International Society of the Learning Sciences

Engaging with Natural Language Processing: An Exploratory Study across Multiple Middle School Grades

Srijita Chakraburty, Indiana University, srichak@iu.edu
Daria Smyslova, North Carolina State University, dysmyslo@ncsu.edu
Cindy Hmelo-Silver, Indiana University, chmelosi@indiana.edu
Krista Glazewski, North Carolina State University, kdglazew@ncsu.edu
Oluwatomisin Obajemu, University of Florida, o.obajemu@ufl.edu
Christine Wise, The Findings Group, christine@thefindingsgroup.org
Toni Earle-Randell, University of Florida, toni@learndialogue.org
Tom McKlin, The Findings Group, tom@thefindingsgroup.org
Kristy Elizabeth Boyer, University of Florida, kristy@learndialogue.org

Abstract: This study examines how an NLP-focused AI curriculum was adapted for 5th and 8th graders with different prior experiences. Classroom video analysis revealed that the 5th-grade teacher used structured, hands-on activities, while the 8th-grade teacher fostered autonomous, peer-supported exploration. Both groups showed significant learning gains. Using activity system analysis, we explored how tools, guidance, and collaboration shaped curriculum enactment within students' Zone of Proximal Development (ZPD). Findings underscore the need for flexible AI curricula, showing that scaffolding and inquiry-driven methods effectively engage students across grade levels.

Introduction

As artificial intelligence (AI) becomes more integrated into daily life, equipping students with foundational AI literacy is increasingly essential (Casal-Otero, 2023). Integrating AI into K-12 education allows students to understand the technologies shaping their world and critically assess AI's societal and ethical implications (Long & Magerko, 2020). With AI-driven tools like smart assistants and recommendation algorithms becoming commonplace, AI literacy is not just a technical skill but a key component of digital citizenship. Recent efforts to introduce AI education in K-12 (Almatrafi et al., 2024) face unique challenges, particularly in middle school, where students' developmental stages and prior exposure influence engagement (Ottenbreit-Leftwich et al., 2022). Younger students often view AI through science fiction and robotic depictions (Williams et al., 2019), highlighting the need for age-appropriate scaffolding (Yim & Su, 2024). Research suggests that hands-on projects facilitate meaningful learning, even for complex topics like machine learning and natural language processing (NLP) (Ng et al., 2023). Early exposure to AI also helps students build technical and critical thinking skills (Glazewski et al., 2023; Lee et al., 2021). While many studies emphasize the benefits of AI education, few examine how the same curriculum adapts across grade levels. This paper addresses that gap by comparing 5th and 8th graders' engagement with an NLP-focused AI curriculum, chosen for its real-world applications in speech recognition, chatbots, and text analysis (Kim et al., 2023). Core NLP concepts like data pre-processing, sentiment analysis, and keyword extraction provide an accessible entry point for students (Alm & Hedges, 2021). However, little is known about how this content is adapted across grades or how teachers adjust instructional strategies to meet diverse classroom needs. This exploratory study examines how the same NLP curriculum was enacted differently in two classrooms. Specifically, we address the following research questions:

RQ1: How do two different teachers enact the same AI and NLP content in their respective 5th and 8th-grade classrooms, and what instructional approaches emerge as they adapt this content for each grade level?

RQ2: What impact do these differences in curriculum enactment have on student learning outcomes in both classrooms?

By comparing these classrooms, we aim to explore how varying instructional approaches influence student engagement and learning outcomes in AI-related tasks across different middle school grades.

