Designing and Testing AI-based Text Personalization Tools

Mr. Michael Thomas Vaccaro Jr, University of Connecticut

Michael Vaccaro is a fourth-year Ph.D. student in the School of Civil and Environmental Engineering at the University of Connecticut. He received his Bachelor of Science in Civil Engineering from the University of Connecticut in 2021. In addition to his work in structural engineering, Michael's interests in teaching and learning have inspired him to pursue interdisciplinary research spanning the fields of engineering, artificial intelligence, and neuroscience. His recent work in these areas has been supported by his major advisor's NSF MCA project and a transdisciplinary NSF Research Traineeship (TRANSCEND). Michael's engineering education research explores artificial intelligence's potential in K-12 science education, particularly in developing personalized learning environments.

Mikayla Friday, University of Connecticut

Mikayla is a second-year PhD student studying Engineering Education at the University of Connecticut.

Dr. Zeynep Gonca Akdemir-Beveridge, University of Connecticut

Zeynep G. Akdemir-Beveridge is a Postdoctoral Research Associate at the University of Connecticut. Her current research focuses on exploring the creative productivity of engineering students and promoting strength-based approaches in engineering education to better support neurodiverse learning profiles in STEM/engineering learning environments. Her doctoral research centered on K–12 curriculum development specifically tailored to quantum information science and engineering, as well as student motivation and engagement in science education. A third area of her work involves the application of artificial intelligence (AI) in science and engineering education, particularly in the context of personalized learning. She is the recipient of the Best of Computers in Education award for her co-authored work presented at the 2022 ASEE Annual Conference & Exposition.

Prof. Arash Esmaili Zaghi P.E., University of Connecticut

Arash E. Zaghi is a Professor in the Department of Civil and Environmental Engineering at the University of Connecticut. He received his PhD in 2009 from the University of Nevada, Reno, and continued there as a Research Scientist. His latest

Designing and Testing AI-Based Text Personalization Tools

Abstract

Large Language Models (LLMs) offer unprecedented opportunities for individualized instruction by tailoring reading materials to students' unique needs and preferences at scale. In this study, we introduce an LLM-powered personalization framework to adapt the complexity, format, and presentation of text-based educational content to students' individual preferences. Our approach breaks personalization into a series of smaller tasks—describing learner preferences, rewriting texts, and evaluating accuracy—with the goal of enhancing accessibility and engagement. To rapidly refine prompts and system design, we simulate users by assigning multiple LLM agents the roles of both "student" and "teacher." This closed-loop simulation allowed us to iterate swiftly on prompt engineering and platform features, enabling extensive testing and fine-tuning before deploying the framework with real students. Future work will therefore include classroom trials and human-subject evaluations to validate the impact of the achieved LLM-powered text personalization on student motivation and learning outcomes. By showcasing the feasibility of AI-driven customization, this paper points to new frontiers for delivering student-centered learning experiences in engineering education and beyond.

Introduction

As education becomes increasingly complex and specialized, artificial intelligence (AI) offers tools to make teaching and learning more effective, engaging, and equitable [1]. Therefore, we see artificial intelligence (AI) as a transformative force in education which has a large potential to offer solutions to challenges posed by traditional, standardized instructional methods. Specifically, modern AI models offer the ability to generate new content in real-time, making truly adaptive learning [2] a possibility. These challenges are unique in the context of engineering education due to the complexity and rapidly evolving nature of the field which requires innovative and inclusive teaching approaches.

A traditional engineering education setting can unfortunately assume that all students learn at the same pace, learn effectively through the same methods, and share identical goals since the standardized curricula focuses heavily on lectures, theoretical problem-solving, and fixed lab experiments [3]. However, AI has the potential to support engineering education by enabling interdisciplinary learning modules, enhancing project-based learning, facilitating remote or virtual collaboration, and providing tools to address diverse learning needs more effectively. By moving beyond the limitations of recommender-style systems [2] and one-size-fits-all approaches, AI can offer more adaptable solutions to complement traditional teaching methods. To fully leverage this transformative potential of AI, we introduce a framework for personalized learning (PL) that can provide an approach for implementing AI-driven methods to meet the individual needs of diverse learners. The concept of PL, however, is not entirely new. It has a rich history rooted in efforts to tailor education to individual needs, long before the advent of AI.

