ARTICLE

Integrating Ethics and Career Futures with Technical Learning to Promote AI Literacy for Middle School Students: An Exploratory Study

Helen Zhang[1] · Irene Lee[2] · Safinah Ali[3] · Daniella DiPaola[3] · Yihong Cheng[1] · Cynthia Breazeal[3]

Abstract

The rapid expansion of artificial intelligence (AI) necessitates promoting AI education at the K-12 level. However, educating young learners to become AI literate citizens poses several challenges. The components of AI literacy are ill-defined and it is unclear to what extent middle school students can engage in learning about AI as a sociotechnical system with socio-political implications. In this paper we posit that students must learn three core domains of AI: technical concepts and processes, ethical and societal implications, and career futures in the AI era. This paper describes the design and implementation of the Developing AI Literacy (DAILy) workshop that aimed to integrate middle school students' learning of the three domains. We found that after the workshop, most students developed a general understanding of AI concepts and processes (e.g., supervised learning and logic systems). More importantly, they were able to identify bias, describe ways to mitigate bias in machine learning, and start to consider how AI may impact their future lives and careers. At exit, nearly half of the students explained AI as not just a technical subject, but one that has personal, career, and societal implications. Overall, this finding suggests that the approach of incorporating ethics and career futures into AI education is age appropriate and effective for developing AI literacy among middle school students. This study contributes to the field of AI Education by presenting a model of integrating ethics into the teaching of AI that is appropriate for middle school students.

Keywords AI ethics · Bias · Career implications · Sociotechnical systems · Middle school education

---

✉ Helen Zhang
   zhangzm@bc.edu

Extended author information available on the last page of the article

Introduction

The rapid expansion of Artificial Intelligence (AI) in industry and academic fields, made possible by enormous amounts of data and computation power, necessitates the development of a workforce that is knowledgeable and capable of working with AI (Manyika et al., 2017; National Science Board, 2020). While AI education curricula and learning opportunities are becoming more prevalent, promoting AI education at the K-12 level, however, is not easy because the curricula must be engaging, relevant, and age-appropriate to young learners. And, there is a long-standing gap in access to computer science and AI education between students who are from minority groups and/or low-income families and their White, more affluent peers (Danyluk et al., 2014). Thus, special actions are needed to ensure that all students are prepared to become successful in the era of AI.

This paper reports an exploratory study that aims to broaden participation in AI by developing AI literacy among middle school students. Identifying the middle school years as a key phase in development of AI literacy, we developed the "Developing AI Literacy" (DAILy) curriculum. The DAILy curriculum engaged students in integrating their understanding of three domains in order to become prepared for participating in AI-infused fields and industries of the future: (1) an age-appropriate technical knowledge and skills in AI, (2) an understanding of AI's ethical and societal implications, and (3) a knowledge of AI's impact on jobs (AI career futures) and how to adapt to the future of work (career adaptability).

The three interwoven domains were established based on our previous experience of working with students (e.g., DiPaola et al., 2020; Payne, 2020) and the notion that AI is not only a technical field but one that has wide-ranging societal and career impacts. The inclusion of ethics in AI education is needed to raise the public's awareness of AI's burgeoning impact on industry and society, and to prepare students to investigate and address ethical issues in AI as critical consumers and potential future creators of AI technologies (Grosz et al., 2019; Payne, 2020). Users of AI systems need to be aware of the potential for bias in their predictions. Future designers of AI, in particular, need to ensure that their products minimize negative impacts due to bias in datasets, models, and predictions. Members of the general public need to be aware of how these systems might be biased against them so they can evaluate their impact, and seek justice for themselves and others.

In this paper we report the design of the DAILy curriculum, assessment tools developed to examine student learning, and the findings from our implementation of the curriculum among middle school students. The results illuminate what middle school students were capable of learning and doing with AI, what challenges they encountered in making sense of AI, and how the DAILy curriculum impacted their attitudes toward AI and future career ideas.

Theoretical Framework

A Need for Integrating Ethics and Career Futures into K-12 AI Education

In the past decade, several initiatives and projects promoting K-12 AI education have emerged and various AI courses, tools, and tutorials have been launched for

teaching AI to students in the USA, China, Europe, Korea and many other countries. For instance, existent curricula aimed at young students, up through 8th grade, include Code.org's AI for Oceans, ReadyAI's AI-in-a-box, AI4kids, and MIT AI Education Initiative's collection of AI curricula and tools including the Media Lab's AI + Ethics curriculum for middle school. Curricula for high school aged students include AI4All's Bytes of AI and full-length Open Learning curriculum, UDC's AI + Curriculum for European High School (Guerreiro-Santalla et al., 2020), ISTE's AI Foundations course, and Reaktor's Elements of AI. Some other programs such as Technovation's AI Family Challenge, are explicitly aimed at families' exploration of concepts in AI and applying AI tools to solve community problems. Meanwhile several K-12 teacher professional development programs in AI curricula have been established such as CSER's Teaching AI in the Classroom, ISTE's AI Explorations and their practical use in schools, and teacher AI-4-All's Open Learning program. These programs are critical to bringing AI into schools as some countries such as China have mandated that all high school students learn about artificial intelligence (Jing, 2018).

