Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #48355

Transforming K-12 STEM Education with Personalized Learning through Large Language Models (Fundamental)

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 the development personalized learning environments.

Mikayla Friday, University of Connecticut

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

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

Transforming K-12 STEM Education with Personalized Learning through Large Language Models (Fundamental)

Abstract

Personalized learning has been a long-standing goal of education in the United States. To date, numerous computer-based tools have been developed to personalize the learning process; however, these tools are often fundamentally limited in their design as rule-based models. This limits the capabilities of computer-based learning platforms to the resources that can be created by human developers. Large language models (LLMs) hold a unique potential to advance personalized learning due to their ability to quickly generate a near-infinite number of texts. Specifically, LLMs like OpenAI's Generative Pretrained Transformer (GPT) models have the potential to adapt their outputs in response to a user's requests and learning preferences. However, the performance of these models is highly dependent on the quality of the system and user prompts developed by the user which, together, define how the model will handle the presented task. As minor changes in their structure can lead to wide differences in performance, these prompts must be carefully designed. To that end, this paper presents a framework through which educational STEM-based textual learning materials can be personalized. Specifically in the framework defined here, GPT-4 first analyzes student choices to identify their learning preferences according to the Felder-Silverman learning style model, and then uses these preferences to personalize STEM-related educational texts. The results of this work may substantially enhance the application of large language models in K-12 and post-secondary science education. In many instances the use of these models in education is discouraged, often due to concerns surrounding academic integrity. In response to these concerns, we discuss student perceptions and the potential benefits and drawbacks of an LLM-powered personalization framework. The hope is that this work will demonstrate the ways in which LLMs can support education by improving the accessibility of educational texts to diverse learners.

Introduction

Recent advances in artificial intelligence have revitalized interest in personalized learning (PL). In particular, large language models (LLMs) have emerged as a promising tool to tailor educational content to the needs of diverse learners in both K-12 and higher education [1]. Although PL has been widely researched for its potential to optimize student engagement and improve learning outcomes [2], its implementation often remains limited by constraints on real-time customization in either computer-based or in-person interventions. With modern LLMs, educators and researchers now have the tools to move beyond static resources or rule-based adaptive tutors towards more dynamic systems that can customize learning materials on demand [3]. This shift not only challenges existing norms in education but also has significant potential to enhance students' learning experiences.

Despite broad agreement on its importance [1], [4], there is currently no universally accepted definition of PL. The National Academy of Engineering, which named PL one of its fourteen grand challenges for the twenty-first century in 2008, captures the variability in these definitions. They explain that PL approaches can "range from modules that students can master at their own pace to computer programs designed to match the way it presents content with a learner's personality" [5, p. 45]. Further capturing this diversity, a systematic review by Bernacki et al. [6] outlined three dimensions that differentiate definitions of PL: the learner characteristics considered (e.g., preferences, existing knowledge), the ways in which the personalization is achieved (e.g., providing choice, modifying the instructional approach), and the outcomes of interest (e.g., student motivation, content mastery). With this variability in mind, we find it necessary to specify what we mean by PL in this paper. We adopt the U.S. Department of Education's (2017) definition, which holds that PL should tailor the pace, approach, objectives, content, and sequencing of instruction to meet individual learner needs [4, p. 9]. Unlike several other definitions [6], it does not focus on student learning outcomes (e.g., performance); rather, it highlights the importance of individual learners' characteristics and the considerations necessary to design an effective PL environment.

In moving towards active content adaptation in PL environments, researchers have designed intelligent tutoring systems (ITSs) that can generate new content in response to student performance. For instance, Price et al. [7] developed iSnap, an ITS that adds dynamic hint generation to the existing *Snap!* coding language. While hints were successful in improving student performance and were generated in real-time, the platform was limited to a specific type of coding exercise and was only capable of moving students towards the final answer. In other words, the hint feature was not capable of evaluating code snippets that students would later integrate into a larger final product. LLMs offer a significant departure from this existing application-specific format owing to their structure as attention-based transformers and the diversity of their training datasets that gives them expertise across several content areas [8].

