# Enhancing Special Education with Generative AI: Suggestions from our Practice[1]

Amy Walter[2]

***Abstract***

*This paper reflects on the practical integration of generative AI, specifically ChatGPT, in a middle school (grades 6-8) special education setting. Initially adopted by teachers to manage an overwhelming workload, the use of AI evolved into a collaborative tool that supported instructional planning, simplified complex texts, created lesson materials, and enhanced student learning experiences. By sharing real-time applications and insights, the paper illustrates the critical role of teacher expertise in the effective use of AI to align with educational standards and support diverse learning needs. It also highlights the broader implications of integrating AI within the Technological Pedagogical Content Knowledge (TPACK) framework, emphasizing the need for thoughtful, informed engagement to maximize the pedagogical value of AI tools. The reflection provides suggestions for educators navigating the complexities of incorporating AI into their practice, underscoring the importance of professional judgment and iterative refinement in optimizing AI outputs for classroom use.*

***Keywords:*** *ChatGPT, generative AI, instructional planning, PLC, SDI, special education, teacher expertise, TPACK framework*

## Enhancing Special Education with Generative AI: Reflections on our Practice

Our journey with generative AI in the classroom began not with a clear research agenda or a strategic plan but out of sheer necessity. Like many educators, we found ourselves overwhelmed by the increasing demands of our roles—demands that far outpaced the time and resources available to us. Initially, we turned to generative AI not because we foresaw its transformative potential but because we needed immediate support to manage our workload and meet the diverse needs of our students. Little did we know that this initial, unplanned foray into using AI would evolve into a significant learning experience, reshaping our approach to teaching and offering insights that we

1. I would like to express my gratitude to Ashley, whose collaborative learning journey with me significantly contributed to the development of this article. Her insights and dedication have been invaluable in shaping the work presented here. I am lucky we get to work together.

2. I have no known conflict of interest to disclose. I used ChatGPT 4.0 https://chatgpt.com/ to improve readability by identifying areas that might be redundant or unclear to unfamiliar readers. I then reviewed the tool's suggestions, double-checked the flagged sections, and revised them where I found the feedback valid.

now share in the hope of adding value to the conversation with other educators who face similar challenges.

This reflection is not a traditional academic paper. It does not begin with a research question or comprehensive literature review. Instead, it chronicles our experience with using generative AI effectively—a process that was not linear but marked by trial and error, discovery, and adaptation. For clarity, we present our story in a more structured format, but our actual learning process was often overlapping and iterative, with multiple lessons occurring simultaneously. To share our learning, we have reviewed almost two years' worth of our interactions with ChatGPT. The following examples explain how we navigated this new territory, learned to leverage generative AI to reduce our workload, enhance instructional planning, and support student learning.

We have organized this reflection into several sections to guide readers through our journey. We begin with the background and context that led us to explore AI, highlighting some challenges unique to special education. Next, we delve into the practical applications of AI, detailing how we used it to create instructional materials, simplify complex texts, foster student critical thinking, and anticipate learning challenges. Throughout these sections, we draw out key lessons and strategies directly applicable to everyday classroom situations. We conclude with broader implications for the effective use of AI in education, emphasizing teacher expertise's crucial role in initiating and refining AI outputs, aligning our reflections with the TPACK framework (Mishra & Koehler, 2006). This framework highlights the intersection of technological, pedagogical, and content knowledge as essential for effectively integrating new tools like generative AI into classroom practice.

## Background & Context

To understand our journey, it is important to consider the context that led us to explore generative AI. After ten years of teaching middle grades English Language Arts, I transitioned into a role as an instructional coach, working closely with teachers across my school. I continued to immerse myself in teaching, often stepping in to lead classes when teachers were out for extended periods. This dual role provided me with a unique perspective—I was not just advising from the sidelines but deeply involved in the day-to-day challenges of teaching. The combination of hands-on experiences teaching students and supportive work mentoring colleagues gave me a deep understanding of the pedagogical and practical challenges teachers face.

One area where these challenges are particularly acute is in special education. Through my work, I was aware of the incredible workload that special education teachers carry. These educators face the challenge of addressing the diverse needs of their students—who often require highly tailored support—while navigating a system that seldom provides the necessary time or resources for teachers and students to achieve their learning outcomes together successfully. My collaboration with Ashley, a dedicated special education teacher, is the focal point for this exploration of generative AI. Ashley often found herself teaching three different grade levels simultaneously in the same room, working with students who required specially designed instruction (SDI) while held to the same general education content standards (Individuals with Disabilities Education Act, 2006).

Ashley's unique teaching assignment led us to form our own Professional Learning Community (PLC), a collaborative model where educators regularly analyze student needs and refine instructional strategies (DuFour & Eaker, 1998). Ashley brought an in-depth understanding of her students' needs, training, and experience as a special education teacher. I contributed my

knowledge of the English Language Arts (ELA) curriculum and instructional strategies. Together, we crafted lessons to make the district's prescribed curriculum more accessible and meaningful for her students. It was in the midst of this collaborative work that our journey with generative AI began.

## Unique Contribution & Broader Implications

Our journey with generative AI is distinct because it chronicles the actual, real-time use of ChatGPT in a middle school setting, specifically within the context of special education. While much of the existing literature, such as the works of Dickey et al. (2023), Choi et al. (2024), and Khazanchi and Khazanchi (2024), focuses on the use of AI in higher education and instructional design, our experience addresses the needs and realities of K-12 teachers. Unlike higher education contexts where course design can be more systematic and long-term, K-12 educators often need to respond with greater urgency, adapting pre-existing curricula to meet the immediate needs of diverse student populations.

The literature emphasizes the potential of generative AI in creating content quickly, brainstorming, refining course materials, and automating parts of the instructional design process (Dickey et al., 2023; Choi et al., 2024). However, it also highlights significant limitations, such as the AI's lack of contextual understanding, its propensity for generating hallucinated or incorrect content, and its need for clear, iterative guidance from human users. Our experience aligns with these findings but adds depth by detailing specific strategies and teacher knowledge required to engage with generative AI in a secondary special education setting effectively.

Our reflection reveals that effective engagement with generative AI depends heavily on teachers' ability to apply pedagogical knowledge, scaffold instruction, adjust materials, and critically evaluate AI outputs. The broader implications of our work suggest that while generative AI holds promise as a tool for reducing workload and enhancing instructional planning, its success in the classroom hinges on the educator's content, classroom, instructional, and technological expertise and the ability to provide clear, iterative guidance. This underscores the need for professional development focused on equipping teachers with the skills to integrate AI effectively into their practice.

We hope that by sharing our experiences, we can provide a practical guide and a source of encouragement for any classroom teacher who feels overwhelmed by the demands of their work. We share what we did and how and why we did it, drawing out key teacher moves that directly apply to everyday classroom situations.

## Integrating Generative AI for Progress Monitoring in Special Education

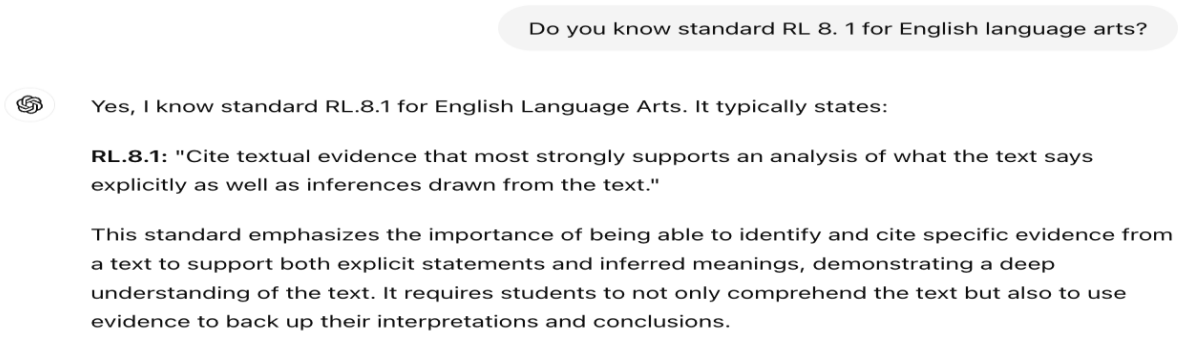
The need for more robust progress monitoring came with new district requirements around special education. Our weekly one-hour PLC sessions, which we dedicated initially to crafting lesson plans and developing student materials, were increasingly consumed by the task of writing more progress monitoring assessments. These assessments needed to be meticulously aligned with individual student goals, tied to state standards, attached to classroom content objectives, and designed to mirror state testing formats. While we recognized the value of such data, our reality as educators—not test writers—meant this process was slow and challenging. Often, given the complexity of aligning questions with specific standards, evaluating vocabulary, and designing plausible distractors for multiple-choice questions. we managed to create only one or two assessments

per hour. This focus on assessment creation left us with little time to develop the instructional activities necessary to foster the learning we were measuring.

