The Journal of Educational Research

ISSN: 0022-0671 (Print) 1940-0675 (Online) Journal homepage: www.tandfonline.com/journals/vjer20

Teaching high school students about generative AI: Cases of teacher lesson design

Victoria Delaney, Ibrahim Oluwajoba Adisa, Christopher Mah & Victor R. Lee

THE JOURNAL OF EDUCATIONAL RESEARCH
https://doi.org/10.1080/00220671.2025.2510415

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Teaching high school students about generative AI: Cases of teacher lesson design

Victoria Delaney[a], Ibrahim Oluwajoba Adisa[b], Christopher Mah[c] and Victor R. Lee[c]

[a]Department of Mathematics and Statistics, San Diego State University, San Diego, CA, USA; [b]Graduate School of Education, Institute for Human-Centered Artificial Intelligence, Stanford University, Stanford, CA, USA; [c]Graduate School of Education, Stanford University, Stanford, CA, USA

ABSTRACT

Teachers who wish to enact lessons about generative AI are required to simultaneously learn about it and develop curricula with activities that align with their discipline. We present two cases of high school teachers, June and Margot, who had different prior experiences, resources, and learning goals related to GenAI instruction. We found that they designed lessons that positioned GenAI as an *object-of-study* or *subject-specific,* but neither lesson solely focused on either approach. Prior disciplinary and lesson planning knowledge and in-the moment student reactions to activities shaped their appraisals of lesson effectiveness. However, we observed that co-design experiences and activities were central for helping to develop teachers' pedagogical design capacity for GenAI. We contribute two cases that illustrate how co-design can support high school teachers who wish to integrate GenAI into their discipline, and by offering contrasting models of pedagogical approaches.

ARTICLE HISTORY
Received 17 February 2025
Revised 2 May 2025
Accepted 20 May 2025

KEYWORDS
AI education; curriculum design; co-design; high school

Introduction

Many K-12 students use generative artificial intelligence (GenAI) applications, such as ChatGPT, in ways that influence what and how they learn (Walton Family Foundation, 2023). Even students who do not use these tools still interact with GenAI chatbots that are embedded in technologies like Google Search, Grammarly, and YouTube. Outside of school, students' online experiences are increasingly mediated by AI in the form of algorithmic recommendations, AI-generated media, and interactions with AI agents. The ubiquity of GenAI in everyday media has underscored the importance of educating students to become conscious and thoughtful AI consumers. It has also raised questions about what GenAI competencies high school students should learn (e.g., Ng et al., 2024), where they should learn about them (e.g., Druga et al., 2022), and who should bear the responsibility of teaching about GenAI. These questions have sparked discussions about how to best prepare K-12 classroom teachers to teach with and about GenAI. Recently they have begun to include teachers outside of computer science, the discipline historically associated with artificial intelligence and AI literacy (Long & Magerko, 2020).

Integrating GenAI tools into classroom lessons can be a deceptively challenging endeavor for teachers. Many do not have well-developed knowledge or experience with GenAI tools as they relate to teaching within their subject discipline (Chiu & Chai, 2020). Some worry about GenAI's potential to exacerbate misinformation (Baidoo-Anu & Ansah, 2023), generate biased or discriminatory outputs, or compromise students' data privacy (Rawas, 2023). Others have ethical concerns (Akgun & Greenhow, 2022), questions about the relationship between students' AI use and academic integrity (Lee et al., 2024) or negative dispositions toward GenAI as tools that may potentially displace aspects of student thinking and reasoning (Abbas et al., 2024). To navigate these challenges, researchers have proposed various approaches, including co-constructing norms around GenAI use with students (Mah et al., 2024), shifting toward authentic, performance-based assessment paradigms (Rudolph et al., 2023); and promoting students' ability to understand, evaluate, and use AI tools (Mills et al., 2024). All of these shifts require significant commitments from teachers to adapt and update their instructional praxis. It is therefore imperative that K-12 teachers receive support to (re)design and enact lessons that teach about GenAI.

Some emergent work has begun to explore what teacher learning about GenAI may entail (e.g., Sperling et al., 2024). However, much less is known about how teachers incorporate GenAI applications into their subject areas or teach about it as a learning tool. To address this need, we constructed a GenAI lesson co-design partnership and sought to provide secondary school teachers with support to learn about and teach with GenAI. We included professional learning (PL) activities to build teacher knowledge about GenAI and offered lesson design support for integrating GenAI into their teaching practice. At the same time, we

CONTACT Victoria Delaney vldelaney@sdsu.edu Department of Mathematics and Statistics, San Diego State University, San Diego, CA, USA

encouraged teachers to design lessons grounded in their subject-matter expertise and aligned with their curriculum standards. We aimed to make the process of integrating GenAI into curricula less cognitively burdensome and increase the likelihood that teachers would teach about GenAI in school.

We began with the premise that our co-design partnership should account for inherent diversities in teaching. High school teachers frequently teach within singular subject areas (e.g., math teachers often only teach math courses), have a range of prior experiences with GenAI, including none at all, and operate in school environments with varying policies, student demographics, and available resources. It was therefore imperative for our team to support teachers' GenAI learning while encouraging them to take ownership for *how* they integrated it into their lessons. We hypothesized that teachers' diverse knowledge resources and approaches to designing lessons could result in vastly different interpretations about how to integrate GenAI into high school teaching. We present two illustrative cases of teacher lesson co-design in this paper. First we illustrate the lesson design and enactment arc of June (an AP research methods teacher), and then of Margot (an AP English teacher), who demonstrated contrasting approaches to integrating GenAI and teaching their students about it. We conceptualize June and Margot as agentic GenAI curriculum designers whose lessons were influenced by their knowledge, disciplinary conventions, and available resources, and further mediated by the activities and participants in our professional learning community. In the sections that follow, we describe the lessons that June and Margot designed and taught, and broader implications for those who wish to support teacher GenAI lesson design.

Co-design as a mediator of the teacher-curriculum relationship

We situate this work within theories of curriculum design (the outcome) and co-design (the process). We argue that teaching with GenAI involves curriculum design work and that co-design is one form of teacher support that can influence what teachers ultimately do when teaching about GenAI. By *curriculum*, we refer to summaries of goals, objectives, and beliefs about what should be taught and how they are represented in written statements (Stein et al., 2007). That teachers would design curriculum in the first place is a relatively novel phenomenon, as traditional classroom teaching positions teachers as receptors of curriculum from written materials, notably textbooks (Choppin & Borys, 2017). More recent work has explored how teachers acquire, adapt, and/or innovate curriculum with digital tools, particularly from their interactions with curricular resources on the internet (Remillard et al., 2024).

For many teachers, designing and enacting lessons involves triangulating curriculum from traditional print materials with novel resources, modes of instruction, or activities acquired on the internet (Pepin et al., 2017). For instance, an English teacher may search for an Op-Ed as an example of argumentative writing, a civics teacher may search for news articles that illustrate examples of governance, and a math teacher may search for online applets that visualize concepts, such as sampling distributions. Some teachers have begun to include GenAI tools as online resources for lesson design (Moundridou et al., 2024). Finding online curricular materials (or prompting GenAI tools to create them), however, is not equivalent to lesson design or teaching. One key challenge that high school teachers face is coordinating online resources with existing curriculum standards, enduring understandings, and epistemic practices. We argue that to do this type of cognitive work, teachers must develop curriculum design skills.

We share Brown's premise that teaching is inherently design work, and that teachers display a range of design and improvizational skills when they bring written curriculum (documents or digitized written materials) to life (Brown, 2011). He suggests that teachers' decisions and skill for enacting curriculum are in large part a function of their *pedagogical design capacity* (Brown, 2002). Pedagogical design capacity refers to teachers' knowledge and curricular resources that are available to them as they interact with materials and make decisions on what to teach (see Figure 1). Pedagogical design capacity can be influenced by teachers' experiences, school environments, support ecosystem, proximity to technology and other resources, colleagues, and other (often interconnected) factors (Remillard & Heck, 2014). Because these factors are complex and diverse, teachers with the same curriculum and curriculum standards can enact lessons that are vastly different from each other, ultimately impacting what students learn.

To date, there is no single full curriculum, online or in print, that neatly maps learning about GenAI with curriculum standards that high school teachers must teach. Yet, there are plentiful examples of teachers who teach about novel, real-world phenomena by designing and enacting their own lessons (Delaney, 2024; Esmonde, 2014; Trinter & Hughes, 2021; Walsh et al., 2023). We hypothesize that whether or not teachers ultimately teach about GenAI is influenced by their skill and stance on curriculum enactment as design work, and their willingness to design curricular activities that incorporate GenAI. Hence, we believe that teachers can and should be equipped to mitigate potential uncertainties that can surface when teaching with GenAI by developing strong lesson design skills. Furthermore, we propose lesson co-design as one fruitful avenue that can build teachers' pedagogical design capacities.

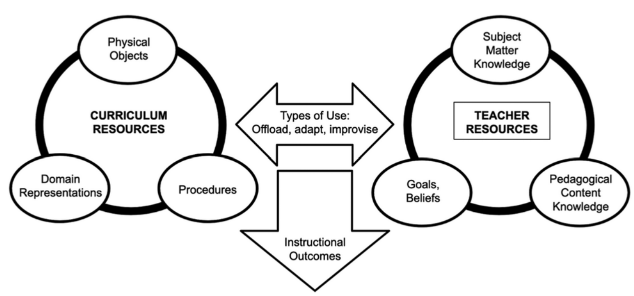


Figure 1. Pedagogical design capacity, reproduced by the authors from Brown's Design Capacity for Enactment framework.

Co-design with professional learning as a teacher resource and mediating design process

Co-design partnerships can coordinate efforts from teachers, researchers, and/or industry members to produce curricula and teaching innovations (Roschelle et al., 2006). In a co-design partnership, members work together for a sustained period of time to design and refine curriculum activities, including implementing and evaluating them in classrooms. Previous co-design efforts in science and AI education suggest that teachers are more likely to continue using novel curricula beyond the scope of co-design partnerships if they have a vested stake in creating them (Lin & van Brummelen, 2021; Severance et al., 2018).