Many of the current approaches of teaching AI focus mainly on the technical aspects of AI learning and few programs emphasize the ethical and societal implications of AI, one of the "Five Big Ideas of AI" as identified by the AI4K12 initiative (Touretzky et al., 2019) that every K-12 student should know and be able to do in AI. In their review of 49 AI curricula and programs for K-12 students, Zhou et al. (2020) found only 13 taught ethics and fewer than 10 addressed the weakness and/or strength of AI. Only one program explicitly focused on ethics where middle school students were engaged in evaluating their Youtube Redesigns from an ethical lens (DiPaola et al., 2020).

Ethics, or moral philosophy, is the study of moral principles that govern a person's behavior. It involves systematizing, defending, and recommending concepts of right and wrong behavior. The inclusion of ethical consideration is crucial in AI because AI has presented substantial ethical and socio-political challenges that call for thorough philosophical and ethical analysis (Coeckelbergh, 2020; Gunkel, 2012; Müller, 2020). The main ethical issues related to the impact of AI on human society are the prevention of existential risks to humanity, the impact of AI technologies on our privacy, the impacts of bias in AI, and the development of AI systems that meet our ethical standards (Gordon & Wrenn, 2020). For instance, algorithmic bias in AI applications such as facial recognition, predictive policing, and credit score assignment has had negative impacts on communities of color (Buolamwini & Gebru, 2018; Kirkpatrick, 2016; Selbst, 2017; Van Brakel, 2016). These systems can be discriminatory if they are trained upon biased datasets. Even worse, experiencing consequences of bias in AI systems may dissuade people of color from participating in the AI. This would hinder attempts to improve AI systems by involving diverse groups of users and developers in identifying and mitigating sources of bias. Thus, ethical analyses and guidelines are necessary to avoid negative repercussions of AI on society. Furthermore, research has raised questions about the extent to which students are aware of AI's impact on their everyday lives and its application in fields and industries of the future (Druga et al., 2017; Hasse et al., 2019). This lack of awareness may limit

students' understanding of the relevance of AI and thus their interest in pursuing learning trajectories that lead toward careers in AI. Therefore, in their proposed design framework for K-12 AI education, Zhou et al. called for recognizing the ethical implications of AI and embedding ethical discussions and activities in all AI curricula (Zhou et al., 2020).

Teaching AI Ethics through a Sociotechnical Systems Lens

Another reason that AI education should include ethics topics is that AI is a sociotechnical system. The term socio-technical system (STS) was originally coined by Emery and Trist (1960) to describe systems that involve a complex interaction between humans, machines and the environmental aspects of the work system. Nowadays STS are systems that span social, cognitive and information systems (hardware, software, personal and societal spaces) (Badham et al., 2000). Common STS examples include communication systems ranging from email, blogs, social media and news media, to consumer products such as shopping and entertainment systems (e.g., YouTube). However, STS are not always neutral sources of information and serve stakeholder (and sometimes political) agendas. They are created by humans and humans decide the goals of the sociotechnical systems they create. As such, it is important that people are educated to become aware of the goals of these systems and be able to distinguish between the advertised and actual goals of corporations in order to make informed decisions on whether or not to actively participate. For example, the YouTube recommendation algorithm is advertised as a way to provide entertainment for users whereas in actuality its aim is to make profit for the company, and gather user information.

Furthermore, from a complex systems perspective, it is important to understand the distal effects that organizations using these systems may have on individuals, including those that do not directly use the system, and thus society. Algorithmic bias in machine learning is one such example (Buolamwini & Gebru, 2018). In AI systems training data is used to build a model capable of making predictions. The composition of the training data affects the predictions made by the system. Training data can be biased and not be representative of populations. It can also utilize proxies for data it does not include—for example using zip code as a proxy for socioeconomic status and race may pose unintended (and potentially harmful) consequences. Individual users receive predictions made by the systems, regardless of whether they are represented in the system, and even those who do not directly use the system may be impacted by its existence. An example of secondary and tertiary effects of a biased system's existence and use include unfair hiring practices that impacts the economic viability of communities, and potential civil unrest that results from widening economic gaps.

