeling the acquisition of procedural knowledge," *User Modeling and User-Adapted Interaction*, vol. 4, no. 4, pp. 253–278, 1995, doi: 10.1007/BF01099821.

[6] M. M. Perera, A. Mahmood, K. E. Wijethilake, F. Islam, M. Tahermazandarani, and Q. Z. Sheng, "A Survey of the State-of-the-Art in Conversational Question Answering Systems," Sep. 06, 2025, *arXiv*: arXiv:2509.05716. doi: 10.48550/arXiv.2509.05716.

[7] T. Scialom and J. Staiano, "Ask to Learn: A Study on Curiosity-driven Question Generation," in *Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, Spain (Online): International Committee on Computational Linguistics, 2020, pp. 2224–2235. doi: 10.18653/v1/2020.coling-main.202.

[8] B. S. Bloom, M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, "Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain," *David McKay Company,* New York, NY, USA, 1956.