Can ChatGPT Help Pre-Service Teachers Analyze Classroom Discourse? Critical Reflections from a Science Methods Course

by Jacob Pleasants, University of Oklahoma

Abstract

Leading productive classroom conversations is an essential part of a science teacher's pedagogical practice, but also difficult to do well. Novice science teachers find it challenging to direct conversations that progress students' conceptual understanding while engaging them in scientific practices. They need to engage in deliberate, reflective practice to improve their discourse practices and sustain that deliberate work as they continue to develop their pedagogy. In my secondary science methods course, I work to build my students' reflective practice by having them analyze transcripts of science instruction. I have used several scaffolding tools and frameworks to support them in that work, but a novel option is using a language model such as ChatGPT to assist with the analysis. In this article, I describe how I incorporated ChatGPT into a sequence of discourse-focused learning activities. My students explored and critiqued the capabilities of ChatGPT as a discourse analyst, and here I share our collective appraisals and insights into how to use it most effectively. I also share the ways in which I saw my students' reflective and analytical practices develop over time.

Introduction

Leading and managing productive classroom conversations is an essential element of science pedagogy. Conversations are an important site for students to engage in scientific practices, which require both scientific ways of thinking and scientific ways of speaking and communicating (Lemke, 1990; National Research Council [NRC], 2007; Oliveira, 2010; Windschitl et al., 2012). Leading productive classroom conversations is regarded as a core part of "ambitious" teaching (Lampert et al., 2013; Windschitl et al., 2012) and a "high-leverage teaching practice" (Ball et al., 2009; Hlas & Hlas, 2012). Both descriptors indicate it to be a practice that novice teachers should prioritize because it is likely to significantly impact their overall effectiveness in the science classroom.

Because of its importance, various tools have been developed to assist novice teachers in leading science classroom conversations (e.g., Michaels & O'Connor, 2012; Windschitl et al., 2018; Wingert, 2016). Michaels and O'Connor (2015), for instance, developed a typology of "talk moves" that teachers can use as linguistic tools to pursue various goals for classroom conversations. At the same time, it is broadly recognized that managing productive discourse is challenging to do well and that developing those skills requires considerable time and

effort (Krist & Shim, 2023; Oliveira, 2010). Experience helps, but more important is "deliberate practice," in which teachers aim to improve their performance through feedback and self-reflection (Ericsson & Harwell, 2019; Schön, 1983). Ideally, deliberate practice would be established during pre-service teacher education and become an integral part of teachers' ongoing professional development.

Science methods courses play a vital role in terms of helping pre-service teachers (PSTs) begin the process of developing their ability to lead productive classroom discourse (Stroupe & Gotwals, 2018). Science methods instructors need to introduce novice teachers to guiding principles and practical tools, provide PSTs opportunities to put those ideas into practice, give meaningful feedback, and support PSTs in engaging in critical self-reflection (Davis et al., 2017; Forzani, 2014; Pecore et al., 2023; Stroupe et al., 2022). Methods courses are essential but are only ever a first step in the broader and ongoing process of teacher professional learning (Stroupe et al., 2022). Methods instructors thus need to give novice teachers the tools to engage in the ongoing process of deliberate practice so that they can continue to improve their classroom discourse (Bronkhorst et al., 2014).

There are many ways methods courses can give PSTs opportunities to engage in deliberate practice for leading classroom discussions (e.g., Pecore et al., 2023; Stroupe et al., 2022). One common approach is requiring PSTs to analyze recordings (audio or video) of instruction, either their own or of another teacher (Rich & Hannafin, 2008). This kind of analysis is a frequent task on performance assessments such as the PPAT and National Board Certification, which emphasize reflective practice (Jay et al., 2002). Recordings can be examined in various ways, but if the focus is on classroom discourse, then analyzing the *transcripts* of those recordings is especially informative. Transcripts make evident the many detailed decisions teachers make when leading classroom discussions. Using a transcript, PSTs can work slowly through the many highly contextual teacher actions that could easily be missed when experiencing or observing the conversation in real time.

I have long included some form of audio/video recording and transcript analysis in my science methods courses. I have consistently recommended to my PSTs that they make those kinds of analyses a regular and ongoing part of their professional practice. As much as I make this case, and as much as my students have generally found the analyses to be valuable, I am also aware that only some of my graduates engage in that kind of self-analysis unless they are required to do so. Close examination of one's own teaching practices is humbling, challenging, and time-consuming. I am, therefore, continuously in search of ways to reduce those barriers. Various technologies have been developed to facilitate analysis, such as annotation tools for video recordings (e.g., Rich & Hannafin, 2009). More recently, there has been a growing interest in the use of artificial intelligence and learning analytics to offload some of the analytical work (e.g., Chen, 2020). Those technologies are not yet readily or freely available to teachers, but an intriguing possibility is using publicly available large language machine learning models such as ChatGPT. Though not specifically designed to analyze discourse, the language processing capabilities of these

technologies are well-suited to the task (Chowdhary, 2020). Given the ability of language models to rapidly summarize and distill large volumes of text, I wondered: to what extent might ChatGPT help teachers analyze transcripts of classroom instruction?

