Early Adopters: Navigating AI Integration in Special Education Teacher Preparation

AUTHORS
Tara Kaczorowski
Angela Stockman
Andrew Hashey
John Kaczorowski

Journal of Special Education Preparation
4(2), 18-29

DOI: https://doi.org/10.33043/9ca46254

ABSTRACT
Exploring the role of artificial intelligence (AI) in education is important as it challenges traditional teaching practices and shapes how educators may approach them in the future. In this article, we document the transformative integration of AI in special education teacher preparation, highlighting how we, as early adopter professors, attempted to navigate this journey, offering practical applications for AI use. Practical applications of generative AI tools include aligning course objectives, developing modules, and creating assignments and assessment measures. Additionally, we describe innovative uses for AI, such as incorporating chatbots in teacher preparation courses, navigating curriculum development, generating case studies, and aligning individualized education program (IEP) goals with curricular standards. We also explore how AI can be employed as a reflective coaching tool for teaching practice. Ethical considerations are emphasized, focusing on transparent communication about AI use and documenting the learning process to humanize assessment experiences and mitigate potential risks.

KEYWORDS
Artificial intelligence, chatbots, curriculum development, special education, teacher preparation, video-based reflection

As artificial intelligence (AI) becomes more accessible to the public, educational institutions worldwide are considering its transformative effects on teaching and learning. This evolving landscape elicits diverse perspectives, ranging from concerns about academic integrity to opportunities for redefining assessment practices. In this article, we advocate for a deliberate approach to AI integration in special education teacher preparation, positing that, like any technology, AI enhances learning when thoughtfully integrated. We emphasize collaborative engagement with AI to navigate its practical and ethical complexities, recognizing it as a tool that, when wielded with intentionality, has the potential to elevate educational practices.

We, the authors of this article, include full-time faculty, adjunct faculty, and administrators in teacher preparation programs across two universities (one public, one private) in the Northeastern United States. At this time, our universities have broad AI policies, allowing instructors the academic freedom to choose how and when to integrate AI into courses, if at all. As AI has become more readily available, we have found ourselves examining ways in which we can learn together with our teacher candidates to use AI productively and ethically to enhance and elevate our educational experiences. As we progress in this journey, we aim to document faculty and student experiences implementing AI in practical and innovative ways within teacher preparation programs to provide insights from multiple angles and promote an openness to unanticipated findings.

A term coined and conceptualized by McCarthy and his colleagues for a conference at Dartmouth College in 1956, AI is based on the "conjecture that every aspect of learning or any other feature of intelligence can, in principle, be so precisely

described that a machine can be made to simulate it" (McCarthy et al., 1955, para. 1). Even at that time, this concept was not a new one. In his famous 1950 article, Alan Turing, the renowned World War II codebreaker, introduced the Imitation Game and proposed exploring the deceptively simple question, "Can Machines Think?" (Turing, 1950, p. 1). As we fast forward through decades of fascinating developments and conversations around AI, we find ourselves trying to understand how machine and human thinking may both complement and supplement one another in the context of teaching and learning.

Humans are uniquely capable of considering context and complex ethical and philosophical conflicts that are paramount in educational decision-making. Teachers serve as situational processors of sorts, loaded with imperatives that must function together - sometimes amidst situational conflict - to achieve the best possible outcomes. Maybe that is why education feels inherently personal, individualized, and situational. It often requires constant introspection, reflection, and (re-)evaluation in order to find a path to knowledge, best practices, and growth. If that feeling is true, how can we, as educators of future teachers, help direct these processes, scaffold effective behaviors, and assist in the development of the next generation who will need to do the same, and how do we do that at the incredible scale and pace required? In short, how might we design and create infinite systems of scalable personalization with finite time and resources, and how do we collaborate with teachers, learners, and AI to make this a reality?

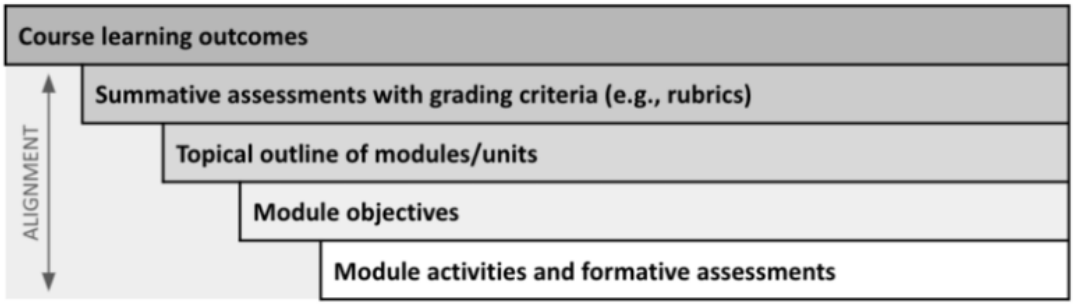
The applications we describe are utilized in programs preparing both special education and general education teachers, as we take an inclusive approach to teacher preparation. While we present specific considerations for special education, these practices are widely generalizable. Research on AI in teacher preparation is in its relative nascency, as most research to date focuses on teachers' professional knowledge and practical guidance and frameworks for AI integration (Sperling et al., 2024). We have not located any intervention research that examines the impact of AI use on the learning of teacher candidates. Therefore, at this time, we are primarily documenting how we are utilizing AI for teacher preparation, drawing connections to educational theory and research. We look forward to exploring ways in which we can systematically study the impact of AI integration, as the scholarship of teaching and learning is a critical vehicle for professional growth in the context of rapid technological innovation. We also affirm that our presentation of the applications below is framed within the same ethical boundaries that must anchor all teaching and learning experiences. Although it should perhaps come as no surprise, we have found the infusion of AI in teacher education work replete with fresh quandaries (e.g., *To what extent is it useful to try to detect AI in student work?*) and thought-provoking wonders (e.g., *Might it be unethical to not provide teacher candidates opportunities to learn about, and use, AI effectively in their work?*). In each example, we attempt to unearth some of the most relevant ethical considerations and offer suggested methods to forge ahead into this brave new world. In so doing, we hope we provide a glimpse at the intriguing and oft transformational potential of AI applications while maintaining a reality-based perspective about the ethical considerations, concerns, and limitations inherent to all technological development.