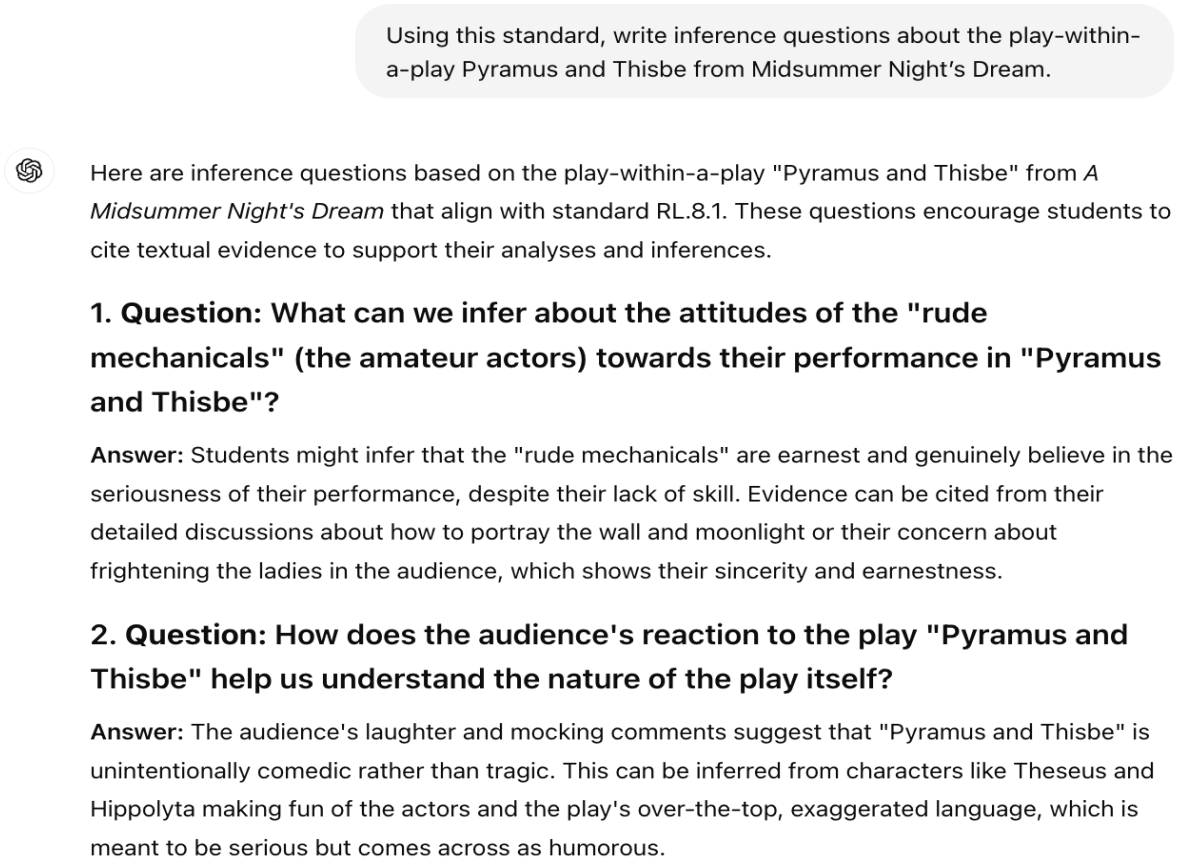
During this period, we discovered ChatGPT. We initially viewed it as a potential solution to lighten our workload, not yet fully understanding generative AI's broader implications or possibilities. Looking back at my ChatGPT history from 2022, the prompts are direct: "Write multiple-choice questions about this," followed by a text excerpt[3]. The responses varied from basic recall questions like "What color is the boy's hair?" to more sophisticated ones like "What does the boy's cloak symbolize?" Each question required careful consideration to ensure it aligned with the learning standards we were assessing, matched the student's Individualized Education Program (IEP) goals, and used appropriate vocabulary. While the AI-generated question, "What does the cloak symbolize?" may have seemed appropriate, it required adjustment to better align with our lesson's focus on inference rather than symbolism. Results like this underscored the importance of revision and adaptation based on our understanding of the end game: state testing, district standards, and student expectations. From the beginning, we found that this nuanced teacher knowledge was critical when working with generative AI.

We realized that including specific standards or key elements from those standards in our initial prompts significantly improved the quality of the AI's output. We learned to ask ChatGPT if it knew a certain standard, and upon receiving a clear and correct response, we could prompt the AI more precisely. As seen in Figure 1, the resulting AI-generated questions aligned more closely with our instructional goals.

**Figure 1:** *Effect of Including Specific Standards in AI Prompts*




3. At the time of this article's writing, copyright laws regarding generative AI remain unsettled, with ongoing legal cases primarily focused on systematic, for-profit AI training. Our use—uploading texts to create educational materials—is consistent with fair use guidelines and does not contribute to AI training or commercial exploitation. https://www.authorsalliance.org/2024/02/29/why-fair-use-supports-non-expressive-uses/ We did collaborate with our organizations legal and library staff to reach this conclusion. We are mindful that these practices may need to change as court cases are settled.




*Note*. This figure shows that when specific standards or key elements from those standards were included in initial prompts, the quality of the AI's output improved. We could create more precise prompts by asking ChatGPT if it knew a certain standard and getting a clear, correct response.

Through trial and error, we learned to request more questions than needed in a single prompt. Initially, we requested a small set of questions, which often necessitated multiple rounds of refinement. We streamlined the process by requesting a larger batch upfront, thereby finding sufficient high-quality questions in the first output. This strategy saved time by reducing the need for multiple rounds of questioning and allowed us to select the most relevant questions from a larger pool.

In some cases, ChatGPT recognized the text we referenced, especially well-known works like Shakespeare's plays, requiring no additional input beyond the title. However, we needed to provide specific excerpts for less common texts not included in its training data. This experience underscored the importance of understanding how generative AI is trained and the implications of its training data on the content it generates. Given the ongoing legal discussions surrounding AI and copyright, we were mindful of using excerpts in ways that align with fair use—specifically for educational purposes rather than for training AI systems. By being aware of these factors, we were more effective in engaging with the AI. However, this also highlighted a significant limitation: AI tools like ChatGPT are trained on large datasets that may not always be representative or bias-free. This can lead to incorrect outputs or reflecting cultural biases or reinforcing stereotypes, which requires educators to critically evaluate and sometimes modify the AI-generated content to ensure it aligns with educational values and standards. Understanding both the capabilities and limitations of AI is crucial for using these tools responsibly in educational settings.

Integrating ChatGPT into our workflow allowed us to devote more time to developing instructional activities by alleviating some of the burdens associated with assessment creation. However, this experience highlighted the role of exercising teacher expertise to effectively guide AI to ensure that its outputs align with educational standards and support student learning. Our deep understanding of ELA standards was crucial in identifying questions that were aligned with curricular goals and meaningful for assessment. Moreover, our knowledge was instrumental in evaluating and revising the AI-generated content. We had to modify ChatGPT's outputs to reflect the instructional needs that we identified as unique to our special education students. This process demonstrated how the intersection of content knowledge and pedagogical strategies is critical to successfully integrating AI tools in instructional planning.

## Key Teacher Moves for Effective AI Engagement to Create Assessments

1. Provide information to generative AI about the skill or specific standards being assessed.
2. Share relevant student details (while protecting confidentiality) to align AI output with specific student needs.
3. Ask for more than you need.
4. Be prepared to evaluate and revise output.
5. Know how generative AI is trained and what it "knows."

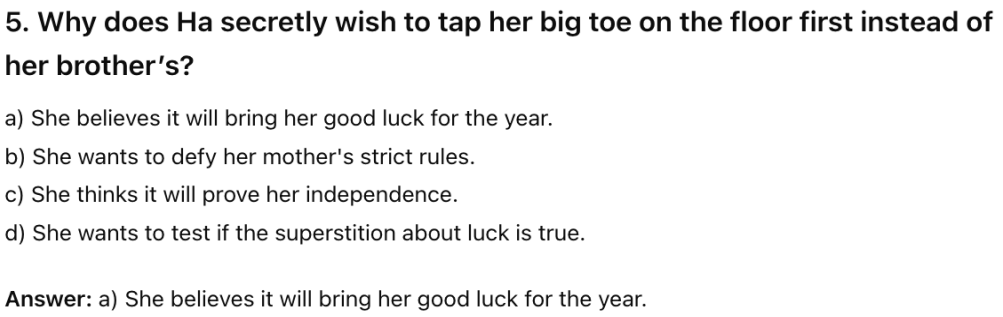
## Refining AI-Generated Content through Conversational Prompting

As we continued to explore the potential of generative AI for our classroom needs, we recognized that the key to effectively using AI-generated content was not simply accepting the initial outputs. We had to have conversations with ChatGPT. We crafted our prompts deliberately, asked questions, gave directions to revise, and evaluated and revised outputs.

After spending a fair amount of time initially revising and editing outputs, we learned the value of being specific in our initial prompts to improve the relevance and quality of AI outputs. For instance, specifying the question format and requesting an answer key became standard practice. Prompts like "Write 10 multiple-choice inference questions with an answer key about *A Midsummer Night's Dream*, Act 2, Scene 1," or "Write 3 short answer, detail-level comprehension questions about Amendment 14 of the US Constitution. Provide sample responses" immediately yielded more targeted results.

Since we started with primarily multiple-choice questions for progress monitoring, a crucial part of this process involved scrutinizing the AI-generated answer choices, particularly the foils (incorrect answers). We often found that some foils could be considered viable answers, potentially confusing students. To address this, we began engaging in dialogue with ChatGPT to clarify ambiguities. This iterative back-and-forth was necessary because generative AI lacks nuanced understanding. We were responsible for ensuring that each question had a clear "best" answer to prevent confusion. In one case, when we reviewed a question about a character's motivation, the initial answer choices included the text shown in Figure 2 (next page).

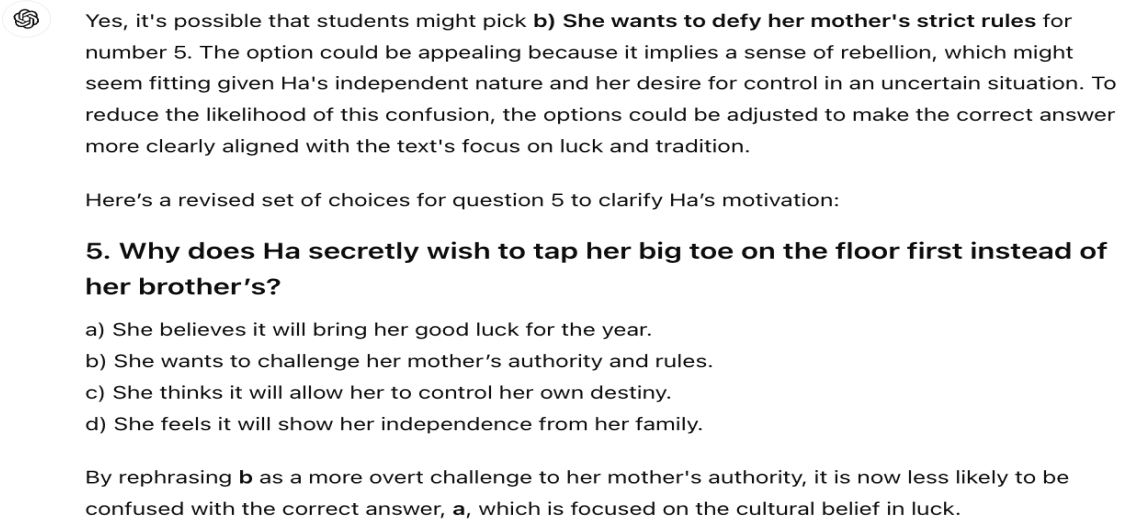
**Figure 2:** *Initial AI-Generated Multiple-Choice Question*




*Note*. This figure shows an example of an AI-generated multiple-choice question, where some foils (incorrect answers) could confuse students. In this case, option b might be misconstrued as a correct answer, highlighting the need for teacher review and clarification to ensure each question has a clear "best" answer.