In a recent example, Sajja et al. [9] describe the development of a GPT-3-powered virtual teaching assistant for students in higher education and its integration with a course website hosted on Canvas, a popular learning management system. Students could ask questions about content or request that study materials (e.g., sample quiz, flashcards) be generated from the resources available on the course site [9]. This integration is particularly promising as it demonstrates the ability of LLMs to synthesize student requests with relevant materials by supplementing—not replacing—existing educational practices. While this system can generate content in real time, it still does not fully meet the requirements of a PL platform due to its inability to address students' unique preferences and interests.

In this paper, we propose a two-agent PL framework for middle school STEM education using OpenAI's GPT-4 platform [8]. The proposed framework enables the LLM to develop a rich understanding of a students' learning preferences from their interactions with text-based materials. We store the data on these preferences in a profile that is unique to each student. Once known, this profile is used to either adapt text to or generate content aligned with an individual's learning preferences. This profile may be updated, allowing the LLM to incorporate changes in preferences over time. We conclude the paper with a brief discussion of middle school students' perceptions of this framework and its potential implications for engineering education.

A Framework for Implementing Personalized Learning Using Large Language Models

In this section, we present a two-agent framework designed to leverage LLMs for PL. As past research has shown, LLMs often exhibit enhanced performance when multi-step reasoning tasks are addressed through a structured sequence of sub-tasks. For instance, Luan et al. [10] achieved enhanced performance in robot task planning using LLMs when one agent first decomposed complex requests into single-objective actions before a second agent refined those actions working with a visual language model. Zhou et al. [11] define this approach as "least-to-most" prompting, wherein a complex task is first decomposed into a series of small, easily interpretable problems. The answers to these smaller problems are then combined, collectively informing the solution to the original, more complex request. The least-to-most strategy is particularly powerful as it allows the user to build context that is critical to model performance [12].

Building on these ideas, our framework divides the task of text personalization into two major steps, (1) profile identification and (2) text personalization, as outlined in Figure 1. While similar to the least-to-most prompting strategy described above, our framework deliberately applies the same two-step process for each student. This is intentional as it ensures consistency between students and reduces variability that might otherwise arise if an LLM agent were to decompose the task differently each time. This consistency is especially beneficial in the design of a controlled study [13] and for predictable implementation.

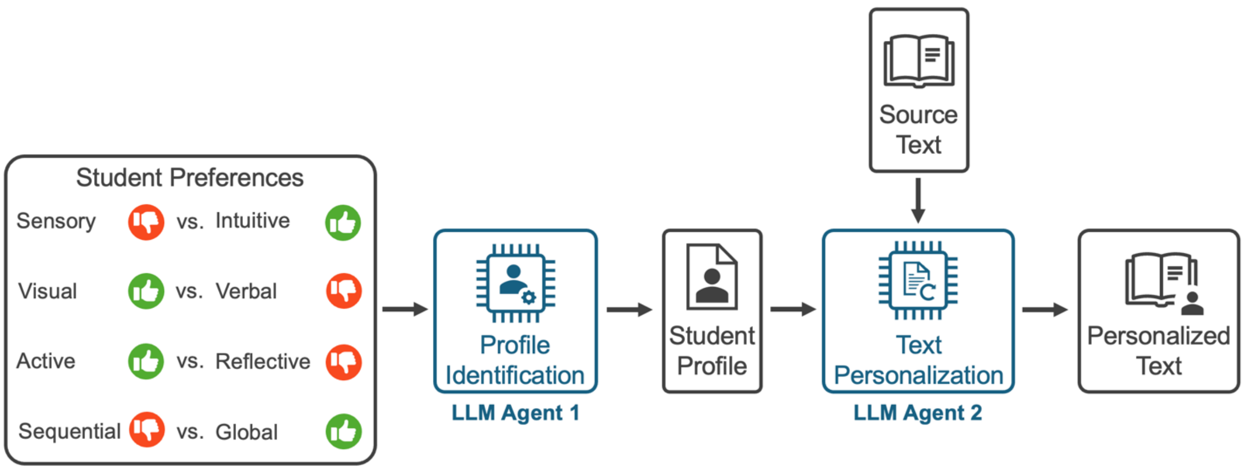

Figure 1. Overview of two-agent text personalization framework.