CURRICULUM ALIGNMENT AND COURSE DESIGN

AI can be an invaluable tool for instructional design and content development. In recent years, faculty at our institutions have undertaken comprehensive updates and redesigns of undergraduate and graduate teacher preparation programs. We generally take a *backward design* approach to curriculum development, beginning with the desired results, determining acceptable evidence, and then planning learning experiences and instruction to match (Wiggins & McTighe, 2005). This effort necessitates substantial curriculum alignment, both at a macro level and within individual courses. Without the use of AI, achieving these changes and launching newly revised curricula would have been significantly more time consuming.

In our curriculum planning processes, after faculty engaged in brainstorming for big picture planning, the next step was to update syllabi and assignments to reflect the revised content. Producing dozens of syllabi in a short time can be overwhelming for full-time faculty who already have many responsibilities. To support us in this process, we utilized generative AI tools (e.g., *ChatGPT*, *Claude*, *Gemini*) to help translate our brainstorming into measurable course outcomes and well-organized topical outlines. This involved providing a single-shot prompt like "*Help me turn this list of ideas into 7-8 learning outcomes,*" followed by a copy and paste of our brainstormed ideas. We then engaged in "dialogue" with the AI to refine these outcomes by asking the AI to *add*, *delete*, *rephrase*, and *combine* ideas, utilizing our content expertise to guide our "conversation." While this required some time to prompt the AI and refine the outputs, the overall time saved in generating quality course outcomes and topical outlines was substantial compared to creating them from scratch. Moreover, we acknowledge that although faculty have the expertise to do this work without AI, using AI in this

FIGURE 1: Funnel Down Approach to Course Design and Alignment with QR Code to an Example Chat



way allows us to be more efficient curriculum designers, offloading lower-order tasks to enhance problem-solving efficiency.

After using generative AI for big picture curriculum alignment, we also applied it to internal course development, aligning course objectives with topical outlines, module-level objectives, assignment descriptions, and assessment criteria and methods (such as rubric development). This was particularly beneficial for developing new online graduate courses. While there are various approaches to using generative AI for this kind of alignment, we found that working with a "funnel down" approach was both productive and efficient (see Figure 1). This approach involved generating course outcomes and summative assessments early on and then making adjustments as we developed the modules/units for the course. Though the approach began linearly, it quickly evolved into a recursive process as we revisited course outcomes and assessments as the modules evolved. Figure 1 also includes a QR code linking to an example AI chat used for course redesign.

Once the framework for each course was developed, some faculty elected to use AI to generate a variety of ideas for in-class and out-of-class activities to support each module, offering students choice in their learning modalities. Generative AI was particularly useful in helping faculty generate case studies, which promote learning through the application of course content in teacher preparation programs (Richman, 2015). We discovered through this process that when utilizing AI for case study generation, it is crucial to provide careful prompting and critically evaluate the output, as generative AI is prone to reinforcing stereotypes, highlighting biases, or even promoting harmful rhetoric about specific groups of people (Howard & Borenstein, 2018). For example, during a professional development session when we were practicing using AI for case study generation, we entered a prompt to generate a description of a student with a disability who was also an emergent bilingual raised by two mothers. ChatGPT responded by stating that it was not biologically possible to have two mothers. Initially troubled by that response, we queried AI again, prompting it to offer answers that were inclusive of LGBTQ couples, and ChatGPT adjusted by eliminating its originally exclusionary perspective. This experience underscored the importance of being actively engaged in the process and not simply copying and pasting from generative AI without critical evaluation. We are pleased to have observed improvements in AI's responses over time, reflecting a more inclusive and accurate understanding of diverse family structures.

SUPPORTING ACADEMIC SKILLS

Gratified by the influence of AI on our curriculum development, we engaged university students in leveraging AI to enhance their learning and development of key academic skills. Rather than forbidding the use of AI, we advocated for discriminating use, creating contexts where students can learn how to use AI in rewarding and ethical ways and in consultation and collaboration with their instructors. This included explicit instruction on prompt engineering and how to use AI as a scaffold that supports and enriches learning rather than supplants it. Though we have not yet begun to research the impact of our instruction with AI, we connect our instruction back to high-leverage practices (HLPs) for students with disabilities (Aceves & Kennedy, 2024) that are grounded in research and an integral part of our teacher preparation programs. Our most utilized HLPs include using explicit instruction (HLP 16), providing scaffolded support (HLP 15), using student assessment data, making adjustments to improve student outcomes (HLP 6), and providing positive and constructive feedback (HLP 22).

Getting Started with Prompt Engineering

While most of our students understand that large language models such as GPT-4 are trained on massive amounts of data and designed to respond to simple queries that are void of much context, learners are not always aware of specific strategies for writing prompts within AI platforms. Few students arrived in

FIGURE 2: Criteria for Quality Prompting

Clarity: Straightforward, unambiguous, and precise language ensures that requests are understood.

Specificity: Defining the scope and details of requested information better enables the model to generate responses that meet users' expectations.

Neutral Tone: Maintaining a neutral tone helps avoid biased responses, as it allows the chatbot to remain objective and factual.

Inquiry Type: Prompts that clearly indicate whether the user seeks a description, explanation, comparison, etc., guides the model in structuring its response accordingly.