We questioned whether option b could be misconstrued as correct and asked ChatGPT, "Do you think students might pick b for number 5?" Examples like this one illustrate the importance of teacher intervention in refining AI outputs to ensure clarity and accuracy. We cannot simply copy and paste generative AI output.

**Figure 3:** *Dialogue with ChatGPT to Clarify Answer Choices*




*Note*. This figure demonstrates the iterative process of engaging with ChatGPT to refine answer choices for clarity and accuracy. It shows a teacher questioning whether students might select option b as the correct answer, illustrating the necessity of teacher intervention to avoid potential confusion in AI-generated content.

Another challenge was maintaining the context of our initial prompts over extended interactions. After multiple exchanges, ChatGPT sometimes lost track of the parameters we had set. We found it necessary to reset the context by providing clarifying prompts like, "We are focusing on inference questions that require minimal background knowledge from outside sources. Use vocabulary accessible to students reading at grade level four. Focus questions on character or theme." These adjustments were crucial for maintaining the quality and focus of the outputs.

Our experience with ChatGPT reinforced the importance of maintaining a strong teacher presence in the AI-assisted content creation process. The key to the effective use of generative AI

lies in providing clear, detailed instructions and engaging in ongoing dialogue to optimize the quality of outputs. Additionally, it is crucial to be mindful of what information is shared with AI tools, ensuring that student confidentiality is protected while still providing enough context for meaningful output. Our content and student knowledge, along with understanding of pedagogical strategies, were essential in refining AI-generated resources to ensure they met our educational standards and objectives.

Despite these challenges and the need for ongoing revisions, we significantly reduced the time spent creating progress monitoring assessments—from about an hour to approximately 15 minutes. This efficiency gain allowed us to focus more on high-impact instructional decisions, such as refining lesson design, analyzing student progress, and scaffolding support. While some may worry that relying on AI for assessment creation could reduce the depth of engagement with the text, standards, and expectations, we found that it instead freed up cognitive space for more strategic instructional planning, deeper analysis of student needs, and targeted intervention design.

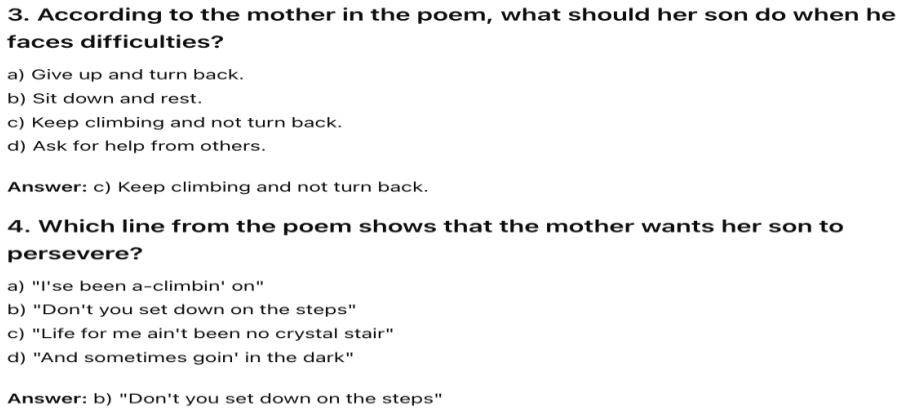
## Key Teacher Moves for Effective AI Engagement

1. Craft clear and specific prompts.
2. Engage in iterative dialogue.
3. Maintain context and focus in AI interactions.
4. Evaluate and revise output.

## Learning More Efficient Prompting

Over time, we discovered a more effective approach to generating useful content from ChatGPT by applying a strategy we taught our students: Role, Audience, Format, Topic (RAFT). This strategy provided ChatGPT with a clear framework to guide its responses, specifying who it was "acting" as, who the audience was, the desired format, and the topic. For instance, a typical prompt might be: "You are an expert special education teacher. You teach sixth graders who need maximum reading support and are working on answering detail-level questions by citing textual evidence. Write ten multiple-choice questions with an answer key using the poem 'Mother to Son' as the text." The generated questions included those in Figure 4.

**Figure 4:** *Using the RAFT Strategy to Generate AI Content*



**3. According to the mother in the poem, what should her son do when he faces difficulties?**

a) Give up and turn back.
b) Sit down and rest.
c) Keep climbing and not turn back.
d) Ask for help from others.

**Answer:** c) Keep climbing and not turn back.

**4. Which line from the poem shows that the mother wants her son to persevere?**

a) "I'se been a-climbin' on"
b) "Don't you set down on the steps"
c) "Life for me ain't been no crystal stair"
d) "And sometimes goin' in the dark"

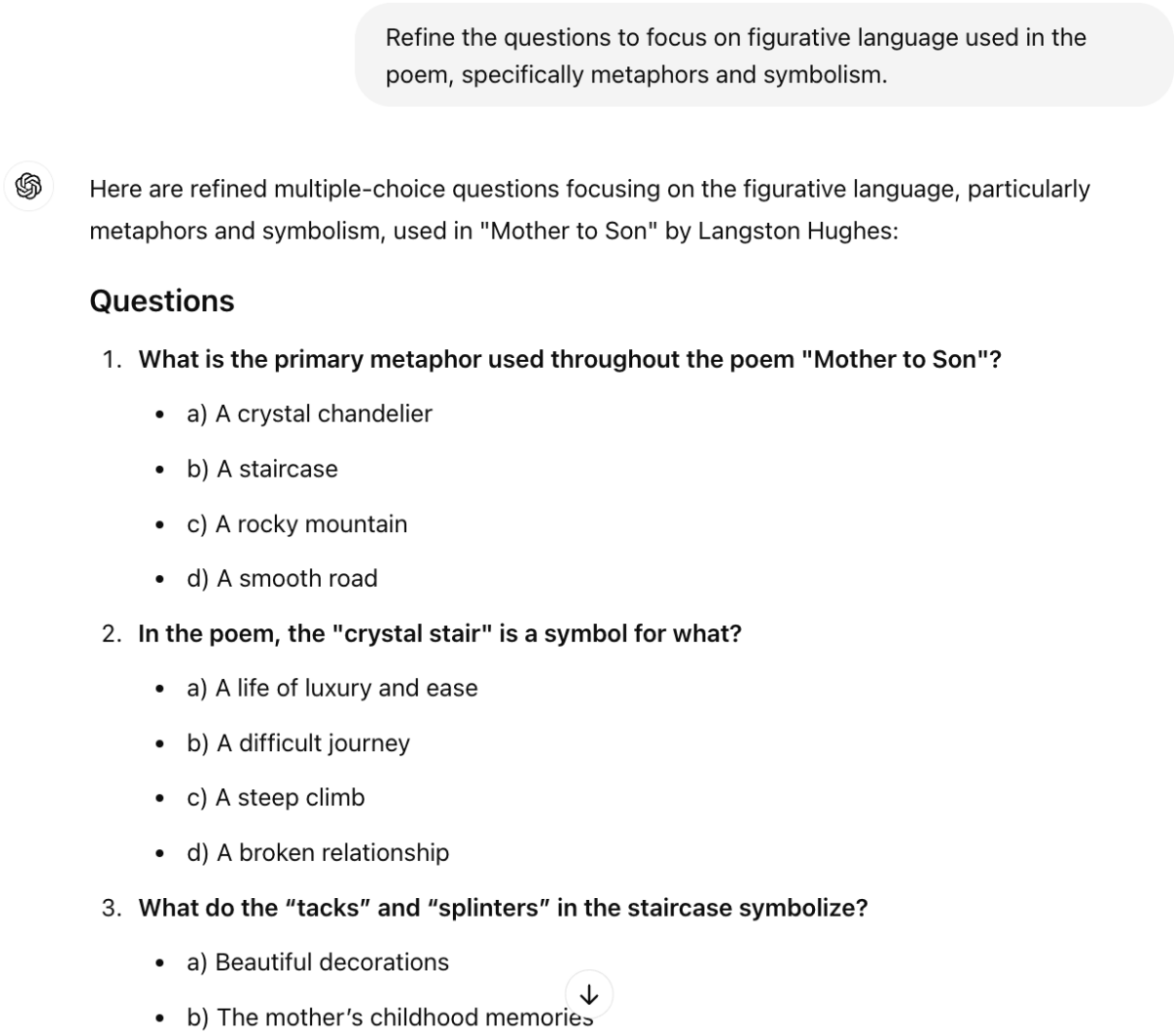
**Answer:** b) "Don't you set down on the steps"

*Note.* This figure illustrates how applying the RAFT (Role, Audience, Format, Topic) strategy improved the quality of content generated by ChatGPT.

The questions generated through this approach were relevant and effectively targeted the comprehension skills we aimed to develop, such as identifying explicit details and interpreting textual evidence. The AI-generated questions aligned precisely with our objectives by centering on key lines from the poem and prompting students to link these details to broader themes. Using RAFT allowed us to achieve these results quickly by streamlining the content creation process.

Even with RAFT, we found that follow-up prompts were often necessary to focus on specific content areas or skills, such as emphasizing figurative language or ensuring the inclusion of particular vocabulary. For example, after generating a set of basic comprehension questions, we might provide more specifics about vocabulary or skills, as seen in Figure 5. In this way, follow-up prompts ensured that AI-generated materials were not only relevant but also aligned with instructional priorities, reinforcing key concepts and scaffolding student learning effectively.

**Figure 5:** *Refining AI-Generated Content with Follow-Up Prompts*




*Note*. This figure shows how follow-up prompts were used alongside the RAFT strategy to fine-tune AI-generated content.