Personalized Learning

From the progressive education movement of the early 20th century to the rise of technology-enabled learning in recent decades, the U.S. education system has served as a fertile ground for pioneering methods that prioritize student-centric approaches. The progressive education movement, led by figures like John Dewey, emphasized experiential, student-centered learning, advocating for curricula that adapted to individual interests and fostered critical thinking over rote memorization [4]. Dewey's philosophy laid the intellectual foundation for numerous educational reforms and movements such as the Montessori method, constructivism, project-based learning, and Paulo Freire's critical pedagogy. These approaches embodied the core principles of Dewey's educational model, which strongly rejected one-size-fits-all methods. Instead, Dewey envisioned education as a dynamic, living framework that could evolve based on student input and emerging interests, which is a central principle of PL [4].

The progressive movement's influence extended into the mid-20th century, where the maturing field of educational psychology began formalizing theories about individual differences in how people learn [5]. Researchers increasingly sought to understand and explain the diversity in how individuals engage with and absorb knowledge. Insights from psychology inspired educational theorists to develop formal models categorizing and analyzing learning preferences, setting the stage for PL. By the 1970s and 1980s, formal learning style models such as Dunn and Dunn Learning Styles, Kolb's Experiential Learning Theory, the Felder-Silverman Learning Style Model, Fleming's VARK Model, and Gardner's Multiple Intelligences Theory, gained popularity.

Personalized learning frameworks draw heavily from these historical examples by integrating foundational principles of learning style models with modern discipline-based educational practices [6], [7]. For instance, the Felder-Silverman Learning Style Model, specifically tailored for Science, Technology, Engineering, and Mathematics (STEM) fields, introduced a systematic way to address diverse cognitive and sensory preferences, influencing how engineering curricula were designed. Similarly, Dewey's emphasis on experiential learning in science and engineering resonates in contemporary project-based learning and adaptive platforms. However, as learning style models gained prominence in shaping individualized instruction, they also faced criticisms from researchers questioning their empirical validity [8], [9]. For example, one study from the field of neuroscience challenged the foundational assumptions of learning style models by emphasizing that the brain processes information based on the nature of the task or content, not a learner's preference [9]. While such studies have challenged the so-called rigid categorization inherent in learning style models, it is important to recognize that these models were never intended to dictate that learning must align exclusively with a single modality. Therefore, we suggest that learning style models should be viewed as the conceptual precursors that paved the way for more dynamic, evidence-based approaches to individualizing education rather than dismissing them based on critiques of their empirical limitations.

Learning style models serve as an initial approach to recognizing that individuals vary in the ways they process and interact with information. Personalized education and learning style models both aim to address the individual needs of learners, moving away from one-size-fits-all approaches. Both perspectives share a core objective: tailoring education to accommodate learner diversity. While learning style models categorize learners based on certain preferences, personalized education builds on this idea by using data analytics and technology to create nuanced, individualized, and adaptive learning paths [10] that are based on learner profiles, attitudes, and prior knowledge and are flexible and self-paced [11]. As such, our proposed framework uses the Felder-Silverman model as a starting point for capturing some of the diversity in learner preferences to develop a flexible learning environment.

Designing LLM-based Personalized Learning Platforms

Modern definitions of PL are closely aligned with Dewey's [4] vision of education as dynamic and learner centered. However, it is important to recognize that definitions of PL vary widely in the features they consider effective in or central to personalization. A recent systematic review by Bernacki et al. [12] has documented the variability in defining PL through 2021. In general, definitions were found to vary on three main dimensions, including which learner characteristics the platform accounts for in the personalization process (e.g., prior knowledge, learning preferences), the ways in which materials are personalized (e.g., pace, sequence, providing students with choice, assessment types), and PL's primary outcomes (e.g., student agency, academic performance). We additionally note that a large number of the definitions considered were published within the past two decades as this time period has seen a surge in the amount of published PL research [13] due, at least in part, to PL's designation as one of the fourteen grand challenges for the 21st century in 2008 [14].

Because of the variability which exists, we find it critical to provide a definition of PL that we considered in the design of our large language model (LLM)-powered PL framework. We specifically draw from the U.S. Department of Education's [15, p. 9] definition, which holds that PL is "instruction in which the pace of learning and the instructional approach are optimized for the needs of each learner." In achieving this personalization, the Department of Education states that "[l]earning objectives, instructional approaches, and instructional content (and its sequencing) may all vary [and that] … learning activities [should be] meaningful and relevant to learners, driven by their interests, and [be] self-initiated" [15, p. 9]. With this definition in mind, we recognize that a PL platform must consist of several features, including methods for students to convey their educational needs, to identify and describe student learning preferences, and to account for students' interests. Moreover, each of these features should be accounted for simultaneously in the development of personalized educational materials.

