Fostering Epistemic Insights into AI Ethics through a Constructionist Pedagogy: An Interdisciplinary Approach to AI Literacy

Ziyan Lin, Yun Dai
The Chinese University of Hong Kong, Hong Kong SAR, China
linziyan@link.cuhk.edu.hk, yundai@cuhk.edu.hk

Abstract

There is a growing consensus on the importance of AI ethics in K-12 education, yet effective teaching remains a challenge. AI ethics requires an interdisciplinary understanding of computer science, philosophy, and the humanities, alongside epistemic insights into how AI systems acquire, process, and apply knowledge differently from humans. To address this challenge, this study presents the design, development, and implementation of three theory-informed activities aimed at fostering epistemic insight and ethical understanding of AI among upper primary school students (ages 10-12). Grounded in constructionism, our pedagogical design leverages hands-on experimentation with guided reflection to concretize complex AI concepts. Students examine rule-based, data-driven, and generative AI systems, employing mathematical reasoning to represent AI decision-making processes and reflect on ethical issues such as fairness, bias, and transparency. The interdisciplinary, constructionist approach encourages learners to discern how AI knowledge construction differs from human cognition, thereby enhancing their ethical reasoning. The findings show that students not only developed a foundational understanding of ethical principles but also gained epistemic insight into AI's relationship with human knowledge and values. This article provides a practical, theory-informed framework and interdisciplinary teaching resources to advance K-12 AI ethics education and support educators in fostering AI literacy.

Introduction

The rapid advancement of AI, particularly recent breakthroughs in Generative AI, highlights the urgent need to integrate AI ethics into K-12 education (Sanusi et al. 2023). As AI technologies permeate everyday life, students are increasingly exposed to AI systems but often develop misconceptions. For example, they may equate AI with computer programs or human-like entities (Kreinsen and Schultz 2022) or perceive AI as either a threat or an entirely objective tool (Mertala et al. 2022). Such misconceptions hinder students' ability to critically and ethically engage with AI. To address this, scholars call for preparing students with the knowledge, skills, and critical attitudes needed for responsible and informed interactions with AI (UNESCO 2022), particularly in understanding basic ethical principles such as fairness and accountability (Long and Magerko 2020).

However, teaching AI ethics is challenging due to the epistemic complexity of AI. That is, students need to understand how AI acquires, processes, and applies knowledge, and how these processes differ from human intelligence (Barelli et al. 2024). Gaining epistemic insights into AI requires knowledge spanning computer science, mathematics, and the humanities (Raji, Scheuerman, and Amironesei 2021). Such interdisciplinary understandings help students comprehend how AI systems function, recognize the role of humans in their development and operation, and consider the ethical and societal implications of AI. By acquiring these insights, students can critically evaluate AI systems, appreciate their strengths and limitations, and manage human-AI relationships ethically.

In response to these challenges, this study presents three constructionist activities designed to promote ethical understandings and epistemic insights among primary school students. While hands-on approaches can support AI ethics education, existing tools often lack opportunities for sustained ethical reflection. Informed by constructionism, our pedagogical framework integrates an interdisciplinary approach with guided reflective questions to encourage continuous ethical consideration. It also makes explicit the processes of AI knowledge construction, showing how these differ from human cognition. This framework prompts students to explore key ethical principles and critically reflect on how AI relates to human knowledge and values. By building and directly interacting with rule-based, machine learning (ML)-based, and generative AI, students learn how these systems acquire and process information in distinct ways. These activities strengthen students' technical understanding and prompts ethical reflections, helping them move beyond misconceptions to gain a nuanced perspective on AI's role in society.