Fact-Based Queries: Requesting fact-based responses rather than opinions, especially in areas where bias is likely to be an issue, prompts AI to rely on verifiable information rather than generalized statements.

Cultural Awareness: Specifying relevant cultural factors helps the AI to tailor its response to the appropriate context, reducing the risk of culturally insensitive or inappropriate responses.

Explicit Instructions to Avoid Bias: Explicitly prompting instructions can avoid specific types of bias that are common in certain disability circles.

FIGURE 3: CRAFTS Acronym for Prompting

C	**Call to Action** Begin with a clear call to action, much like a basic zero-shot prompt. *e.g., Explain the concept of motif in literature.*
R	**Role** Ask AI to assume a specific role as it completes the request. *e.g., Assume the role of a seventh-grade special education teacher who supports students with learning disabilities, including those with dyslexia, dysgraphia, and language processing disorders.*
A	**Audience** Consider how the output will be used, and by whom, taking care to ensure that AI is aware of who it will be framing its response for. *e.g., Include descriptions and examples of motif that students with learning disabilities would understand.*
F	**Format** Direct AI to produce the output in a meaningful format. *e.g., Compose your explanation in three paragraphs, and follow with examples of motifs found in popular Taylor Swift songs. Use bullet points to separate these examples from one another.*
T	**Technicalities** Include parameters and other specifics relevant to the desired conventions. *e.g., Use words that most seventh graders born and raised in Buffalo, New York would understand.*
S	**Sociocultural Lens** Include social and cultural contexts and direct the bot to mitigate cultural biases in its output. *e.g., Use neutral language that does not assume characteristics, roles, or preferences based on protected characteristics like race, gender, age, etc. For example, avoid gender-specific terms when gender is irrelevant.*

our classrooms equipped to use generative AI in Fall 2023, shortly after our universities had established policies that supported ethical use. In fact, most of our students were admittedly distrustful of AI, and the limited experiences they had produced less-than-ideal results. Explicitly coaching these learners to add necessary context, examples, non-examples, and criteria of quality in their initial or subsequent queries was an important first step that improved their results and inspired them to try and try again. Invitations to use chatbots to define course concepts, unwrap standards, design learning progressions, and craft lessons, units, and assessments wound their way through many of our courses. As we were new users ourselves, these formative experiences enabled us to study our students' typical prompt-crafting approaches, notice common missteps, and help them work through the resulting knots in their processes. It was through trial and error, in the context of learning alongside our students, that we began to define these criteria for quality prompting (see Figure 2).

Attending to these criteria prepared most students to conduct simple and slightly more complex queries well. In general, prompt quality improved

through explicit instruction, practice, assessment, feedback, and revision. The CRAFTS acronym (see Figure 3) was a useful scaffold for those honing their prompting skills. Eventually, some learners became sophisticated users, particularly those who were likely to test more complex inquiries, such as few-shot or chain-of-thought prompting (Wei et al., 2022), when simple queries returned unsatisfying results. For example, prompting chatbots to produce simple responses to queries like, "Design a lesson that teaches students with dyslexia how to identify the main idea in informational text," produced less useful results than querying bots to "Assume the role of an eighth-grade special education teacher. Design a 10-minute mini-lesson that includes an explicit strategy for identifying the main idea in a news story. Ensure that this strategy is one that is aligned to best practices for supporting comprehension instruction with dyslexic learners." It was our experience that pre-service special educators, in particular, relied on these advanced approaches most often. This made sense, as using AI in service to students with learning disabilities requires nuanced and heavily contextualized queries.

Instructional AI for Scaffolding Academic Behaviors and Writing

Our examples thus far have involved the use of generative AI, which produces content for the user. Another way we have supported academic skills is through the use of *instructional* (or *assistive*) AI, which guides writing and thinking processes rather than producing written content for the learner. Instructional AI can be used to appropriately self-level feedback and provide support based on students' current skill sets, thereby reducing the need for an expert "on-demand" in the form of a professor or writing center tutor. Once again, we successfully leveraged instructional AI to offload behavioral/skill-building feedback and coaching and devoted the newly generated time capital to engage in higher-level thinking and facilitate both the acquisition and demonstration of content knowledge and classroom-specific applications.

As an example of how newfound/reallocated time can invigorate faculty instruction, some of us utilized a tool called *Packback* to help craft formative and summative discourse and writing assessments. Its built-in behavioral and writing tutor, powered by AI, supports question and argument development, thesis construction, and the alignment of supporting ideas. Packback also provides coaching, guiding students to assess the credibility of sources. This provides personalized, real-time support to students and ensures they can preserve their own authentic voice as they demonstrate understanding of content, rather than having that demonstration obscured by limitations of their current academic skill sets. This strengths-based emphasis also supports ethical best practices for supporting agency. Internalized stigmas associated with asking for help, or even self-identifying as a student who needs help, create barriers that Packback enables learners to overcome. First-generation college students, as well as those from low socioeconomic backgrounds and other historically marginalized communities, tend to be more affected by such stigmas (Winograd & Rust, 2014), and the implications for increasing equitable access in this way were also a major factor in our decision to implement Packback. This application of AI allowed faculty users to pay more attention to the applicative context and content being created by each individual student while using the course design, scope and sequence, and assessments to provide sequential structure. Additional reductions in such direct professor intervention for procedural and remedial skill support meant we could pace the course to include more deep processing of ideas.

While Packback's AI engine is designed to meet the students at their current level with appropriate feedback as the student works, we also had to create a course structure and assessments customized to meet course objectives within the platform. In that way, we crafted formative and summative assessments in a cadence meant to support the scaffolding of not only assignments, course deliverables, and collaborative discourse but also the scaffolding of the behaviors and skills necessary to be more responsive to individual student needs. We have used the AI-supported tools Packback provides – live Socratic discussions, multiple polling features, extended-form writing, and professor analytics and communication tools built for scaling purposes – to provide a sequential course structure that reinforces procedural and process-based skills and behaviors.