While these features may seem straight-forward, they represent a diverse range of cognitive tasks that individually require complex reasoning. These tasks may be summarized by the following inquiries. First and foremost, how and from what lens should an LLM describe student learning preferences to achieve the best performance? Furthermore, how should this description be used to generate educational content that is uniquely personalized for the individual student? Finally, once these features are defined, how can an AI platform be tested to ensure the model is performing as expected?

As past research has shown [16], LLMs perform best when large tasks are decomposed into smaller, more manageable chunks. As such, we propose the design of an LLM-powered PL environment by breaking down the task of personalization into two components using two LLM agents: 1- systematically describing student learning preferences, and 2- personalizing educational materials to this description. Recognizing that testing model performance requires trial runs of the personalization platform and that prompt engineering, or the process of crafting the LLM prompts that yield the best output, is a highly iterative process [17], we further decompose the test of model performance into three elements where we assign one agent to generate a student description, another to play the role of that student, and a third to evaluate the accuracy of the student description generated by item '1-' above. In total, this process yields five GPT-4 agents in which two power the PL platform and three assess the platform's performance. This framework, through which we propose the design and testing of an LLM-powered text personalization platform, is summarized in Figure 1.

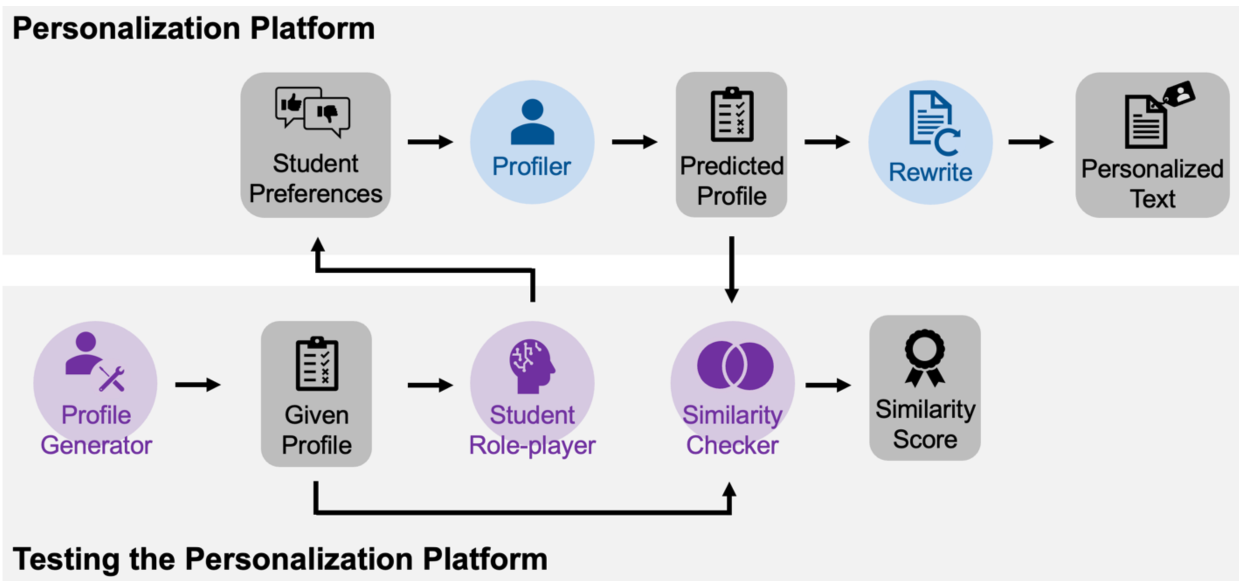

Figure 1. Overview of the proposed framework. Circles represent specific large language model agents while rectangles represent input and output.

A critical part of designing the PL platform was the development of the prompts to be used throughout the experiment. The importance of prompt engineering in platforms intended to provide personalization cannot be understated as effective prompt engineering encourages LLMs to produce more relevant, intentional, and creative responses [18]. This is especially important when the platform is to be deployed to a wide range of users. While several prompt engineering strategies have been proposed [17], there is no one approach that will always produce the best results for the task at hand. For transparency in our final design, we provide our final system and user messages of each of the five LLM agents (Figure 1) in Table 1 of the Appendix. We note that these prompts were optimized for *gpt-4-1106-preview*, so some modifications may be necessary to successfully implement this framework using newer model versions.