Theoretical framework

This study is grounded in sociocultural learning, emphasizing how tools and discourse mediate learning (Cole, 1999). Vygotsky's Zone of Proximal Development (ZPD) (1978) informs our approach, highlighting that learning is most effective when students engage with tasks slightly beyond their abilities, supported by appropriate scaffolding. Inquiry-based learning plays a central role, enabling students to construct knowledge through exploration, particularly in AI and NLP education (Minner et al., 2010). Engaging students early in AI practices, such as data processing and model training, builds technical, critical thinking, and problem-solving skills (Kafai & Proctor, 2022). Tailoring instruction to students' technological familiarity ensures meaningful engagement, from foundational understanding to independent inquiry. To examine how AI and NLP instruction unfolds across grade levels, we apply activity systems theory to analyze how tools, discourse, and interaction shape learning in two distinct classrooms. This framework allows us to contextualize teacher guidance, student collaboration, and technological resources as interconnected components of learning (Engeström, 1999). By framing classroom activity through subject (students), object (learning goals), tools, community, division of labor, and rules, we explore how curriculum enactment differs across 5th and 8th graders. This approach reveals how students' prior experience and exposure to technology influence their engagement, emphasizing the need for tailored scaffolding in AI education.

Methods

This project integrated AI into science classrooms through an inquiry-based NLP curriculum aligned with Next-Generation Science Standards. We co-designed the curriculum and an interactive platform for data preprocessing, keyword extraction, and sentiment analysis (Katuka et al., 2023). To prepare, teachers completed a four-week PD workshop, meeting weekly to build NLP knowledge before teaching the five-day curriculum, which included pre- and post-tests. The study involved a 5th-grade class (n = 16) taught by Kathy and an 8th-grade class (n = 12) taught by Emma (pseudonyms). Student learning was assessed using a validated NLP test ($\alpha = 0.77$). For RQ1, we analyzed classroom video recordings (Jordan & Henderson, 1995) to examine teacher adaptations and student engagement. For RQ2, we used paired t-tests to compare pre- and post-test scores, evaluating the impact of instructional strategies.

Results

RQ1: Enactment of the curriculum

We analyzed 5-minute video segments from two key episodes in both classrooms: (1) how teachers prompted students to share initial AI concepts and guided their understanding, and (2) how they introduced a key NLP concept from the curriculum. These episodes provided insight into students' starting points and how teacher scaffolding and instructional approaches shaped their engagement with AI content.

Episode 1 - Introduction to AI

For the first episode, we analyzed classroom videos capturing how teachers initially gauged students' understanding of AI and its applications in everyday life. The teachers prompted students to share their initial ideas about AI and guided the discussion to connect these concepts to real-world examples.

Kathy's 5th grade classroom: Kathy began by asking, *"What do you think of when you hear the term Artificial Intelligence?" Students initially responded with "Aliens?"* and *"robots."* To push them toward deeper thinking, she asked, *"What kind of robots? Be specific."* This led to responses like *"robots that we program"* and *"to take photos"* before students arrived at a concrete example: *"like chatting on Amazon or a shopping website."* To further clarify AI concepts, Kathy displayed a slide with popular robot depictions, prompting students to discuss whether Wall-E was AI. When one student said, *"It tells you the future,"* Kathy asked, *"It predicts the future?"* leading to a clarification that *"No, the movie is about the future."* She then asked, *"So, is it AI?"* prompting hesitation. A student reasoned, *"Robots can just be machines… not always AI."* This discussion highlighted students' early understandings and the need for further exploration of AI versus general technology.

Emma's 8th grade classroom: Emma's students demonstrated a broader initial understanding of AI, mentioning *"Snapchat AI"* and *"Siri"* as real-world examples. One student even shared, *"I use ChatGPT for homework."* Emma facilitated rather than directed, probing deeper when a student mentioned a *"smart toaster."* She asked, *"What would make the toaster 'smart'?"* leading to the realization that it *"determines the right level of toastiness on its own."* She then showed AI-generated and real images, asking students to justify their classifications. One student noted, *"AI images don't have reflections,"* demonstrating a critical analysis of AI-generated content. Similar to Kathy, Emma used robot depictions, leading to a debate about Star Wars characters: *"Which one has a higher level of AI, R2D2 or C3PO?"* Students reasoned that C3PO *had "thoughts and feelings"* and *"could translate and talk,"* reflecting a deeper grasp of AI's capabilities.