Teaching AI Ethics through a Career Futures Lens

AI is rapidly changing the industry and shaping the future labor market. Exclusion of topics around career futures in K-12 AI education would leave the students

unprepared for the future workplace. A 2017 report by McKinsey Global Institute predicted that 75 million to 375 million workers (3–14% of the global workforce) may be required to change occupations and/or upgrade their skills by 2030 because of the adoption of AI (Manyika et al., 2017). All future workers will be required to adapt to working with increasingly intelligent machines, necessitating their preparation to learn in-demand skills and be flexible in setting expectations about work. Therefore, learning how AI will impact jobs (career futures) and developing skills of how to adapt to changes in working environments (career adaptability) is critical to preparing our youth for a future with AI. Further, learning AI's impact on jobs and employment helps further students' understanding of societal changes driven by AI technologies.

Embedding Ethics into CS/AI Education

Research on computer science education has revealed that the traditional approach of teaching ethics as distinct from the subject content often fails to prepare students for real-world work (Boss, 1994; Gardner, 1991). Fiesler et al. (2020) analyzed 51 university-level AI and ML courses and found that a majority of the courses cover ethics-related topics within the last two classes. Ethics topics are often considered as a part of these technical courses "if time allows." Such an approach of ethics instruction fails to translate into experiences outside the classroom and leaves students unprepared for the current and future work in technology. Recently many CS educators have realized the potential benefits of ethics education and attempted to integrate ethics across computer science curriculum (Loescher et al., 2005). For instance, Skirpan et al. (2018) integrated ethics throughout an undergraduate Human Centered Computing course and demonstrated a meaningful increase in students' consideration of ethics in technology design. Narayanan and Vallor (2014) argued that in order to see ethics as an important part of the building process, students must be able to practice ethical decisions during in-class collaborations. Overall, researchers agree that ethics education should be integral to CS education and it can motivate students and prepare them to create ethical designs. Yet there are still many debates about how to best integrate ethics education into CS courses (DuBois & Burkemper, 2002).

Teaching AI Concepts at the K-12 Level through Interactive Approaches

Different from the math-first, theoretical approach (McGovern et al., 2011; Torrey, 2012) that has been shown successful in teaching AI at the undergraduate, graduate, or professional level, AI curricula and tools aimed at K-12 audiences often take a non-mathematical and interactive approach. For example, Narahara and Kobayashi (2018) proposed utilizing hands-on educational modules to introduce ideas in AI and Robotics. Students first build a physical prototype of a toy car, then play and test the car in a virtual reality (VR) environment, train an AI model based on the dataset acquired from the virtual testing, and run the toy car using a trained AI model on a physical track. In another study of teaching primary school students AI concepts, Ho

et al. (2019) engaged students in building and training a self- learning lawn bowling robot.

One reason for the shifted approach is the high barrier presented by the prerequisite mathematical knowledge of the math-first approach. Many K-12 students, particularly those underrepresented in STEM/Computing, may feel unprepared and not be able to persist and succeed in completing these courses. Marques and colleagues, in their systematic mapping study, noted that many curricula teaching Machine Learning at the K-12 level employed the interactive approach where students are exposed to AI concepts through interactions with AI tools as an end-user followed with some degree of knowledge building and reasoning about how AI works. Some of the underlying complex AI processes are black-boxed as hands-on activities to prevent students from being cognitively overwhelmed. Marques et al. (2020) highlighted the importance of identifying a balance between the black-boxed processes and uncovered processes as well as a learning sequence based on the complexity of the concepts when developing AI curricula.

The DAILy curriculum utilizes such an interactive approach where students experiment with a collection of AI technologies such as the Teachable Machine and tools developed by the research group to help them learn and practice AI concepts and processes. Further to help students understand abstract AI processes, DAILy employs participatory simulations where students act out the roles of individual elements of an AI system and then see how the system as a whole becomes "intelligent" at performing certain tasks. The direct, personal participation in a simulated game has been used in math and science education and proven to be effective in terms of promoting student engagement, supporting collaboration, and helping students of different genders and races/ethnicities to develop deeper understandings of the underlying scientific patterns and processes (Klopfer & Yoon, 2005; Squire & Klopfer, 2007). This approach also aligns well with the use of embodied interaction proposed in the AI Literacy framework (Design Consideration 2: Embodied Interactions, Long & Magerko, 2020) where learners can "put themselves in the agent's shoes as a way of making sense of the agent's reasoning process" (p. 598). By acting out how neural networks work, students develop a concrete idea of how these AI processes work, reflect on their behaviors, and collectively figure out the underlying mechanisms.