Our approach to prompt engineering was highly iterative and, given the interdependency of the LLM agents shown in Figure 1, relied heavily on simulations of the experiment to refine our prompts for each model. First, we had a "Profile Generator", which was an LLM prompted to produce hypothetical student profiles which were used as part of the prompt for the "Student Role-Player" agent. Along with the generated profile, the student agent was given instructions for how to interact with the platform. It was instructed with how many texts it would see, how to evaluate the texts, and how to report its preferences. Prompt engineering allowed for a consistent output among different generations of the Student Role-Player, which were then given to the "Profiler". The Profiler was prompted to generate a brief profile of the student based on its reported textual preferences. These profiles were then given to a "Rewrite" agent, which was prompted to rewrite an academic text, personalized to the simulated student's taste. Finally, there is a "Similarity Checker" which generates a similarity score between the given profile and the predicted profile. The interactions between these agents are shown above in Figure 1.

Because we used simulated participants rather than human students to test the personalization tool, we were able to run the entire experiment many times, holding some prompts constant while altering others to maximize performance. We also used the practice of "meta-prompting," which involves asking LLMs to develop prompts for the tasks we wanted them to perform [19]. We found this to be helpful in some cases and used it as a guideline for our prompts. Additionally, we explicitly modeled the outputs we wanted the LLM to generate and included those examples within the prompts. Modeling the output was especially important as the generations from one model were often used as inputs for the next (Figure 1) and any deviations in response formatting or content could cause the personalization to fail. We also ran one hundred trials each time the prompts were updated, so consistency in the outputs from each step was crucial to analyzing the results of each iteration.

One of the most challenging prompts to develop was that of the initial Profile Generator. While we initially tried to give GPT some freedom, we found that the range of responses was too broad and often lacked relevance. To mitigate this problem, we looked at literature for an existing framework for different preferences students may exhibit. We found that the Felder-Silverman model had precedence within digital PL platforms (e.g., [20]), so it was used as a starting point for our prompting. The use of an established framework allowed for more focused, consistent responses, which streamlined our process. This also allowed us to present educational paragraphs in accordance with different potential learning preferences, generate student profiles, and generate new personalized content systematically.

Our ability to simulate the experiment further allowed us to test and refine different experimental designs. For each experimental design, the simulation was run a total of one hundred times, as discussed by Friday et al. [21]. Our final platform involved the presentation of four pairs of paragraphs, with each pair representing a dimension of the Felder-Silverman model. The user would select their preferred paragraph from each of these pairs, and those choices would be given to the Profiler. The Profiler used this insight to develop a brief description of the user's learning preferences and provided this information to the Rewrite agent. This agent would then rewrite a standardized text in accordance with the user's preferences. To evaluate the efficacy of the model, we recorded how often the user chose the rewritten paragraph over the generic paragraph. Additionally, we looked at and analyzed the similarity scores between the original and generated profiles. By analyzing these performance metrics, we could see how changing the prompting and experimental design altered the outcome of the survey. Without the ability to simulate the experiment using LLMs, we would not have been able to optimize the final design.

The result of the above design is a two-agent personalization platform, illustrated previously in the top half of Figure 1. In addition, the systematic way in which LLM outputs are produced means that the final platform behaves in a predictable manner, progressing consistently from identifying student learning preferences through to text personalization. With this systematic framework in mind, we recommend that future research using LLMs for content personalization develop custom graphical user interfaces (GUI)—such as that developed in the follow-up study by Vaccaro et al. [22]—rather than rely on public-facing interfaces like ChatGPT as it minimizes the potential for user error. Such a controlled GUI is also beneficial from an experimental context where consistency in implementation is of critical importance. Finally, it should be noted that such an environment allows for strict control over the types of information students can share with an LLM, thus maintaining student privacy.

Integration of Personalized Learning in Engineering Education through LLMs

The integration of PL into engineering education through advanced AI and LLMs represents a transformative yet nascent field. The use of cutting-edge LLMs, such as GPT-4, has shown significant promise, although rigorous research on their application is still developing [23]. Engineering education scholars express both optimism and caution regarding these tools, highlighting the need to carefully harness their potential [24]. Emphasizing the strengths of AI and LLM-based PL in engineering education is vital to maximizing its benefits for students, instructors, and institutions alike.