In this paper, I describe how I incorporated ChatGPT (Version 3.5, OpenAI, 2023) into my secondary science methods course and the critical explorations I conducted with my students to examine its utility as a discourse analyst. By sharing this work, I aim to engage fellow science educators in conversations about ChatGPT and other machine learning technologies that extend beyond issues of plagiarism or whether they can write competent lesson plans or assessments (e.g., van den Berg & du Plessis, 2023). I am no optimist when it comes to technologies like ChatGPT, but nor am I pessimistic; I instead advocate for an inquisitive and skeptical approach that prioritizes critical inquiry into emerging technologies (Pleasants et al., 2023). That means not only evaluating the technical strengths and limitations of new technologies but also their potential to cause harm and undermine educational goals and values (Krutka et al., 2022; Williamson et al., 2023). Asking PSTs to evaluate as well as use ChatGPT helps them practice taking up a critical stance toward educational technologies. In addition, evaluating ChatGPT requires PSTs to clarify their thinking about what constitutes a quality analysis of classroom discourse. It also requires them to consider whether the kinds of talk valued by ChatGPT aligns with their own educational principles. In these ways, PSTs' evaluations of the technology can also support the development of their discourse analysis skills.

Analyzing Classroom Discourse in Science Methods

In this section, I will describe how I used ChatGPT in my methods course and the broader sequence of learning activities into which I incorporated it. In several places, I provide "boxes" with technical suggestions and practical advice for those interested in trying similar activities with their students.

Context

The sequence of learning activities I describe here occurred during a semester-long Advanced Science Methods course that is the second in a two-part methods sequence taken by all students pursuing a secondary science teaching license (both graduate and undergraduate). In the semester that I describe here, there were seven students in the course, several of whom were not seeking secondary licensure but rather had an interest in postsecondary science instruction (for simplicity, I refer to all students taking the course as PSTs). The course meets once per week for three hours, and students also complete an associated field experience as part of the course (30 hours over the semester, either in a local secondary science classroom or in a university science classroom). The Advanced Methods course addresses many aspects of science teaching, including laboratory instruction, assessment, and approaches to building coherent units of instruction. Underpinning the exploration of those various topics is the guiding principle that science

instruction should provide students the opportunity to “play the game of science” and inhabit scientific ways of thinking and acting (Pleasants, 2023). The course also includes practice-based elements (e.g., microteaching sessions) to provide PSTs opportunities to engage in science teaching and reflect on their emerging practice (Forzani, 2014; Matsumoto-Royo & Ramírez-Montoya, 2021).

Learning Activities Related to Discourse

Discourse in the science classroom is woven throughout the methods course rather than as a standalone topic, reflecting the essential role it occupies in the “game of science.” Early in the course, I model several science investigations that utilize whole class and small group conversations to support students’ engagement in a wide range of scientific practices, from making observations to interpreting data to developing investigations and crafting explanations. In subsequent class sessions, I explicitly address the critical role of the teacher in managing those conversations and return to this theme often. Table 1 provides an overview of the class learning activities most relevant to the theme of classroom discourse. During the initial class sessions, PSTs engage in three different science investigations as learners, which I audio record and transcribe for later analysis (see Box A for technical considerations of creating transcripts). Each investigation features productive classroom discussions, and PSTs often recognize the importance of those discussions when reflecting on the activities. Still, discourse is not deeply addressed until the fifth class session.

BOX A: Technical Notes on Transcript Creation

Transcribing classroom discourse can be prohibitively labor-intensive and time-consuming, especially for a classroom teacher. The technical approach that I have found most useful is using the automatic transcript function within Zoom. I use a tablet or a laptop to open a Zoom room that I then record to capture a class session. I then use the Zoom-generated transcript of that recording as a starting point. The Zoom transcript requires cleaning and formatting, which can still be time-consuming. It begins as a continuous “wall” of text and so line breaks need to be made for speakers, and the speakers need to be labeled. If there is crosstalk or if a speaker is not very audible, the transcript tends to become very messy. However, for whole-class settings in which there is generally only one speaker talking at a time, the machine-generated transcript is usually quite serviceable. For the purposes of my methods class, the transcript simply needs to be intelligible rather than “pristine.” The Appendix gives an example of the kind of transcript that I create for the class.

Table 1

Overview of the Learning Activities Used to Address Classroom Discourse

Class Session	Relevant Class Activities	Recordings and Assignments
1	Tricky Tubes Activity (Modeled by Instructor)	Used to Create Recording #1
3	Electrostatics Investigation (Modeled by Instructor)	Used to Create Recording #2
4	Rate of Reaction Lab (Modeled by Instructor)	Used to Create Recording #3
5	Introduction to Discourse Analysis *Examination of Recording #1*	
6	Making Sense of Discourse, Continued *Examinations of Recordings #2 & #3*	Discourse Analysis Task Assigned (Introduces ChatGPT as analyst)
7	Debrief of ChatGPT's analysis performance	
8-10	Microteaching Sessions (10-15 minutes per student)	Video recordings and transcript sent to students for self-analysis and exploration using ChatGPT
11	Debrief of ChatGPT's utility in doing self-analysis	
13-14	Microteaching Sessions Round 2 (15-20 minutes per student)	Video recordings and transcript sent to students for self-analysis and further ChatGPT explorations
15	Final debriefing of ChatGPT's utility	

During the fifth class, I work with PSTs to describe why classroom conversations are an important part of science instruction and identify the features that make conversations *productive*. PSTs typically focus on ensuring that conversations engage students in scientific practices and that they help students make progress in their understanding of science concepts. I then turn PSTs' attention to what the teacher must do to make conversations productive. I begin with a contrasting case of less-than-productive practice by sharing examples of the Initiation-Response-Evaluation (IRE) pattern of classroom discourse (Lemke, 1990). We discuss why this familiar pattern of interaction is inconsistent with the aims of science instruction, although it might have some limited utility. To identify and describe more productive approaches, we then begin to examine transcripts from the initial class sessions. To scaffold PSTs' inquiries into the transcripts, I give them a table to complete in small groups that asks them to identify moments in the transcript that they found particularly productive and unproductive in promoting the teacher's goals. They describe what the teacher was trying to do in that part of the transcript, why they found it (un)successful, and what they would change (if needed). The basic template is included in the Appendix, and Box B provides further details about this activity.