Agent 1: Profile Identification

As illustrated above, the first of these tasks is profile identification where the goal is to identify students' individual learning preferences. Focusing on the presentation of educational texts, our framework begins with an informal assessment of student learning preferences. To do so, we present students with four pairs of short paragraphs. Each pair covers the same topic, but the two paragraphs are differed to reflect distinct aspects of each of the four dimensions of the Felder-Silverman model. This model seeks to explain how students best perceive (sensory/intuitive), receive (visual/verbal), process (active/reflective), and understand (sequential/global) new information [14]. Providing diverse pairs of training texts in this way helps us capture various learning preferences in the students' profiles. Doing so is critical as the training texts provide the initial data for each profile.

Note that traditionally termed learning style models, including the one developed by Felder and Silverman, have garnered significant criticism for reducing students to exclusive categories and for their lack of empirical evidence [15]. We therefore stress that our use of the Felder-Silverman model in this framework is intended only to provide a diverse set of training examples to initiate a student profile. In this way, the Felder-Silverman dimensions act as an informative proxy to elicit an initial spread of learning preferences (e.g., visual vs. verbal wording, sequential vs. global structuring, etc.), not to place any student into a fixed category. We note one major benefit of using a widely known model is that GPT-4 already "knows" the construct vocabulary from its training data. Therefore, we can structure our prompts around the Felder-Silverman model without the need for a specialized or a fine-tuned LLM. This is advantageous as it increases our framework's accessibility to both teachers and researchers.

The training data used to form a student's learning profile therefore consists of four choices corresponding to the text a student liked most within each of the four pairs of paragraphs (see Figure 1). This data is provided to the first LLM agent which is responsible for developing a concise yet descriptive profile of the student's learning preferences. In designing and testing this text personalization framework [16], [17], we used OpenAI's Chat Completions application programming interface (API), which accepts both a system and user message as input. In short, the system message provides the model with high-level instructions (e.g., the model's goal and the type of output it should generate), while the user message provides request-specific context [18]. The system and user messages used for the profile identification agent are provided in Table 1 of the Appendix. Note that these prompts were initially developed for OpenAI's *gpt-4-1106-preview* model, so some modification may be needed to use them with other models or versions.

The profile identification agent outputs a short, three- to four-sentence paragraph that describes the student's learning preferences. This profile is unique to each student because the model can generate slightly different outputs each time even with identical inputs [19], providing flexibility in the personalization. Multiple profile options can therefore be developed, enabling students to pick one or a set thereof that yield the best results. An example of the type of student profile to be generated is given at the end of the system message in Table 1 of the Appendix. This profile is used as a component of the input to the text personalization agent, which is discussed next.

Agent 2: Text Personalization

Once the student profile has been created, the next step is to generate a personalized text aligned with the student's preferences. This must be done without removing critical information from the text which a student may need to complete an assignment, for instance. Rather than relying entirely on GPT-4's text-generation capabilities, our personalization agent takes two inputs (see Figure 1): (1) an original educational text, and (2) the student profile created by the first agent. The personalization agent then uses these to produce a text that is adapted to the student's learning preferences. By providing the LLM a source text, we ensure that the adapted version remains focused on the subject matter. Doing so also enables the framework to be applied broadly across STEM-related texts, making it possible to personalize a wide range of scientific articles and textbook passages.