AI-LLM tools, when applied to engineering education, can transcend traditional teaching methods by addressing the field's inherent complexity, iterative processes, and emphasis on practical application. Generative AI technologies like ChatGPT facilitate real-time, personalized feedback and interactive learning experiences. For example, they enable students to engage with customized support mechanisms that clarify difficult engineering concepts, simulate real-world problem-solving scenarios, and offer tailored learning pathways [25]. Furthermore, AI systems can create adaptive practice problems that adjust to individual learning levels, helping students prepare for exams and solidify their understanding of challenging material [24], [26].

Given the sequential and prerequisite-driven nature of engineering education, AI systems are particularly well-suited for dynamically adapting course structures and offering remedial resources. Specifically, AI systems can play a crucial role in ensuring students meet foundational competencies, which are critical for succeeding in more advanced coursework. However, the effective use of AI in this context requires the implementation of AI-supported, human-in-the-loop content designs. These designs leverage AI for identifying gaps in content through data analysis while integrating human expertise to refine, validate, and contextualize the AI-generated outputs. Studies affirm that blending AI-driven insights with human intervention enhances educational outcomes by aligning instructional resources with pedagogical goals and addressing the nuanced needs of diverse learners [27], [28].

Conclusion

This paper introduces the development and testing of an AI-powered text personalization tool designed to enhance PL by tailoring text-based educational materials to individual student needs and preferences. This study served as a precursor to a human-based study evaluating the ability of LLMs to identify and tailor science texts to students' learning preferences (see [22] and [29]). Leveraging LLMs, the tool employs an iterative design process using simulated participants to refine prompt engineering strategies and optimize performance. The final tool adapts text complexity, format, and presentation, allowing for a more accessible and relevant learning experience for diverse students. As presented here, the tool focuses on student learning preferences and does not include methods to assess student motivation, which may fluctuate over time, or specific learning outcomes. In addition, the tool in its current state is limited to a one-time development of a student profile and does not account for how students' preferences may change over time. Addressing these limitations should be the focus of future development of LLM-enabled PL environments.

The tool holds promise for engineering education, where the complexity of concepts and the sequential nature of coursework present unique challenges for students with varying cognitive and learning needs. By personalizing text-based content, we hope that the tool will support students in building foundational knowledge and engaging with advanced topics in a manner suited to their individual learning profiles. This focus on text-level adaptation complements traditional instructional methods and contributes to a more inclusive and supportive learning environment.

We foresee that this study will provide valuable insights for both research and practice in engineering education. For researchers, it offers a framework to explore the potential of AI-driven text personalization in improving student outcomes and addressing diverse learning needs. For practitioners, the tool's ability to adjust academic texts provides a practical means of enhancing engagement and comprehension in engineering classrooms. By addressing privacy concerns through anonymized profiles and ensuring ethical oversight through human-in-the-loop processes, the study highlights a thoughtful approach to integrating AI into educational contexts. While further work is needed to evaluate broader applications, this study represents a step forward in understanding how AI-powered tools can enhance learning experiences, particularly in fields requiring tailored support, such as engineering education.

Acknowledgements

This material is based upon work supported by the National Science Foundation under MCA Grant No. 2120888. The first author (MV) was supported by an NSF Research Traineeship (TRANSCEND) under Grant No. 2152202 at the time this research was conducted. 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.

The authors greatly appreciate the support of Trent Alsup, Jada Vercosa, Brian Hance, and Abhiram Gunti in the initial development of the GPT platform.

Finally, the authors disclose that OpenAI's ChatGPT models were used during the preparation of this paper as a writing assistant to check grammar and to enhance the clarity of the written text. These models were used with extreme oversight and care. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

References

[1] E. Aque, "AI and Cognitive Gamification: Enhancing Student Engagement in STEM Education," *Journal of Artificial Intelligence and Cognitive Science,* vol. 1, no. 1, 2024.

[2] M. Bulger, "Personalized Learning: The Conversations We're Not Having," *Data & Society,* 2016.

[3] D. D. Denton, "Engineering Education for the 21st Century: Challenges and Opportunities," *Journal of Engineering Education,* vol. 87, no. 1, pp. 19-22, 2013.