Figure 4 shows how we implemented a guided, structured, and sequential course design that efficiently covered the necessary content acquisition and skill development outlined in our course objectives, including a variety of deliverable assessments to match. We found the structuring of weekly, formative, learner-outcome reflections on in-class activities and crowdsourcing student-generated applications were helpful to build up to summative assessments of modular concepts. These assessments provided the majority of the content for the final, more formal, summative deliverable to demonstrate achieved learning outcomes aligned with course objectives. Packback provided the AI assistance to build requisite skills where there was a need for individualized, remedial support while not monopolizing in-class time and resources to do so. We found the resultant

FIGURE 4: Weekly Course Structure with Packback Instructional AI Tool

MONDAY asynchronous activity before class	**Prep Day** Automated Poll (*Packback Questions* tool - opens at 12:00 AM) *Identify the most important main idea take-away from lecture prep materials*
TUESDAY In-class activity	**Modified Lecture Day (with time to write “crowdsourced content”)** Socratic Application Question (*Packback Questions* tool) *Find an application of a concept we talked about and ask students how they would implement in their own classroom*
WEDNESDAY asynchronous activity before class	**Prep Day** Select a peer question you’d like to present to begin the next in-class activity *Look for a peer’s posted application question in Packback you’d like to discuss/explore as a class*
THURSDAY In-class activity	**Class Content Processing and Related Activity** 100-word learner outcome reflection (*Packback Deep Dives* tool) 2 Responses to peer questions (*Packback Questions* tool) *Now that you have digested and clarified content knowledge and engaged in application activity, answer peer questions and inventory your learning*
FRIDAY asynchronous activity after class	**Reflection or Work on Summative/Modular Deliverables** “What? So What? Now What?” Format Journal or Summative Deliverable (*Deep Dives* tool) *Synthesize your learning and be sure to include how you’ll use it in your studies or future classroom*

efficiency freed us up to clarify content-based confusion without sacrificing instruction, cultivate excitement about identified student interests, and utilize the influx of time to prepare meaningful experiential activities within our full-class setting.

While most educators strive to maintain such learner-centered postures, special education teachers must be especially attuned and responsive to the unique needs of the children they serve. It was our goal to demonstrate careful course planning for our students and provide a series of assignments and assessments built to accurately reflect the learning outlined in the course outcomes. Reflective discussion of this type of instructional design enabled our students to recognize the benefit of efficient and cohesive planning and execution with AI support.

SUPPORTING TEACHING SKILLS

In addition to using AI to support broad learning skills, as educators in teacher preparation programs, we also used it to support teacher candidates in their development of skills specifically related to instructional planning, delivery, and assessment.

Using AI to Elevate Instructional Planning and Preparation

Traditionally, teacher candidates learn to write lesson and unit plans during their teacher preparation programs. However, there is a nationwide shift in expectations that turns pre-service teachers toward high-quality instructional materials and encourages adaptation rather than the development of original units and lesson plans (Council of Chief State School Officers, 2022). Generative AI has also become increasingly skilled at writing detailed lesson plans. We need to ask ourselves: Is our instructional time best spent coaching pre-service teachers to write original lessons or, instead, to analyze and internalize lessons provided to them? Comparably, we acknowledge that many of our teacher candidates will work in schools where they are not given access to high-quality instructional materials, and, as future special educators, teacher candidates may still need to develop specially designed instruction for learners with disabilities that is aligned to individualized education programs (IEPs) as well. We have found it necessary to invite critical discourse with our teacher candidates about their roles, responsibilities, threats, and opportunities in given contexts when considering the use of AI.

Early in our programs, long before they attempt to write a lesson of their own, teacher candidates first learn to lesson plan by analyzing example lessons,

looking for key components like standards, objectives, materials, procedural descriptions, instructional supports, opportunities for student engagement, and formative and summative assessments. In class, they discuss which practices are supported by research and what barriers to learning may be present within the lesson plan. We are beginning to expand this analysis to include lessons produced by AI. For example, a detailed prompt into generative AI – *Write a 45-minute lesson plan for third-grade students around this English Language Arts Standard: "3R6: Discuss how the reader's point of view or perspective may differ from that of the author, narrator or characters in a text." The lesson should embrace principles of universal design for learning, offering multiple and flexible means of representing the content, student engagement, and student action and expression* – will produce a detailed lesson plan, complete with most of the key components we would be looking for on our lesson plan templates. If we do not explicitly address AI for instructional planning in our courses, teacher candidates may be inclined to simply copy and paste the AI output into the template without critically analyzing that output to ensure the recommendations by AI are actually based on research and best practice.

As we move forward with AI integration in our programs, we plan to give teacher candidates time to generate lessons with AI, putting into practice some of the aforementioned prompting strategies to improve upon the initial AI-generated lesson. For example, ChatGPT rarely writes lesson objectives in a measurable/observable way initially. A teacher candidate could type the following prompt to adjust the lesson objectives: *Rewrite the learning objectives so they are written in an observable/measurable way with clear criteria that align with the lesson assessment.* Depending on the initial prompt, ChatGPT may give some initial ideas for differentiation or universal design, but they are often general ideas. If the teacher candidate is given a case study or is generating a lesson for students at one of their field placements, they could add specificity with follow-up prompts:

- *I have a student who is an English Language Learner in this course; what are at least three ways I could support them during this lesson?*
- *There are two students with IEPs in my class. One student has ADHD and, while they enjoy reading, they have difficulty staying focused on a single task. The other student has a learning disability and is not yet reading at a third-grade level. This student is receiving Tier 3 reading support focused on decoding and phonological awareness. How can I support each of these students during this lesson?*

Not every response from AI will be high quality, but it will generate some ideas that can then be used to spark discussion in teaching methods courses. Then, if and when teacher candidates need to write their own lessons, we hope they will be able to use AI to launch ideas so they can spend their time applying critical thinking and analysis skills to improve the lesson and prepare for quality instructional delivery.