Kathy's structured questioning helped 5th graders refine their abstract ideas into concrete AI applications, aligning with their Zone of Proximal Development (ZPD) by scaffolding their understanding step by step. In contrast, Emma's 8th graders, with more prior exposure to technology, engaged in a facilitative, inquiry-driven discussion, using their background knowledge to critically examine AI concepts. Both approaches demonstrated that students, regardless of age, can engage meaningfully with AI when instructional strategies align with their prior knowledge and learning context.

Episode 2: Data pre-processing

Next, we analyzed classroom videos where students were introduced to data pre-processing, a key NLP concept that transforms text into a simplified form for analysis through three steps: (1) removing stop words, (2) removing punctuation and special characters, and (3) converting text to lowercase.

Kathy's 5th-Grade Classroom: Kathy introduced data pre-processing in a step-by-step manner, making abstract AI concepts more tangible. She connected stop words to writing conventions: *"You know when you write a title, we don't capitalize unimportant words?",* helping students relate to the idea of filtering out less meaningful words. For punctuation removal, she pointed to familiar symbols on their keyboards*: "Look at the top of your keyboard—@, #, $, %, hyphens—these are special characters."* She then explained lowercasing, reinforcing that *"Even if it's the name of a country or city, computers treat it the same way."* For the classroom activity, students took turns crossing out stop words on the whiteboard, following Kathy's instructions. She asked, *"Do you see any uppercase letters? Any more punctuation?"* guiding them to complete the task accurately. While the focus was on applying the steps, there was little discussion on why these processes are important for AI.

Emma's 8th-Grade Classroom: Emma's classroom adopted a more exploratory approach, encouraging students to think critically about why pre-processing matters. She connected stop words to American Sign Language (ASL): *"Words like 'A,' 'THE,' and 'AND' aren't always needed. In ASL, you ask questions with your eyebrows instead of using 'WHAT' and 'WHY.'"* She then explained that removing stop words prevents them from skewing AI's analysis: *"We want an even playing field, so we don't let 'THE' be counted as more important than other words."* For punctuation and special character removal, she explained: *"Anything that's not a letter—numbers, pound signs, exclamation marks—gets removed."* Students worked hands-on with a color-coding task, marking different word types in a Google Doc. When some used the "find and replace" function to speed up the process, Emma reminded them, *"You can't cheat the system—manual pre-processing helps us understand why automation is useful."* Later, they compared their work with AI-generated outputs, leading Emma to ask*, "Notice how the meaning changes when stop words and punctuation are removed?"* helping them recognize the impact of pre-processing on NLP models.

While Kathy's structured approach helped younger students focus on accuracy and procedural understanding, Emma's inquiry-driven approach encouraged critical thinking about why these steps matter in AI. Both approaches scaffolded learning in developmentally appropriate ways, demonstrating how AI concepts can be adapted across different grade levels. Next, using an activity system lens, we examined how the curriculum was adapted in Kathy's and Emma's classrooms, focusing on two episodes: the introduction to AI and data pre-processing. While both classrooms followed the same curriculum, instructional structure and enactment differed, shaped by tools, teacher guidance, and student interactions (Table 1).

Table 1
Activity System Comparison

Component	Kathy	Emma
Tools/ Mediating Artifacts	**Episode 1:** Slides showing robots from movies ("Wall-E") to help students relate abstract AI concepts to familiar images. **Episode 2:** Physical tools like slides projected on whiteboard; references to everyday objects (e.g., special characters on keyboards) to simplify abstract AI concepts.	**Episode 1:** Digital tools like AI-generated and real images to encourage critical analysis of AI-generated content. **Episode 2:** Digital tools like Google Docs and color-coding to simulate real-world AI tasks through text pre-processing.
Subject	5th-grade students with minimal direct exposure to AI, relying on media-influenced concepts of AI ("robots," "aliens")	8th-grade students with more hands-on experience with technology and AI, offering examples of real-world AI applications ("snapchat," "chatgpt")
Object	**Episode 1:** Help students move from abstract notions of AI (e.g., robots and aliens) to concrete applications like chatbots. **Episode 2:** Introduce students to data pre-processing and developing proficiency in basic procedures.	**Episode 1:** Encourage students to critically discuss AI's role in everyday technology, drawing on prior knowledge. **Episode 2:** Engage students in critical thinking about how and why AI systems process data, emphasizing deeper understanding.