Engaging Middle School Aged Students in AI Ethics Education

The major impetus for educating middle school students about AI is driven by their increasing contact with AI in daily life. First, since many students acquire their first mobile device during middle school, they start consuming data on social media websites such as Twitter and Instagram where they are exposed to AI moderated and generated content (Pew Research Center, 2018). Second, middle school students are already creators of media generated with AI through social media apps such as Snapchat and Instagram. According to a survey conducted by the Pew Research

Center (2018), 45% of teens admit they are on social media almost constantly and YouTube, Instagram and Snapchat are the most often used social media tools by teenagers. The youth view and create content with tools such as photo filters that integrate generative modeling techniques. Thus, they may be using AI-enabled technologies without realizing it. Third, students upload personal data, such as images, videos and text, on social media sites and may unwittingly be contributing to datasets used to train AI models. Finally, students are witness to and could be targeted by fake media that are generated by applications of AI such as Deepfakes, like in the case of FaceApp. This exposure to AI, whether direct or indirect, can impact students. While some impacts can be relatively harmless, such as entertainment or art, other exposure could be harmful. Students may unwittingly be persuaded to.

think that a fake event, image, or text is real, and act accordingly. Because students are vulnerable to these manipulations, they need to be knowledgeable about AI. Their awareness of manipulated media has ramifications for democracy, trust, security, and privacy. Knowledge of the existence of AI-moderated and generated media would empower students to take information that they witness online with essential skepticism. Thus, AI literacy focusing on machine learning techniques used in media curation and generation is imperative for students to be informed citizens and critical consumers of online media.

Another impetus for engaging middle school students in AI education is that young adolescence is an important time for students to form identity in STEM (i.e., considering themselves as a STEM person) and start thinking about their future career interests (Dabney et al., 2012; Maltese & Tai, 2010). Through analyzing data of a longitudinal study that followed 12,000 students from 8th grade to college, Tai et al. (2006) found that students reporting an interest in science careers in eighth grade were three times more likely to obtain a college degree in a science field than those who did not show that interest after controlling for differences in background and academic history. The DAILy curriculum engages students in finding out future STEM jobs that they are interested in and exploring how AI has been and will continue impacting their jobs of interest. This career training leverages students' existing interests to further spark their enthusiasm in AI and at the same time reinforces the notion that AI is a sociotechnical system with enormous societal impacts.

Finally research on child development has suggested that ethical concepts are age-appropriate to middle school students. Students in the middle childhood years typically have developed a sense of morality and a conscience based on values (Wood, 1997). Many young teens recognize the importance of societal norms and of following rules for the good of society (Crain, 1985). They are capable of reflecting on their behavior and its impact on others. In the context of AI technology, DiPaola et al. (2020) found that middle school students were capable of understanding various stakeholders and values in a technical AI system, a preliminary step in recognizing how AI can be biased in its design. In the DAILy curriculum, students are invited to reflect on how the systems they use and create might impact themselves and the society at large.

DAILy Curriculum

The design of the 30-h DAILy curriculum was based on our definition of AI literacy that students must learn three core domains to become AI literate citizens: technical concepts, ethical and societal implications, and AI related careers. The three core domains were established by the project leadership team and informed by our team's previous research on promoting AI education among young learners (e.g., Ali et al., 2019, 2021; DiPaola et al., 2020). In the DAILy curriculum, we sought to bring together a set of domains that would engage students in linking AI concepts to ethical issues through considering how datasets and models contribute to bias in AI. To counteract the potential for students to feel dissuaded from participating in AI brought on by discussions of negative impacts of bias in AI, we sought to provide time for students to consider alternative future scenarios in which AI can be helpful and beneficial to humans in the future. The career adaptability framing was employed to help build students' confidence of dealing with emerging changes to fields and careers due to AI. Furthermore, the generative AI activities added a creative exploration of AI that demonstrates how AI can be used to innovate and co-construct artifacts. This creative use of AI, in addition to classification and prediction, helped students envision additional opportunities for the integration of AI in students' future and career plans.

Experienced middle school teachers and AI researchers were involved in the development of the curriculum and reviewed the initial drafts of the curriculum. We also pilot tested key activities with a group of middle school students in an after-school setting and asked for their feedback and suggestions to make the curriculum more exciting for middle schoolers.