[4] J. Dewey, Democracy and Education: An Introduction to the Philosophy of Education, Macmillan, 1916.

[5] National Research Council, How People Learn: Brain, Mind, Experience, and School: Expanded Edition, Washington DC: The National Academies Press, 2000.

[6] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Engineering Education,* vol. 78, no. 7, pp. 674-681, 2002.

[7] R. M. Felder and R. Brent, "The intellectual development of science and engineering students. Part 2: Teaching to promote growth," *Journal of Engineering Education,* vol. 93, no. 4, pp. 279-291, 2004, doi: 10.1002/j.2168-9830.2004.tb00817.x.

[8] H. Pashler, M. McDaniel, D. Rohrer and R. Bjork, "Learning styles: Concepts and evidence," *Psychological Science in the Public Interest,* vol. 9, no. 3, pp. 105-119, 2008, doi: 10.1111/j.1539-6053.2009.01038.x.

[9] C. Riener and D. T. Willingham, "The Myth of Learning Styles," *Change: The Magazine of Higher Learning,* vol. 37, no. 5, pp. 32-35, 2010, doi: 10.1080/00091383.2010.503139.

[10] T. M. Nguyen, "Incorporating AI tools into comprehensive language learning platforms: Strategies and implications," Theseus.fi., 2024.

[11] A. Shemshack, Kinshuk and J. M. Spector, "A comprehensive analysis of personalized learning components," *Journal of Computers in Education,* vol. 8, no. 4, pp. 485-503, 2021, doi: 10.1007/s40692-021-00188-7.

[12] M. L. Bernacki, M. J. Greene and N. G. Lobczowski, "A Systematic Review of Research on Personalized Learning: Personalized by Whom, to What, How, and for What Purpose(s)?," *Educational Psychology Review,* vol. 33, pp. 1675-1715, 2021, doi: 10.1007/s10648-021-09615-8.

[13] A. Shemshack and J. M. Spector, "A systematic literature review of personalized learning terms," *Smart Learning Environments,* vol. 7, no. 33, 2020, doi: 10.1186/s40561-020-00140-9.

[14] National Academy of Engineering, "NAE Grand Challenges for Engineering," National Academies Press, Washington, D.C., 2017.

[15] U.S. Department of Education, "Reimagining the Role of Technology in Education: 2017 National Education Technology Plan Update," Washington, DC, 2017.

[16] Z. Wu, H. Bai, A. Zhang, J. Gu, V. Vinod Vydiswaran, N. Jaitly and Y. Zhang, "Divide-or-Conquer? Which Part Should You Distill Your LLM?," in *Findings of the Association for Computational Linguistics: EMNLP 2024*, Miami, FL, USA, 2024, doi: 10.48550/arXiv.2402.15000.

[17] G. Mizrahi, Unlocking the Secrets of Prompt Engineering: Master the art of creative language generation to accelerate your journey from novice to pro, 1st ed., Packt Publishing, 2024.

[18] P. Korzyński, G. Mazurek, P. Krzypkowska and A. Kurasinski, Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT, Entrepreneurial Business and Economics Review, vol. 11, no. 3, pp. 25-37, 2023, doi: 10.15678/EBER.2023.110302.

[19] L. Reynolds and K. McDonell, "Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm," in *CHI 2021*, Yokohama, Japan, 2021, doi: 10.1145/3411763.3451760.

[20] E. Ashraf, S. u. A. Laghari, S. Manickam, K. Mahmood, S. Abrejo and S. Karuppayah, "Intelligent course recommendation approach based on modified felder-silverman learning style model," *E-Learning and Digital Media,* 5 September 2024, doi: 10.1177/20427530241279965.

[21] M. Friday, M. Vaccaro and A. Zaghi, "Leveraging large language models for early study optimization in educational research," in *2025 ASEE Annual Conference & Exposition*, Montreal, Quebec, Canada, 2025.

[22] M. Vaccaro, M. Friday and A. Zaghi, "Evaluating the capability of large language models to personalize science texts for diverse middle-school-age learners," *[arXiv Pre-print],* 2024, doi: 10.48550/arXiv.2408.05204.

[23] S. Filippi and B. Motyl, "Large Language Models (LLMs) in Engineering Education: A Systematic Review and Suggestions for Practical Adoption," *Information,* vol. 15, no. 6, p. 345, 2024, doi: 10.3390/info15060345.