BOX B: Considerations for Analyzing Transcripts in Science Methods

For an initial foray into transcript analysis, I have found that giving PSTs a transcript of a lesson that *they experienced as learners* is a useful place to start for several reasons. First, they are already oriented to what the activity is about, its goals, and the science content being addressed. They know where the activity is headed and why. Second, they can use their experiences as learners to explore the impact of various teacher decisions, in terms of both their conceptual thinking and emotional experience (e.g., "Was I nervous to answer that question? Was I hesitant or excited to contribute my ideas?"). Third, it allows me to model the self-analytical work that I expect PSTs to do. By putting my own teaching practices on display for scrutiny, I show students how to embrace the vulnerability that comes when engaging in self-study.

PSTs might be reluctant to critique the practices of their methods instructor. To mitigate this, I usually begin by pointing out several of the missteps that I can see in the transcript (e.g., poorly worded questions, missed opportunities to get students talking with each other). I generally find that my students quickly abandon any hesitance and eagerly offer up their critiques. Importantly, the transcripts I use are not for the full activity but rather of selected portions during which there were rich and interactive classroom conversations, typically representing 10-15 minutes of instructional time. I typically select portions that represent a variety of instructional goals that might be pursued during classroom conversations, such as: eliciting students' initial ideas, sharing and identifying interesting observations, interpreting data, and working with newly introduced science ideas.

Supporting Students' Analysis of the Transcripts

My PSTs generally need extensive in-class support during their initial analyses of the transcripts. What they often find most challenging is describing the teacher's actions in precise, concrete terms. Like many teachers who are new to this kind of analysis, they often perceive the gist of what the teacher is trying to do (e.g., have the group come to consensus about key observations) but have more difficulty explaining *how the teacher is doing it* (Tekkumru Kisa & Stein, 2015). I thus spend considerable time working with PSTs in small groups to describe how the teacher phrases questions that get students to engage in the targeted mental work, how the teacher responds to students in ways that continue to push students' thinking, and how the teacher introduces new ideas and lines of inquiry. To give PSTs the language to describe those "talk moves," I provide them with the typologies created by Michaels and O'Connor (2012) as well as Windshitl et al. (2018). I also help PSTs examine the transcripts for evidence that students are, in fact, doing what the teacher intended them to do (e.g., to what extent were the students using evidence to make claims about the phenomenon?).

After the sixth class session, I give PSTs an analysis assignment to complete on their own. The assignment presents them with a transcript from one of the same activities they analyzed but with a different group of students. PSTs then work through the same task of identifying productive and unproductive moments and analyzing the teachers' actions. This is when I introduce the possibility of using ChatGPT to facilitate their analysis. In the

assignment materials, I provide a record of a chat session where I fed ChatGPT the transcript used in the assignment and then asked it a series of questions to get it to analyze the discourse. I invite PSTs to use the chat record in any way they wish, and I also invite them to do their own chat sessions if they think it would be helpful (further details and considerations are provided in Box C). Regardless of how they use it, the assignment asks PSTs to critically evaluate the analysis that ChatGPT produced, and we discuss those evaluations during the following class session.

BOX C: Introducing and Supporting PSTs' Use of ChatGPT

I introduce students to the use of ChatGPT by providing them my own set of prompts and output. Those prompts illustrate that asking ChatGPT follow-up questions is necessary to get any sort of useful analysis. My first prompt yields a very broad-level summary of the transcript, and only through the follow-ups am I able to get ChatGPT to link its evaluations of the teacher's actions to specific turns of talk. For those who have not had much experience using ChatGPT, these methods of writing prompts are not obvious, so the model is helpful.

Although I give PSTs records of my own interactions, during the first assignment I also invite them to do their own chat sessions if they would like to explore other aspects of the transcript. I would not require anyone to make a ChatGPT account, which requires giving personal information to OpenAI, so I have created a shared account that my students can use for the purposes of the course. For later course tasks, I ask PSTs to do their own analyses using ChatGPT and I do not provide them with my own prompts. A benefit of the shared account is that they can see prior chat sessions to serve as models to craft their own prompts and follow-ups.

Further Applications

The final set of discourse-focused class activities is a series of "microteaching" sessions (Arsal, 2014; Griffiths, 2016) that occur midway through the course and again near the end of the course. Each PST is tasked with leading a short portion of a science lesson that aligns with their subject area and grade level specialization. The rest of the class (instructor included) plays the role of students in those lessons. After each microteaching session, I lead the class in a reflective conversation about the learning experience to provide feedback to the presenter and work through any unexpected issues that arose during the session. I then share with each presenter a video recording of their lesson, along with a transcript, which they use to conduct a self-analysis of their discourse (structured similarly to the prior analysis tasks). I also ask each PST to use ChatGPT to analyze their transcript. They are free to decide how to use its output when writing up their self-analysis. Whatever their choice, I ask them to explain how they used it and critically evaluate the analysis it produces.