AI Technical Concepts Addressed in DAILy Curriculum

The DAILy curriculum is organized by a hierarchy of key technical concepts that have been suggested age appropriate based on our previous work (Ali et al., 2019), including (1) an Introduction to AI, (2) Logic Systems, (3) Machine Learning, (4) Supervised learning and (5) Unsupervised learning. Each curricular module focuses on one key technical concept (see Fig. 1). The first module provides a broad introduction to AI and engages students in distinguishing technologies that use AI or not and identifying key features of AI technologies. Then in Module 2, students learn logic systems and practice building decision trees to sort out different types of pastas. In Module 3, students learn about machine learning in general and supervised learning. They use Teachable Machine to train supervised learning models. They also experiment training the models with different datasets (e.g., datasets with heavily one type of pictures) and discuss the outcomes and their ethical implications. In Module 4, students learn artificial neural networks (NN). In Module 5, students learn another type of machine learning: unsupervised learning, and how Generative Adversarial Networks (or GANs) work. Table 1 shows key activities in each DAILy module.
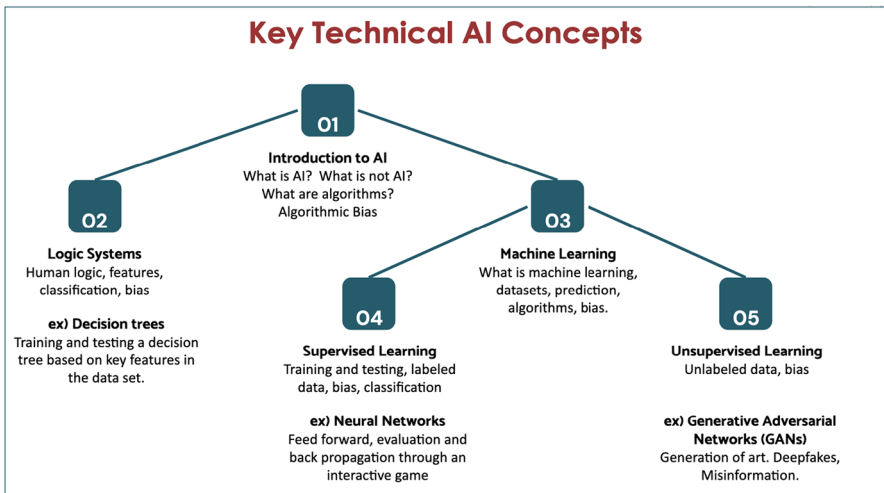
Fig. 1 The organization of key AI concepts covered in DAILy curriculum

To make AI learning engaging and accessible, the DAILy activities utilize everyday contexts and interactive activities (e.g., hands-on, kinesthetic) to explicate AI processes and implications and emphasize the relevance of AI to students' lives. For instance, the Pastaland activity in Module 2 is designed to engage students in learning decision trees and experiencing the construction of a decision tree. After they learn the definition of decision tree (a flowchart-like structure that looks sort of like an upside-down tree) and discussed a few examples of decision trees, students are introduced to a problem:

> "*There is a land of pasta known for most excellent cuisine with a queen who wants to classify all the dry pasta in her land and store them in bins. She wants to be able to quickly find the pasta she needs to cook the dishes she desires.... YOU, as a subject in PastaLand, are tasked with building a classification system that can be used to describe and classify the pasta so the pasta can easily be found when the queen wants a certain dish. Please help the Queen of Pasta-Land!!*"

Student groups are given 12 types of pastas (e.g., farfalle, ravioli, gemelli) and need to create a decision tree to sort them out. Afterwards, they test their decision tree using a new type of pasta and discussed whether and how their decision tree could sort out the new pasta. Due to the pandemic, this activity was implemented virtually and students worked with images of different types of pastas and created the decision trees using Google Drawing and placing the images of pastas on the branches of the tree. Figure 2 shows an example of the Pastaland activity. Another example of interactive activities is the participatory game designed to teach artificial neural networks in Module 4 where students play the role of nodes in a 3-layer NN, reflect on how they were modeling an NN, and connect between their actions in the

Table 1 DAILy curriculum: interweaving of AI concepts, ethics, and careers

Module	Category	Description
Module 1: Introduction (3 h)	AI concepts	Learn an age-appropriate definition of AI; Discern instances in everyday life that utilize AI
	Ethics	Design an algorithm to make "the best" peanut butter and jelly sandwiches; Exercise the understanding of algorithms and stakeholders in algorithmic decision making
Module 2: Logic Systems (4 h)	AI concepts	Learn how to create Decision Trees by classifying pasta by its key features; Test their tree's ability to sort new pasta type; Imagine examples of using classification in real life
	Ethics	Identify how the features selected impact the decision trees' overall structure and pasta classification capabilities; Discuss how bias can seep into decision trees by people who make different choices of features and how to mitigate bias
	Careers	Imagine and share what the future careers might be
Module 3: Supervised Learning with Teachable Machine (6 h)	AI concepts	Learn the definition of machine learning; Train a supervised learning model of image classification in Teachable Machine; Experiment with developing different classifiers of image, sound, or gesture data in Teachable Machine
	Ethics	Experiment with using imbalanced datasets to train models using Teachable Machine and compare with the models trained using more balanced datasets; Discuss how to mitigate the bias and generalize to other examples of algorithmic bias
	Careers	Create an inventory of themselves by identifying their strengths, weaknesses, and interests; Find and explore careers that match their strengths and interests
Module 4: Supervised Learning with Neural Networks (3 h)	AI concepts	Participate in a game where students play the role of nodes in the NN and "train" the NN to caption images; Reflect the process involved in the game and learn the structure and functioning of neural networks
	Ethics	Explore where and how the NN system can be biased; Watch a video of an interview with experts working on AI ethics
	Careers	Watch a video about how AI transforms a truck-driver's work; Explore how their matched careers are/will be impacted by AI; Share and discuss exploration findings with peers

Table 1 (continued)