Community	Kathy and the 5th-grade students involved in teacher-led discussions.	Emma and the 8th-grade students engaged in independent, discussion-driven learning.
Division of Labor	**Teacher:** Leads structured step-by-step activities, offering explicit instructions and close guidance to move students from abstract to concrete concepts. **Students:** Follow teacher-led, structured activities with close guidance, focusing on procedural accuracy.	**Teacher:** Facilitates open-ended discussions, encouraging autonomy and critical thinking. Provides guidance only as needed. **Students:** Engage in independent, peer-supported exploration with minimal guidance, directed towards critical thinking through collaborative tasks .
Rules	**Episode 1:** Structured questioning to guide students' responses, helping them connect abstract AI ideas to real-world examples. **Episode 2:** Focus on procedural accuracy, ensuring students master the foundational steps of pre-processing.	**Episode 1:** Flexible rules encouraging inquiry and student-driven exploration of AI examples **Episode 2:** Flexible rules allowing student-driven inquiry, with a focus on critical thinking and deeper understanding.

The activity system analysis revealed how classroom dynamics shaped curriculum implementation. In Kathy's 5th-grade classroom, scaffolding focused on foundational skills, using physical tools (e.g., keyboard references) to make AI concepts more tangible. Her step-by-step instruction fostered collaboration and skill development. In contrast, Emma's 8th-grade classroom emphasized independent exploration, using digital tools (e.g., Google Docs) for hands-on learning. Emma played a facilitative role, prompting students to analyze AI-generated data with minimal but targeted support. This analysis highlights how instructional strategies aligned with students' prior knowledge. Kathy's structured approach built foundational AI literacy, while Emma's autonomy-driven approach leveraged students' familiarity with technology for inquiry-based exploration. These findings reinforce the need for adaptive scaffolding, grounded in ZPD principles, to effectively engage students across grade levels.

RQ2: Grade-level instructional strategies and student learning outcomes

Pre- and post-test analyses showed significant learning gains in both classrooms ($p < .05$). Paired *t*-tests confirmed statistically significant progress, though the 8th-grade class's higher initial scores ($M = 4.80$, $SD = 2.77$) limited further growth (post-test $M = 6.70$, $SD = 3.03$, $d = 0.66$). In contrast, the 5th-grade class started lower ($M = 3.13$, $SD = 1.20$) but showed greater improvement (post-test $M = 5.56$, $SD = 2.28$, $d = 1.34$). Despite different instructional strategies, both were effective: Kathy's structured support aided 5th graders with less AI exposure, while Emma's exploratory approach encouraged independent NLP engagement.

Discussion and conclusion

This study demonstrates how an AI and NLP curriculum can be adapted for students at different learning stages, emphasizing the importance of aligning instruction with students' prior experiences and developmental levels. Findings revealed that 8th graders had stronger foundational knowledge of AI than 5th graders, which influenced their engagement during classroom discussions and activities. This highlights the role of prior exposure to technology in shaping students' interactions with AI concepts. Consistent with ZPD-aligned scaffolding, 5th graders benefited from structured support that connected abstract AI concepts, such as data preprocessing, to familiar, everyday experiences. Meanwhile, 8th graders, with greater digital fluency, engaged more autonomously in an inquiry-based approach, allowing for deeper exploration. These findings reinforce the need for thoughtfully designed scaffolding in AI curricula, ensuring all students can engage with complex material in ways that align with their prior experiences. This study also underscores the practical need for curriculum flexibility. The same AI and NLP curriculum was interpreted and enacted differently in each classroom. This adaptability allowed teachers to meet students at their respective levels, demonstrating that a one-size-fits-all approach is inadequate for AI education. These findings align with broader discussions on personalized learning environments, where instruction is tailored to students' readiness and experience levels. Using activity systems as a framework provided insight into how tools, collaboration, and instructional strategies mediate learning. By framing learning as an interconnected system of tools, learners, and goals, we observed how the curriculum was adapted to different developmental stages, shaping students' engagement with AI. This highlights the crucial role of teacher agency in adapting instruction, ensuring that scaffolding and inquiry-driven approaches effectively support diverse learners. As AI continues to shape education, future research should further refine instructional strategies to address diverse learning levels and contexts. This study contributes to understanding how flexible AI curricula can be effectively implemented, reinforcing the importance of tailored scaffolding, inquiry-based learning, and teacher adaptability in AI education.