Instructional Outcomes and Reflections

The Development of PSTs' Analytical Practice Over Time

Table 2 illustrates the development of PSTs' analytical practice over several weeks of the course. The entries in the table were made by various students and are representative of the kinds of work that most students in the class were doing. As Table 2 shows, their analyses became more precise over time. In their initial analyses, PSTs tended to provide general descriptions and discussions of what the teacher was doing that made the discussion productive. As their analytical practice developed, they began noticing more specific talk moves and could better articulate how and why those discursive actions moved the conversation in productive directions.

To be clear, many PSTs still had difficulty with this kind of analysis even after several weeks. When conducting self-analyses of their microteaching, they sometimes focused on structural choices in their lesson (e.g., the visuals they used or the organization of the activity) rather than their discourse. They continued to wrestle with articulating in concrete terms how their actions moved or did not move the conversation in the direction they intended. They could identify what they were trying to do in the conversation and describe what they were doing in broad terms (e.g., asking students thought-provoking questions, asking students to share their reasoning) but did not always tie those claims to specific talk moves in the transcript. Specificity is essential because as much as the PSTs might have *wanted* to get students to engage in critical thinking or scientific reasoning, the extent to which that *actually* occurred depends on what exactly was (or was not) said by the teacher and the students. The need for greater specificity was a recurring point that I raised during class and in my feedback on my students' submitted work. By their second round of self-analysis, most managed to do this reasonably well. As described in the following section, we found that ChatGPT was not necessarily very assistive with this particular issue.

Table 2

Representative Examples of PSTs' Discourse Analyses Over the Semester

What the teacher was trying to do	How the teacher did it (Be as concrete as possible)	Why it was productive
Tubes Transcript (Class Session 5)		
Provide evidence for claims	Literally asked for evidence/reasoning to be provided	Making us think about our claims and being able to confidently support or disprove an idea
Trying to get students to articulate the higher-order process they are using to rule out an option for how the tube strings were connected	Asking students to describe the logic underpinning eliminating an option	Use current knowledge to make an observation Make predictions Provide counter arguments
Analysis of Electrostatics on the Individual Assignment (After Class Session 6)		
Pushes students to engage with their peers' ideas.	**Teacher**: Okay, so our other groups… I'd like you to ask questions of the group that just presented to clarify things they said or ask them questions about their thinking. What questions would you like to pose?	This gives several messages that are important for a productive conversation, particularly in the science context. Most prominently, that students' ideas are taken seriously in themselves, and not just when validated by the teacher.
Direct students to think about their models and how they can be described or amended	Asked students to clarify their diagrams and pointed out an important part of the student models (i.e. follow-ups, focusing)	Set up a basis for them to discuss the concept of charge and how you could end up with a +, -, or neutral charge
Microteaching Self-Analysis Round 1 (Class Sessions 8-10)		
Fostering critical thinking	Asked thought-provoking questions and provided a demonstration to students to encourage them to consider what happens during electrolysis. Also asked students to provide evidence for their conclusions.	Students were able to visualize the process being taught and draw their own conclusions about how electrolysis compares to boiling. By providing evidence, students are able to adequately demonstrate their understanding of the concept.
Get students to develop and defend hypotheses based on their observations	Students were asked to justify or explain their hypotheses using evidence from their observations: - "How can you tell?" - "Why is it hard?" - "Say more"	Students have to convey their thoughts and practice defending their views using evidence
Microteaching Self-Analysis Round 2 (Class Sessions 13-14)		
Asked probing questions to better assess students' understanding	**Teacher:** So why again does there have to be some K [Kinetic Energy]? (**Students** describe why there has to be some based on the energy pie charts they have drawn) **Teacher:** But in terms of some more physical quantity, so what does it mean that there's more K, more K than when?	The questions generated students' descriptions of their thinking process (using the pie chart consistently) and of their understanding of the meaning of the elements of the representation, in this case kinetic energy and later potential energy.

Reflections on ChatGPT

My students' first encounter with ChatGPT was on their individual analysis assignment (after Session 6). Most reported that they reviewed the ChatGPT analysis after they had conducted their own, though a few said that they consulted it beforehand. They all found the analysis rather impressive and useful in certain respects, though limited in others. Students agreed that it was most useful as an "idea generator" in that it identified interesting moments in the transcript. Its analyses, however, were often too generic to be of much help. For instance, it identified a part in the transcript where the teacher "promoted peer interaction" by "creating a dynamic environment" where students could "compare and discuss their observations." The transcript certainly showed peer interactions, but the analysis did not describe any specific actions the teacher took to make that happen. Overall, the ChatGPT analyses lacked the sort of concrete details and specificity that I had been working with PSTs to notice.

My students were disappointed by the "suggestions for improvement" that ChatGPT generated. They wrestled with how to improve the lesson, and they thus hoped that ChatGPT would provide some useful suggestions. ChatGPT did generate ideas that were unobjectionable at first glance, such as to "provide clearer instructions" and "use more formative assessment." My students recognized, however, that those suggestions did not seem to resolve any specific issues in the lesson. There were not any moments in the transcript that seemed to indicate unclear instructions, and the conversation as a whole provided numerous opportunities for the teacher to gauge students' thinking, obviating the need for the "quick quizzes or polls" that ChatGPT suggested as formative assessments.