Module 5: Unsupervised Learning and Generative AI (14 h)	AI concepts	Learn about Generative Adversarial Networks (GANs) through examples of generated art, images, faces, videos and sounds; Play a simulation game where they play the roles of generators and/or discriminators to see how generator and discriminator network in a GAN works; Experiment with a web tool to train their own text generator in a text style of their choice; Learn about how misinformation and fake news is generated and how it spreads through a digital game
	Ethics	Discuss the ownership of machine-generated art; Collectively play a game to understand societal consequences of AI-generated media
	Careers	Watch videos of people working in arts and media who use AI in work; Discuss fields that AI skills are applicable and how creativity is applied; Create a career roadmap to outline steps, resources, and potential barriers along their pathway to a career in the future AI era
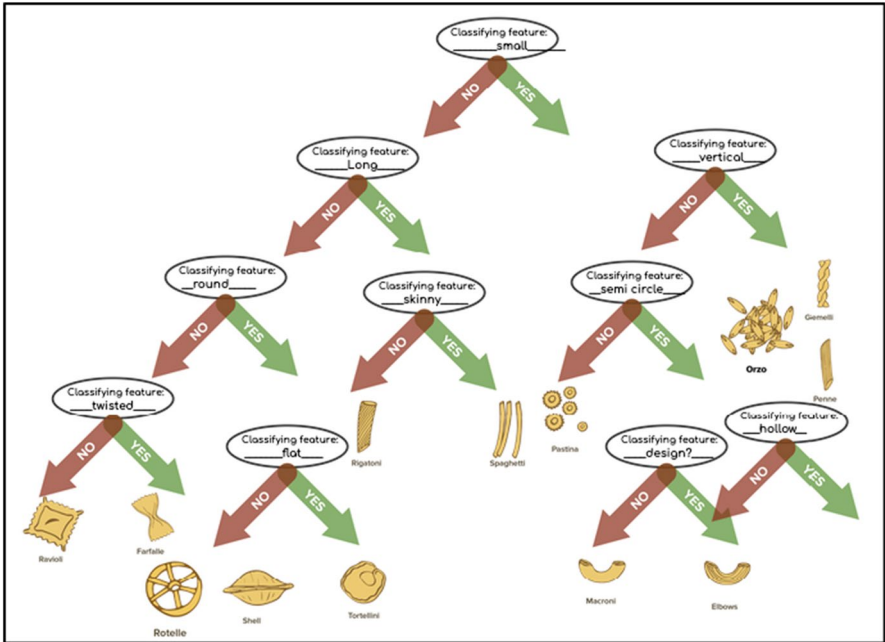
Fig. 2 A student's construction of a Decision Tree to uniquely classify pasta during the Pastland activity (in progress)

game and the processes in NNs. Detailed information about the game can be found in later Section "Virtual Implementation of DAILy Workshop" and Fig. 3.

Further nearly half of the DAILy activities are devoted to generative AI and introduce how AI has been integrated with arts, media, and social media (Module 5).
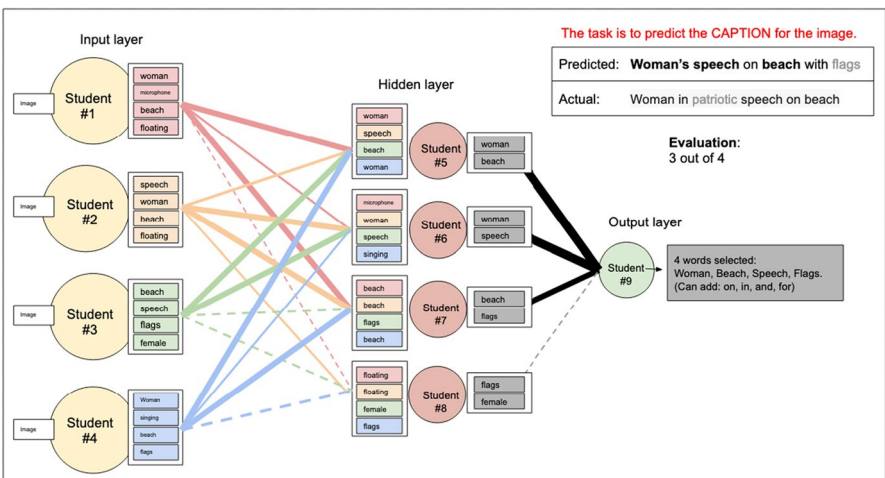

Fig. 3 The online Artificial Neural Network Game during play

These activities not only leverage students' interest and experience in these fields to engage them but provide a low entry into AI for students who may be more interested in topics other than STEM or CS and those who may not be confident about their technology skills. The technical aspect addressed in Module 5 includes that generative AI algorithms learn to create novel data that could belong to the training dataset; and algorithms such as Generative Adversarial Networks (or GANs) have the ability to produce novel visual art, text, music and even videos. GANs consist of a generator network that creates new data instances, and discriminator network that classifies the generated media based on a training dataset, and provides feedback to the generator network. GANs have found applications in art, education, robotics and healthcare. Module 5 engages students in understanding how GANs work, exploring AI tools that generates text, images, and videos, and discussing the ethical and societal implications of Generative models such as Deepfakes or potentially fueling the spread of misinformation.