This iterative approach mirrored the dynamic nature of classroom teaching, where instruction is continually adjusted based on student needs and responses. By prompting ChatGPT to include specific vocabulary or focus on particular themes, we could ensure the AI-generated content was targeted and comprehensive, thus better supporting our students' learning progression. Engaging in this back-and-forth exchange with ChatGPT also mimicked the collaborative nature of lesson planning, where refining and revising are essential to creating high-quality educational materials. Ashley noted that ChatGPT had become a third non-human member of our PLC.

By refining our prompts with strategies like RAFT and engaging in a continuous dialogue with ChatGPT, we transformed it from a basic question generator into a customized assessment co-designer. This collaborative process allowed us to leverage the AI's capabilities to ensure the generated content was relevant and aligned with our instructional goals. Our ability to iteratively refine the prompts and revise the output ultimately enabled us to create more precise, meaningful assessments and instructional materials that directly supported our students' learning, enhancing comprehension and critical thinking skills in previously time-consuming and challenging ways.

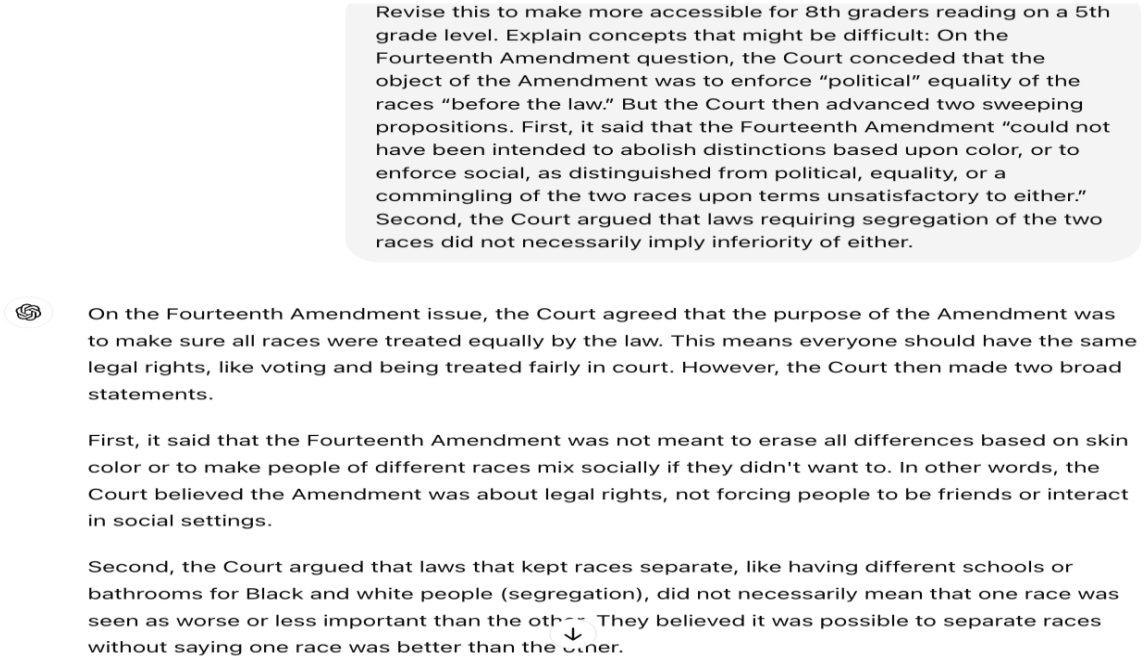
### Key Teacher Moves for Effective AI Engagement to Refine Assessment Items

1. Provide Role, Audience, Format, and Topic (RAFT) in initial prompts.
2. Name specifics for what you want.
3. Engage in iterative dialogue with GenAI.
4. Evaluate and revise output.

## Using AI to Simplify Complex Texts

As we became more proficient in using ChatGPT for assessment creation, we recognized its potential to simplify complex texts and create more accessible learning materials for our students. This practice was particularly valuable when working with historical and legal texts, which often contained dense language and intricate concepts. For our students reading below grade level, accessing these materials in their original form posed a significant challenge. Simplifying these texts while maintaining the core ideas was crucial for ensuring all students could engage with grade-level content and participate in higher-order thinking. As seen in Figure 6 (next page), this simplified version retained the original text's core concepts while translating legal jargon into language accessible to students reading at a lower grade level. By making the text more approachable, we enabled our readers to build foundational knowledge that prepared them for engaging with more sophisticated ideas. Before using ChatGPT, adapting materials for our readers was a labor-intensive process that could take several hours. With ChatGPT, this process was streamlined, allowing us to reallocate time toward more dynamic and interactive instructional activities.

**Figure 6:** *Simplifying Complex Texts for Accessibility*




*Note*. This figure illustrates how ChatGPT was used to simplify complex texts to make them more accessible for students reading below grade level.

Our dialogue with the AI iteratively refined outputs until they were optimized for our instructional purposes. For example, in the text about the Fourteenth Amendment above, we appreciated how the introductory paragraph highlighted the presence of "two broad statements," reinforcing this structure with clear linguistic cues—"First" in one paragraph and "Second" in the next. These signals would support students' comprehension. Additionally, we valued how the revised text incorporated the word "segregation." In our next iteration, we prompted ChatGPT to include "abolish" and "inferiority," as these were essential terms for students to understand and use. By leveraging our professional judgment and experience, we guided ChatGPT, transforming it from a basic text simplification tool into a strategic partner in enhancing student learning.

We applied our understanding of the content and our students' specific learning needs to use ChatGPT to simplify complex texts. In our PLC meetings, we used RAFT to thoughtfully design prompts that communicated our instructional objectives and specified the format of the outputs. We evaluated the AI-generated outputs, ensuring they were simplified, educationally sound, and aligned with our goals.

### Key Teacher Moves for Effective AI Engagement for Simplifying Complex Texts

1. Create prompts with careful consideration of the desired output.
2. Use iterative refinement.
3. Apply professional judgment—drawing on your expertise in pedagogy, content, and student needs—to review and adjust AI-generated output to align with your instructional goals.

## Utilizing AI for Strategic Instructional Planning & Student Support

As we worked with ChatGPT, we realized that our approach was evolving from using it as a product production platform to a creative thinking tool. Initially, we viewed ChatGPT much like an assistant—generating questions and simplifying texts. However, as we became more confident, we engaged in a "conversation" with the AI, refining and adjusting its outputs much like in a dialogue with a teaching colleague. This evolution was not automatic; it required us to continuously apply our knowledge of the curriculum, effective instructional strategies, and student needs—and how to protect student confidentiality—to guide ChatGPT effectively. The AI was a powerful tool, but it only became pedagogically sound when we actively directed and refined its responses to align with our instructional goals.

### Scaffolding Instruction

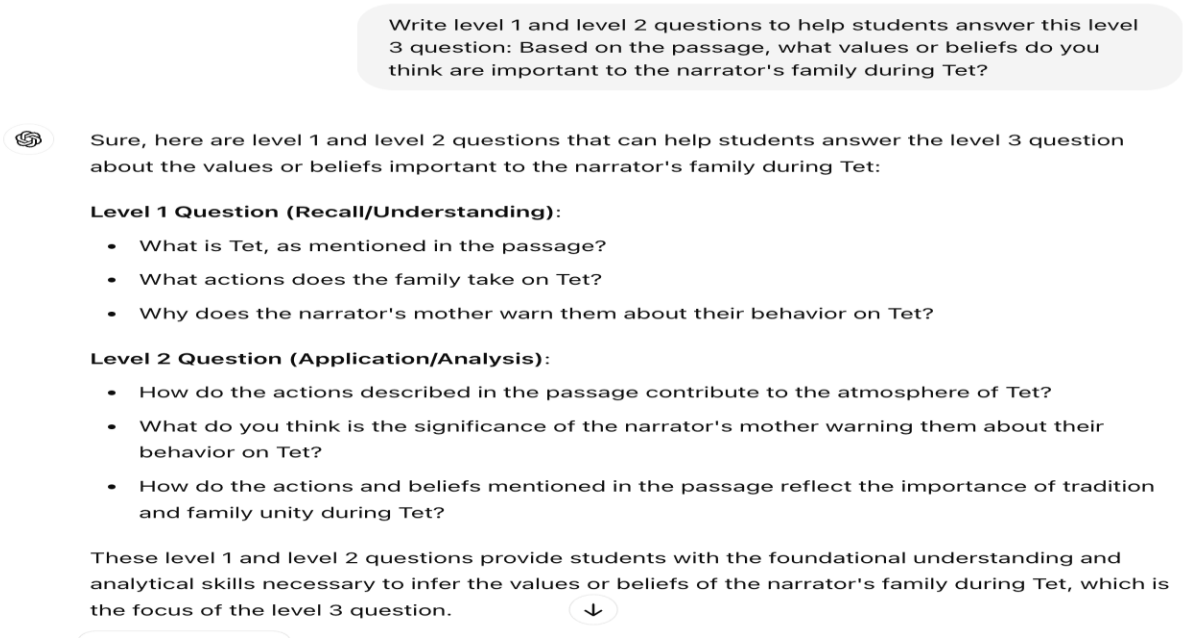
One of the first instructional strategies we explored with generative AI focused on enhancing students' inferencing skills. Many of our sixth-, seventh-, and eighth-grade students struggled in this critical area. While we had become proficient in using ChatGPT to generate questions for comprehension, we noticed that students often found it challenging to make accurate inferences, especially when they had varying levels of background knowledge. Inferencing requires students to read between the lines, drawing on textual evidence and their understanding to form conclusions or predictions, and we realized this was an area that needed more focused development.

During several PLC meetings, we discussed the specific knowledge and cognitive processes students needed to answer inference questions successfully. We also reflected on the steps we followed when making inferences. These conversations led us to revisit a strategy we had recently encountered in professional development: three-level comprehension guides (Vacca, Vacca, & Mraz, 2013). This strategy involves using three levels of comprehension—literal, interpretive, and applied—to scaffold students' understanding and inferencing skills.