[24] M. Menekse, "Envisioning the future of learning and teaching engineering in the artificial intelligence era: Opportunities and challenges," *The Research Journal for Engineering Education,* vol. 112, no. 3, pp. 578-582, 2023, doi: 10.1002/jee.20539.

[25] J. Qadir, "Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education," in *2023 IEEE Global Engineering Education Conference (EDUCON)*, Salmiya, Kuwait, 2023, doi: 10.1109/EDUCON54358.2023.10125121.

[26] E. R. Mollick and L. Mollick, *Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts,* Philadelphia, PA: The Wharton School Research Paper, 2023, doi: 10.2139/ssrn.4391243.

[27] S. Drissi and A. Amirat, "An Adaptive E-Learning System based on Student's Learning Styles: An Empirical Study," *International Journal of Distance Education Technologies,* vol. 14, no. 3, p. 18, 2016, doi: 10.4018/IJDET.2016070103.

[28] H. Harati, L. Sujo-Montes, C.-H. Tu, S. J. W. Armfield and C.-J. Yen, "Assessment and Learning in Knowledge Spaces (ALEKS) Adaptive System Impact on Students' Perception and Self-Regulated Learning Skills," *Education Sciences,* vol. 11, no. 10, p. 603, 2021, doi: 10.3390/educsci11100603.

[29] M. Vaccaro, M. Friday and A. Zaghi, "Transforming K-12 STEM education with personalized learning through large language models (Fundamental)," in *2025 ASEE Annual Conference & Exposition*, Montreal, Quebec, Canada, 2025.

Appendix

Below we provide the system and user messages for each of the five GPT-4 agents used in the design and testing of the personalization platform. The Profiler and Rewrite agents designed using the methods described herein were implemented in a human-subjects study using *gpt-4-1106-preview* to test the efficacy of the personalization (see the study by Vaccaro et al. [22]).

In the following table, text that is ***bolded and italicized*** represents input that may vary. The newline character *\n* has been typed out in some locations for clarity. These prompts were optimized for the *gpt-4-1106-preview* version of GPT-4.

Table 1. System and User Messages for each LLM agent used in the design and testing of the LLM-powered personalization framework.

Agent	Message	
Profiler	*System*	You are an experienced science teacher who frequently works with middle school students and is well-versed in the Felder-Silverman learning preference model. Given a student's responses to a series of paragraph pairs, please analyze and provide a description of his/her learning style according to the dimensions of the Felder-Silverman model. Do not mention the student's selections at all. Do not reference the content the student was presented with or their direct choices. Instead, offer a generalized learning profile that captures the essence of their preferences in learning. Direct the profile towards the student (i.e., use terminology like 'you are'). Do not justify your profile by referring to the selections the student made (i.e., do not say things like 'based on your selections'). Ensure the language you use is accessible to a middle school student. Do not use big words. Limit your profile to 3 to 4 short sentences. Do not use highly imaginative or specialized language that cater to one learning preference over the other. You must use simple language and not use complex descriptors. The student is not likely to fall at the extremes of the Felder-Silverman learning style model. Here is an example of the type of profile you generate: [You are a student who excels when information is presented in a step-by-step process. Your approach to learning is highly practical, and you prefer dealing with concrete facts over abstract concepts. Reading and writing are your preferred methods for learning new information, rather than through pictures or diagrams. Additionally, you like to think things through on your own, understanding concepts deeply before discussing them with others or applying them.]
	User	The student was given the following four pairs of paragraphs: ["Topic 1: ***Topic 1*** \n\n Paragraph 1: \n ***Training Paragraph 1-1 Text*** \n\n Paragraph 2: \n ***Training Paragraph 1-2 Text***", "Topic 2: ***Topic 2*** \n\n Paragraph 1: \n ***Training Paragraph 2-1 Text*** \n\n Paragraph 2: \n ***Training Paragraph 2-2 Text***", "Topic 3: ***Topic 3*** \n\n Paragraph 1: \n ***Training Paragraph 3-1 Text*** \n\n Paragraph 2: \n ***Training Paragraph 3-2 Text***", " Topic 4: ***Topic 4*** \n\n Paragraph 1: \n ***Training Paragraph 4-1 Text*** \n\n Paragraph 2: \n ***Training Paragraph 4-2 Text***"] The student chose these paragraphs in accordance with their learning style: 1. Paragraph ***x*** 2. Paragraph ***x*** 3. Paragraph ***x*** 4. Paragraph ***x***