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Background

AI Ethics in K-12 Education

AI ethics is concerned with the ethical challenges associated with the development, deployment, and use of AI systems in K-12 educational settings (Siau and Wang 2020). With AI's rapid advancement and widespread integration into society, fundamental principles such as fairness, bias, transparency, accountability, and non-maleficence have come to the forefront (Jang, Choi, and Kim 2022). Fairness and bias arise from issues such as biased training data and algorithm bias (Barbierato et al. 2022), while transparency makes AI decision-making processes transparent and understandable (Jobin, Ienka, and Vayena 2019). Accountability ensures that those involved in designing and using AI are held responsible for its outcomes (Memarian and Doleck 2023), and non-maleficence emphasizes minimizing the harm caused by AI systems and gaining a critical attitude toward AI in society (Lane, Minkkinen, and Mäntymäki 2019).

However, helping students understand the ethical issues of AI requires more than just knowing how the technology works (Dai et al. 2023). Students need to understand the epistemological differences—different ways of knowing, creating knowledge, and making decisions—between AI systems and humans. AI systems, whether rule-based or data-driven, process information in ways that are fundamentally different from human reasoning, but the design and function of AI systems often reflect the values and biases of the people who create them (Barelli et al. 2024). Understanding these differences can help students critically evaluate the ethical and societal impact of AI technologies. To fully grasp such epistemic complexities, students need to draw on knowledge from multiple disciplines, including computer science, mathematics, and societal concerns in AI ethics (Raji, Scheuerman, and Amironesei 2021).

Epistemic Insight

Epistemic insight offers a theoretical lens to addressing the interdisciplinary complexities of AI ethics. It encourages students to understand how knowledge is created, validated, and applied within AI systems and, thereby, fosters their ability to critically engage with the ethical dimensions of AI (Billingsley et al. 2018). Specifically, epistemic insight helps students integrate knowledge from mathematics, computer science, and ethics, enabling them to tackle real-world challenges, such as evaluating fairness and transparency in AI (Dai 2024a). For example, students can examine how binary systems and probability form the basis of AI's decision-making processes and how these processes can introduce bias when datasets are incomplete or unbalanced.

Furthermore, epistemic insight encourages students to reflect on the relationship between humans and AI (Billingsley and Hardman 2017). It enables them to explore how AI systems can reflect human biases and values and how AI differs fundamentally from human intelligence. By engaging students in creative activities using generative AI, students can compare how humans and AI learn and generate new content while also considering the potential risks of misuse with such tools. In addition, epistemic insight allows students to recognize how their own decisions shape the development of AI systems, including the biases and assumptions embedded in algorithms and datasets. As a result, students develop a more critical perspective, realizing that AI is not neutral or objective but influenced by human decisions. This awareness helps students examine the strengths and limitations of AI and evaluate its broader ethical implications (Bilstrup, Kaspersen, and Petersen 2020).

Constructionism as a Pedagogical Approach

We propose a constructionist approach to cater to young students' need for age-appropriate resources and limited technical backgrounds in K-12 AI ethics education (Kilhoffer et al. 2023). Scholars have warned that teaching AI ethics through traditional methods like case studies or ethical dilemma discussions often oversimplifies complex ethical and technical issues, leading to superficial understanding (Li et al. 2024). Constructionism, in contrast, engages students in actively designing and building AI systems, making abstract concepts like AI ethics and epistemic insight more accessible (Harel and Papert 1991). This hands-on approach allows students to interact directly with AI, critically reflecting on its decision-making processes and understanding how human values shape its outcomes (Weng et al. 2024). By engaging with tangible projects, students develop technical skills and deepen their ethical awareness in a meaningful and playful way (Dai 2024b).

Recent studies underscore the effectiveness of constructionist learning for AI ethics. For example, high school students have successfully explored fairness and transparency by building AI systems using real-life data (Alonso 2020; Krakowski et al. 2022). However, such approaches remain underexplored in primary education. In a notable exception, Irgens et al. (2022) engaged children (aged 9-13) in a critical ML program to explore biased datasets, but the tools lacked embedded opportunities for continuous ethical reflection. Our approach builds on this by consistently offering students opportunities to explore the epistemological differences between various AI approaches, promoting critical engagement and responsible interactions with AI.