Reflecting on Instructional Delivery with AI Coaching

Alongside a growing emphasis on candidates' ability to adapt high-quality instructional materials, the integration of video-based coaching in preparation programs is gaining traction as a means to foster self-reflection and improve candidates' pedagogical skills. Reflective ability is multifaceted and refers to teachers' ability to (a) describe important teaching decisions, (b) analyze the reasons behind those decisions, (c) evaluate the impact of those decisions on student learning, and (d) apply insights to create a plan for extending effective or changing ineffective practices in the future lessons (Nagro et al., 2017). Teacher preparation programs have increasingly turned to technology platforms to support teacher reflection (e.g., Vosaic, GoReact) as they provide candidates with the ability to upload videos of their teaching, mark or tag important moments, and then compose reflective annotations tied to those moments. Vosaic recently integrated a new AI feature, *AI Mate*, to "enhance teacher coaching and improve video analysis for research." Using transcripts from teaching videos, AI Mate can analyze an uploaded video lesson and provide time-stamped feedback based on prompts customized by the user. AI Mate can be used to support faculty members as they guide and evaluate teacher candidates' performance. It can also help teacher candidates independently engage in self-reflection, helping them identify strengths and opportunities for improvement and set goals for future teaching.

AI-enhanced coaching offers a predictable mix of advantages and challenges, and benefits from thoughtful decision-making by those well-versed in teacher preparation pedagogy. Instead of primarily serving as a summative measure of teaching proficiency (e.g., videos of lessons as part of a capstone project), we view video as holding far greater promise when wielded as a tool for formative assessment and growth (Kaczorowski & Hashey, 2020). The key to this work is scaffolding teacher candidates' video-based reflective ability (e.g., Reichenberg, 2022; Nagro et al., 2022), so they enter the profession with the skills to continuously enhance their teaching practices. As with other applications discussed herein, the infusion of AI into

video-based reflection activities holds the potential to upend the traditional roles of both teachers and learners in this space. When scaffolding video-based reflection, teacher educators typically begin by developing candidates' ability to use annotation tools to notice and describe important instructional decisions by watching videos of other teachers, highlighting observed strengths, and making recommendations for further enhancements. The annotation features allow users to mark important lesson segments and annotate those marked moments using text, audio, or video commentary. Once foundational knowledge about evidence-based pedagogy is established and familiarity with the technology tool is achieved, candidates can capture their own teaching with video and use it to describe important moments, evaluate the effectiveness of their teaching, and apply insights from their reflection to set goals for continual growth.

Faculty feedback in a video-based reflection process is critical to nurturing candidates' reflective ability and growth, helping them identify strengths, notice areas for improvement, and set goals for the future. Much like the examples described above that highlight how AI can offload time required for the initial stages of developing assignments or grading, AI Mate offers a similar affordance in video coaching as it can analyze a video lesson and generate initial feedback based on a prompt. For example, in a course where candidates are developing their ability to incorporate specific questioning techniques, a faculty member could prompt AI Mate to, "*Identify all moments where the teacher posed questions to students in this lesson*." Working from the automatically generated video transcript, AI Mate creates an annotation with hyperlinked time stamped moments for each question posed in the lesson (e.g., 1:34 – 1:42), and provides the text of questions asked. In this way, AI enables faculty to efficiently access the most pertinent moments related to an instructional skill being taught, freeing up time for the instructor to more deeply analyze these important moments based on criteria they may have taught in class and used in their assignment rubric. The nuanced expertise of the faculty perspective is preserved in this example as the application of AI heightens, rather than subverts, the faculty-student learning relationship. A student receiving detailed feedback from this faculty member about their questioning techniques is then better able to understand their relative areas of strength and areas for growth. Likewise, the faculty member is afforded more time to analyze the frequency and quality of questioning techniques, helping them provide better feedback and thereby gauge their students' questioning ability. Just as we have illustrated in other examples herein, AI Mate can capably provide a foundation for faculty to build upon as they analyze candidates' teaching videos and provide input about the strengths and areas for growth in specific instructional delivery skills.

With more complex prompts, AI Mate can also perform overarching evaluations of teaching videos. One example is the prompt, "*Analyze this lesson using Danielson's Framework for Teaching (FFT). Provide an overall summary of observed strengths and opportunities for improvement, and mark three specific moments where you provide recommendations for growth.*" While AI Mate will execute this prompt, we see a far more diminished role for faculty expertise and insight, and it raises questions about the quality of feedback for the candidate. In the best-case scenario, an instructor could again use this AI-generated analysis as a base as they view the video and provide feedback. However, given the nuanced nature of classrooms in which teaching and learning are enacted, videos of actual teaching (as opposed to simulated in-class microteaching) are replete with myriad contextual factors that require a holistic view of the teaching event. These factors are essential for faculty to consider in their evaluation of candidates' teaching performance. In evaluating teaching performance, faculty must always take into account contextual factors such as an understanding of student characteristics (e.g., disability status, multilingual learners), classroom variables (e.g., grade level, individual behavior intervention plans, school-wide practices), curricular focus (e.g., content area(s), relationship to larger unit), and pedagogical orientation (e.g., explicit instruction, inquiry-based approaches, project-based learning), to name just a few. Thus, while AI applications like AI Mate can respond to prompts that attempt to capture the "big ideas" about evidence-based teaching practices in the provision of video-based coaching and also offer suggestions about means of improving instruction, we assert that a faculty member's ethical use of AI for evaluating teaching will always be anchored in a deep understanding of the realities of a given instructional context. As such, AI-generated assessments of specific or general teaching practices, as captured in videos, are most valuable as a starting point for faculty input and less valuable as a standalone end product.