The system and user messages for this LLM agent are provided in Table 2 of the Appendix. Note again that these prompts were initially developed for OpenAI's *gpt-4-1106-preview* model, but can be updated to work with newer models. A close read of the system messages for both agents reveals the prompt structure we found to be optimal. First, both system messages (for agents 1 and 2) define a role for the LLM to assume. As this framework was developed for a middle school science intervention, we define the model's role as a middle school science teacher. From there, each agent's system message specifies the main task—either describing the students learning preferences or personalizing a given text. Following this, we constrain the LLMs output. For instance, we found that GPT-4 would often use ornate language or would generate unnecessarily long responses. As such, we emphasize that the output should be clear, concise, and appropriate for a middle school student. We end the system message of the profile identification agent with an example. This is not necessary for the text personalization agent as the structure of the rewritten text will vary based on the source text.

Student Perceptions of Educational Artificial Intelligence

We implemented the two-agent framework described above in a small-scale (n = 23) randomized controlled pilot study at one middle school. For the interested reader, more details about the exploratory study are provided by Vaccaro et al. in [20]. Students were asked to indicate their preference between the personalized text output by agent 2 and an original text to assess the framework's effectiveness. Although evaluating this implementation is not the focus of this paper, we briefly summarize illustrative student feedback from the exploratory study below.

In general, students responded positively to the AI personalization framework. One participant commented that the process was quick and easy to complete, aligning with the average completion time of ten minutes for this portion of the study. This participant related their experience with the personalization framework to their past experiences with adaptive platforms:

> *It didn't seem too long… If I were to use this kind of thing in a class… like it wouldn't, it wouldn't be like a lot of the adaptive tests we take where it takes, you know, three, four class periods.*

A powerful, yet implicit, benefit of the framework described earlier is that students do not need to get just one personalized text back from the LLM. Rather, the model can provide a near limitless number of options. The same is true of the student profile which, if desired, can be renewed each day or varied with subject area. One student highlighted this benefit when asked to describe what it was like to participate in the research study:

> *It was like choosing an answer even though knowing that you're not going to get it wrong because they're all right.*

Five other students built on this idea, describing how they enjoyed the freedom to choose between multiple text options. When asked what they liked about the PL platform, one of these students described how having multiple options—rather than being forced to read a single provided text—was beneficial:

> *I feel like the freedom of being able to choose from like the two choices, instead of just having this one article that tells you about like what the information means is like...*

Another one of these five students emphasized that there were times when one of the options was clearer to them than the other, despite the texts covering the same content. This disparity highlights that it is not always the content itself that leads to confusion:

> *Oh, I liked, um, choosing because they were like, sometimes the paragraphs were very similar, but then other times, like, they had two very different tastes of like… some of them were just clearer.*

While most students appreciated having multiple options, one of the students who participated in the study was concerned that the different texts can vary in the details they include. This was of special concern should the students be tested or otherwise evaluated on the reading(s). The student who expressed this point-of-view explained that:

> *Well, [each of the options] had the main topic of the text, but I feel like if you were going to ask a question about it, because there was, like, differences between the texts. Like, some had more detail and some didn't. So maybe there was, like, a question asked that wasn't asked in one of the other texts, so it wasn't shown…*

This sentiment highlights some of the potential tension between individualization and the current model of education which often focuses on learning outcomes and summative assessment (e.g., ability to answer a specific question correctly or incorrectly) rather than formative assessment. Although this larger debate is beyond the scope of the present paper, it is worth noting for future research in PL.

Finally, we note that there were times when GPT-4 generated words or phrases that were beyond students' vocabularies, despite attempts to control the language the model used (see the Appendix). One participant suggested:

> *I feel like it would be very helpful for the next people, if there was some sort of glossary within the… activity. There were more [confusing words] in some of them [than in others], but yeah.*

We include this feedback to illustrate the types of resources students may need to use LLMs successfully on their own. We note that it may be beneficial to purposefully integrate them into future platforms—especially those that are used outside of a controlled research environment.

Discussion

The framework described above was developed for middle school science education and, as such, the prompts discussed here are limited to this use case. However, these prompts can be adapted to new contexts, grade levels, or subject areas, offering future researchers, educators, and students a starting point for LLM-powered PL. We hope that the series of system and user messages we provide here can serve as a valuable resource for those currently unfamiliar with language models and for future applications of LLMs in education.