We decided to experiment with this approach using ChatGPT. We investigated ChatGPT's familiarity with the three-level comprehension strategy, and upon receiving a clear, accurate description that aligned with our understanding, we set out to create instructional materials. We began by having ChatGPT generate three-level comprehension statements using familiar and accessible texts—like the classroom environment as the "text." For instance, a literal statement might be, "Ms. Ashley posts students' birthdays on the wall," which students could confirm with visual evidence. This would then lead to an interpretive statement such as, "Ms. Ashley values her students as individuals," based on the evidence of personalized engagement and classroom activities. The final applied statement could require students to predict how Ms. Ashley might respond if a new student joins the class mid-year, drawing on their understanding from the earlier levels.

As we worked with ChatGPT to generate three-level statements, we realized that the generated statements often did not always logically build from one level to the next. To address this, we reversed our approach, starting with the applied or inference-based statement and asking ChatGPT to work backward to create the interpretive and literal statements. This method proved more effective in ensuring that each level of comprehension logically supported the next, guiding students step-by-step from basic understanding to deeper inference.

**Figure 7:** *Refining Three-Level Statements for Deeper Comprehension*




*Note*. This figure shows how reversing the process of generating three-level statements with ChatGPT improved logical progression from one comprehension level to the next. The AI generated a more coherent set of statements by starting with an inference-based statement and working backward to create interpretive and literal statements.

As we refined our approach, we experimented with mixing statements students could affirm or refute based on the text with questions they had to answer, all at the same comprehension level. For example, we asked students to affirm or refute the statement: "Puck, rather than fate, is responsible for the chaos in the play." We then followed up with ChatGPT generated: "If Puck causes the chaos, what does this say about Shakespeare's view on fate versus choice? If he doesn't, what else creates the disorder, and why does it matter?" This combination helped students build a more layered understanding of the text, applying knowledge to more complex inference questions. The AI's ability to quickly generate multiple practice opportunities enabled us to provide students with exercises that gradually increased in complexity, reinforcing their skills over time.

This approach allowed students to engage in regular practice with challenging texts and to become familiar with the structure of the three-level comprehension guides. Students began to recognize these patterns, often referring to them as "3CG sets," and understood that each question was a stepping stone toward answering the final, most challenging question. They developed a clearer appreciation for returning to the text—a practice embedded in the level one questions—as a foundational strategy for answering higher-level inference and application questions. Over time, we observed that students improved in answering literal-level questions correctly, which provided a strong baseline for tackling more complex questions. Their accuracy increased across all three question levels, and they demonstrated greater confidence and willingness to engage in classroom discussions around level three questions, using textual evidence to support their reasoning.

By leveraging generative AI, we created targeted, scaffolded practice opportunities more efficiently than ever. Through an iterative process of refining and adjusting the AI's outputs, we developed a series of comprehension exercises that built on students' existing knowledge, guiding them from basic understanding to more complex inferencing. This approach enhanced our students' skills and fostered a deeper understanding of navigating texts systematically, drawing on evidence and reasoning to make informed inferences. Using ChatGPT for this type of lesson planning, we could tailor these activities to support our students' growth.

**Supporting Meta-cognitive Skills & Student Empowerment**

As we continued experimenting with ChatGPT, we recognized its value not just as a content generator but as a tool to support the development of meta-cognitive skills. Through our classroom observations, we noticed that while our students were adept at comprehending concrete information and literal meanings and were finding success with the three level-level questions, they often faced challenges when inferring deeper motivations from more less common cues. Our students struggled with higher-order thinking skills, like those described by Theory of Mind (ToM)—the ability to understand that others have different thoughts, beliefs, and intentions (Dore et al., 2018). This gap was particularly evident during discussions that required our students to interpret character motives, anticipate actions, or understand irony and implied meanings. For instance, when asked why a character might behave in a certain way, students frequently provided surface-level responses that overlooked the character's internal states or broader contextual factors. For example, when discussing Egeus's motives in *A Midsummer Night's Dream*, a common response was simply, "Egeus is mad at Hermia because she won't listen to him." While technically correct, this response failed to consider his belief in patriarchal authority, his need for control, or his personal stake in Hermia's marriage to Demetrius. We realized that for our students to engage with and comprehend narrative texts fully, they needed to develop stronger ToM and related meta-cognitive skills.

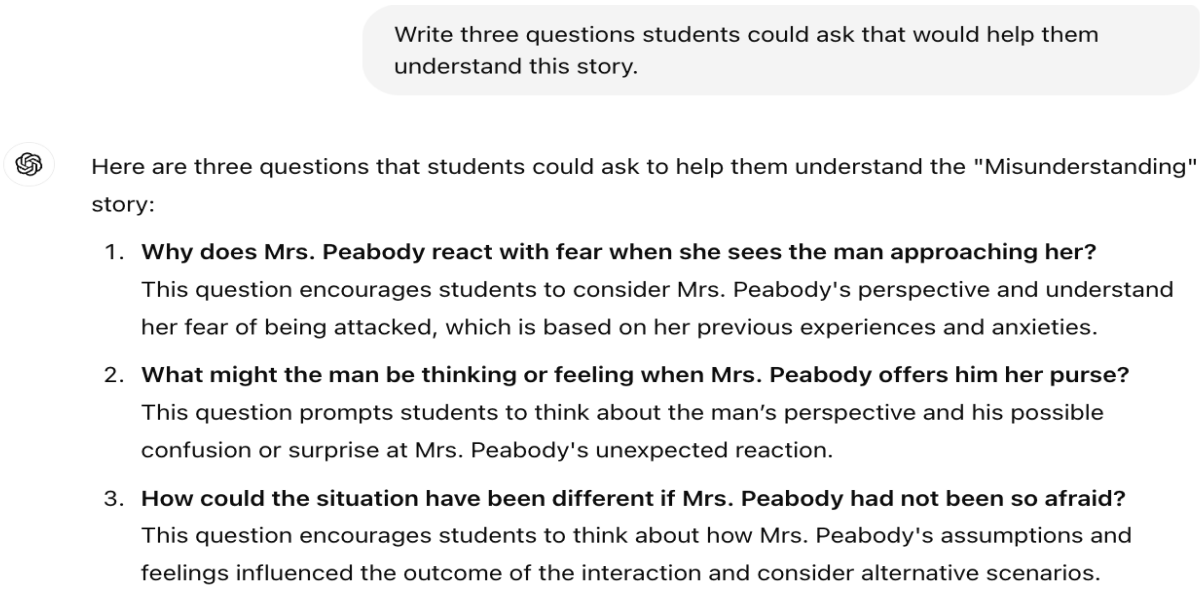
To address this challenge, we used ChatGPT to create materials to enhance our students' ToM and critical thinking abilities. One effective approach involved introducing students to Happe's (1994) "Strange Stories," which are crafted to test and develop ToM by presenting scenarios that require students to predict or explain a character's behavior based on their mental state. Recognizing the importance of varied and repeated practice to strengthen these skills, we turned to ChatGPT to generate additional stories similar in style and structure to Happe's.

Our process began by engaging ChatGPT in a dialogue about ToM, prompting it to explain the concept and provide relevant examples. This exchange helped confirm our understanding and allowed us to ensure that ChatGPT's responses aligned with current educational research. Although the version of ChatGPT we used at the time did not have direct access to several specific articles we had studied (e.g., Dore et al., 2018), its responses were generally consistent with the concepts presented in the literature, reinforcing its utility for our purposes.

We provided ChatGPT with "Strange Stories" excerpts found on the Internet (used for educational purposes under fair use principles) to serve as examples to guide the generation of new scenarios. Using examples rather than abstract descriptions led to more relevant and accurate outputs. ChatGPT's ability to quickly generate multiple new "strange stories" allowed us to offer our students a broader range of materials as well as personalized stories to practice their ToM skills.

In addition to building ToM, we aimed to strengthen our students' questioning skills, a critical component of reading comprehension and meta-cognition. To facilitate this, we used ChatGPT to generate a series of questions related to each story, as seen in Figure 8. Rather than having students simply answer these questions, we focused on fostering meta-cognitive reflection by asking them to discuss which questions would be most helpful in understanding the stories and why. This approach encouraged students to think critically about the purpose and value of different types of questions.

**Figure 8:** *Enhancing Questioning Skills & Meta-Cognitive Reflection*




*Note*. This figure demonstrates the use of ChatGPT to generate questions aimed at strengthening students' questioning skills and meta-cognitive reflection.

We utilized these ChatGPT-generated questions to engage students in discussions about which questions would deepen their understanding of the characters and which might be less useful. Interestingly, students exceeded our expectations by thoughtfully debating the relevance of certain details, such as Mrs. Peabody's age, in understanding her reaction to a man's approach. They reasoned that her response could vary significantly depending on whether she was very old or young, demonstrating a deeper level of inference and engagement than we had initially anticipated. Building on this foundation, we then had students generate their own questions about the stories. To support varying needs, some students received question stems as scaffolds, while others generated questions independently.

By emphasizing higher-order thinking, we continued to move beyond using ChatGPT as an efficiency tool. Instead, we engaged in a dynamic, iterative process with the AI, continuously refining its outputs to design richer, more meaningful learning experiences. This collaborative approach allowed us to tailor materials specifically targeting metacognitive skills. Through this dialogue with ChatGPT, we adapted our instructional strategies to better support our students' unique needs.