While we affirm an integral role for faculty expertise and insight in the context of class assignments within teacher preparation programs, AI-supported coaching might also be leveraged independently, without faculty involvement, by reflective, self-driven students and by practicing teachers. In this case, we see value in leveraging a tool like AI Mate as a means by which teachers can

receive individualized feedback on their teaching, especially in the absence of a mentor, learning partner, or community of practice. Although some districts support robust approaches to individual teacher professional development, many do not; therefore, facilitating a self-directed approach to professional development can an invaluable tool. This is perhaps especially true for novice teachers without access to mentorship, whose schools may not attend to important aspects of teacher induction, or those who are simply struggling to improve their instructional skills. As long as an individual user can leverage relevant criteria for desired teaching performance (e.g., Danielson's FFT, explicit instruction, HLPs), they can create prompts for video-based coaching tools like AI Mate to help evaluate important aspects of their teaching and provide suggestions for improvement. A teacher experiencing difficulty providing timely and behavior-specific feedback, for example, could upload a video and use the prompt "*Evaluate all the instances where the teacher gave feedback in this lesson. Use Archer and Hughes' description of effective feedback skills to identify moments where the teacher could improve the quality of the feedback.*" Upon reviewing the moments highlighted by the AI coach, the teacher can prompt it to provide suggestions for improvements in specific areas. Here, AI as a pedagogical coach may be viewed as an essential ally in service of better teaching and, thus, improved student outcomes in the classrooms of teachers seeking to improve their instruction. Although a more knowledgeable other (e.g., instructional coach, mentor teacher, administrator) is an ultimate collaborator in teacher induction, such individuals are not always available or accessible in ways that allow novice teachers to benefit from their guidance. AI-enhanced video coaching, as illustrated in the examples above, offers an important avenue for enhancing the provision of feedback in formalized teacher preparation spaces, as well as for those who seek input on their own teaching performance in a more individualized way.

Ethical Assessment Design with AI

Assessing learning in AI-augmented special education settings demands diverse perspective-taking, the intentional interrogation of biases, and a willingness to apply critical lenses to popular but perhaps superficial and even problematic uses of these tools within the field. For example, special educators can prompt AI to provide rapid, abundant, and meaningful feedback on student work alongside a more quantitative evaluation. However, employing AI for this purpose without the consent of each learner and their guardians pushes the boundaries of ethical use. It is important to establish protocols and norms for AI-augmented assessment practices in order to prevent this sort of harm.

Chatbots may also be used to develop differentiated summative assessments, rubrics, examples and anchors, and related learning progressions that integrate multimodal formative assessment opportunities. For example, special educators might use the CRAFTS protocol (see Figure 3) to generate project based learning experiences and related assessments in alignment with grade-level standards. They could prompt AI to generate aligned rubrics and other assessment tools next. Finally, they might direct the application to produce examples and anchors aligned to each performance level. Once complete, teachers might input these data and challenge the machine to produce learning progressions and differentiated approaches that are attentive to the specific needs of learners with disabilities.

As the learning experience unfolds, learners and their guardians can be coached to use AI to receive just-in-time feedback and just-right instructional support at times when the teacher is unavailable. Inputting task directions, aligned rubrics, examples, and anchors and then tasking AI to use these data to generate actionable feedback that also addresses learners with disabilities' specific needs offers a powerful lever for teachers. Establishing norms and protocols for seeking student and parent consent to use AI for such purposes is an important first step for any educator seeking to use AI in this way. For ethical reasons, it makes sense that whenever possible, teachers coach students and their guardians to engage AI for feedback purposes themselves rather than inviting instructors to submit student work to the machine on their behalf. Such practices better protect ownership and privacy. They also build the capacity of students and their guardians to ethically engage AI in service to their unique interests and needs.

It is important for educators to know that AI algorithms are informed by the same social, political, and racial biases that permeate society, and this has implications for assessment design. Taking care to evaluate AI outputs and consider where biases might manifest in the feedback and content generated is a critical first step. Establishing assessment design protocols and learning how to refine prompts to mitigate such biases are essential to ethical use. We must also work to diversify the data sources that AI tools are trained on by ensuring that learners with diverse abilities are shaping the algorithms. Submitting related feedback samples that are free from gender, racial, ethnic, and other biases is also useful. All of these efforts require the consent of students and their guardians, as well as their collaboration.

While it is true that chatbots run on

biased algorithms, it is also true that AI-augmented tools can enable educators to mitigate the personal biases they bring to their own work. This is particularly useful when teachers engage in research that is intended to uncover and address complex learning needs. AI helps educators align and balance their assessment systems, triangulate the data that drives their decision-making, and rely on formative assessment to gain much-needed perspective about performance. While standardized test results and summative local assessments offer meaningful generalizations about what students may know or be able to do relevant to learning outcomes, it is formative assessments and, more importantly, assessments that invite multimodal expression that enable a much clearer view of why students might perform the way they do. Such assessment practices rely on qualitative data, and because they fail to quantify learning or performance, interpretations are more subjective, and analyses are more time-consuming.

Integrating AI tools within such processes can attenuate the influence of bias and increase efficiency by providing data analyses that are consistently and rapidly applied across different contexts and student groups. AI can also highlight patterns and anomalies in student performance that might not be obvious to educators alone and recommend intervention approaches that rely on an array of high-level, research-based best practices, which teachers may not recall on demand or have time to investigate themselves. When appropriately prompted and engaged, AI encourages a more objective review of students' work and learning behaviors, enabling educators to make more informed, equitable decisions.