Description of Resources

This study presents three structured constructionist activities designed to foster upper primary students (ages 10-12). The first two activities center around a real-life scenario where students create an AI system to classify edible bananas with different AI approaches. Activity 1 explores rule-based AI, while Activity 2 focuses on ML-based AI, enabling students to compare and contrast these methods and their strengths and weaknesses. Activity 3 enables an examination of human-AI comparison on learning and generating knowledge. Students are encouraged to actively engage with AI systems' technical and ethical dimensions throughout these activities. The **Learning Outcomes** are as follows:

- **Understanding the technical process of AI systems:** Students will differentiate between rule-based and ML-based AI systems, understanding their essential components and decision-making processes through hands-on programming tasks.
- **Raising ethical awareness:** Students will recognize and apply key ethical principles (fairness, bias, accountability, transparency, non-maleficence) in AI systems, considering their implications in technical and societal contexts.
- **Gaining critical evaluation:** Students will build and critically evaluate simple AI models, identifying potential biases or errors and proposing improvements based on ethical considerations.
- **Developing epistemic insight:** Students will develop an interdisciplinary understanding of AI ethics, reflecting on the epistemological differences between human and machine intelligence and recognizing the human role in shaping AI behavior and impacts.

Structured Activities

Activity 1: Scratch-based Rule Classifier

This activity introduces students to the key concepts and mechanisms behind rule-based AI. Using Scratch, a beginner-friendly, block-based programming environment, students create a simple rule-based classifier to differentiate between "good" and "bad" bananas. This task highlights the transparency inherent in rule-based AI systems, demonstrating how decisions are made based on pre-set rules.

Students begin by observing three bananas with varying characteristics, such as color and size (see Figure 1a). They discuss and share their observations about given bananas and determine which features make a banana good or bad. Students are asked to identify at least two features (e.g., color and size) of a ripe banana. Whether a banana demonstrates such features is represented using binary numbers (1 for "yes," 0 for "no"). With this rule, students program a robot (a Spirit) in Scratch to recognize "good" or "bad" bananas by evaluating the scores assigned to each feature. For example, students may decide that a banana with a bright yellow color (1 point) and a large size (1 point) is a "good banana" because its total score is 2 (Figure 1b). Conversely, a banana that is small and green (0 points for both features) will be classified as a "bad banana" because its score is 0.

Using if-else blocks in Scratch, students implement the robot's decision-making process, programming it to select bananas that meet the predefined score threshold of 2 or above. As students work through the task, they test and debug their codes, which helps them understand the iterative nature of rule refinement in AI systems. Each debugging step strengthens their grasp of the direct relationship between programmed rules and the system's output.

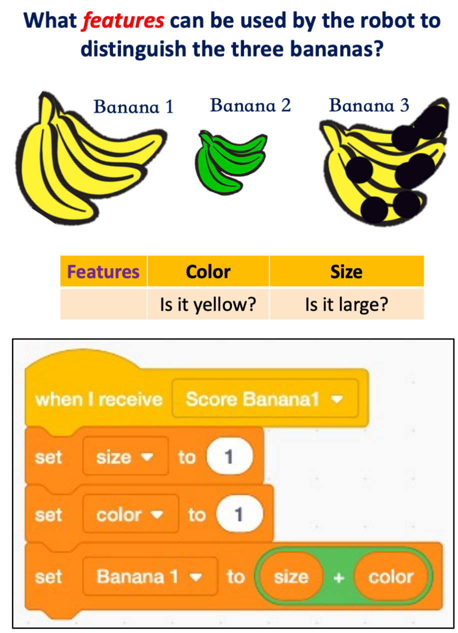

Figure 1: Students Explore Rule-based AI Using Scratch.
a (top): Students define the key features (color and size) to classify a ripe banana.
b (down): The recognition resulst of features are represented using binary numbers (1 for "yes", 0 for "no") in the working process of a classifier.