Based on this conversation, most of my students were intrigued but also somewhat wary of the utility of ChatGPT moving forward. One point that we established was that more carefully crafted prompts and follow-ups could resolve some of the issues. Based on our conversation, we started the "user guide" shown in Table 3, which we updated after our subsequent explorations and conversations.

Table 3

User Guide for Transcript Analysis with ChatGPT

Known Limitations	Ways to Mitigate	Lingering Issues
ChatGPT will assume and impose generic goals of the conversation/lesson that might not be appropriate.	In the prompt, specify what the teacher was trying to achieve before asking ChatGPT to analyze how the teacher was/wasn't successful.	The way you conceptualize a specific goal (e.g., promote critical thinking) may not align with how ChatGPT uses that term. You might need to clarify.
ChatGPT will typically begin with broad characterizations of a transcript rather than concrete teacher actions.	Ask ChatGPT to refer to specific parts of the transcript that illustrate its analyses. Alternatively, just ask it to identify "important" or "interesting" parts of the transcript.	It is pretty good at identifying parts of the transcript that are worth examining. Getting it to analyze those parts at the level of specific talk moves, though, is very difficult.
ChatGPT can identify interesting parts of the transcript, but its interpretations rarely address specific "talk moves" used by the teacher.	Prompt ChatGPT to explain how the specific words used by the teacher led to the results it has identified (e.g., what the teacher said that invited student contributions).	Even with this prompting, it may not generate anything specific, and sometimes the specific turns of talk that it identifies don't align with its analysis; further prompting does not help.
ChatGPT cannot infer what's going on in the parts of the activity that are not evident in the transcript.	Before asking for an analysis of the transcript, give an overview of the activity and the context.	ChatGPT might still make assumptions about what was going on that are not appropriate.
The syntax of the transcript can cause ChatGPT to confuse the teacher's words for a student's.	Structure the transcript as clearly as possible; use line breaks to separate speakers.	Even when unambiguously structured, ChatGPT can still make errors for reasons that are unclear.

User guide in hand, my students' next opportunity to use and evaluate ChatGPT was during their analyses of the transcripts from their microteaching sessions. As I debriefed those experiences with them and gathered their critical reflections, several points of consensus emerged. The main point of agreement was that ChatGPT was most useful for identifying interesting parts of the transcript, but that its analyses of those sections were less insightful, even with additional prompting. When working with a relatively long transcript, which was the case for several students' second microteaching session, ChatGPT could quickly point out genuinely productive moments as well as ones that did not go as well. A few of my students found that it identified parts of the transcript that they might have otherwise missed had they done the analysis wholly on their own. The analyses that ChatGPT provided for those moments, however, were consistently weak.

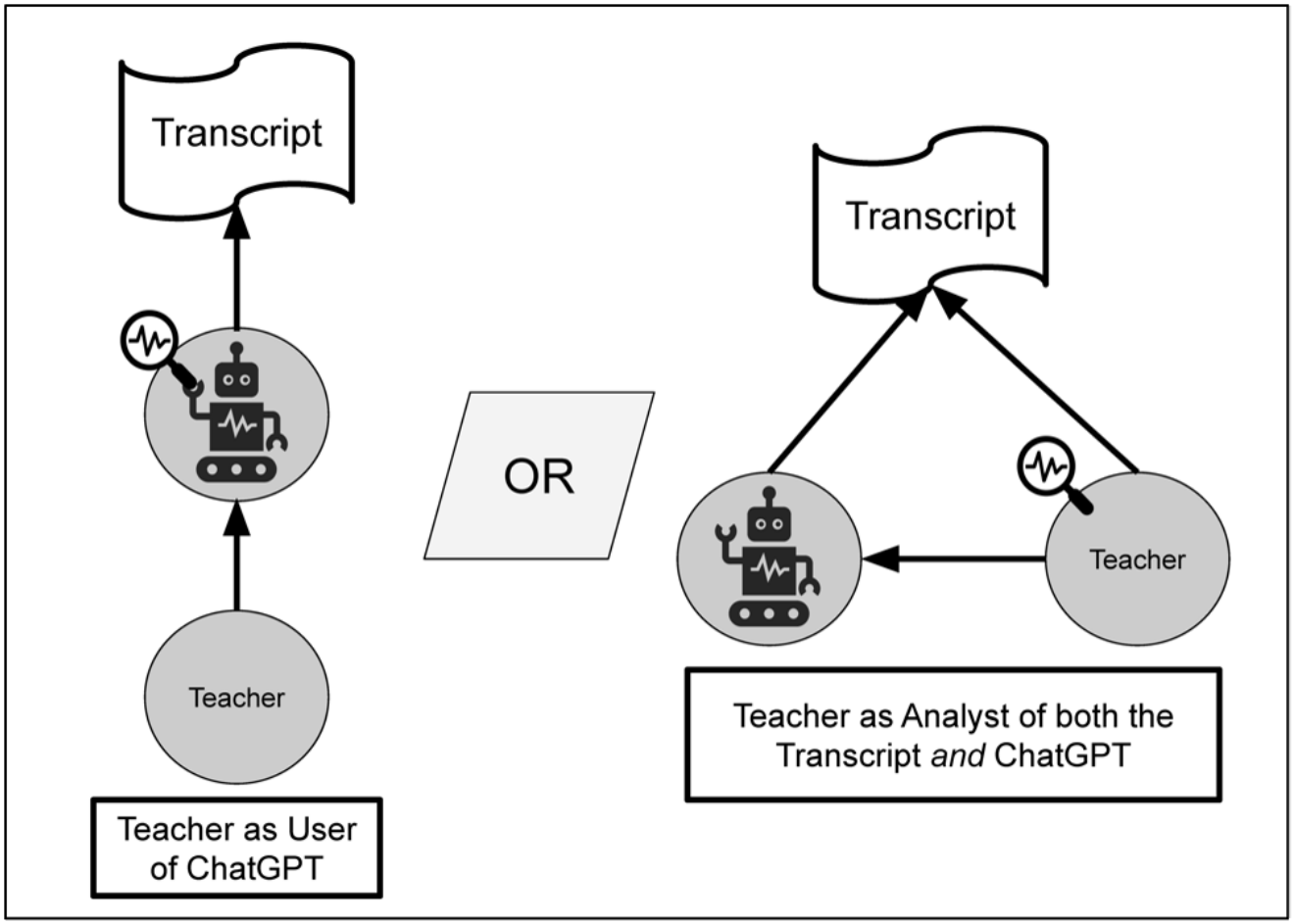
As noted above, one of the aspects of discourse analysis that proved difficult for my students was the need for specificity and concrete connections between talk moves and conversational goals. This, it turns out, is also an area with which ChatGPT struggles mightily, even if the goals for the lessons are clearly specified. When prompted to describe what exactly the teacher was doing well, ChatGPT often used vague and generic language, such as that the teacher was "promoting critical thinking" or "student collaboration." When my students tried to use follow-up prompts to get its analyses to be more concrete and specific, they often found that it became increasingly "off target and confused." The analyses seemed to be consistently superficial and follow-up prompts could not resolve the issue.