We focused specifically on middle school because early engagement in STEM is critical for maintaining student interest through high school and beyond. As some participants noted in the study described above, being able to choose a text to read was beneficial, suggesting that any given text may appeal to one student's learning preferences more than others. We note that middle school represents a critical time for fostering students' long-term academic interests, as demonstrated by Maltese et al. [21] who found that students who reported interest in STEM from middle school through college were more likely to earn STEM degrees.

As evidenced by the system and user messages included in the Appendix, the student profiles and, by extension, the performance of the text personalization model are dependent on the Felder-Silverman model. While this was necessary to initialize a description of student learning preferences, future work with larger datasets could look to reduce or eliminate the profile's dependence on this learning style model. For instance, profiles can be updated with each text passage that a student selects or rewrites. Additional information—such as student interests and areas where they need more support—can also be provided to generate increasingly descriptive and student-specific profiles with time. Finally, by refining the methods and prompts presented in this paper, future research can explore the effects of LLM-powered PL on student learning outcomes.

Conclusion

This paper describes a two-agent approach to developing an LLM-powered PL environment. The following points highlight the major findings and contributions of this work:

- We developed two LLM agents to implement the PL framework. The first of these agents focuses on the development of a profile to describe a student's learning preferences, while the second uses this profile to personalize a given text.
- A pre-defined learning preference model was critical to the development of a framework that does not require a specialized or find-tuned LLM, improving the accessibility of the work to researchers, educators, and students alike. The Felder-Silverman model, which is well-known and is covered by in GPT-4's training data, was used for this framework.
- Most students responded positively to the GPT-4-powered text personalization framework. Students enjoyed the simple structure and quick response times in addition to the freedom to choose between multiple readings. Students valued the absence of a right or wrong choice, suggesting that this would be beneficial to their academic growth.
- Future work should aim to apply this framework in larger studies, across subject areas, and for different grade levels. Work may also aim to evaluate the use of LLMs on students learning outcomes and motivation, among other measures.

Acknowledgments

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.

During the preparation of this paper, the authors used OpenAI's ChatGPT models 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.

Ethics Statement

The study regarding human subjects was reviewed and approved by the University of Connecticut's Storrs-campus Institutional Review Board (IRB) under decision number H23-0348. A parent or legal guardian of all participants completed written informed consent (permission) forms for their child to participate in this research study. All participants signed an assent form prior to completing study activities.

References

[1] J. W. Keefe, "What is personalization?," *Phi Delta Kappan*, vol. 89, no. 3, pp. 217–223, 2007.

[2] J. F. Pane, E. D. Steiner, M. D. Baird, and L. S. Hamilton, "Continued Progress: Promising Evidence on Personalized Learning," *Rand Corporation*, 2015, doi: 10.7249/RR1365.

[3] M. Bulger, "Personalized learning: The conversations we're not having," *Data and Society*, 2016, [Online]. Available: https://datasociety.net/library/personalized-learning-the-conversations-were-not-having/

[4] U.S. Department of Education Office of Educational Technology, "Reimagining the Role of Technology in Education: 2017 National Education Technology Plan Update," 2017. Accessed: Jan. 05, 2025. [Online].

[5] National Academy of Engineering, "NAE Grand Challenges for Engineering," Washington, D.C., 2017. Accessed: Jan. 05, 2025. [Online]. Available: https://www.nae.edu/187212/NAE-Grand-Challenges-for-Engineering

[6] 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)?," *Educ Psychol Rev*, vol. 33, pp. 1675–1715, 2021, doi: 10.1007/s10648-021-09615-8.

[7] T. W. Price, Y. Dong, and D. Lipovac, "iSnap: Towards Intelligent Tutoring in Novice Programming Environments," in *Proceedings of the 2017 ACM SIGCSE Technical Symposium on computer science education*, Association for Computing Machinery, 2017, pp. 483–488. doi: 10.1145/3017680.3017762.