Throughout the activity, students can trace every step of the AI's decision-making process, which emphasizes the transparency of rule-based systems. This hands-on experience builds foundational programming skills while giving students a tangible understanding of how rule-based AI functions. Introducing students to rule-based systems sets the stage for future comparison with ML-based AI in the next activity. At the end of the task, students reflect on how the AI system they created operates and how their involvement in defining features impacts the AI's behavior.

Reflection Questions for Activity 1

- How did you decide which features (e.g., color, size) the robot should use to classify the bananas?
- What other features could you have considered including?
- Can you explain how the robot made its decisions?
- If someone else were to use your program, do you think they would understand how it works? Why or why not?

- Who do you think is responsible for any mistakes the robot makes? Is it the robot itself or the person who programmed it?

Activity 2: Google Teachable Machine Image Classifier In the second task, students engage with Google Teachable Machine (GTM), an intuitive platform that simulates the processes involved in ML-based AI. In this task, students are tasked with training an image classifier to recognize different types of bananas, with 60 given images (https://teachablemachine.withgoogle.com/models/J50oBRso4/). To train the classifier, students go through three core steps of ML: data labeling, model training, and evaluating input/output. Specifically, students begin by labeling each banana as either "good" or "bad" based on its appearance in the images. Then they drop and upload the images into the GTM platform and train a classifier model to recognize different categories of bananas. With the trained model, they use three images to test the classifier and document the results in an activity sheet, as shown in Figure 2.

5. Report the outcome!

Test image (number)	Classification result	Confidence (value)
1. 		
2. 		
3. 		

Figure 2: The Activity Sheet Students Use to Test the Classifier and Document the Result.

A central part of this task is helping students interpret these confidence scores. For example, if the model gives an image a 60% confidence score for being labeled a "good" banana, students should understand that this means the model is 60% sure that the image represents a good banana. This exploration helps them understand the probabilistic nature of ML-based AI. Unlike rule-based AI (as discussed in Activity 1), which makes deterministic decisions, ML-based AL relies on probabilities to make predictions rather than certainties.

Students test their models on additional banana images from online sources to deepen their understanding, evaluating how well their model generalizes beyond the initial dataset. Then, they compare their model's performance with their peers using the same test images. This comparison encourages students to think critically about how training data impacts model performance and accuracy, fostering discussions around fairness, bias, and transparency.

At the same time, students reflect on the ethical implications of their models' outputs, considering how misclassifications or biases could result in unfair outcomes. This reflection prompts students to consider the real-world impact of AI systems, exploring the principle of non-maleficence. They also discuss the role of human involvement in AI, reflecting on how their labeling decisions may have influenced the AI's performance and fairness, linking back to the idea of accountability discussed in Activity 1.

Reflection Questions for Activity 2

- What does it mean if the confidence score is 60% for a good banana?
- Can you explain how the model made its decisions? How is this different from the robot in Activity 1?
- Did your AI make any mistakes when classifying images? What do you think caused these mistakes?
- What did you do to improve the accuracy of your model?
- Could the mistakes in your AI model cause any harm? What could you do to reduce the risk?
- Did you notice any differences in the outputs between your and your classmate's model? What do you think caused these differences?

Activity 3: Generative AI and Poetry Creation with Poe This activity introduces students to generative AI through a hands-on experience with Poe, a language model platform. The objective is for students to understand how generative AI functions and explore the epistemological differences between generative AI and supervised machine learning.

The activity begins with a class discussion about Siri, a voice assistant many students are familiar with. Students discuss its working mechanisms based on what they learned in Activity 2 about ML-based AI. Then, they are encouraged to consider whether AI can generate new content, such as poetry, rather than just performing tasks like voice recognition based on a given database. After the discussion, students are guided to the hands-on task of using Poe to generate a poem that includes their names.