In terms of providing suggestions for improvement, my students found ChatGPT's output to be of mixed quality. When given clearly specified the goals of the conversation, it could occasionally provide reasonable and useful improvements. One of my students only used ChatGPT after conducting her own self-analysis and was surprised when it made the very same suggestion for improvement that she came up with. However, my students also noticed that ChatGPT seemed to overestimate the value and effectiveness of dubious pedagogical approaches. Most notably, it seemed to place substantial value on lecture-style teaching and would often suggest that the teacher deliver detailed explanations of scientific ideas to students. That preference for teacher-driven discourse is a consequence of the training data used to develop ChatGPT. OpenAI has not disclosed the precise makeup of the training data, but it is generally believed to include large swaths of freely available online content (Vincent, 2023). Presumably, that would include many science lesson plans, articles on science pedagogy, and similar digital items. It is not terribly surprising that those training items would foster a preference teacher-driven talk, but it is important to recognize all the same.

The reflective conversations that my students and I had about ChatGPT served multiple educational goals. To a limited extent, they supported my students in being savvy users of ChatGPT, as illustrated in the left side of Figure 1. More valuable, though, was the way that these conversations positioned my students as analysts of both the transcripts and of ChatGPT itself, as shown in the right side of Figure 1. As my students went through multiple rounds of using ChatGPT and critically evaluating its output, the depth of their critiques steadily increased. Their observations about ChatGPT's preference for certain forms of pedagogy, for instance, emerged only during conversations later in the semester. As they deepened their analyses of ChatGPT, my students necessarily had to add depth and clarity to their own analyses of classroom discourse. As they compared their own thoughts to the suggestions offered by the technology, they were forced to consider what constitutes a valid argument regarding which parts of a classroom conversation were successful and which needed to change. I cannot say with certainty to what extent critically examining ChatGPT contributed to my students' development as discourse analysts (see Table 2). Yet the quality and depth of our conversations strongly suggest that it played a valuable role in their learning.

Figure 1

Positioning of PST in Relation to ChatGPT



Implications for Future Practice

Based on their explorations, my students overall found that ChatGPT has some worthwhile capabilities in terms of analyzing classroom discourse, but it needs to be approached with care. For those who have followed conversations around ChatGPT and other generative AI technologies, this advice likely sounds familiar (Han et al., 2023; Tamkin et al., 2021; Thorbecke, 2023). As impressive as ChatGPT can seem at times, it also has significant limitations. The prompt-crafting strategies shown in Table 3 mitigate certain issues and make ChatGPT more assistive, but limitations remain. ChatGPT can provide starting points, but a deep examination of a transcript requires the work of a knowledgeable human. Of course, our findings may just reflect the current state of the technology or even the specific version of ChatGPT that we used (the free 3.5 version rather than the 4.0 version, which requires a paid subscription). Without a doubt, ChatGPT and other language model technologies will continue to be developed, and versions will likely be created customized to the educational context. Specialized systems could be designed to not require such carefully structured prompts and follow-ups. In addition, the preference that ChatGPT showed for lecture-based, traditional forms of instruction could potentially be resolved by using more carefully curated training data.