AI Ethics and Career Futures

Each module utilizes a similar structure to engage students in interweaving their learning of technical, ethics, and career of AI. Students typically start with learning of the key technical AI concept, investigate potential for bias in the datasets and algorithms (and potential mitigation strategies), discuss the societal and ethical impacts of biased AI systems, and connect AI to their daily lives and future selves by engaging in career exploration activities. For instance, in Module 1, after learning that AI uses algorithm, students learn and experience how different stakeholders make different algorithmic decisions. They first create algorithms for making the best peanut butter jelly sandwiches and compare the algorithms. Then they bring their understanding of stakeholders like doctors, parents, and classmates (e.g., doctors probably care more about nutrients over the taste of the sandwich, classmates may care more about the taste, and parents may care both), to determine what "the best" peanut butter jelly sandwiches would include based on the perspectives of the stakeholders. Afterward, they discuss how the algorithms of "the best" peanut butter jelly sandwiches made by a parent would differ from the algorithms made by a doctor or a teenager. In Module 2, students learn about various bias issues of AI through investigations of technologies (e.g., doing a Google search for images of "outdoor recreation", exploring the face dataset from QuickDraw) and discuss the implications of the bias they found in existing technologies such as who may be impacted and how they may be impacted.

The career exploration activities were designed to reinforce students' ideas of the ethical and societal implications of AI. Drawing upon the Psychology of Working Theory (Duffy et al., 2016) and the Tools for Tomorrow curriculum (Kenny et al., 2004) that fosters career and self-exploration as well as knowledge about resources and barriers, the DAILy career activities seek to facilitate students' exploration of potential career trajectories in the AI age. The core elements of the career component focus on enhancing 1) critical reflection and action (e.g., finding out what students' interests and strengths are, discerning factors that contribute to accessing

meaningful and decent work); 2) proactive engagement (e.g., investigating how their desired careers will be impacted by AI through YouTube videos, developing the skills and orientation to take actions to maximize one's volition and initiatives); and 3) social support and community engagement (e.g., examining whether students have access to relational and community support). A set of videos of interviews with AI experts working in different fields (entertainment, law, automotive, etc.) were also included to further inspire students.

Methods

Research Questions

This exploratory study aimed to examine whether and to what extent the DAILy curriculum helped middle school students develop AI literacy. Using a mixed-methods convergent design (Creswell & Plano Clark, 2018), this paper focused on the following research questions:

1. What are students' conceptions of AI? Do these ideas change after the DAILy curriculum?
2. Does learning with the DAIly curriculum change students?

   a) technical knowledge and skills of AI?
   b) attitudes toward AI and AI careers?
   c) ideas about AI's ethical and societal implications?

Participants

Twenty-five middle school students in a summer STEM program participated in this study. The STEM program was an urban outreach initiative of a local university, with a focus on recruiting urban students who have a grade of C or lower in science/technology courses, are from low-income families, and may become the first person to go to college in their family. The recruitment was conducted through the partnership between the university and local urban schools: the students were first identified and recommended by their teachers and/counselors. Then the STEM program coordinator reached out to the students and their families and held information sessions to inform them about the program and to encourage them to fill in the application form.

A majority of the participants of the DAILy workshop were from groups underrepresented in STEM/CS, with 32% being African American, 56% Hispanic/Latino, and 56% female. Most of the participating students (92%, $n=23$) were in the grade 7–9 range. Students were asked about their prior experience with using Scratch as a proxy for programming/coding experience: 58% had no prior experience and 42% reported as having some experience.

Virtual Implementation of DAILy Workshop

Due to the pandemic, the DAILy workshop was implemented online. Students met online for ten hours every week for three weeks. All the activities were implemented synchronously via Zoom and all curricular materials were accessible through Google Classroom. The workshop was taught by a team of educators (2 educators and 3 teaching assistants). The sessions typically started with the instructor introducing the unit's topic, followed by a whole-class activity, a small group or individual activity, a discussion relating to ethical implications, and its connections to AI careers. Participants were randomly grouped into three groups of 7 or 8 individuals for small group discussions and hands-on activities. We found that students in this study completed all curricular activities with high attendance rates (85%—95% of all participants attended each day).