The level of agency that teachers are afforded in co-design processes varies with respect to organizational and structural considerations that are unique to classrooms (Lee et al., 2024). Early efforts positioned teachers as curriculum "testers" who had little say in what activities or tools were produced. Their primary role was to implement curricular innovations and report feedback to researchers and developers (Roschelle et al., 2006). More recent efforts in the learning sciences have examined power dynamics in co-design and, in response, have crafted activities and organizational structures that position teachers with authority and agency in the lesson design process (Lee et al., 2024; Potvin et al., 2024; Swanson & Lawrence, 2024).

Co-design activities that grant teachers design authority can mediate curriculum design (see Figure 2). Features of co-design, such as lesson rehearsals and access to disciplinary expertise, may offer teachers opportunities to revise their written curriculum materials to better account for the range of student thinking and anticipate challenges that may arise during classroom instruction. Activity structures that give teachers iterative feedback can improve their lesson designs and build their skill in anticipating what to expect when they enact their lessons with students. However, open questions remain about how teachers leverage resources from co-design experiences and coordinate them with their internal design resources (as theorized by Brown's pedagogical design capacity). Furthermore, co-design for GenAI curriculum integration remains underexplored, particularly with respect to professional learning activities that give teachers design authority. Our study investigates these areas and contributes an in-depth examination of teachers' design resources as they integrate GenAI into their teaching practice. The research questions that we explore are, *"How are teachers integrating GenAI into their disciplinary lessons?"* and "*What lesson activities did teachers design and teach about GenAI?"*

Methods

We present descriptive case studies of June and Margot (pseudonyms), two high school teachers who participated in our co-design partnership and sought to teach their students about GenAI. The following sections provide further detail of the research context, including researcher positionality, professional learning (PL) activities embedded within our co-design partnership, data collection, and data analysis methods.

Researcher positionality

Our team consists of faculty members, support staff, and graduate students with knowledge of AI, experience developing K-12 curricula, and classroom teaching. The first author is a former high school mathematics teacher and computer science graduate whose program of research focuses on AI, teaching, and curriculum. The second author is a postdoctoral scholar with computing expertise who studies AI and data science education. The third author was an advanced graduate student at the time of this writing, is a former high school English teacher, and now works at an AI technology

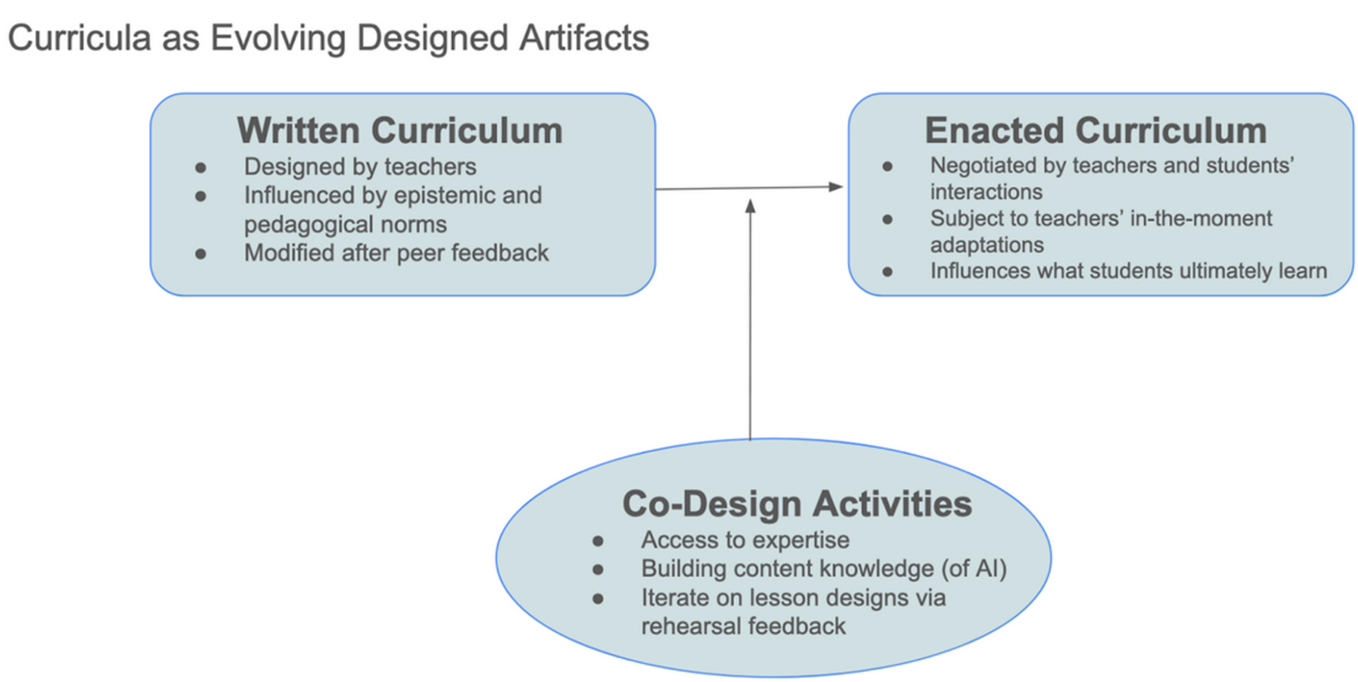


Figure 2. Co-design activities as mediating factors of curriculum enacted with students.

company. The last author is a mid-career faculty member who manages the CRAFT (Classroom-Ready AI resources For Teachers) project and has substantial experience in AI and STEM education. Taken together, our collective subject-matter knowledge and teaching experience enabled us to facilitate CRAFT sessions that complemented the diverse forms of knowledge and expertise that teachers held.

Research context

High school teachers were recruited to participate in a GenAI curriculum co-design effort, CRAFT. We aimed to study (a) how and to what extent activities in co-design supported teacher learning about AI and GenAI, and (b) how teachers designed lessons that integrated GenAI into their teaching practice.

Five teachers were selected to participate in CRAFT based on their interest in learning about AI, integrating GenAI into their lessons, and their availability to participate in a year-round fellowship in 2023–24. All were familiar with commercial GenAI tools (such as ChatGPT) and some had formal AI learning experiences (for instance, June completed a Coursera course about AI), but none had experience teaching with or about GenAI. The cohort, which included the five teachers, two researchers who facilitated CRAFT's activities, and one support staff member with knowledge of the research project, met for 60–80 min from October-May on Zoom for six sessions. Holding six sessions gave us sufficient time to enact our PL and lesson co-design goals with teachers and take up as little of their free time as possible. All sessions took place on Saturday to accommodate time zone differences and teachers' schedules.

CRAFT

CRAFT's professional learning curriculum was composed of synchronous and asynchronous activities. During synchronous Zoom sessions, we aimed to build teachers' knowledge of AI and GenAI tools, then prepare them to teach with or about it. Our first two sessions focused on building teachers' AI knowledge and situating it within classroom activity. In the first session, teachers examined AI standards from AI4K12.org (Touretzky et al., 2023) and found commonalities with curriculum standards from their discipline. In the second session, they reflected on AI and ethics by discussing academic integrity in the context of writing an essay about Shirley Jackson's "The Lottery" with a chatbot. Teachers rehearsed a portion of the GenAI lesson that they designed and intended to teach in sessions three and four, and shared student work and teaching reflections in sessions five and six (see Table 1). Teachers completed asynchronous tasks outside of Zoom PL sessions (see Figure 3).

The primary aim of CRAFT was for teachers to design then enact one lesson that taught students about GenAI, broadly conceived. We suggested that teachers design either

Table 1. Summary of CRAFT professional learning activities.

Session Number	PL Goals	PL Activities
Session 1 (Oct.)	• Help fellows make connections between AI standards and their own curriculum	1. Icebreaker: Exploring Image Generators 2. Research and scheduling logistics 3. 5 Big Ideas of AI (Touretzky et al., 2023) Curriculum Crosswalk
Session 2 (Dec.)	• Deepen understanding of what AI is/is not • Consider multiple perspectives on ethical and responsible use of GenAI tools in school	1. Icebreaker: AI or Not AI? 2. Ethics of Generative Writing 3. (Asynchronous) AI Lesson Critique
Session 3 (Jan.)	• If teaching: practice one component of your lesson, receive feedback, and reflect	1. Icebreaker: AI Discoveries 2. GenAI Lesson Rehearsal (2 fellows) and Peer Feedback
Session 4 (Mar.)	• If observing: give feedback to the fellow to aid their lesson implementation	1. Icebreaker: AI in the News 2. GenAI Lesson Rehearsal (2 fellows) and Peer Feedback
Session 5 (Apr.)	• Reflect on learning from enacting lesson and analyzing student work	1. Icebreaker: Resources for AI Learning 2. GenAI Lesson Rehearsal (1 fellow) and Peer Feedback* 3. GenAI Lesson Debrief and Reflection (2 fellows) and Peer Feedback
Session 6 (May)		1. No icebreaker due to time constraints 2. GenAI Lesson Debrief and Reflection (3 fellows) and Peer Feedback 3. PL Closure: What do you wish your colleagues knew about GenAI?