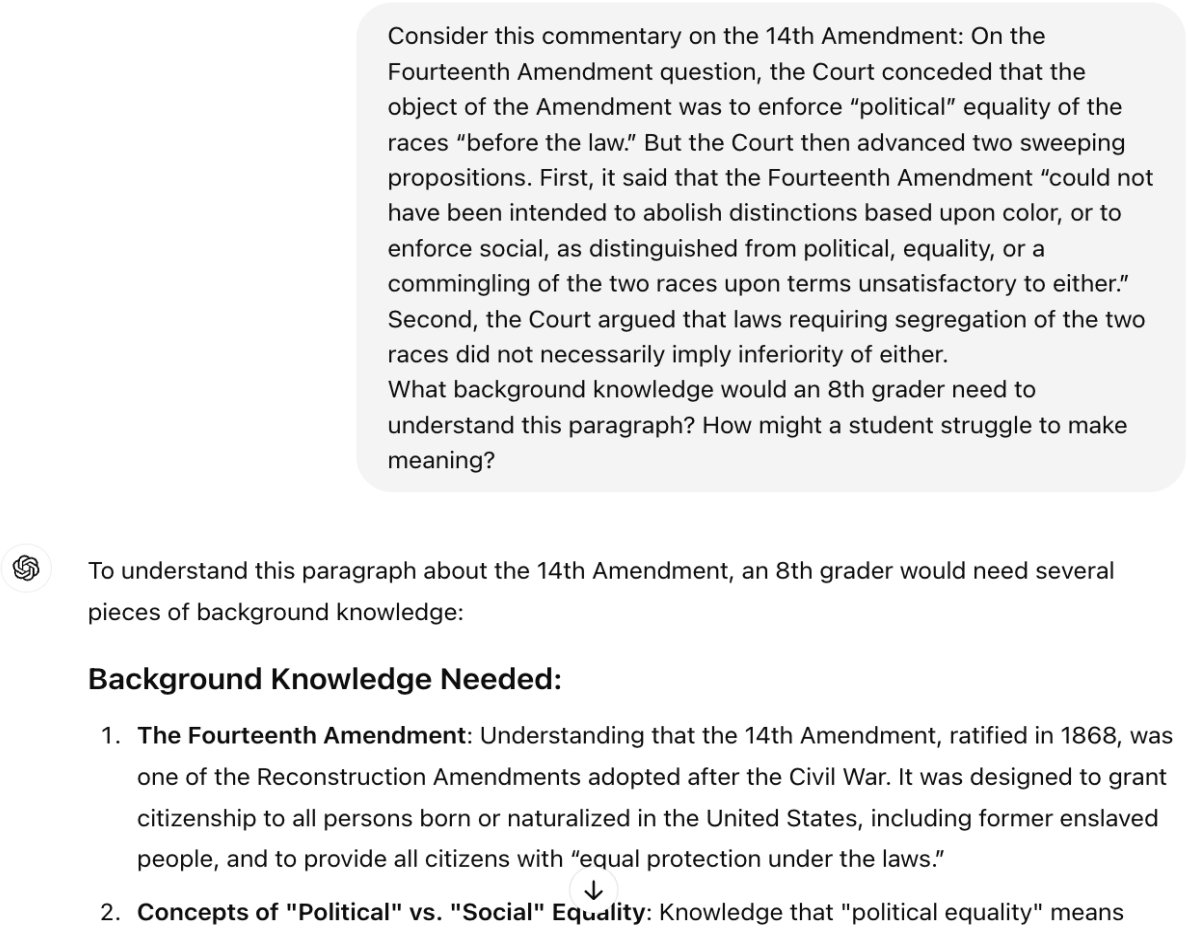
## Identifying & Anticipating Learning Challenges

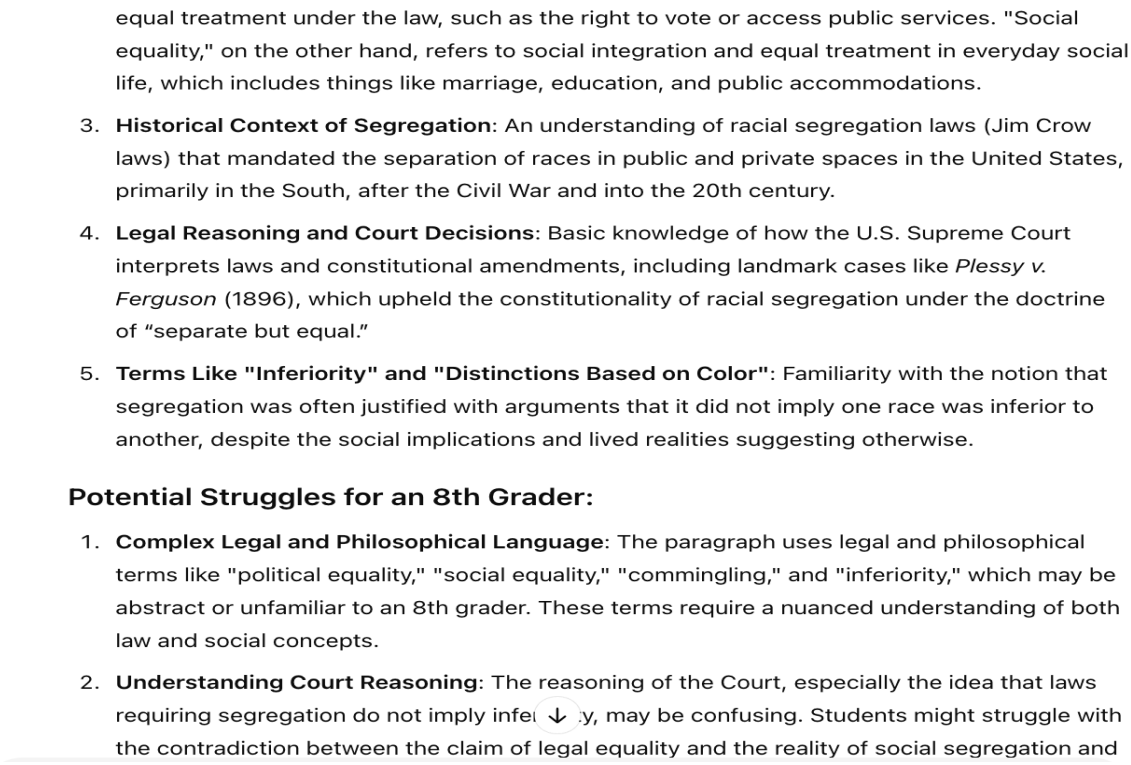
We discovered that ChatGPT could be a valuable tool for proactively identifying and anticipating potential learning challenges, allowing us to develop targeted teaching strategies. Understanding where students might struggle was crucial in crafting lessons that effectively supported their learning needs. For instance, when we introduced a complex commentary on the Fourteenth Amendment, we anticipated that the dense language and intricate historical context could create significant barriers to comprehension for our students.

To address this, we used ChatGPT to help pinpoint specific areas where students were likely to face difficulties. See Figure 9. The AI provided insights into potential vocabulary hurdles and conceptual misunderstandings that could arise from the text. With the feedback from

ChatGPT, we engaged in discussions during our PLC meetings to determine the most impactful starting points for instruction. We decided to focus on building students' vocabulary as a foundational step, recognizing that a strong grasp of key terms was essential for understanding the broader context.

**Figure 9:** *Using AI to Identify Instructional Focus Areas*

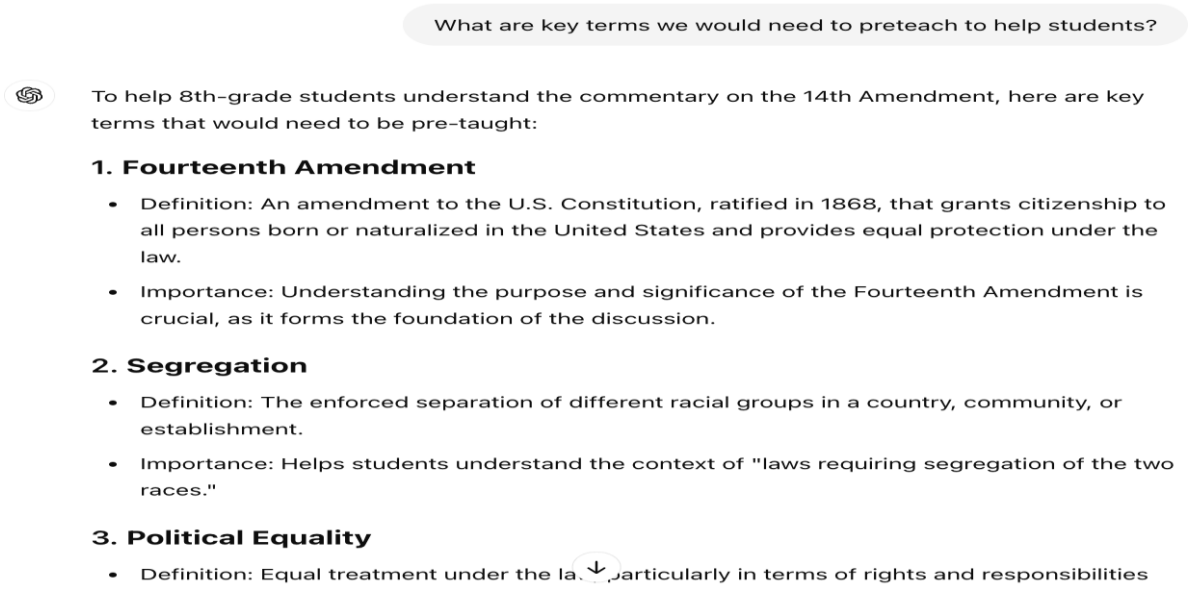





*Note*. This figure illustrates how ChatGPT was used to identify specific areas where students might face difficulties, such as challenging vocabulary and potential conceptual misunderstandings.

As we refined our instructional approach, we continued to use ChatGPT to generate vocabulary lists and definitions tailored to the text's context. As illustrated in Figure 10, using AI to generate and differentiate vocabulary lists helped scaffold learning.

**Figure 10:** *Generating & Differentiating Vocabulary Lists Using AI*



such as voting, access to public services, and legal protection.

- Importance: Clarifies the Court's use of "political equality" as distinct from "social equality."

**4. Social Equality**

- Definition: Equal status and integration of all people in social settings, which includes rights related to social interactions, access to public places, and intermingling in society.
- Importance: Understanding this term helps students grasp the distinction made by the Court between legal rights and everyday social interactions.

**5. Inferiority**

- Definition: The condition of being lower in status or quality than another or others.
- Importance: Helps students understand the Court's argument that segregation laws did not imply one race was inferior to the other.

**6. Distinctions Based on Color**

- Definition: Legal or social separations or categorizations made on the basis of skin color or race.
- Importance: Key to understanding the Court's stance that the Fourteenth Amendment was not intended to abolish racial distinctions.

**7. Commingling**

*Note*. This figure shows how ChatGPT was utilized to generate vocabulary lists and definitions tailored to the context of the text.