References

Alm, C. O., & Hedges, A. (2021). Visualizing NLP in undergraduate students' learning about natural language. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(18), 15959-15961.

Almatrafi, O., Johri, A., & Lee, H. (2024). A Systematic Review of AI Literacy Conceptualization, Constructs, and Implementation and Assessment Efforts (2019-2023). *Computers and Education Open*, 100173.

Brown, A. L., Ash, D., Rutherford, M., Nakagawa, K., Gordon, A., & Campione, J. C. (1993). Distributed expertise in the classroom. *Distributed cognitions: Psychological and educational considerations*, 188-228.

Casal-Otero, L., Catala, A., Fernández-Morante, C. *et al.* AI literacy in K-12: a systematic literature review. *IJ STEM Ed* 10, 29 (2023).

Cole, M. (1999). Cultural psychology: Some general principles and a concrete example. *Perspectives on activity theory*, 87-106.

Engeström, Y. (1999). Activity theory and individual and social transformation. *Perspectives on activity theory/Cambridge University Press*.

Glazewski, K., Ottenbreit-Leftwich, A., Jantaraweragul, K., Jeon, M., Hmelo-Silver, C., Scribner, J. A., ... & Lester, J. (2022, July). PrimaryAI: Co-Designing Immersive Problem-Based Learning for Upper Elementary Student Learning of AI Concepts and Practices. In *Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education Vol. 2* (pp. 628-628).

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The journal of the learning sciences*, *4*(1), 39-103.

Katuka, G. A., Chakraburty, S., Lee, H., Dhama, S., Earle-Randell, T., Celepkolu, M., ... & Mcklin, T. (2024, March). Integrating natural language processing in middle school science classrooms: An experience report. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1* (pp. 639-645).

Kim, K., Kwon, K., Ottenbreit-Leftwich, A., Bae, H., & Glazewski, K. (2023). Exploring middle school students' common naive conceptions of Artificial Intelligence concepts, and the evolution of these ideas. *Education and Information Technologies*, *28*(8), 9827-9854.

Kafai, Y. B., & Proctor, C. (2022). A Revaluation of Computational Thinking in K–12 Education: Moving Toward Computational Literacies. *Educational Researcher*, *51*(2), 146-151.

Lee, I., Ali, S., Zhang, H., DiPaola, D., & Breazeal, C. (2021, March). Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM technical symposium on computer science education* (pp. 191-197).

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).

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *47*(4), 474-496.

Ng, D. T. K., Su, J., Leung, J. K. L., & Chu, S. K. W. (2023). Artificial intelligence (AI) literacy education in secondary schools: a review. *Interactive Learning Environments*, 1-21.

Ottenbreit-Leftwich, A., Glazewski, K., Jeon, M. *et al.* Lessons Learned for AI Education with Elementary Students and Teachers. *Int J Artif Intell Educ* **33**, 267–289 (2023).

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019, July). Envisioning AI for K-12: What should every child know about AI?. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 9795-9799)

Williams, R., Breazeal, C., & Gray, J. (2019). Hands-on AI projects for the classroom: A guide for elementary and middle school teachers. *MIT Media Lab*.

Yim, I. H. Y., & Su, J. (2024). Artificial intelligence (AI) learning tools in K-12 education: A scoping review. *Journal of Computers in Education*, 1-39.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Acknowledgements

This research was supported by the National Science Foundation through grant DRL-2147810 and DRL-2147811. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the authors, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.