**Notes:* One fellow needed to rehearse their lesson during session 5 due to illness.

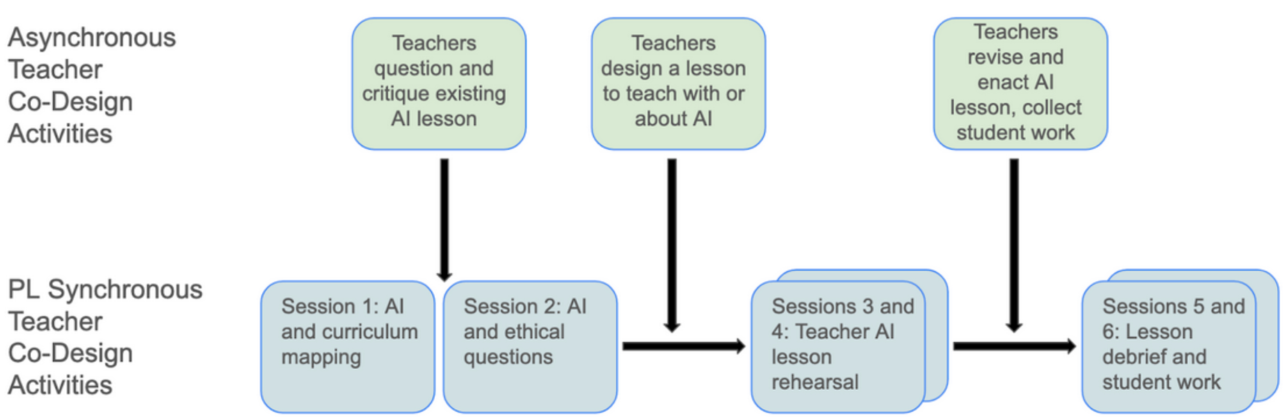


Figure 3. CRAFT's synchronous and asynchronous professional learning activities.

a lesson that teaches students about how AI or GenAI tools work (e.g., the process of training a model) *or* a lesson that integrates concepts about AI into their subject areas. To aid their preparation, we asked them to critique an example AI lesson plan created by a previous CRAFT fellow by adding questions and comments to its documentation. By showing teachers a worked example of a GenAI lesson, we hoped to stimulate their thinking about activities that could be compatible with their teaching practice.

Data collection

Our data sources consist of PL recordings, participant interviews, and teachers' lesson design artifacts. We obtained institutional approval and teachers' permission to audio and video-record each Zoom PL session and to retain written records from Zoom's chat log. We auto transcribed each video recording and verified by hand that each transcription was accurate. We collected exit tickets (open-ended reflection questions about teachers' reflections, questions, confidence, and goals, distributed through Google Forms) at the end of each PL session. We also retained teachers' CRAFT application materials, where we asked about their learning goals, motivation to participate, and what they hoped to learn.

Two researchers interviewed each teacher for approximately 90 min at the end of CRAFT. Interviews took place within 2–3 wk of the final PL session. We wished to provide teachers with time to reflect on their experience and form their own interpretations of the co-design, and simultaneously increase the likelihood that teachers would remember their experience (Nardon et al., 2021). The interviewing researcher asked teachers about their goals, the relative value of PL activities, how they designed and revised lessons, and what they learned about teaching with GenAI. Each interview was autotranscribed then corrected for errors by one researcher, who listened to each recording and corrected errors based on her familiarity with the participating teachers' voices.

We collected all lesson materials that were designed by teachers. This included teachers' written lesson plans on CRAFT's template and supplementary materials that teachers used to enact their lessons. Most teachers created Google Slide decks and student-facing written activity guides on Google Documents to accompany their lessons. We noted which GenAI tools teachers used and their function (e.g., image generator versus text generator). Finally, we noted instances in video recordings from PL sessions five and six where teachers reflected on their lesson enactments and their perceptions of what students learned.

Data analysis

Our data analysis followed an iterative process and was composed of several phases. First we selected the subgroup of teachers to compare. We gathered their lesson materials and performed content analysis (Krippendorff, 2018) to determine how teachers designed lessons to teach about GenAI, and used thematic analysis (Saldaña, 2012) to characterize their perceptions of student learning and resources that influenced their lesson designs. We built analytic cases (Yin, 2009) of each teacher that illustrated the full design arc of their lessons. We discuss each phase in detail.

Phase 1: Case selection

We selected June, an AP Research methods teacher, and Margot, an AP English Language and Composition teacher, as cases to study in greater detail. Both had 3–4 years' of teaching experience during our study. We selected them because they designed lessons that foregrounded contrasting approaches to teaching with GenAI, and they drew from different pedagogical resources to design their lessons. Their different approaches represent contrasting models for integrating GenAI tools into the classroom.

June taught ninth graders on the U.S. East Coast in a large, urban school district that received Title 1 funding. Student GenAI use was banned in June's district during the year she participated in CRAFT. However, June wrote in her application materials that teachers should grant students opportunities to learn about, "AI hallucination and possible benefits to their students, and… their future career development." June learned about GenAI tools prior to CRAFT through a Coursera course titled, "AI for Everyone." She reported frequently using ChatGPT to prepare lesson plans and worksheets.

Margot taught eleventh graders in a large, suburban school district on the U.S. West Coast. She held professional curriculum authoring experience prior to CRAFT and experimented with creative writing and chatbots during the previous summer. Margot maintained a vested interest in developing students' GenAI skills:

> …AI and machine learning will dramatically impact the skills and literacies all students will need to obtain well-paying jobs in the future and to be conscious, considerate informed citizens… it's important for students to learn how to use AI in a way that will support their writing skills and development of critical thinking skills rather than using it to replace the learning of these skills. (Margot's CRAFT application).

Phase 2: GenAI lesson plan content analysis

We identified all data sources that directly and indirectly provided evidence of how teachers' GenAI lessons materialized. This included all lesson materials from teachers, PL curricular materials, transcripts from CRAFT sessions where June and Margot rehearsed (session 3) and reflected (session 5) on their lesson and students' work, interview transcripts, and exit tickets collected after June and Margot's lesson enactments (sessions 5 and 6).

We performed content analysis (Krippendorff, 2018) on the primary documents that June and Margot designed and used to teach their lesson: written plans, Google Slides, supplementary activities guides on Google Documents, and student work. Their lessons approximately followed the following categorical components: opening, main activity (which June and Margot rehearsed with the CRAFT cohort in PL session 3), synthesis activity, and extension. We partitioned teachers' lesson materials into these four components. To supplement our understanding of what teachers actually taught during their lessons, we used CRAFT recordings, participant interviews, and exit surveys to discern how teachers enacted lessons beyond their written materials. We then developed a deductive coding scheme to systematically examine how teachers integrated GenAI into each lesson component.

Based on how we framed teachers' task of designing an AI lesson at the end of CRAFT session 2, we conjectured that they would either (1) teach GenAI in ways that supported student learning *within their subject area*, or (2) teach a lesson about *GenAI as a technological object*: either as a standalone subject (e.g., how model training works) or as embodied within novel tools (e.g., exploring an image generator). Each researcher analyzed June and Margot's lesson plans individually to determine the role of GenAI (*subject-specific* or *object-of study*) in each activity that contained it (some activities did not). For instance, we coded an activity where Margot's students gave ChatGPT feedback to improve their writing as an instance of *subject-specific* GenAI integration, because the aim was for students to reflect on high-quality aspects of human-generated texts. We coded an activity where June's students reflected on the affordances and limitations of image generators as an instance of *object-of-study*, because students explored how well the tools performed independent of June's subject discipline.

We then met collectively twice to discuss and resolve disagreements. Between the first and second meeting, we realized that we needed to familiarize ourselves with the AP Research Capstone and AP English Language and Composition curriculum standards and practices (The College Board, 2024) so that we did not incorrectly juxtapose our own conceptions of research and writing with the curriculum standards described in AP Research Capstone. We continued discussing the role of GenAI in each curricular activity in both teacher cases until we reached intersubjective agreement (Krippendorff, 2018).

Phase 3: Triangulation with a subset of CRAFT data sources

Next, we identified all data sources where teachers described how they designed their lessons. These include data sources from CRAFT sessions 3 and 5, teacher interviews, and the final two exit surveys. We considered two categories of design resources that could have potentially supported teachers' lessons: CRAFT features, such as members of the community or PL activities, or teachers' individual pedagogical design capacities as theorized by Brown (2002). Each researcher individually reviewed the data sources and identified utterances where June or Margot described their lesson design. We collectively discussed and reached intersubjective agreement on which design resources influenced their lesson designs. In many cases, more than one design resource informed teachers' lessons. For example, when June spoke about how she used feedback from rehearsal to re-organize her lesson's activities, we identified that as an instance of CRAFT community members and PL activities. We met and continued discussion until we reached agreement on (1) the corpus of data sources where teachers spoke about their lesson designs and (2) design resources that supported teachers' ability to create GenAI lessons.

Following that, we applied open coding (Saldaña, 2012) to data sources that supported teachers' characterizations of what students learned about GenAI after their lesson enactments. This methodology is appropriate because we did not have a strong preconception of what teachers desired their students to learn about GenAI. Each researcher individually open-coded data from CRAFT session 5, the final two exit tickets, and teachers' interviews. We then met for several rounds of discussion to first develop conjecture maps (Sandoval, 2014), then themes, about how June and Margot described their students' GenAI learning (see Figure 4).

Phase 4: Case development

We iteratively developed a series of analytic memos that answered our research question for both teachers (Creswell & Creswell, 2022). We used findings from our collaborative research sessions to create narrative descriptions of June and Margot's lesson enactments (Yin, 2009). When constructing these descriptions, we took into account primary and supplementary data sources to illustrate, as closely as possible, what student activity occurred during the lessons. We reviewed our memos iteratively as a team and performed data reduction (Krippendorff, 2018) to systematically construct narrative cases of June and Margot's lessons. The cases form the basis of our findings.

Findings

We describe the lessons created by June and Margot using findings from our analytic cases. We present June's lesson first, followed by Margot's, and compare features of their lessons and enactments in the discussion.

June's lesson: How can we leverage GenAI tools to enhance our learning experience?