With those possible future improvements in mind, can ChatGPT (or a similar technology) be the labor-saving analyst that helps future teachers commit to sustaining a deliberate, reflective practice? Its ability to rapidly sift through lengthy transcripts and provide reasonable starting points for self-analysis and reflection can potentially save time and energy. However, while it operates extremely quickly once given a transcript and prompts, there is considerable labor that is required to prepare and format a transcript before it can be given to ChatGPT. Notably, preparing a transcript is often far more tedious and less insightful than the "saved" labor of analyzing the transcript. A time-strapped teacher might be better served by reflecting on a messy transcript than taking the time to clean it and feed it into ChatGPT. Rather than save labor, then, ChatGPT might simply *displace* it in undesirable ways – an unintended but also not necessarily atypical outcome of digital technologies (Jarrett, 2022). Future development might bring improvements to language model technologies, but a technological "solution" is not likely to be right around the corner (Morozov, 2013).

Nevertheless, using ChatGPT in a methods course can serve several valuable teacher education goals. First, it provides an opportunity for PSTs to critically evaluate an emerging technology. In addition to a practice of self-reflection when it comes to discourse, I also want PSTs to develop a practice of critical inquiry when it comes to education technologies (Krutka et al., 2022; Pleasants et al., 2023). Whether or not they ultimately choose to use ChatGPT in the future, the experience of rigorously evaluating it is worthwhile in developing that critical practice. When the next iteration of the technology arrives, they will be in a better position to make informed choices about it. Second, evaluating ChatGPT's analyses forces PSTs to think critically about what an analysis ought to do and the qualities that make it informative. When my students shared their evaluations during class, the ensuing conversations were generative and helped them think through their own analyses more deeply. In this respect, the limitations of ChatGPT are pedagogically useful. If it could truly do all the analytical work *for* my students, opportunities for critical reflection would be lost.

One open question is when PSTs should use and critique ChatGPT as they conduct their analyses of transcripts. My students were free to decide whether they wanted to use it right away to provide starting points for their own reflections, or do their own analyses first and then compare what ChatGPT generated, or take some other approach. I left that choice open because I was genuinely curious to know what approaches my students found to be most useful. From a teacher education standpoint, there are potential advantages to having students follow a specific approach in the methods context. Having students do their own analysis first, for instance, allows them to have a concrete point of comparison when they appraise the analysis that ChatGPT provides. I hope that as other teacher educators incorporate ChatGPT and similar technologies into their methods course, we can continue to iterate and explore the potential value of different possibilities.

References

Arsal, Z. (2014). Microteaching and pre-service teachers' sense of self-efficacy in teaching. *European Journal of Teacher Education, 37*(4), 453–464. https://doi.org/10.1080/02619768.2014.912627

Ball, D. L., Sleep, L., Boerst, T. A., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *The Elementary School Journal*, *109*(5), 458–474. https://doi.org/10.1086/596996

Bronkhorst, L. H., Meijer, P. C., Koster, B., & Vermunt, J. D. (2014). Deliberate practice in teacher education. *European Journal of Teacher Education*, *37*(1), 18-34. https://doi.org/10.1080/02619768.2013.825242

Chen, G., Chan, C. K., Chan, K. K., Clarke, S. N., & Resnick, L. B. (2020). Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. *Journal of the Learning Sciences*, *29*(4-5), 642-680. https://doi.org/10.1080/10508406.2020.1783269

Chen, G. (2020). A visual learning analytics (VLA) approach to video-based teacher professional development: Impact on teachers' beliefs, self-efficacy, and classroom talk practice. *Computers & Education*, *144*, 103670. https://doi.org/10.1016/j.compedu.2019.103670

Chowdhary, K. (2020). Natural language processing. In K. Chowdhary, *Fundamentals of artificial intelligence* (pp. 603-649). Springer. https://doi.org/10.1007/978-81-322-3972-7

Davis, E. A., Kloser, M., Wells, A., Windschitl, M., Carlson, J., & Marino, J.-C. (2017). Teaching the practice of leading sensemaking discussions in science: Science teacher educators using rehearsals. *Journal of Science Teacher Education, 28*(3), 275–293. https://doi.org/10.1080/1046560X.2017.1302729

Ericsson, K. A., & Harwell, K. W. (2019). Deliberate practice and proposed limits on the effects of practice on the acquisition of expert performance: Why the original definition matters and recommendations for future research. *Frontiers in Psychology*, *10*, 2396. https://doi.org/10.3389/fpsyg.2019.02396

Forzani, F. M. (2014). Understanding "core practices" and "practice-based" teacher education: Learning from the past. *Journal of Teacher Education, 65*(4), 357–368. https://doi.org/10.1177/0022487114533800

Griffiths, J. (2016). Bridging the school placement gap with peer micro-teaching lesson study. *International Journal for Lesson and Learning Studies, 5*(3), 227–238. https://doi.org/10.1108/IJLLS-11-2015-0035

Han, Z., Battaglia, F., Udaiyar, A., Fooks, A., & Terlecky, S. R. (2023). An explorative assessment of ChatGPT as an aid in medical education: Use it with caution. *Medical Teacher* Preprint. https://doi.org/10.1080/0142159x.2023.2271159

Hlas, A. C., & Hlas, C. S. (2012). A review of High-Leverage Teaching Practices: Making connections between mathematics and foreign languages. *Foreign Language Annals*, *45*(1), 76–97. https://doi.org/10.1111/j.1944-9720.2012.01180.x

Jarrett, K. (2022). *Digital labor.* Polity Books.

Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education*, *18*(1), 73–85. https://doi.org/10.1016/S0742-051X(01)00051-8

Klassen, R. M., Rushby, J. V., Maxwell, L., Durksen, T. L., Sheridan, L., & Bardach, L. (2021). The development and testing of an online scenario-based learning activity to prepare preservice teachers for teaching placements. *Teaching and Teacher Education, 104*, 103385. https://doi.org/10.1016/j.tate.2021.103385

Krist, C. & Shim, S. (2023). Which ideas, when, and why? An experienced teacher's in-the-moment pedagogical reasoning about facilitating student sense-making discussions. *Journal of Research in Science Teaching*. Online First Article. https://doi.org/10.1002/tea.21908

Krutka, D. G., Heath, M. K., & Smits, R. M. (2022). Toward a civics of technology. *Journal of Technology and Teacher Education*, *30*(2), 229-237. https://www.learntechlib.org/p/221182/

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., Cunard, A., & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226–243. https://doi.org/10.1177/0022487112473837

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Ablex.

Matsumoto-Royo, K., & Ramírez-Montoya, M. S. (2021). Core practices in practice-based teacher education: A systematic literature review of its teaching and assessment process. *Studies in Educational Evaluation*, *70*, 101047. https://doi.org/10.1016/j.stueduc.2021.101047

Michaels, S., & O'Connor, C. (2012). *Talk science primer.* TERC.

Michaels, S., & O'Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussions. In L. B. Resnick, C. S. C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through academic talk and dialogue* (pp. 347–361). American Educational Research Association. https://doi.org/10.3102/978-0-935302-43-1_27

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism*. Public Affairs.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8.* R. A. Duschl, H. A. Schweingruber, & A. W. Shouse (Eds.). The National Academies Press.

Oliveira, A. W. (2010). Improving teacher questioning in science inquiry discussions through professional development. *Journal of Research in Science Teaching, 47*(4), 422–435. https://doi.org/10.1002/tea.20345

OpenAI (2023). ChatGPT 3.5 [computer software]. Retrieved from https://chat.openai.com/chat

Pecore, J. L., Nagle, C., Welty, T., Kim, M., & Demetrikopoulos, M. (2023). Science teacher candidates' questioning and discussion skill performance in a virtual simulation using experiential deliberate practice. *Journal of Science Teacher Education*, *34*(4), 415–435. https://doi.org/10.1080/1046560X.2022.2111775

Pleasants, J. (2023). Embracing the game-like character of science and engineering activities: A perspective to guide teachers' instructional decisions. *Journal of Science Teacher Education, 34*(7), 751-769. https://doi.org/10.1080/1046560X.2022.2148857

Pleasants, J., Krutka, D. G., & Nichols, T. P. (2023). What relationships do we want with technology? Toward technoskepticism in school. *Harvard Educational Review, 93*(4), 486-515. https://doi.org/10.17763/1943-5045-93.4.486

Rich, P. J., & Hannafin, M. J. (2008). Decisions and reasons: Examining preservice teacher decision-making through video self-analysis. *Journal of Computing in Higher Education*, *20*(1), 62–94. https://doi.org/10.1007/BF03033432

Rich, P. J., & Hannafin, M. (2009). Video annotation tools: Technologies to scaffold, structure, and transform teacher reflection. *Journal of Teacher Education*, *60*(1), 52-67. https://doi.org/10.1177/0022487108328486

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action.* Basic Books.

Stroupe, D., & Gotwals, A. W. (2018). "It's 1000 degrees in here when I teach": Providing preservice teachers with an extended opportunity to approximate ambitious instruction. *Journal of Teacher Education*, *69*(3), 294-306. https://doi.org/10.1177/0022487117709742

Stroupe, D., Gotwals, A., Christensen, J., & Wray, K. A. (2022). Becoming ambitious: How a practice-based methods course and "macroteaching" shaped beginning teachers' critical pedagogical discourses. *Journal of Science Teacher Education*, *33*(6), 683–702.

https://doi.org/10.1080/1046560X.2021.1988037

Tamkin, A., Brundage, M., Clark, J., & Ganguli, D. (2021). Understanding the capabilities, limitations, and societal impact of large language models. *ArXiv* preprint. https://doi.org/10.48550/arXiv.2102.02503

Tekkumru Kisa, M., & Stein, M. K. (2015). Learning to see teaching in new ways: A foundation for maintaining cognitive demand. *American Educational Research Journal*, *52*(1), 105-136. https://doi.org/10.3102/0002831214549452

Thorbecke, C. (2023, August). AI tools make things up a lot, and that's a huge problem. *CNN Business.* https://www.cnn.com/2023/08/29/tech/ai-chatbot-hallucinations/index.html

van den Berg, G., & du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences*, *13*(10), 998. https://doi.org/10.3390/educsci13100998

Vincent, J. (2023, March 15). OpenAI co-founder on company's past approach to openly sharing research: 'We were wrong.' *The Verge.* https://www.theverge.com/2023/3/15/23640180/openai-gpt-4-launch-closed-research-ilya-sutskever-interview

Williamson, B., Macgilchrist, F., & Potter, J. (2023). Re-examining AI, automation and datafication in education. *Learning, Media and Technology*, *48*(1), 1-5. https://doi.org/10.1080/17439884.2023.2167830

Windschitl, M., Thompson, J., Braaten, M. (2018). *Ambitious science teaching.* Harvard Education Press.

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, *96*(5), 878-903. https://doi.org/10.1002/sce.21027

Wingert, K. (2016). Student talk flowchart and protocols. https://stemteachingtools.org/assets/landscapes/FullSet_StudentTalkProtocolsandFlowchart1_c.pdf.