This creative process allows students to observe how a large language model (LLM) generates text based on input. To foster more profound reflection on human and machine intelligence, students compare their experiences creating poems in Chinese language classes — following rules for tone, rhyme, and structure — against how the AI synthesizes text from patterns in large datasets. By engaging with this comparison, students explore the similarities and differences between how humans and AI produce creative works.

After creating their poems, students reflect on the differences between generative AI and supervised learning models, like those they worked with in Activity 2. They discuss how each AI system acquires, processes, and generates knowledge and compare this with how humans approach similar tasks. Students are also encouraged to reflect on the ethical implications of generative AI, including potential challenges like misuse of generative AI.

Reflection Questions for Activity 3

- How does Poe create a poem? How is it different from how you make a poem in your Chinese class?
- Can you find similarities and differences between how Poe generates a poem and how GTM classifies images?
- What knowledge does Poe rely on to generate poems? How is this different from the knowledge you use when you create poems?
- What might happen if Poe generated a poem that contained offensive or inappropriate language?
- Who would be responsible for that, and what could be done to prevent it?

Setup and Resources

For these activities, each student requires a laptop or tablet with internet access to Scratch, Google Teachable Machine, and Poe platforms. Additionally, students will need a dataset of banana images for classification in Activity 2 and optional reflection worksheets to document their learning and responses to the designed reflection questions.

Research Methods

This study was part of an eight-lesson AI curriculum conducted in August 2023 at a Hong Kong public primary school (Dai et al. 2024b). The curriculum was innovatively designed using constructionist theory and the cultivation of epistemic insight, integrating guided reflective questions with hands-on, experiential activities. This theoretical foundation directed the selection of unique tasks, such as interactive AI projects and continuous ethical reflection exercises, as well as the data collection and analysis methods. Participants were 34 fifth-grade students (17 boys, 17 girls, average age 11) with no prior AI experience. Informed consent was obtained from students and parents. The curriculum introduced AI concepts and ethics through constructionist activities. Activities 1 and 2, focusing on rule-based and ML-based AI, each lasted 1 hour. Activity 3 on generative AI was a shorter 15-minute activity integrated into a Natural Language Processing (NLP) lesson. By combining active knowledge creation with ongoing ethical contemplation, the curriculum ensured that students engaged deeply with both the technical and ethical dimensions of AI, distinguishing it from traditional AI ethics education programs.

Data collection included video recordings of classroom interactions, screen recordings of tasks, student worksheets, and cognitive interviews. All recordings were transcribed verbatim. Thematic analysis (Braun and Clarke 2006) was performed to identify patterns in students' understanding and application of ethical principles (fairness/bias, accountability, transparency, non-maleficence) and the development of epistemic insight. Findings were triangulated with observational data to validate the designed activities' effectiveness and ensure our interpretations' robustness.

Results

Cognitive Understanding of AI Ethics

Fairness/bias

Students recognized how biased datasets led to unfair classifications in the classifier during the ML-based AI activity. One student noted, "At first, I only put eight or nine (images) in good bananas, but a lot more in bad bananas, which led to good bananas being identified as bad ones," highlighting their understanding of the importance of balanced data enabled by direct interaction with GTM. They also observed how superficial features like color could lead to biased outcomes. One student explained why misclassifications happened, reasoning, "Maybe because these bananas were all black or with dark spots, the machine thought that dark colors represented a bad banana. And when the person who's holding the banana happens to be dark-skinned, then the machine thought it's a bad banana." The interview quotes show the student grasped the connection between data bias and unethical decisions.

Accountability

With deep engagement with Scratch, students could easily identify their responsibilities as developers to identify features, manually set rules, and program correctly for the smart robot to classify bananas. Similarly, the hands-on project with GTM allows students to grasp that improving AI models requires human intervention. They reflected, "We should keep improving the data, add more labels so that the machine can get it right." Their response indicates a developing understanding of the importance of data curators and designers.