June sought to teach how GenAI can potentially support students' research activities. Her lesson was exploratory in nature. She instructed students to analyze either a text chatbot or an image generator, evaluate its usefulness, and use it to create an artifact that served as the basis for asking a STEM-based research question. June hoped that students would draw connections between the affordances of GenAI tools and the process of doing research. At the same time, she aimed to draw students' attention to constraints of GenAI tools, both in the context of the research process and generally as a class of tools that intersect with learning writ large. June's goal was for students to identify where GenAI tools can aid their research (and by extension, their learning) versus aspects of research where human judgment, cognition, and problem-solving skills are needed.

June's lesson activities

June began with a short activity that introduced AI's generative features. Students examined two images then responded to the questions, "Can you tell the difference between which image was real and which one was generated by AI? What clues made you choose that image as fake?" Following that, June asked students, "If you were an engineer, what

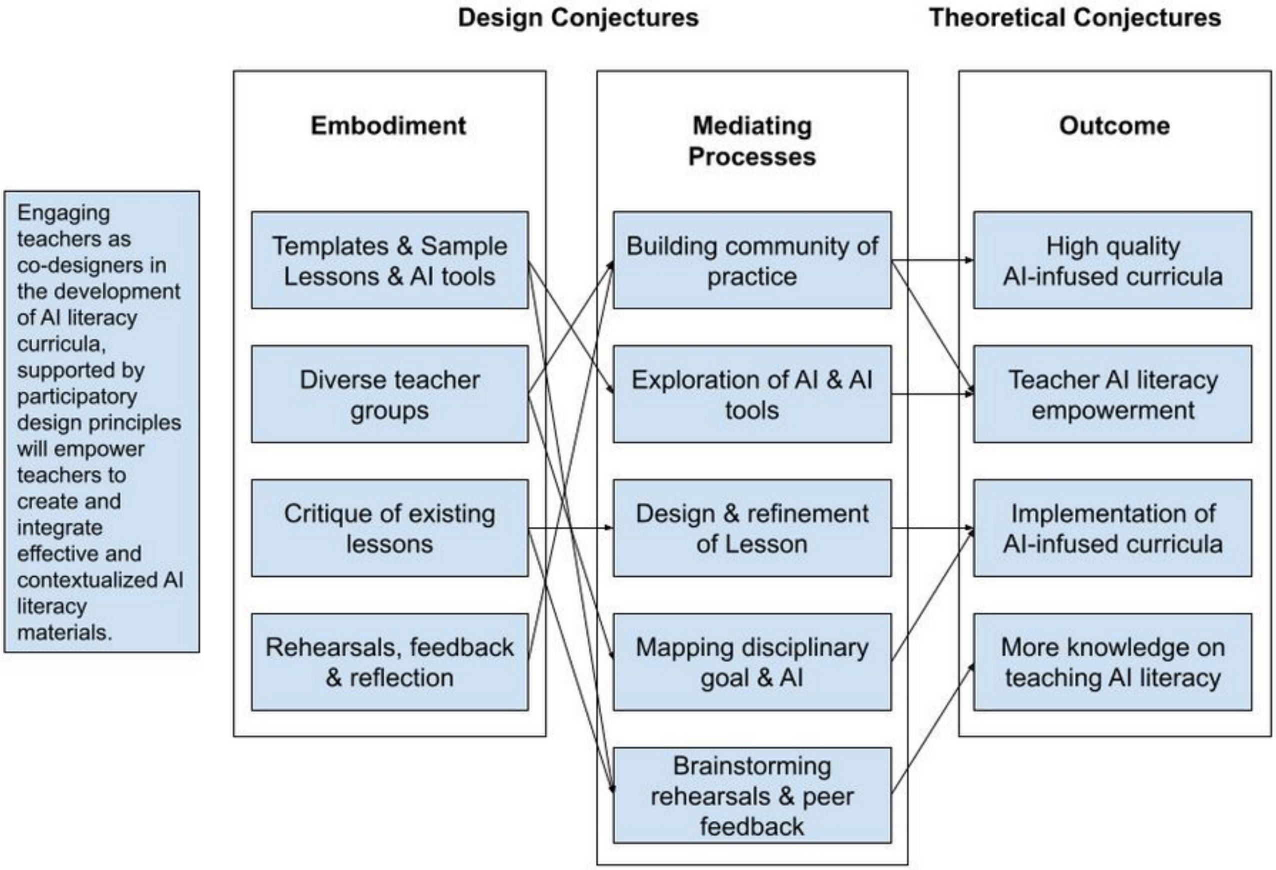


Figure 4. Conjecture Map of CRAFT.

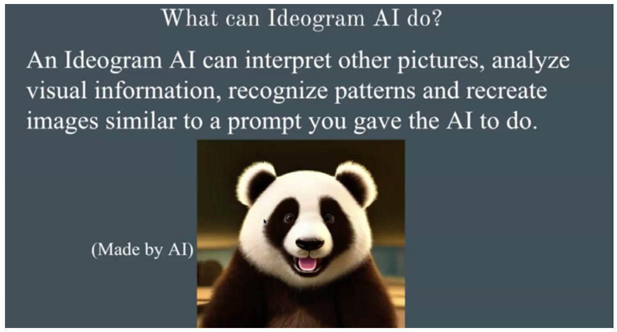
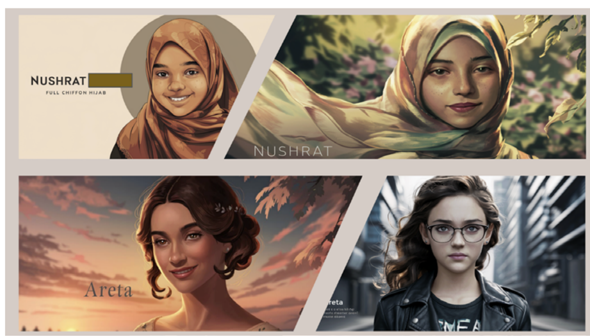
functionalities would you aim to design for AI to support high school students' learning?" June's aim was to introduce GenAI systems and output as objects of study. June mentioned in her CRAFT application materials that many mainstream GenAI tools, such as ChatGPT, were banned in her district. It is possible that June designed the opening activity to collectively introduce students to GenAI and focus their attention on features that could support learning.

Next June launched the main lesson activity. She framed the activity as an in-depth study where students would choose one of two GenAI tools that she selected in advance for them: Ideogram (an image generator) or Liner (a text generator). Students experimented by prompting their GenAI tool with queries of their choosing and answered questions about its output in a "Know, Want-to-Know, Learn" graphic organizer. The substance of June's questions asked students to reflect on the affordances of their GenAI tool ("*Which functions of this AI tool do you believe are beneficial for enhancing the learning experience?*"), limitations of their GenAI tool ("*Can you suggest any improvements that would make the integration of this AI tool into your learning experience?*") and their experience as learners with the GenAI tool ("*Did the AI tool provide personalized assistance in your learning journey?*"). This activity positioned GenAI as an object of study. Students were instructed to observe features of GenAI output by exploring their tool and reflecting on what they wished to learn from it. June wrote the questions that accompanied this activity to highlight GenAI's potential impacts on learning as a whole, rather than focusing on aspects of learning specific to students' abilities to conduct research.

Following the exploration, June told students to create a GenAI product with their tool from the previous activity and present it to the class. Students were instructed to share both their GenAI products and the prompts that they used to create it. In the CRAFT PL session that followed June's lesson enactment, she explained that she held whole class discussions where students made connections between the prompt and GenAI output. For instance, when June's students prompted Ideogram to generate an image of a "person named Nushrat," Ideogram generated two images of a human with feminine features wearing a hijab (see Table 2). We identified this activity as both object of study and GenAI within June's discipline. By holding class discussions about the relationship between GenAI input and output, June enabled students to notice patterns between the language in the prompt and inferences made by the underlying AI model. We also characterized this activity as GenAI in research because the "T" in the AP Research QUEST framework (Question, Understand, Evaluate, Synthesize, Transmit) implores students to practice communicating and disseminating findings as part of the research process (The College Board, 2024).

June's lesson included an extension activity for students who finished their GenAI product before the whole class discussion. Students were instructed to, "Generate one or more scientific questions that you have about AI, then…go

Table 2. An overview of the curriculum and enactment of June's lesson.

<table>
<tr><th>Lesson activity</th><th>Curriculum artifacts</th><th>Student learning
(as described by June)</th></tr>
<tr><td>Opening Activity: What does AI generate, and how can you imagine it helping your learning?</td><td>
</td><td>Students said GenAI images were "too perfect." For instance, in the panda image, they said its fur was "too white," as real pandas often roll around in the dirt and have off-white fur.</td></tr>
<tr><td>Main Activity: Exploring a GenAI tool and reflecting on its affordances and limitations</td><td><table><tr><th>Questions</th><th>Answer</th></tr><tr><td>Did this AI tool meet your expectations? Why or why not?</td><td></td></tr><tr><td>How would you describe your overall experience with this AI tool?</td><td></td></tr><tr><td>Did the AI tool provide personalized assistance in your learning journey?</td><td></td></tr><tr><td>Which functions of this AI tool do you believe are beneficial for enhancing the learning experience?</td><td></td></tr><tr><td>Can you suggest any improvements that would make the integration of this AI tool into your learning experience?</td><td></td></tr></table></td><td></td></tr>
<tr><td>Synthesis Activity: Crafting and presenting a GenAI product</td><td>
</td><td>Students realized that AI systems can make inferences about their cultures and identities from their names.</td></tr>
<tr><td>Extension Activity: Conducting independent research on AI in school library</td><td>- If you complete this project already, Start to generate one or more scientific questions that you have about AI. Then, go to check out the one relatable AI research journey paper from school library database.</td><td>Students began to research AI on their own as part of the synthesis task.</td></tr>
</table>

Note: Images in the center column, row 1 and 3 were generated by Ideogram.

check out an AI research paper from the school library" (June's Lesson, Slide 8). It is unclear how many of June's students worked on the extension activity. However, June mentioned during her lesson debrief that students "…actually went ahead and did the research by themselves, and incorporated it into their presentation" (CRAFT session 5), suggesting that some students completed the research extension activity as part of their GenAI product presentation. We labeled this activity as both AI as an object of study and AI in research, as students were asked to research AI as a topic and understand AI by reading more about it (the "U" in the AP Research QUEST framework).