This use of ChatGPT allowed us to foresee learning obstacles that might have otherwise required extensive time, observation, and trial-and-error to identify in the classroom. By anticipating potential stumbling blocks well in advance, we could design lessons that directly addressed these gaps. For example, after reviewing the AI-generated list of vocabulary terms, we selected the most relevant and impactful words—such as segregation, equality, political, social, inferiority, and commingling—and identified terms that might pose comprehension challenges. Instead of waiting to see where students might struggle, we preemptively asked ChatGPT to generate student-friendly explanations and examples, particularly for more abstract words like inferiority and commingling. AI-generated analogies, such as comparing commingling to mixing different colored beads and explaining inferiority through self-esteem in competitive sports, helped us prepare targeted explanations and activities.

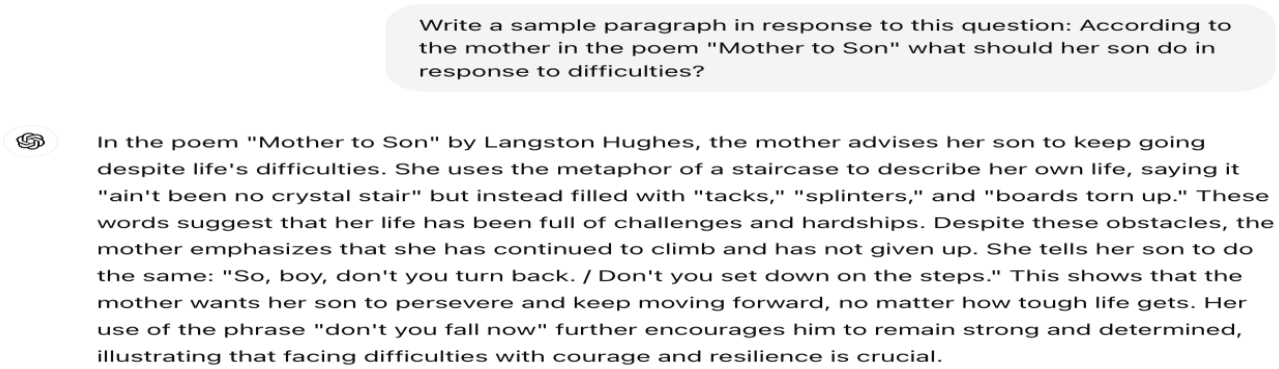
While ChatGPT provided numerous options and suggestions, our role was to critically evaluate each one, considering its relevance and applicability to our specific learning context and student needs. When the AI initially suggested a traditional word-definition matching exercise, we knew this would not effectively engage our students. Instead, we used this output as a starting point and asked ChatGPT to create a more interactive group game to reinforce word meanings. The AI-generated vocabulary charades activity, where students acted out words while peers guessed, would engage our students during class. We also prompted ChatGPT to generate sentence stems incorporating key vocabulary, ensuring that all students, including those needing more support, could use the words in context. This interplay between AI assistance and teacher expertise enabled us to create more responsive and targeted lesson plans.

## Generating Exemplars to Model Skills

Creating exemplars was another area where we effectively utilized generative AI to enhance our instructional practices. Exemplars—high-quality examples demonstrating key skills—are invaluable for helping students understand expectations and develop their abilities. While we saved as many exemplars from previous students as possible, each class's unique and individualized needs meant we frequently needed to create new examples tailored to specific learning goals. This process was time-consuming, especially when catering to the diverse skill levels within a single classroom.

Recognizing the potential of ChatGPT to assist in this area, we began leveraging the AI to generate exemplars that modeled critical skills, such as citing textual evidence to support a claim. For instance, when teaching our students to analyze the poem "Mother to Son," we used ChatGPT to create multiple sample exemplar paragraphs demonstrating how to integrate and effectively explain textual evidence, as shown in Figure 12.

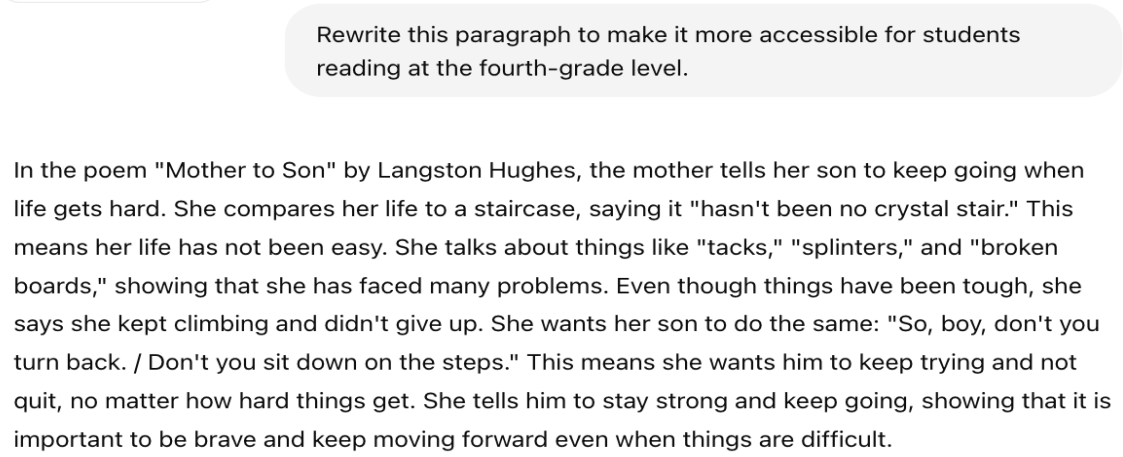
**Figure 11:** *Using AI to Generate Exemplars for Critical Skill Development*




*Note*. This figure shows how ChatGPT was utilized to create exemplar paragraphs that modeled critical skills, such as citing textual evidence to support a claim.

While this exemplar was useful, we realized that some students, particularly those reading at lower levels, might find the language inaccessible. To address this, we iteratively refined the AI-generated exemplars, prompting ChatGPT to simplify the language while preserving the key ideas and analytical rigor. See Figure 13. This process involved providing feedback to the AI and revising its outputs. This approach allowed us to produce multiple versions of each exemplar, tailored to different reading levels. By scaffolding the exemplars in this way, we ensured that all students could engage with the material meaningfully, reinforcing their understanding of constructing a well-supported paragraph.

**Figure 12:** *Refining AI-Generated Exemplars for Accessibility & Skill Development*




*Note*. This figure illustrates the process of refining AI-generated exemplars to make them more accessible to students reading at different levels.

Our ongoing interaction with ChatGPT allowed us to refine the AI's outputs continuously. We provided specific prompts and feedback, asking the AI to emphasize particular skills or aspects of analytical writing, such as the clarity of thesis statements or the depth of evidence analysis. To do this effectively, we had to rely on our expertise in writing instruction—understanding not just what makes strong analytical writing but also how to scaffold it for different levels of learners. This iterative process ensured that the exemplars modeled high-quality writing and aligned closely with our instructional goals and learning objectives. The ability to quickly generate and adapt these exemplars freed up valuable time, again enabling us to focus more on how to actively engage students with the content and guide them through the process of analyzing and interpreting texts.

## Designing Multi-sensory Activities to Enhance Engagement

As our collaboration with ChatGPT deepened, we saw its potential for designing multisensory learning experiences. Multisensory activities are exceptionally vital for students in special education, as they help reinforce learning by engaging multiple senses, which enhances memory retention and understanding (Tomlinson, 2014). Despite recognizing the importance of these activities, we often struggled to find the time and resources to develop them fully. ChatGPT offered support, allowing us to explore creative and engaging ways to cultivate our students' learning through a variety of sensory modalities.

Initially, we used ChatGPT to brainstorm general ideas for multisensory activities. The AI provided a range of ideas, but some were too complex or required resources beyond our reach. However, as we iteratively refined our prompts and provided more specific information about our classroom context, ChatGPT's outputs became more targeted and practical, allowing us to develop high-quality multisensory activities.

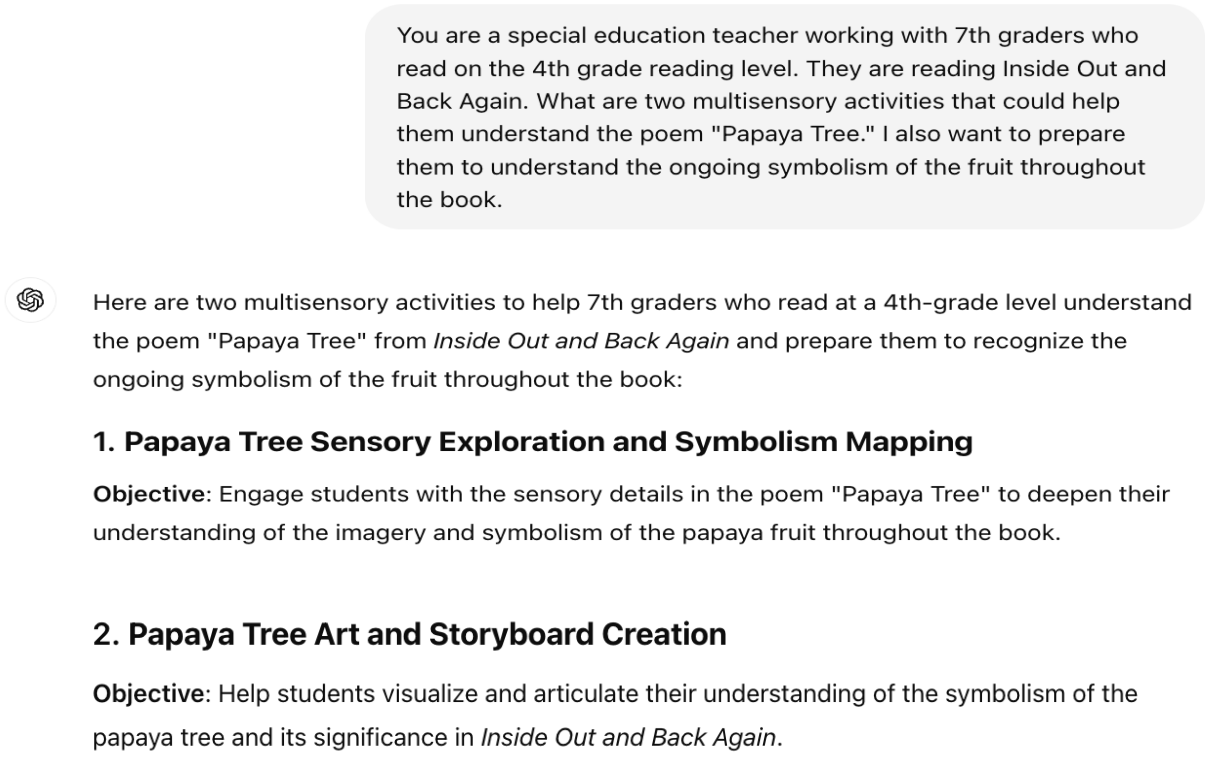
One successful example of this collaborative design process was our work with the book *Inside Out and Back Again* by Thanhha Lai. We had identified the papaya as a significant symbol throughout the text, recognizing that it represented more than just a fruit—it was a powerful symbol of home, memory, and change for the protagonist, Ha. Given that our students had little familiarity with papayas, their growth stages, or their cultural significance and knowing our students