Transparency

Students exhibited different levels of understanding regarding the transparency of rule-based AI, ML-based AI, and generative AI. During the Scratch activity with rule-based AI, many students appreciated its transparency, noting that they could easily follow how decisions were made based on explicit rules. One student remarked, "It's clear because we can know exactly what scores are for each banana and what rules the robot applies." In contrast, students expressed frustration with the lack of transparency in ML-based AI. One student shared, "Even when the banana is perfectly yellow, with no spots at all, the machine could still recognize it as a bad banana." This comment highlights their awareness of the opaque nature of ML models, where decisions are made based on patterns learned from data rather than visible, traceable rules. Students found it more challenging to understand why the model made specific predictions, revealing the difficulties in ensuring transparency in ML-based systems.

After exploring generative AI, students recognized that its decision-making process was even less transparent. One student noted that while Poe seems to "think on its own" by generating new responses, it is more complex to trace how it arrives at these outputs than systems like Siri, which simply "asks and checks" information online. The reflection highlights students' growing understanding of how generative AI operates based on complex learned patterns, making its internal workings harder to explain or fully understand than rule-based and ML-based AI.

Non-maleficence

During the interviews, students demonstrated their emerging critical thinking towards AI regarding how technical limitations could impact society. For instance, one student reasoned that if the classification model they built was used in a supermarket or garbage classification center, it may lead to undesired societal impacts as "if the machine thinks some bananas are bad just because of color, it could make people waste food."

Epistemic Insight

Interdisciplinary Thinking

Students demonstrated interdisciplinary thinking when reflecting on how incomplete or limited data can lead to lower confidence and misclassifications in AI systems. Some students applied mathematical reasoning, explaining that smaller sample sizes increase each data point's impact. One student noted: "With 30 images, each image has a larger percentage, so the impact of an apple (picture) is greater, whereas the impact is smaller in a 300-image dataset." Shifting to an AI perspective, other students highlighted the importance of features in ML. For instance, one explained, "If you add an apple to a dataset with 30 bananas, it has fewer features, so the machine would easily pick up on its features. But with 300 images, adding it won't affect its feature judgment as much". These reflections demonstrate students' developing epistemic insight, connecting AI knowledge acquisition to ethical implications across disciplines.

Human Intelligence vs Machine Intelligence

Students gained epistemic awareness of the distinct cognitive processes underlying human intelligence and machine intelligence, particularly how each system represents and processes knowledge. One student highlighted the difference in knowledge representation when comparing how generative AI and humans create poems, stating, "Machines use 1s and 0s. They use different combinations of 1s and 0s to understand and create poems. Humans directly look at what is written and create from that." His response reflects the student's understanding that while AI processes data using binary code and predefined patterns, humans rely on language and contextual awareness to generate knowledge.

Furthermore, this fundamental understanding enabled students to develop a critical attitude towards the limitations of AI. For example, one student critiqued that machine's inability to accurately classify a person holding a banana, noting, "The machine can be really too stupid; it could mistake a person with bruised spots on the body for a banana—just like how the banana has dark spots all over." Additionally, students reflected on the differences in learning processes between humans and AI. One student observed, "Humans can learn faster with a few examples and apply knowledge flexibly, but machines need a lot of images and data."

Human Values in AI Development

Students in this study developed a critical attitude toward AI, recognizing that its outputs are shaped by human values and are not entirely objective. One student explained, "The data the machine uses is input by humans. If people have bad examples, the machine will learn the wrong things". Another noted, "The machine has no thoughts of its own; it does things because we tell it what to do," reinforcing the understanding that AI lacks independent decision-making and is deeply influenced by human values and instructions. By realizing that AI systems are not inherently neutral but reflect human values and biases, students gained a deeper understanding of the role humans play in AI development.