June's description of what students learned

June reflected that students were both curious to explore GenAI tools as objects that could support research and simultaneously cautious about inferences that GenAI tools make when generating products from human descriptions. She observed that "every student brings their own perceptions and sets of skills when it comes to using AI" which created opportunities for students to gain multiple perspectives on how GenAI tools behaved in different real-world contexts. We saw this play out during CRAFT session five, when June presented student work that evoked questions about GenAI, culture, and identity (see Table 2, row 3). Students who presented this project stated their goal: "To see how well the Ideogram can create images that resemble us and identify the features it picks up from our name." Using prompts like "Person named Nushrat" and "Person named Areta," the group discovered that, for "Nushrat," Ideogram generated two artistic depictions of a woman wearing a hijab (Table 2, row 3), and two light-skinned women with wavy brown hair for "Areta." June's students concluded that Ideogram was likely trained on data representative of people like themselves, as in their opinion, it accurately inferred that "Nushrat" would be wearing a hijab.

This interaction exemplified students' emerging abilities to question and evaluate GenAI's functionality.

June shared that the GenAI product creation activity naturally led to student reflections on issues concerning GenAI and society, even without her explicit guidance. Students raised ethical questions about data privacy and ownership when reacting to the contents of their peers' presentations. She elaborated:

> One of the things that is most fascinating to me is, I didn't talk about ethics or ethical question related to AI, especially Ideogram. One of the groups went into it, and then kind of [asked], 'Oh, how come we type [a first name], and then [Ideogram] can generate an image that looks like [someone we know]? They portray our personal life and took our … data and then generate images that exactly look like us.' (June, CRAFT session 5 recording)

Additionally, June described how students made connections between the inputs (prompts) they provided to GenAI and its subsequent output. She noted how their interactions fostered both creativity and critical thinking during whole-group discussion. For example, she characterized students' interactions with chatbots (during the lesson synthesis activity, see Table 2, row 3) as instances of creativity with music: "Some people actually copy the lyrics. They got it from one of the song[s] like the original song, and then they put it into [the] AI generator, and then they generate a new song, which is kind of cool." While students in this group demonstrated curiosity and creative use of GenAI, June also encouraged the whole class to reflect on matters of copyright and ownership. She challenged them to consider who owns songs and lyrics that are GenAI products. Furthermore, she discussed with students how unresolved issues of ownership in creative pursuits might influence the ways they craft prompts and distribute GenAI products responsibly.

Finally, June articulated that it was important for students to learn how to identify material produced by GenAI. In her opening activity where students differentiated between AI-generated and non-AI-generated images (see Table 2, row 1), June highlighted how collaborative discussions led to shared insights about cues for identifying AI-generated visuals. She noted:

> … Students started generating what … [an] AI-generated image looks like. They said, 'If you take a look on this panda, their eyes are kind of symmetrical. And then, pandas never get this pure white fur in reality, they are always getting dirty… the fur [turns] into like yellowish, brownish colors, you know, because they like to mess it up. (June, CRAFT session 5 recording)

Through exchanges like these, the students developed their ability to evaluate GenAI products by honing their observational skills. They examined the image of the panda and compared it to their mental models of visual features that they expected pandas to exhibit "in the wild." By drawing contrasts between pandas in the GenAI image and real-world pandas, students became more attuned to visual features characteristic of GenAI.

June concluded that student-driven GenAI explorations contributed significantly to their learning. She said, "I feel like teachers shouldn't set a very rigid goal… it's more fun to also learn from students and then let them explore and … choose their own interest project … This seems more engaged than the traditional way of teaching" (June's interview) In sum, she believed that her lesson should contain problems that students find personally meaningful and interesting, because they provided opportunities for whole-group discussion and ethical debate. Her emphasis on creativity, collaboration, and ethical reflection supported students in developing critical thinking and observational skills that extend beyond her AP Research class.

How June described resources that influenced her lesson design

June reflected that CRAFT activities and discussions among cohort members helped to shape and refine her lesson. She summarized structural changes that she made to her lesson between her rehearsal and class enactment:

> I removed the activate [AI] prior knowledge, removed the… frontload of giving them the ethics perspective before we jumped into our activity. I actually placed exploration first, just to give them a chance to mess around [*with GenAI tools*]. (June, CRAFT session 5 recording)

June expressed interest to design a lesson that enabled students to explore GenAI tools and connect their affordances to the process of learning. After rehearsing with the cohort, June narrowed the scope of her lesson from a constellation of GenAI topics (exploring tools, conversations about ethics, sharing their prior experiences) to a connected sequence of activities that built upon one another. June reflected during her post-interview that feedback from the cohort impacted her lesson design by "support[ing] me to… make my lesson more organized, and make it more logical and better than before." (June, interview).

June reported that CRAFT PL activities likewise supported her learning, and by extension, her lesson design. She said:

> …because as I didn't use GenAI tools beforehand…practicing it as a student helped me to understand how students would fill in [*AI prompts*], and then how they were going to think. And then…what kind of conversations they may start to generate when they were using those tools. (June, interview)

As someone who regularly used GenAI tools for lesson planning, but never taught with GenAI in the classroom, June reported that PL activities that instructed teachers to use GenAI "like one of your students" helped her to develop insights about what students experience when they encounter GenAI. As a result, the PL activities allowed her to practice anticipating student conversations about GenAI and how she could respond to and support students.

June's teacher and curriculum design resources (Brown, 2002) informed the overall focus of her lesson. As an AP Research teacher, June was responsible for developing students' skills as independent researchers, regardless of their chosen topic (The College Board, 2024). We observed that her lesson activities adhered to curriculum "big ideas" from AP Research: that students practiced research by asking questions, understanding new source materials, and communicating their findings. Her pedagogical content knowledge of the subject of research is reflected in her lesson design, where she first narrowed students' focus to one of two GenAI tools (to purposefully constrain their choices), then enabled them to explore GenAI openly and in alignment with their research interests. June said,

> I want them to utilize [GenAI] first…kind of loading them with prior knowledge. Then I want to load them with the ethics part, cautious of, what are some AI policies? And then [they] should be progressing towards [AI] like a co-designer, as after they understand… how they want to use it, in their lives or academically. (June, interview).

In summary, June primarily used GenAI tools as an object for students to examine in the context of AP Research. She gave them latitude to explore ideas of their own interest and develop inquiries, grounded by their use of GenAI. Although she described her initial knowledge of GenAI as limited, she drew heavily from support and feedback from the CRAFT cohort to iterate upon her lesson and build confidence to teach about GenAI. While she did not intentionally plan for student discussions of ethics and AI, those issues came up organically as a result of students' research findings. June reflected on students' personal AI use and policy implications as a result of her lesson enactment.

Margot's Lesson: How can we use GenAI to make us better writers?

Margot enacted her CRAFT lesson during one 60-minute class period, although she mentioned that it was one lesson within a longer sequence she designed about AI literacy and writing skills. One of her primary aims was for students to identify differences between human and AI-generated writing. This was important to Margot because she aimed to position students with agency over GenAI tools. To convince students that they held writing expertise over GenAI, Margot encouraged students to identify weaknesses in AI-generated writing. At the same time, she created opportunities for students to notice positive aspects of AI-generated writing and incorporate them if they desired. Finally, Margot wished to inform students about opposing perspectives on GenAI tools from popular media. She selected articles for students to analyze that gave contrasting perspectives and encouraged students to use evidence from the texts to make their own judgments about GenAI tools for learning.

Margot's lesson activities

Margot opened with a "quick write" and a small group activity that did not directly involve GenAI. She first asked students to reflect individually by answering, "Does good writing follow a formula? Why or why not? What distinguishes good writing from great writing?" Students did not share their responses with each other or Margot. Following this, Margot gave small groups of students three samples of writing. She instructed them to grade each one along several dimensions - thesis, evidence, commentary, sophistication - using the AP English Language essay rubric (The College Board, 2024). Neither of these activities required students to use GenAI tools or have explicit discussions (in the lesson materials) about GenAI. Margot then revealed that two of the three sample texts were AI-generated. She asked students to detect which text was human-created and gather evidence that informed their decisions. We characterized this activity as GenAI as the object of study because students were tasked with identifying features of language that they believed resembled output from GenAI machines.

Following this activity, Margot asked students to answer four reflection questions about AI and human-generated writing (see Table 3, row 3). The first question, "*What was different about the writing published by humans? Why? Justify your answer with evidence*," further probed students to speculate about differences between GenAI and human-generated writing and was similar to Margot's question in the small group activity. Two of the later questions asked, "*What about the AI generated text worked/did not work well? How could you apply this/avoid this in your own writing?*" We labeled these questions as GenAI in writing, as Margot asked students to think about how studying the writing from GenAI could impact their own writing practice.

Margot's final activity asked students to give ChatGPT feedback on how to improve one of their sample AI-generated texts from the small group activity. She then asked them to observe and evaluate ChatGPT's revised output by answering the question, "*How well did ChatGPT incorporate your feedback into its revision?*" This activity was both AI as an object of study and AI in writing. By evaluating ChatGPT's output, students were asked to observe qualities of a "revision" performed by a GenAI agent. By asking, "How well…" students made judgments as to whether the quality of the writing improved, which relied on their knowledge of good writing.