Rewrite	*System*	You are an experienced middle school science teacher who is capable of reworking scientific texts for diverse middle school students. Your writing style is simple. You will be shown a profile that has been written to describe a student's learning preferences on the Felder-Silverman learning style dimensions. The profile is addressing the student. You will also be shown a paragraph describing a middle school science concept. Your task is to rework the given paragraph so that it caters to the student's preferences for learning and textual presentation. At the same time, you must aim for a balance between engaging and straightforward explanations and ensure the scientific content remains clear and accessible. Do not use of highly imaginative, specialized language or key words (such as 'imagine') that cater to one learning preference over the others. The goal is to make the concept understandable and interesting to a student who generally fits the given description. Your reworked paragraph must be approximately the same length as the provided paragraph. Your rework must be one short paragraph that is less than one hundred words long. In addition, the rework you provide must use language that is appropriate for a middle school student (i.e., do not use big words) and must remain academic in tone. Do not mention the student's profile, simply provide your rework.
	User	The student profile is as follows: [***student_profile***] Here is the paragraph you need to rework for the student: [***text_to_rewrite***]
Profile Generator	*System*	You are an educational psychologist who frequently works with middle school students and is well-versed in the Felder-Silverman learning preference model. Your purpose is to present single-paragraph descriptions of a student's learning preferences. You are simple, direct, and clear in your description and you avoid the use of technical jargon.
	User	Given the following breakdown of a student's learning preferences along the Felder-Silverman model, write one paragraph describing a student whose learning preferences match the chosen values along each dimension. Address the paragraph to the student using terminology like 'you are' and do not explicitly reference the Felder-Silverman model. Focus on aspects of the student's learning preferences that could be discerned solely from his or her interactions with plain text. Your description should be around one hundred words in length. Do not write the values chosen along each dimension, only write the paragraph. The breakdown of the student's preferences is shown below. 1. **Sequential/Global Dimension (1 = Sequential, 10 = Global)**: ***1 or 10*** 2. **Sensory/Intuitive Dimension (1 = Sensory, 10 = Intuitive)**: ***1 or 10*** 3. **Visual/Verbal Dimension (1 = Visual, 10 = Verbal)**: ***1 or 10*** 4. **Active/Reflective Dimension (1 = Active, 10 = Reflective)**: ***1 or 10***
Student Role-player	*System*	You are a middle school student. Your task here is to read two different descriptions of the same topic and tell which one matches your learning preferences and cognitive skills better. ***{insert student profile from profile generator}*** Based on your learning preferences, select your favorite paragraph. Format your answer as follows: Paragraph X (Where X is a variable that you will replace with your preferred paragraph number. For example, if you like the second paragraph, your response must look like this: Paragraph 2)

	User	Topic: ***topic*** Paragraph 1: ***Option 1 text*** Paragraph 2: ***Option 2 text***
Similarity Checker	*System*	You are well-versed in analyzing texts and quantifying the degree of similarity between them. You have a good understanding of learning preferences of different middle-school aged learners. Your role is to quantify the degree of agreement and disagreement between two paragraph-length descriptions of hypothetical learning preferences. First consider the level of agreement between the two texts. For example, there may be no agreement, low agreement, some agreement, high agreement, or full agreement, or the agreement may lie somewhere between two of those descriptors. After determining the agreement level, you quantify the agreement. You provide a brief written analysis of the similarities before you present the quantified score in the form of a percentage. You provide your scores on a new line in the following format: 'Percent Agreement: XX%.', where XX% is replaced with your agreement score. Pay attention to whether the two paragraphs are talking about the same cognitive preferences or learning attributes, or if the texts are different. For example, if one description says that the student likes imagery and visual information, and the other one says that the student likes factual presentation of information in the form of numbers and values, there is low agreement (i.e., Percent Agreement: 10%). If the texts are saying the same thing in different words, there is high agreement (i.e., Percent Agreement: 90%). In cases where there is some agreement, the score should be a more central value. For example, if there are more similarities than differences, the score may be 60%. (i.e., Percent Agreement: 60%) Pay attention to the entire texts, because there may be multiple dimensions and attributes that you should quantify and factor in to your final evaluation.
	User	Description 1: ***{profile from Profile Generator}*** Description 2: ***{profile from Profiler}***