Conclusion and Discussion

This study designed and implemented three constructionist activities that fostered AI ethical understanding and epistemic insight among upper primary students. By engaging with rule-based AI, ML-based AI, and generative AI, students gained a deeper understanding of AI's technical processes while critically reflecting on its ethical implications. These hands-on activities provided immediate feedback and facilitated the integration of technical learning with ethical considerations. A key strength of this approach is its ability to encourage consistent ethical reflection throughout the process. By guiding students to explore and reflect on the epistemological differences between AI and humans, students developed a holistic understanding of AI and the role of human values in shaping AI outputs. Additionally, the multiple disciplinary perspectives in these activities also enhanced their critical engagement and ethical reasoning.

Practically, this study offers a model for designing age-appropriate AI ethics activities that balance technical and ethical learning. These structured activities can also be incorporated into different disciplinary contexts, such as mathematics and the humanities, providing K-12 educators with a framework to promote interdisciplinary learning.

References

Alonso, J. M. 2020. Teaching Explainable Artificial Intelligence to High School Students. *International Journal of Computational Intelligence Systems* 13(1): 974. doi.org/10.2991/ijcis.d.200715.003.

Barbierato, E.; Vedova, M. L.; Tessera, D.; Toti, D.; and Vanoli, N. 2022. A Methodology for Controlling Bias and Fairness in Synthetic Data Generation. *Applied Sciences* 12(9): 4619. doi.org/10.3390/app12094619.

Barelli, E.; Lodi, M.; Branchetti, L.; and Levrini, O. 2024. Epistemic Insights as Design Principles for a Teaching-Learning Module on Artificial Intelligence. *Science and Education.* doi.org/10.1007/s11191-024-00504-4.

Billingsley, B.; and Hardman, M. 2017. Epistemic insight: Teaching and learning about the nature of science in real-world and multidisciplinary arenas. *School Science Review*, 98(365): 57–58.

Billingsley, B.; Nassaji, M.; Fraser, S.; and Lawson, F. 2018. A Framework for Teaching Epistemic Insight in Schools. *Research in Science Education* 48(6): 1115–1131. doi.org/10.1007/s11165-018-9788-6.

Bilstrup, K.-E. K.; Kaspersen, M. H.; and Petersen, M. G. 2020. Staging Reflections on Ethical Dilemmas in Machine Learning. *Proceedings of the 2020 ACM Designing Interactive Systems Conference,* 123–131. New York: Association for Computing Machinery. doi.org/10.1145/3357236.3395558.

Braun, V.; and Clarke, V. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3(2): 77–101.

Dai, Y. 2024a. Dual-Contrast Pedagogy for AI Literacy in Upper Elementary Schools. *Learning and Instruction*, 91, p.101899. https://doi.org/10.1016/j.learninstruc.2024.101899

Dai, Y. 2024b. Integrating Unplugged and Plugged Activities for Holistic AI Education: An Embodied Constructionist Pedagogical Approach. *Education and Information Technologies*, 1–24. doi.org/10.1007/s10639-024-13043-w.

Dai, Y.; Lin, Z.; Liu, A.; and Wang, W. 2023. An Embodied, Analogical, and Disruptive Approach of AI Pedagogy in Upper Elementary Education: An Experimental Study. *British Journal of Educational Technology* 55(1): 417–434. doi.org/10.1111/bjet.13371.

Harel, I. E.; and Papert, S. E. 1991. *Constructionism.* Norwood, NJ: Ablex Publishing.

Irgens, G.; Vega, H.; Adisa, I.; and Bailey, C. 2022. Characterizing Children's Conceptual Knowledge and Computational Practices in a Critical Machine Learning Educational Program. *International Journal of Child-Computer Interaction* 34: 100541. doi.org/10.1016/j.ijcci.2022.100541.

Jang, Y.; Choi, S.; and Kim, H. 2022. Development and Validation of an Instrument to Measure Undergraduate Students' Attitudes Toward the Ethics of Artificial Intelligence (AT-EAI) and Analysis of Its Difference by Gender and Experience of AI Education. *Education and Information Technologies* 27(8): 11635–11667. doi.org/10.1007/s10639-022-11086-5.