Margot's description of what students learned

Margot believed that students improved their analytical writing skills and acquired experience interacting with GenAI as a writing tool. She described moments in the lesson where students demonstrated their abilities to evaluate and give feedback as they examined the strengths and weaknesses of human and AI-generated writing. Students noticed that GenAI lacks a "human touch," meaning that it lacks knowledge about the real-world context that often allows for shared experiences between writers and their readers. For example, students observed that while human and AI-generated essays both "had some of the required aspects of an essay," the human-written essay was "less stiff" and included more personal and relatable stories where the writer was "yapping about their daily and yearly things" (Margot, interview). In contrast, they felt the AI-generated essays were "straight to the point" but included evidence that was "vague" and lacked counterclaims. These observations sparked discussions about key elements of "great writing," which students had written about in the quick write opening activity. Students concluded that good essays require "counterclaims, specific evidence, and personal experiences for better support" (Margot, interview). Margot felt that students' discussions deepened their understanding of AP

Table 3. An overview of the curriculum and enactment of Margot's lesson.

Lesson Activity	Curriculum Artifacts	Student Learning (as described by Margot)
Opening Activity: Quick write, Does good writing follow a formula? Why or why not? What distinguishes good writing from great writing?	**Quickwrite (3 min)** Do you think good writing follows a formula? Why or why not? What distinguishes good writing from great writing?	Students reflected that great writing contains a style and tone that is unique to the author.
Main Activity: Small groups evaluate three texts (two are AI-generated) with the AP English rubric, and determine which were written by GenAI.	**Step 2**: Guess which one of your 3 essays was written by a human. Which one was written by a human? How do you know? Use evidence to justify your position.	
Synthesis Activity: Students reflect on the characteristics, positive qualities, and negative qualities of AI-generated writing.	**Part III: Reflection questions** Reflection Questions 1. What was different about the published writing by humans? What was similar? 2. Which of the sample essays you read had the strongest writing? Why? Justify your answer with evidence. 3. What about the AI generated text worked well? How could you apply this to your own writing? 4. What about the AI generated text did not work well? How could you avoid this in your own writing?	Students shared that GenAI writing is more organized than human writing, but is boring, has no stylistic elements.
Extension Activity: Students give ChatGPT revision advice and evaluate its performance in creating a revised text.	**Extension Activity: Group or Individual Activity** **Let's give ChatGPT some feedback.** Copy and paste the model text into ChatGPT. Then use the following prompt to provide peer feedback. See what response ChatGPT gives you (Not: This can be a group activity, a whole class activity, or an individual activity). Prompt: Dear ChatGPT, Pretend that you are a student who has written the essay above. Here is some peer feedback we have for you on this essay. Please rewrite the essay to incorporate the feedback we give you.	Students realized that they need to be specific when giving chatbots advice. For instance, in argumentative essays, they need to instruct it to choose a side.

English writing practices and encouraged them to reflect on high-quality features in their own writing.

Margot further reflected on student learning by debriefing how students provided constructive writing feedback to ChatGPT. As they observed how ChatGPT incorporated their feedback, students noted that it often responded in "excessive and exaggerated" ways. For example, when asked to add more transition words, ChatGPT overcompensated by inserting numerous "however(s)" and other transitions. Margot emphasized that while this was not her lesson's primary goal, it became one of the most prominent activities where students pointed out limitations of GenAI writing. She explained during the interview,

> We would tell something like 'add more transition words,' and all of a sudden it would have like 13 [*transition words*] added. It was just the way it reacted to the feedback, that was so excessive, and the students were really entertained by that part.

She believed this experience gave students agency over GenAI tools because it enabled them to celebrate positive aspects of their own writing, such as style, that machines struggled to produce.

Additionally, Margot noted that students were able to identify AI-generated texts despite not being told which were human or AI-created. This demonstrated their growing ability to identify and attend to humanistic features of language – tone, voice, and style – when evaluating writing. Margot remarked during her final interview, "I didn't tell them which ones were human and which ones were AI-generated. I think it was kind of obvious. So, most of them figured it out." Students' abilities to distinguish between the two types of writing further underscored their attunement to analytic features of text.

How Margot described resources that influenced her lesson design

In contrast to June, Margot's lesson design was informed more by aspects within her pedagogical design capacity than features of the co-design. The activities in her lesson that combined GenAI with writing skills were supported by her pedagogical content knowledge. Margot referenced her knowledge of teaching writing or of the AP English Language curriculum in majority of the questions where we asked her to reflect upon her lesson design. She related the design of her lesson activities to her disposition that today's students should know about what GenAI can do, and moreover, have agency to judge the quality of its writing:

> I wanted [*students*] to see that they could critique what the AI produces. And that the AI is not just producing something that

> you have to accept as automatically good. You can look at it, and we can pick it apart and see how it actually isn't writing that well. It's just like, if you're interacting with a peer, they're going to have strengths and weaknesses. And I wanted them to see that because, I don't want them to put ChatGPT on a pedestal and think that's what your writing is supposed to be like. (Margot's interview)

Margot's knowledge of and experience with GenAI prompting also played a role in how much time she spent revising and refining portions of her lesson. When generating the AI sample texts, for instance, she reported spending considerable time and effort modifying her ChatGPT prompts to obtain sample essays with features that appeared to have humanlike qualities:

> I had to spend… I mean, probably a couple of hours…I kept track of every iteration, and I explained what prompts I gave it. Because a bunch of [*prompts*] were duds! Like I gave [*ChatGPT*], "Okay, write an essay about this…" And I didn't tell it to take a side. And then it wrote both sides of the argument. And I didn't show that to the students, but I told them about it…You guys know that you have to take a side, when you write an argument, especially when the rubric says you have to take a defensible thesis. You guys know better than ChatGPT. (Margot's interview)

Margot's experience redesigning sample texts provided her with additional evidence of human authority and agency over artificial agents. By sharing this experience with students, she illustrated a tangible quality within their own writing - knowing to choose a side and defend it with evidence when building an argument - that made it higher quality than argumentative writing from ChatGPT.

Although Margot designed most of the lesson activities herself, she recalled that two important decisions were influenced by members of the CRAFT cohort. She originally planned to show students the human and AI-generated writing separately, but a teacher suggested in the third PL session that Margot allow students an opportunity to guess which writing sample was created by a human. She reflected that by taking the teacher's advice and giving students side-by-side writing samples, students were able to better contrast stylistic features of writing, one of the Big Ideas in the AP English Language curriculum:

> And then someone suggested, why don't you put [*the human-generated writing*] in there with the AI writing to get them to guess?… it allowed them to put it side-by-side and really see the differences. And to see that sometimes the human writing didn't really match what was supposed to be on the rubric. Yet, it was more effective than the AI writing, which might have matched the rubric, but had no voice or personality to it. (Margot's interview)

Margot added the ChatGPT feedback extension activity after her lesson rehearsal with the cohort. She stated that observing cohort members' lesson rehearsals gave her ideas about how to integrate GenAI into her lesson using interdisciplinary teaching methods. By experiencing others' lessons, Margot gained a greater repertoire of ideas about how to apply GenAI when improving aspects of student writing. In summary, Margot already held substantial curriculum design expertise and had a clear vision for what she wished to teach students about *subject-specific* GenAI use in writing. Yet, she adapted portions of her lesson from CRAFT members' feedback in ways that further enabled students to build agency over GenAI.

Discussion

As advancements in GenAI reshape teaching and learning in K-12, teachers are increasingly faced with choices about if and how to design GenAI learning experiences. The teachers in this study demonstrated how they navigated this demand by designing and enacting activities that drew from their own knowledge resources and the collective knowledge of the CRAFT community. June and Margot's cases illustrate a range of activities, uses of design resources, and reflections about the experience of integrating GenAI tools into disciplinary lessons. Though both teachers held 3–4 years of teaching experience during the year they participated in the CRAFT fellowship, their lessons diverged in how they integrated GenAI into activities and how closely they aligned GenAI's uses to their discipline's epistemic norms. We discuss implications of these findings below.

CRAFT as a curriculum mediator

While June and Margot varied in their level of experience with GenAI and curriculum design, both lessons were shaped by their participation in CRAFT. Other teachers in the cohort gave June feedback that helped her refine learning goals and resequence lesson activities. Observing other teachers' lesson rehearsals provided Margot with ideas for smaller adaptations to existing activities that aligned with her GenAI learning goals. CRAFT's co-design environment provided a structured yet flexible setting that allowed teachers to explore GenAI integration using their pedagogical and subject-matter expertise. Rather than prescribing a rigid GenAI curriculum, the co-design process encouraged teachers to be creative, take ownership of lesson design, and develop innovative AI-integrated activities—regardless of their prior expertise with GenAI. Given that teachers' comfort and knowledge of GenAI are expected to vary widely, we argue that future AI-integration efforts in K-12 should consider a teacher-centered approach that prioritizes their pedagogical content knowledge and pedagogical design capacity.

A key contribution of CRAFT was the shared knowledge-building that occurred among researchers and teachers, which served as a foundation for developing teachers' confidence to incorporate novel (and often unpredictable) GenAI tools into their teaching. Though efforts are being made to improve AI curriculum standards and guidelines, such as AI4K12, (Touretzky et al., 2023), a majority of teachers do not have extensive knowledge of how AI or GenAI works (Mulyani et al., 2025), which can discourage them from designing even small activities that integrate it. Yet, through PL activities designed to gradually increase GenAI exposure while leveraging their teacher knowledge, teachers built confidence to teach about GenAI in ways that aligned with their goals, beliefs, and to some extent, their disciplinary curriculum. June's case exemplifies this. While she did not have a technical understanding of how GenAI models function, her participation in CRAFT allowed her to explore GenAI tools, experiment with their affordances, and eventually develop a lesson that met her students' needs and aligned with her instructional goals. This suggests that teachers do not necessarily need to be AI experts to integrate GenAI effectively. Instead, providing a foundational understanding of GenAI—paired with guided exploration and peer support—can empower teachers like June to design meaningful AI-integrated lessons that are relevant to their subject areas (Kong et al., 2024).