For instance, *Otter.ai* is a transcription tool that is particularly useful to special educators who invite learners to demonstrate what they know and can do through various means of multimodal expression. This application records audio, captures images, extracts action items, summarizes discussions, and defines patterns and themes in the resulting data. Teachers and learners may rely on the tool to record and analyze small and full group discussions as well as one-on-one conferences with one another. This enables progress monitoring in multimodal contexts where learners may not rely on written words to express themselves. The potential for Otter.ai to rapidly capture, analyze, code, and summarize volumes of qualitative assessment data, documented in process, as students with learning disabilities leverage the modes of expression that serve them best democratizes the experience while ensuring that special educators are grounding instructional decisions in robust data.

> "Explicitly coaching new users to leverage these tools productively without compromising empathy, understanding, and human connection is especially crucial in special education contexts."

The rise of AI reminds us that the best evidence of learning is gathered in-process. Research suggests that such formative assessment data improves student achievement. This is because formative assessment intentionally illuminates and inspires instructors to be responsive to students' instructional needs. Moreover, formative assessment has been found to particularly benefit striving learners, thereby reducing the achievement gap (Black & William, 2010). Relying on the products of learning, such as test or quiz results, final papers, and other summative measures, does not necessarily help us understand how conceptual knowledge and critical skills are acquired or where within the process meaning breaks down. It is formative assessment data that guides timely instructional decision-making. With the dawn of AI, learners are also able to rapidly generate robust content that is not reflective of their own thinking, compositional or design capacities, or skill mastery. We simply cannot rely on summative assessments alone to help us understand learners or facilitate learning well. This is why pedagogical documentation is beginning to play an increasingly essential role in our practice.

DOCUMENTING LEARNING WITH AI

Aware of the biases and significant false-positive rates inherent in AI detection tools, some of our colleagues began decentering summative assessments and requiring students to document their learning processes in order to mitigate academic dishonesty. We know that what makes pedagogical documentation truly powerful is the reflective nature of this work. We document learning in order to study and improve it. We are eager to understand and better serve our students, which means using AI tools to increase natural intelligence and skills as well as learner agency. Pedagogical documentation helps us realize this vision. It involves coaching students to notice when learning is happening; holding space for them to pause and capture images, video or audio recordings, and artifacts of these moments;

teaching them how to interpret and reflect upon this data; and inviting them to use various tools to curate and share their learning stories.

Pedagogical documentation is a multimodal assessment practice that generates more holistic portraits of each learner, better attuning us to their strengths and needs. This is because each mode of expression enables learners to make their thinking and processes visible in ways others do not (Cope & Kalantzis, 2022). For instance, images capture much that words alone cannot describe, and video recordings of learning in-process help us notice things we would not otherwise. This makes for a far more textured reflection of learning and a nuanced assessment of learners. Experience has taught us that documenting and sharing our collective learning processes fuels feedback loops that improve engagement with course content, between learners, and between learners and instructors. Finally, pedagogical documentation invites us to study the impact of teaching and learning on student performance in ways that summative assessment alone cannot (Hattie, 2023). This is how we humanize the learning process in ways that AI cannot. It is also how we might better mitigate bias and ensure integrity within the assessment process.

For example, AI tools are perfectly capable of quickly producing high-quality literature reviews in various academic contexts, leaving all instructors grappling with the true provenance of any student's submitted work. Alternatively, pedagogical documentation turns our collective attention toward the research and writing *process*. Learners might be invited to document how they conducted their search. They might be challenged to gather audio and video reflections that make their analyses of gathered sources clear, and they might also be asked to document how each phase of the learning experience changed their thinking about what it means to engage in a discriminating inquiry and complex compositional process. Learners might record their peer review engagements, speak to how their findings inform their thinking about unique personal experiences, and use artifacts from their note-making, drafting, or prototype creation to speak to the evolution of their ideas and work. In this way, pedagogical documentation is a perfect fit for classroom cultures where performance is a byproduct of critical and creative thinking. Pursuing this greater vision enables all of us to bring a far more critical lens to each AI engagement.

CONCLUSION

AI tools rapidly tailor learning experiences to the diverse needs and interests of individual students by creating multimodal, differentiated content. Our classroom experiences have taught us that pre-service special educators who possess heightened levels of digital literacy, understanding of AI functionalities, and the willingness to tinker, play, adapt, and critically edit AI-generated content enjoy more satisfying results. Explicitly coaching new users to leverage these tools productively without compromising empathy, understanding, and human connection is especially crucial in special education contexts, and this challenges us to pursue curricular alignment with care.

Successful integration of instructional AI shifts our expectations beyond quick content generation or even complex problem-solving in specific classroom contexts. Instead, we suggest that integration is only truly successful if it func-

ABOUT THE AUTHORS

Tara Kaczorowski, Ph.D.

Tara Kaczorowski, PhD, is an Associate Professor and Executive Director of Education Programs at Daemen University. She teaches college courses related to instructional methods in STEM fields and inclusive education. Her primary research areas are in teacher/faculty preparation, instructional technology, video-enhanced reflection, and math/STEM instruction for students with high-incidence disabilities.

Angela Stockman, MS.Ed.

Angela Stockman is an adjunct instructor at Daemen University in Amherst, New York. The author of several books on multiliteracy instruction, she is regularly invited to lead curriculum and assessment design work in K-12 schools, where she also conducts lesson studies upon request. A former public school teacher, Angela spent 15 years in the classroom before becoming a professional learning facilitator.

Andrew I. Hashey, Ph.D.