Jobin, A.; Ienca, M.; and Vayena, E. 2019. The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence* 1(9): 389–399. doi.org/10.1038/s42256-019-0088-2.

Kilhoffer, Z.; Zhou, Z.; Wang, F.; Tamton, F.; Huang, Y.; Kim, P.; Yeh, T.; and Wang, Y. 2023. How Technical Do You Get? I'm an English Teacher: Teaching and Learning Cybersecurity and AI Ethics in High School. *2023 IEEE Symposium on Security and Privacy (SP).* doi.org/10.1109/sp46215.2023.10179333.

Krakowski, A.; Greenwald, E.; Hurt, T.; Nonnecke, B.; and Cannady, M. 2022. Authentic Integration of Ethics and AI Through Sociotechnical, Problem-Based Learning. *Proceedings of the AAAI Conference on Artificial Intelligence* 36(11): 12774–12782. doi.org/10.1609/aaai.v36i11.21556.

Kreinsen, M.; and Schulz, S. 2021. Students' Conceptions of Artificial Intelligence. *The 16th Workshop in Primary and Secondary Computing Education,* 1–2. doi.org/10.1145/3481312.3481328.

Laine, J.; Minkkinen, M.; and Mäntymäki, M. 2024. Ethics-based AI Auditing: A systematic literature review on conceptualizations of ethical principles and knowledge contributions to stakeholders. *Information and Management*, 61(5): 103969. https://doi.org/10.1016/j.im.2024.103969

Li, L.; Yu, F.; and Zhang, E. 2024. A Systematic Review of Learning Task Design for K-12 AI Education: Trends, Challenges, and Opportunities. *Computers and Education: Artificial Intelligence* 6: 100217. doi.org/10.1016/j.caeai.2024.100217.

Long, D.; and Magerko, B. 2020. What Is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* 1–8. New York: Association for Computing Machinery. doi.org/10.1145/3313831.3376727.

Memarian, B.; and Doleck, T. 2023. Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and Higher Education: A Systematic Review. *Computers and Education: Artificial Intelligence* 5: 100152. doi.org/10.1016/j.caeai.2023.100152.

Mertala, P.; Fagerlund, J.; and Calderon, O. 2022. Finnish 5th and 6th Grade Students' Pre-Instructional Conceptions of Artificial Intelligence (AI) and Their Implications for AI Literacy Education. *Computers and Education: Artificial Intelligence* 3: 100095. doi.org/10.1016/j.caeai.2022.100095.

Raji, I. D.; Scheuerman, M. K.; and Amironesei, R. 2021. You Can't Sit with Us. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency,* 1–10. New York: Association for Computing Machinery. doi.org/10.1145/3442188.3445914.

Sanusi, I. T.; Oyelere, S. S.; Vartiainen, H.; Suhonen, J.; and Tukiainen, M. 2023. A Systematic Review of Teaching and Learning Machine Learning in K-12 Education. *Education and Information Technologies* 28(5): 5967–5997. doi.org/10.1007/s10639-022-11416-7.

Siau, K.; and Wang, W. 2020. Artificial Intelligence (AI) Ethics. *Journal of Database Management* 31(2): 74–87. doi.org/10.4018/jdm.2020040105.

UNESCO. 2022. K-12 AI Curriculum: A Mapping of Government-Endorsed AI Curricula. https://unesdoc.unesco.org/ark:/48223/pf0000380602. Accessed: 2024-12-19.

Weng, X.; Ye, H.; Dai, Y.; and Ng, O. L. 2024. Integrating Artificial Intelligence and Computational Thinking in Educational Contexts: A Systematic Review of Instructional Design and Student Learning Outcomes. *Journal of Educational Computing Research.* doi.org/10.1177/07356331241248686.