[8] OpenAI, "GPT-4 Technical Report," Mar. 2023, doi: 10.48550/arXiv.2303.08774.

[9] R. Sajja, Y. Sermet, M. Cikmaz, D. Cwiertny, and I. Demir, "Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education," *Information*, vol. 15, no. 10, p. 596, Sep. 2024, doi: 10.3390/info15100596.

[10] Z. Luan *et al.*, "Enhancing Robot Task Planning and Execution through Multi-Layer Large Language Models," *Sensors*, vol. 24, no. 5, p. 1687, Mar. 2024, doi: 10.3390/s24051687.

[11] D. Zhou *et al.*, "Least-to-Most Prompting Enables Complex Reasoning in Large Language Models," in *The Eleventh International Conference on Learning Representations*, 2023. doi: 10.48550/arXiv.2205.10625.

[12] P. Xu *et al.*, “Retrieval meets Long Context Large Language Models,” in *The Twelfth International Conference on Learning Representations*, 2024. doi: 10.48550/arXiv.2310.03025.
[13] K. Stanley, “Design of Randomized Controlled Trials,” *Circulation*, vol. 115, no. 9, pp. 1164–1169, Mar. 2007, doi: 10.1161/CIRCULATIONAHA.105.594945.
[14] R. M. Felder and L. K. Silverman, “Learning and teaching styles in engineering education,” *Engineering education*, vol. 78, no. 7, pp. 674–681, 1988, Accessed: Jan. 05, 2025. [Online]. Available: https://engr.ncsu.edu/wp-content/uploads/drive/1QP6kBI1iQmpQbTXL-08HSl0PwJ5BYnZW/1988-LS-plus-note.pdf
[15] P. A. Kirschner, “Stop propagating the learning styles myth,” *Comput Educ*, vol. 106, pp. 166–171, Mar. 2017, doi: 10.1016/j.compedu.2016.12.006.
[16] M. Vaccaro, M. Friday, Z. G. Akdemir-Beveridge, and A. Zaghi, “Designing and testing AI-based text personalization tools,” in *2025 ASEE Annual Conference & Exposition*, Montreal, Quebec, Canada, 2025.
[17] 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.
[18] OpenAI, “OpenAI Platform API Reference,” Create chat completion. Accessed: Jan. 05, 2025. [Online]. Available: https://platform.openai.com/docs/api-reference/chat/create
[19] B. Atil, A. Chittams, L. Fu, F. Ture, L. Xu, and B. Baldwin, “LLM Stability: A detailed analysis with some surprises,” *ArXiv*, Aug. 2024, doi: 10.48550/arXiv.2408.04667.
[20] 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 [Preprint]*, 2024, doi: 10.48550/arXiv.2408.05204.
[21] A. V. Maltese, C. S. Melki, and H. L. Wiebke, “The nature of experiences responsible for the generation and maintenance of interest in STEM,” *Sci Educ*, vol. 98, no. 6, pp. 937–962, 2014, doi: 10.1002/sce.21132.

Appendix

The following tables present the system and user messages used to define the two large language model agents used in this framework. Text that is ***bolded and italicized*** represents input that may vary by application. Samples are not included here for conciseness but are available on the authors' GitHub page for download at https://github.com/m-vaccaro/LLMs-and-Personalized-Learning. The newline character *\n* has been typed out in some locations of the tables for brevity and ease of interpreting the prompts. The interested reader is directed to the GitHub page linked above for the full prompts in their unedited states.

Table 1. System and User Messages that define the Profile Identification agent (Agent 1). *Note:* These prompts were optimized for OpenAI's *gpt-4-1106-preview* model.

Message Type	Content
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***

Table 2. System and User Messages that define the Text Personalization agent (Agent 2). ***Note:*** **These prompts were optimized for OpenAI's *gpt-4-1106-preview* model.**

Message Type	Content
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***]