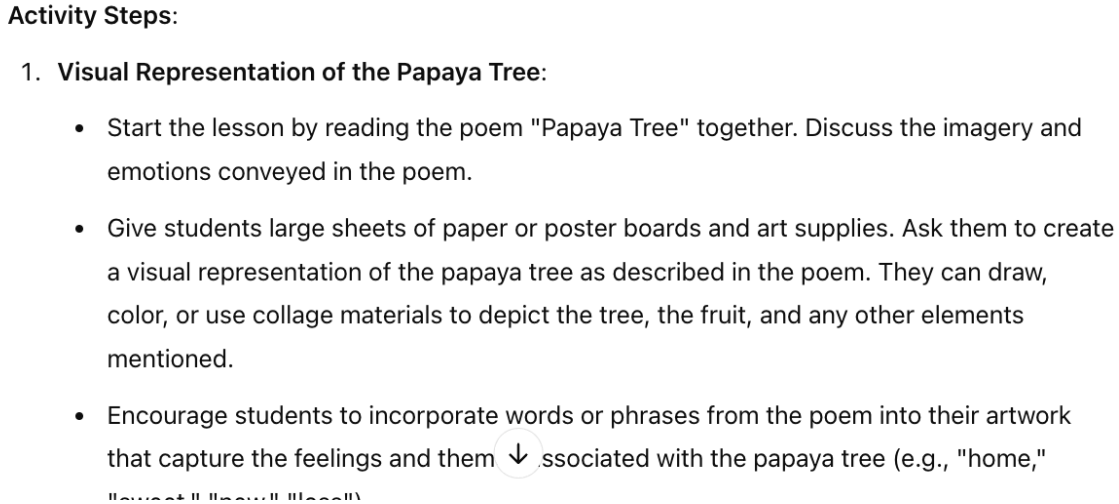
themselves, we knew that a lecture would not suffice to convey the depth of this symbolism. We needed to create a multisensory experience to engage students and help them build a richer understanding.

Ashley had already planned to bring fresh papayas and dried papaya slices into the classroom, allowing students to touch, smell, and taste the fruit. This direct engagement provided a tangible connection to the text, making the papaya more than just an abstract concept. Alongside this, ChatGPT recommended using images and videos showing the stages of papaya growth, from seedling to ripe fruit. These visual aids helped students visualize the fruit's development, connecting it to Ha's memories and experiences as described in the poem "Papaya Tree."

To support students' understanding of symbolism, we collaborated with ChatGPT to develop a multi-sensory activity centered around the papaya tree. ChatGPT suggested an art-based mapping activity, prompting students to create visual representations of the tree and incorporate their interpretations of its meaning as presented in the text. This activity engaged our learners, allowing them to express their understanding creatively while reinforcing their grasp of literary symbolism. We further refined this idea by leveraging ChatGPT's capabilities to generate discussion prompts. These prompts encouraged students to explore how the papaya might symbolize different concepts, such as hope, loss, or resilience, in various parts of the story.

**Figure 13:** *Enhancing Symbolism Understanding Through Art & Role-Playing Activities*






*Note*. This figure shows excerpts from a ChatGPT conversation, demonstrating how it was used to suggest multi-sensory activities that deepened students' understanding of the symbolism of the papaya in a text.

Building on this foundation, ChatGPT helped us develop an interactive storytelling and role-playing exercise. Students worked on storyboarding key scenes from the book where the papaya is mentioned, discussing its evolving symbolism. The AI provided suggestions for role-playing scenarios where students could act out these scenes, allowing them to embody the characters' emotions and experiences.

By collaborating with ChatGPT to design multisensory activities like these, we engaged our students far more deeply than through traditional methods. The sensory experiences helped students build background knowledge and facilitated a deeper comprehension. After the papaya activities, students were able to describe the fruit's significance in relation to Ha's personal journey, a notable improvement from earlier discussions where responses were often vague or limited to surface-level observations. When we revisited the poem "Papaya Tree" later in the unit, students demonstrated greater retention and textual recall.

This experience exemplifies a larger shift in our instructional approach, demonstrating how ChatGPT allowed us to expand our lesson planning in new, creative directions. The AI's ability to quickly generate diverse, multisensory activities tailored to our specific classroom context enabled us to plan more effectively and engage our students more fully.

## Key Teacher Moves for Effective AI Engagement

1. Engage in iterative dialogue
2. Provide clear and specific prompts.
3. Collaborate with AI as a co-designer.
4. Evaluate and revise AI outputs.

## Conclusion

This paper has reflected on our journey integrating generative AI, specifically ChatGPT, into instructional planning within a special education context. Throughout our experiences, we

identified several critical teacher moves essential for effectively leveraging AI tools, such as engaging in iterative dialogue, providing clear and specific prompts, and rigorously evaluating and refining AI outputs. These strategies are central to maximizing the pedagogical value of AI and align closely with the broader framework of Technological Pedagogical Content Knowledge (TPACK) (Mishra & Koehler, 2006). The TPACK framework emphasizes that effective engagement with generative AI depends on a solid foundation of teaching expertise across content knowledge, pedagogical knowledge, and technological knowledge. Key Teacher Moves aligned with the TPACK Framework are listed in Table 1.

**Table 1:** *Key Teacher Moves within the TPACK Framework*

| Key Teacher Moves within the TPACK Framework: | Description |
|---|---|
| **Engage in Iterative Dialogue** | Continuously interact with AI, refining prompts and responses to ensure outputs align with instructional goals. This process mirrors refining lesson plans in dialogue with a colleague, blending technological tools with pedagogical strategies to enhance learning. |
| **Provide Clear and Specific Prompts** | Develop prompts that clearly communicate instructional objectives to enhance the relevance and quality of AI-generated content. Understanding both the technological tool (AI) and the content area is essential to generate appropriate and accurate materials. |
| **Define the Output Structure** | Clearly specify the desired format and content structure of AI outputs to ensure alignment with educational standards and instructional goals. This requires knowledge of both the technological capabilities and the specific content requirements. |
| **Ask for More than Needed** | Request more AI-generated examples than necessary to provide a broader selection and reduce the need for multiple refinement rounds. This strategy leverages an understanding of AI's capabilities to maximize its utility. |
| **Evaluate and Revise AI Outputs** | Critically assess AI outputs for quality and appropriateness, making adjustments to align with learning objectives and pedagogical soundness. This move requires an integrated understanding of how technology can support pedagogical practices and content delivery. |

Within and beyond these Key Teacher Moves, our experience demonstrates that effective integration of AI in education is not merely about using technology but involves a nuanced application of teacher knowledge to guide AI's use. Understanding the curriculum and standards was crucial for prompting the AI accurately and ensuring that the generated materials aligned with our

instructional goals. Similarly, knowing our students' diverse needs allowed us to tailor AI outputs to make them accessible and meaningful.

Our reflection underscores the importance of recognizing AI's limitations, including potential biases in outputs, and the critical role of teacher expertise in mitigating these challenges. It also highlights the necessity for ongoing professional development to equip educators with the skills to effectively integrate AI tools into their teaching practice, ensuring they are used appropriately, ethically, and effectively to support diverse learners.

Ultimately, our exploration of generative AI in the classroom reveals a significant shift—from using AI as a tool for efficiency to engaging it as a non-human co-designer in creating richer, more meaningful learning experiences. This journey has shown us that AI's true potential in education is unlocked when teachers apply their expertise to thoughtfully guide and refine its outputs.

## References


Choi, G. W., Kim, S. H., Lee, D., & Moon, J. (2024). Utilizing generative AI for instructional design: Exploring strengths, weaknesses, opportunities, and threats. *TechTrends, 68*(832–844). https://doi.org/10.1007/s11528-024-00967-w

Dickey, E., Bejarano, A., & Garg, C. (2024). AI-Lab: A framework for introducing generative artificial intelligence tools in computer programming courses. *SN Computer Science, 5*, 720. https://doi.org/10.1007/s42979-024-03074-y

Dore, R. A., Amendum, S. J., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). Theory of mind: A hidden factor in reading comprehension? Educational Psychology Review, 30(3), 1067–1089. https://doi.org/10.1007/s10648-018-9443-9

DuFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement.* National Educational Service.

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children. *Journal of Autism and Developmental Disorders, 24*(2), 129-154. https://doi.org/10.1007/BF02172093

Individuals with Disabilities Education Act, 34 C.F.R. § 300.39(b)(3) (2006). *Code of Federal Regulations.* Retrieved from https://www.ecfr.gov/current/title-34/subtitle-B/chapter-III/part-300

Khazanchi, R., & Khazanchi, P. (2021). Artificial intelligence in education: A closer look into intelligent tutoring systems. In *Handbook of research on critical issues in special education for school rehabilitation practices* (pp. 256-277). IGI Global.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for integrating technology in teacher knowledge. *Teachers College Record, 108*(6), 1017-1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.

Vacca, R. T., Vacca, J. L., & Mraz, M. E. (2013). *Content area reading: Literacy and learning across the curriculum* (11th ed.). Pearson.