Andrew I. Hashey, PhD, is Associate Professor and Chair of Exceptional Education at SUNY Buffalo State University. His scholarship in special education teacher preparation centers on high-leverage practices, school-university partnerships, teacher reflection, and evidence-based literacy instruction. His ongoing research projects take aim at building faculty capacity to implement video-based reflection within their teacher preparation programs.

John Kaczorowski, MFA, MS

John Kaczorowski holds an MFA in Creative Writing, an MS in Sport Management, and was a doctoral scholar in Sport Management (ABD) at University of Illinois. A former English Language Arts teacher, he is now an Experience Manager at Packback and adjuncts in the Education Department at Daemen University.

tions in service to culturally sustaining and deeply human pursuits. When we remember that AI is not merely another tool but a paradigm shift that is reshaping industries, societies, and our own perceptions of learning and cognition, we understand that every instructor and department has a unique role in ensuring students are not only adept users of AI but also critical thinkers, capable of navigating and influencing this rapidly evolving landscape. Collaborating with our colleagues and students to establish a clear vision that makes our shared values and boundaries clear can set a foundation for a more informed, adaptive, and forward-thinking generation of special educators. To accomplish this, we must also remain aware of inherent bias in the algorithms that power AI, their tendencies to erase BIPOC, LGBTQ+, female scholars, and others who live on the margins, and the steps we must take to mitigate harm and teach our students to do the same.

Further research is needed to understand the long-term impacts of AI integration on student learning outcomes and educator practices. Experts in special education technology identify some critical areas for research on AI integration, including personalized learning approaches, accessibility, early identification of learning disabilities, and teacher support (Center for Innovation, Design, and Digital Learning, 2024). The U.S. Department of Education (2024) also emphasizes the need to establish evidence of AI's impact on educational outcomes. They acknowledge the value of documenting the learning process and utilizing broad methods that capture what students know and can do rather than only relying on standardized measures. We posit that studying the pedagogical documentation of pre-service teachers can provide deeper insights into the learning process with AI. By systematically studying and reflecting on AI integration in educational settings, educators can help establish evidence-based practices that enhance learning and teaching. As we explore and evaluate the impact of AI-enhanced learning, again and always, the questions matter: *What should we ask of ourselves in an AI-augmented learning environment? What should we ask one another? What should we ask the machine? And how?*

REFERENCES

Aceves, T. C., & Kennedy, M. J. (Eds.) (2024). *High-leverage practices for students with disabilities* (2nd ed.). Council for Exceptional Children and CEEDAR Center.

Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *92*(1), 81–90. https://doi.org/10.1177/003172171009200119

Center for Innovation, Design, and Digital Learning. (2024). *Inclusive intelligence: The impact of AI on education for all learners*. https://ciddl.org/inclusive-intelligence-the-impact-of-ai-on-education-for-all-learners/

Cope, B., & Kalantzis, M. (2022). *Making sense: Reference, agency, and structure in a grammar of multimodal meaning*. Cambridge University Press. https://doi.org/10.1590/1678-460x202238256749

Council of Chief State School Officers. (2022, January). *High quality instructional materials and professional development network case study: Impact of the CCSSO IMPD network.* https://753a0706.flowpaper.com/CCSSOIMPDCaseStudyImpact/#page=14

Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement* (1st ed.). Routledge.

Howard, A., & Borenstein, J. (2018). The ugly truth about ourselves and our robot creations: The problem of bias and social inequity. *Science and Engineering Ethics*, *24*, 1521–1536. https://doi.org/10.1007/s11948-017-9975-2

Kaczorowski, T. L., & Hashey, A. I. (2020). Using video-enhanced performance feedback for student and instructor reflection and evaluation. In E. Alqurashi (Ed.), *Handbook of research on fostering student engagement with instructional technology in higher education* (pp. 94–115). IGI Global. https://doi.org/10.4018/978-1-7998-0119-1.ch006

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (1955). *A proposal for the Dartmouth summer research project on artificial intelligence*. https://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Nagro, S. A., DeBettencourt, L. U., Rosenberg, M. S., Carran, D. T., & Weiss, M. P. (2017). The effects of guided video analysis on teacher candidates' reflective ability and instructional skills. *Teacher Education and Special Education*, *40*(1), 7–25. https://doi.org/10.1177/0888406416680469

Nagro, S. A. (2022). Three phases of video-based reflection activities to transition teacher candidates from understanding to examining practice. *Journal of Special Education Preparation*, *2*(1), 28–37. https://doi.org/10.33043/JOSEP.2.1.28-37

Reichenberg, J. S. (2022). Video reflection with a literacy coach: The mediation of teacher agency. *Reflective Practice*, *23*(5), 607–621. https://doi.org/10.1080/14623943.2022.2093845

Richman, L. (2015). Using online case studies to enhance teacher preparation. *Journal of Technology and Teacher Education*, *23*(4), 535–559. https://www.learntechlib.org/primary/p/148669/

Sperling, K., Stenberg, C. J., McGrath, C., Akerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open*, *6*, Article 100169. https://doi.org/10.1016/j.caeo.2024.100169

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *LIX*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

U.S. Department of Education, Office of Educational Technology. (2024). *Designing for education with artificial intelligence: An essential guide for developers*. https://tech.ed.gov/designing-for-education-with-artificial-intelligence/

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, *35*, 24824–24837. https://doi.org/10.48550/arXiv.2201.11903

Wiggins, G. & McTighe, J. (2005) *Understanding by design*. Association for Supervision and Curriculum Development.

Winograd, G., & Rust, J. P. (2014). Stigma, awareness of support services, and academic help-seeking among historically underrepresented first-year college students. *The Learning Assistance Review*, *19*(2), 17–41.