Supportive professional learning communities will be crucial as teachers navigate the integration of GenAI into their classrooms. Observing how other educators incorporate GenAI, exchange ideas, and receive peer feedback can help teachers redesign lesson plans, construct classroom GenAI policies that align with their learning goals, and combine GenAI with curricular activities in sensible ways (Reilly, 2017). Both June and Margot adapted their lesson designs based on insights gained from peers during rehearsal. Our CRAFT rehearsal protocol required the non-presenting teachers to "act as students," and thus, June and Margot developed skill in anticipating challenges unique to GenAI that their actual students might encounter. They also brainstormed strategies for addressing areas of challenge with the cohort, which may have contributed to their confidence. This process led June, and to a lesser degree Margot, to refine their lesson plans. June initially included discussions about AI ethics but, after feedback from the cohort, shifted her focus to GenAI's role in STEM research. Margot, on the other hand, expanded her GenAI lesson after discussing with peers, incorporating additional activities that prompted students to critique AI-generated writing.

We believe these structured collaborations and co-design feedback loops are critical, as novel opportunities and challenges are expected to accompany new curricula. However, opportunities for piloting lessons with other educators can help teachers mitigate these issues and support thoughtful design adjustments before implementation in the classroom. Our findings highlight the importance of designing professional learning that emphasizes collaborative, teacher-driven AI exploration (Swanson & Lawrence, 2024). Teachers do not need to be AI specialists to integrate GenAI, but they do need access to structured opportunities that build their confidence, provide hands-on exploration with GenAI tools, and allow for peer-to-peer knowledge exchange. Giving teachers access to these resources may increase their pedagogical design capacity for teaching about GenAI, and by extension, may increase the likelihood that they will design and enact GenAI lessons.

Our findings also raise questions as to whether traditional frameworks for modeling teachers' pedagogical design capacities require an expansion in some cases. While Brown's conception of *pedagogical design capacity* captures how teachers' internal resources shape lesson design (2011), it does not fully account for the influence of external design resources, such as communities of practice, online and digital resources, and AI agents. In the case of GenAI, and particularly for teachers with low prior knowledge of AI, we believe that external design resources, such as participating in a community of practice, are essential for teachers to develop the confidence and skill needed to teach with these novel technologies.

Lesson design with or about GenAI is not a clear bifurcation

We hypothesized that CRAFT teachers would create lessons that taught their students either about GenAI in their discipline (*subject-specific*), or about GenAI as a new tool (*object-of-study*). However, by analyzing June and Margot's lesson activities, we realize that teaching with or about GenAI may be an artificial distinction. In Margot's case, in order to teach students how GenAI could improve their writing, she first needed to teach students how to identify AI-generated writing, a skill unique to studying AI itself. In June's case, she wished to teach students about AI in relation to their learning, but in doing so, integrated her discipline by asking students to research a topic of their choosing with GenAI. From these findings, we gather that teaching with or about GenAI in the classroom is not necessarily a choice about what design and integration approach to take, but rather what aspects of AI to foreground in lesson activities.

This finding holds implications for the contents and design of professional learning activities about GenAI lesson design and integration. PL should incorporate multiple types of interpersonal activities that enable teachers to explore GenAI, develop conjectures about how it relates to their practice and disciplinary content, and receive feedback from others. This stands in contrast to typical forms of technology professional development; for instance, when new tools are demoed for teachers or PL that are single occasion meetings and conversations among teachers may not give opportunities to explore, revise, and reflect on technology. Both June and Margot held AI experience prior to CRAFT, yet their conceptions about how to teach GenAI crystallized only after they participated in discussions with their cohort. Margot integrated more GenAI activities, for instance, into her lesson after receiving feedback from the group rehearsal. Conversely, June removed *AI-as-object* activities (namely, explicit discussions about AI and ethics) from her lesson upon the cohort's suggestion, and replaced them with activities that extended GenAI as a tool to support STEM research.

A second finding is that even though June and Margot had an *a priori* conception of which AI integration strategy to choose - *AI-as-object* or *AI-within-subject* - neither teacher adhered to their chosen strategy for the entire lesson. In fact, both teachers designed at least one activity that leveraged GenAI as both a disciplinary tool and an object of study at the same time. This implies that PL should contain a robust catalog of activities that prepare teachers for both lesson design circumstances. That is, teachers should know how GenAI relates to their discipline, and they should develop baseline proficiency for using AI tools for instruction. Based on June and Margot's lesson designs, we believe it could be fruitful for teacher learning if PL activities accomplish both at once; for instance, using GenAI prompts to generate and evaluate artificial writing, then evaluating the quality of the writing using human rubrics. Future work could explore the design of PL activities as they relate to teachers' knowledge and curriculum design skills for developing and teaching AI lessons in their subject areas.

Teacher resources shaped lesson design and implementation

Our findings indicate that teachers' lesson design and enactment were shaped by both their epistemic practices and their pedagogical content knowledge. The skills they intended for students to develop influenced the sequence and structure of their lessons and how they designed activities that enabled students to explore GenAI. While both teachers sought to integrate GenAI meaningfully into their curriculum, their approaches varied based on their subject area, instructional goals, and teacher knowledge resources.

As an AP Research teacher, June aimed to develop students' abilities to question, evaluate, and make informed decisions about when and how to use GenAI tools. For June, GenAI tools themselves were the primary object of study. Her lesson provided students with opportunities to investigate tools, assess their affordances and limitations in different contexts, and form opinions about their usefulness. By encouraging students to experiment with GenAI and report their findings to their peers, June fostered an inquiry-driven learning environment that positioned both the teacher and students as co-learners.

In contrast, Margot's lesson design gradually introduced students to GenAI in the context of high-quality writing. Students interacted with GenAI only at the end of the lesson, and learned that their task contained AI-generated texts at the lesson's halfway point. Margot spent a large portion of the lesson motivating *great* writing and how to identify it with non-GenAI curriculum resources (e.g., the AP English Language rubric). Margot's primary GenAI aim was to give students agency and authority over tools. Unlike June's lesson, where students observed relationships between GenAI output and the prompts that they gave it, Margot's students evaluated the writing quality of GenAI output.

We attributed the degree of difference between June and Margot's lessons to be in some part related to differences in their pedagogical design capacity. Margot was an experienced designer of English curriculum who showed evidence of English PCK when describing factors that influenced her lesson design. She also used curriculum resources in her discipline: rubrics, descriptors of high-quality writing, and references to the writing process. It is possible that she may have integrated GenAI less than June because she spent more time framing why and how GenAI interacts with the practice of writing. June's GenAI as object-of-study approach enabled her to spend more time on activities that directly involved GenAI tools, but comparatively less time connecting it to the epistemic norms and practices of AP Research. It is possible that June drew design inspiration from CRAFT more than Margot, and by doing so, used her teacher and curriculum resources less.

One important implication of these findings is that classroom AI integration requires teachers to repurpose their PCK in service of GenAI learning: holding whole-group discussion, guiding student inquiry, and helping students navigate unexpected or even problematic AI-generated outcomes. Another implication is that teachers may need to shift their mindsets about authority, as we observed both Margot and June co-learn about GenAI's capacities alongside their students. These shifts suggest that teachers who create spaces for shared inquiry and remain open-minded about learning can better provide students with opportunities to explore and develop skill with the technologies that are defining their generation. Moreover, teaching with GenAI will inherently involve taking risks and improvization skills (Brown, 2002), especially if teachers use it in live class demonstrations. As AI integration in education is still in its early stages, collaborative knowledge-building among teachers, school policies that allow teachers to cultivate lesson design skill and apply it to GenAI, and willingness of teachers to take risks will be crucial. Schools and professional learning communities can and should support teachers in making these opportunities available.

Limitations

While our study benefits from a diverse group of teachers in terms of age, discipline, location, and experience with AI, some limitations warrant discussion. First, we clarify that the small sample size of the CRAFT cohort was influenced by our belief and desire to support these teachers fully on an ongoing basis throughout the year-long program, which we could not have done with a larger teacher cohort. For instance, it was only possible to hold online individual lesson rehearsals and debriefs, and give high-quality feedback, with a small group. Naturally, the small sample size limits generalizability of our findings. However, the primary aim of our paper is to illustrate in-depth contrasting cases about what lesson design that includes GenAI within high school subjects might entail through the experiences of Margot and June. It is likely that the individual experiences of the participating teachers, as well as the unique contexts of their classrooms, influenced how they designed instruction about GenAI.

Additionally, no member of our research team observed and recorded teachers' lessons within their classrooms as they were enacted. Observing teachers' lessons would have given us much richer accounts of student's interactions with the designed curriculum. We also may have gained a more complete view of what students learned about GenAI by witnessing class discussions, rather than relying on teachers' recollections of them. However, the teachers in CRAFT were dispersed across four time zones and none were in close geographic proximity to the research team. To mitigate this limitation, we contacted both teachers (and held a follow-up meeting with June) to ensure that we had the most accurate version of their slides, handouts, and activities. We triangulated data between teachers' interview accounts, lesson design artifacts, and PL sessions to produce the most accurate description that we could about the sequence and duration of their lesson activities. We also collected some evidence of student learning in example work that teachers presented in sessions five and six, and we analyzed how teachers interpreted their students' work.

The online format of CRAFT also limited teachers' interactions in the study, which was held remotely and synchronously. Sessions were limited to 60–90 min, and dialogue between participants was moderated by the facilitators. We would have liked for teachers to interact with each other more freely, share ideas outside of the professional learning sessions, and seek feedback on their lessons outside of our official meetings. These constraints did not seem to hinder Margot or June's lesson designs, but we are curious how their lessons may have differed if the cohort was able to meet in person and for longer durations.

Conclusion

Despite the novelty of GenAI tools and the uncertainties they pose for teaching and learning, we have found that teachers can integrate GenAI into their practice through supportive lesson co-design. Our study contributes two cases of lesson design and enactment where teachers taught GenAI within their subject or as an object of study, although some portions of their lessons contained both aspects. We present these cases as evidence that teachers with a range of skills and experiences can design GenAI lessons, and that co-design may serve as a promising avenue for future GenAI integration efforts.

While the form and focus of GenAI lessons vary depending on teachers' subject area, prior knowledge, and instructional goals, our findings suggest that collaborative design and structured peer feedback play an important role in shaping teachers' practices of GenAI curriculum integration. As schools and educators continue to navigate decisions about GenAI, we recommend that teacher educators and districts offer sustained professional learning opportunities that (1) introduce foundational AI concepts that are discipline-agnostic, (2) support teachers in aligning AI concepts and GenAI tools with discipline-specific curriculum standards, and (3) provide examples and time for teachers to experiment, receive feedback, and revise their existing lessons. In addition, our findings suggest a need for research that extends data collection and analysis of teachers' designed lessons into the classroom, as we were unable to gain a complete sense of what students learned (including detailed accounts of AI literacy that they may have gained) as a result of participating in teachers' lessons. Future studies should also examine how teachers' confidence and GenAI integration practices evolve over time, especially as tools and institutional expectations shift.

Acknowledgements

This project was funded by the Stanford Accelerator for Learning. We express our sincere thanks to Christine Bywater, who managed much of the infrastructure for CRAFT professional learning sessions, to June, Margot, and the rest of the teachers in the CRAFT cohort, and to parth sarin, Benji Xie, and Jacob Wolf, who launched a CRAFT pilot study that served as a model for our work in 2022–23.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Victoria Delaney http://orcid.org/0000-0002-1907-0303
Ibrahim Oluwajoba Adisa http://orcid.org/0000-0003-1657-2030
Christopher Mah http://orcid.org/0009-0004-6805-6876
Victor R. Lee http://orcid.org/0000-0001-6434-7589

References

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*, *21*(1), 10. https://doi.org/10.1186/s41239-024-00444-7

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, *2*(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *SSRN Electronic Journal*, *7*(1), 52–62. https://doi.org/10.2139/ssrn.4337484

Brown, M. W. (2011). The teacher–tool relationship: Theorizing the design and use of curriculum materials. In *Mathematics teachers at work* (pp. 37–56). Routledge.

Brown, M. W. (2002). *Teaching by design: Understanding the intersection between teacher practice and the design of curricular innovations*. Northwestern University.

Chiu, T. K., & Chai, C. S. (2020). Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability*, *12*(14), 5568. https://doi.org/10.3390/su12145568

Choppin, J., & Borys, Z. (2017). Trends in the design, development, and use of digital curriculum materials. *ZDM*, *49*(5), 663–674. https://doi.org/10.1007/s11858-017-0860-x

Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Delaney, V. (2024). *The mathematics of facial recognition: Statistics Teachers' adaptations to a machine learning curriculum*. Stanford University ProQuest Dissertations & Theses. 2024. 31591776.

Druga, S., Christoph, F. L., & Ko, A. J. (2022, April). Family as a Third Space for AI Literacies: How do children and parents learn about AI together?. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1–17). https://doi.org/10.1145/3491102.3502031

Esmonde, I. (2014). "Nobody's rich and nobody's poor… it sounds good, but it's actually not": Affluent students learning mathematics and social justice. *Journal of the Learning Sciences*, *23*(3), 348–391. https://doi.org/10.1080/10508406.2013.847371

Kong, S. C., Yang, Y., & Hou, C. (2024). Examining teachers' behavioural intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model. *Computers and Education: Artificial Intelligence*, *7*, 100328. https://doi.org/10.1016/j.caeai.2024.100328

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Lin, P., & Van Brummelen, J. (2021, May). Engaging teachers to co-design integrated AI curriculum for K-12 classrooms. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–12). https://doi.org/10.1145/3411764.3445377

Long, D., & Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–16). https://doi.org/10.1145/3313831.3376727

Lee, V. R., Pope, D., Miles, S., & Zárate, R. C. (2024). Cheating in the age of generative AI: A high school survey study of cheating behaviors before and after the release of ChatGPT. *Computers and Education: Artificial Intelligence*, *7*, 100253. https://doi.org/10.1016/j.caeai.2024.100253

Lee, V. R., Robillard, S., Recker, M., Clarke-Midura, J., & Shumway, J. (2024). Negotiating Inherent Asymmetries of Co-Design: A Case of Integrative Elementary Mathematics and Computer Science Instruction. *The Journal of Applied Instructional Design*, *13*(1), 1. https://doi.org/10.59668/723.13039

Mah, C., Walker, H., Phalen, L., Levine, S., Beck, S. W., & Pittman, J. (2024). Beyond CheatBots: Examining tensions in teachers' and students' perceptions of cheating and learning with ChatGPT. *Education Sciences*, *14*(5), 500. https://doi.org/10.3390/educsci14050500

Mills, K., Ruiz, K., Lee, K., Coenraad, M., Fusco, J., Roschelle, J., & Weisgrau, J. (2024, June) AI Literacy: A framework to understand, evaluate, and use emerging technology. *Digital Promise*. https://digitalpromise.org/2024/06/18/ai-literacy-a-framework-to-understand-evaluate-and-use-emerging-technology/

Moundridou, M., Matzakos, N., & Doukakis, S. (2024). Generative AI tools as educators' assistants: Designing and implementing inquiry-based lesson plans. *Computers and Education: Artificial Intelligence*, *7*, 100277. https://doi.org/10.1016/j.caeai.2024.100277

Mulyani, H., Istiaq, M. A., Shauki, E. R., Kurniati, F., & Arlinda, H. (2025). Transforming education: Exploring the influence of generative AI on teaching performance. *Cogent Education*, *12*(1), 66. https://doi.org/10.1080/2331186X.2024.2448066

Nardon, L., Hari, A., & Aarma, K. (2021). Reflective interviewing – Increasing social impact through research. *International Journal of Qualitative Methods*, *20*, 33. https://doi.org/10.1177/16094069211065233

Ng, D. T. K., Su, J., Leung, J. K. L., & Chu, S. K. W. (2024). Artificial intelligence (AI) literacy education in secondary schools: a review. *Interactive Learning Environments*, *32*(10), 6204–6224. https://doi.org/10.1080/10494820.2023.2255228

Pepin, B., Choppin, J., Ruthven, K., & Sinclair, N. (2017). Digital curriculum resources in mathematics education: Foundations for change. *ZDM*, *49*(5), 645–661. https://doi.org/10.1007/s11858-017-0879-z

Potvin, A. S., Teeters, L. P., Penuel, W. R., & Dimidjian, S. (2024). Humanizing Co-design through attention to educators' affective and relational experiences. *Journal of the Learning Sciences*, *33*(1), 41–79. https://doi.org/10.1080/10508406.2024.2318557

Rawas, S. (2023). ChatGPT: Empowering lifelong learning in the digital age of higher education. *Education and Information Technologies*, *29*, 1–14. https://doi.org/10.1007/s10639-023-12114-8

Reilly, K. A. (2017). Observing peers develops practice, changes culture. *Phi Delta Kappan*, *98*(6), 13–18. https://doi.org/10.1177/0031721717696472

Remillard, J., Cusi, A., Clark-Wilson, A., & Van Steenbrugge, H. (2024). Research on the relationships between mathematics teachers' practices, knowledge, and skills and the use of digital resources. In *Handbook of digital resources in mathematics education* (pp. 1–14). Springer International Publishing.

Remillard, J. T., & Heck, D. J. (2014). Conceptualizing the curriculum enactment process in mathematics education. *ZDM*, *46*(5), 705–718. https://doi.org/10.1007/s11858-014-0600-4

Roschelle, J., Penuel, W., & Shechtman, N. (2006). Co-design of innovations with teachers: Definition and dynamics.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6*(1), 9. https://doi.org/10.37074/jalt.2023.6.1.9

Saldaña, J. (2012). *The coding manual for qualitative researchers* (2nd ed.). SAGE.

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*(1), 18–36. https://doi.org/10.1080/10508406.2013.778204

Severance, S., Penuel, W. R., Sumner, T., & Leary, H. (2018). Organizing for teacher agency in curricular co-design. In *Cultural-historical activity theory approaches to design-based research* (pp. 45–78). Routledge. https://doi.org/10.1080/10508406.2016.1207541

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

Stein, M. K., Remillard, J., & Smith, M. (2007). How curriculum influences student learning. In F. K. Lester, Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, (pp. 319–369). Information Age Publishing.

Swanson, H., & Lawrence, L. (2024). A preliminary model for responsive co-design. In *Proceedings of the 18th International Conference of the Learning Sciences-ICLS 2024*, pp. 1550–1553. International Society of the Learning Sciences. https://doi.org/10.22318/icls2024.879545

Touretzky, D., Gardner-McCune, C., & Seehorn, D. (2023). Machine learning and the five big ideas in AI. *International Journal of Artificial Intelligence in Education*, *33*(2), 233–266. https://doi.org/10.1007/s40593-022-00314-1

Trinter, C. P., & Hughes, H. E. (2021). Teachers as curriculum designers: Inviting teachers into the productive struggle. *RMLE Online*, *44*(3), 1–16. https://doi.org/10.1080/19404476.2021.1878417

Walsh, B., Dalton, B., Forsyth, S., & Yeh, T. (2023). Literacy and STEM teachers adapt AI ethics curriculum. Proceedings of the AAAI Conference on Artificial Intelligence *37*(13), 16048–16055. https://doi.org/10.1609/aaai.v37i13.26906

Walton Family Foundation. (2023). Teachers and students embrace ChatGPT for education. Retrieved from https://www.waltonfamilyfoundation.org/learning/tea chers-and-students-embrace-chatgpt-for-education

Xie, B., Sarin, P., Wolf, J., Garcia, R. C., Delaney, V., Sieh, I., Fuloria, A., Dennison, D. V., Bywater, C., & Lee, V. R. (2024). Co-designing AI education curriculum with cross-disciplinary high school teachers. Proceedings of the AAAI Conference on Artificial Intelligence *38*(21), 23146–23154. https://doi.org/10.1609/aaai.v38i21.30360

Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). SAGE.