The major change for the virtual implementation of the DAILy workshop was that several activities were converted from live-action to virtual formats in order to be accessible during the COVID-19 pandemic. One example is the participatory simulation called the "Artificial Neural Network Game." In the original game, students played the role of nodes in a neural network formed by students sitting in predefined rows representing layers of the neural network. Input nodes are provided with an image and tasked with writing four words on individual cards that describe the image. Each of the words is distributed to each of the four hidden layer nodes. Hidden layer nodes select two words from their set of four to pass on to the output node. The output node creates a caption using four of the eight words it has received from the hidden layer nodes. Subsequently the "unveiling" takes place wherein the original image and its caption are exposed for all to see. Next, we introduce the processes of back propagation and gradient descent to mimic the training process in supervised learning. After the unveiling of the original image and its caption, students come up with an evaluation function to assess how well the network performed on captioning, then feedback is provided to nodes by passing circled words (if the word appeared in the original caption) or uncircled words back to their originators. After a discussion of the feedback and possible adjustments to the node's word selection behavior, the students can play additional rounds with new images and captions to see if/how the neural network learns to get better at captioning. During wrap-up discussions, facilitators reinforce that students are modeling an artificial neural network and review the analogies made between the actions in the game and processes in supervised learning. In this version of the game, back propagation was described as sending information back through the network wherein each node reflected on which nodes passed them "good" information, and using this information in subsequent rounds when picking words to send on.

To port the Artificial Neural Network game to an online activity, the network diagram was laid out in a Google Drawing and layered with data boxes to hold words. Figure 3 shows the Google Drawing for this NN activity with student sample data. Players in an online platform (typically Zoom) were given access to the Google Drawing and assigned the role of nodes as before. The input nodes were moved to a breakout room to privately view the original image to be captioned prior to returning to the main room to generate descriptive words and enter them into the data boxes.

In the feed forward phase, words were distributed as before to the nodes in the next layer. A drawback of this instantiation is that all of the selected words were visible to all players during game play thus reducing the element of surprise and possibly impacting the selections made by those acting as hidden layer nodes. An affordance of the digital version is the ability to represent and change the weighting of the links between nodes thus making it easier for players to remember which nodes provided them with "good" information in prior rounds. This online version was tested with students participating in an after-school club before the DAILy workshop.

Data Collection

We collected quantitative and qualitative data to make sense of elements of AI and AI processes that students were able to grasp or not and the impact of DAILy on the affective domains of student AI learning. The data were collected from students who attended at least 25 h of the DAILy workshop (we eliminated data of students who missed two or more days of the workshop). The quantitative data include responses to a pre and posttest, administered before and after the camp to participating students. The pre/posttest included three instruments developed by the research group: the AI Concept Inventory (AI-CI); Attitudes toward AI survey; and AI Career Futures survey. More details of the instruments and scoring are described later.

The qualitative data include students' final presentations completed at the end of the workshop, observation notes, and semi-structured interviews with a group of purposefully selected students ($n = 19$) after the workshop. The interviewees were selected to represent a range of variety in terms of age, gender, ethnic background, and prior technology experiences. The interview questions focused on student learning experiences and their understanding of the ethical and social implications of AI. In this paper we focused on student responses to three semi-structured interview questions, *(1) What are the benefits of AI?; (2) What are the harms of AI?; and (3) If you are going to build an AI system, what would you do to ensure it's fair?*. Audio recordings of interviews were transcribed and coded using grounded theory (Strauss & Corbin, 1994) to reveal student understanding of the AI implications to the society. The findings were triangulated with observation notes and other data collected to generate an in-depth understanding of how students learn about bias issues in AI.

Instruments

Three instruments were developed to investigate students' learning of AI with the DAILy workshops: AI Concept inventory, Attitudes toward AI survey; and AI Career Futures survey. To ensure the validity of these instruments, we first conducted reviews of the items with AI researchers and STEM educators to gather their feedback of the content validity. Afterwards two rounds of cognitive interviews were conducted with six middle school aged students (not included in this study) to determine if the items meet the targeted construct and assessment objectives. Then a pilot testing was conducted with 30 middle school aged students (not part of this study population) to investigate if there were any confusions in the wording of

questions and overall language appropriateness of the items in the instruments for middle schoolers. We revised the instruments based on the feedback from the cognitive interviews and the pilot test. In this study, participants completed the identical revised instruments before and after the DAILy workshop.
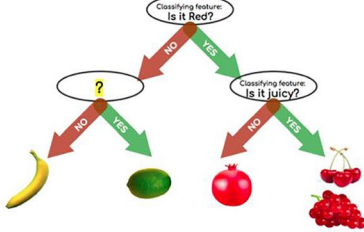
AI Concept Inventory (AI-CI)

Informed by Porter and colleagues' work on modular concept inventory (Almstrum et al., 2006; Porter et al., 2014; Taylor et al., 2014), AI-CI focuses on measuring student understanding of a set of core concepts and processes covered in the DAILy curriculum. In total, AI-CI consists of six scales. Table 2 shows example questions of AI-CI.

1) AI general concepts, including 12 multiple-choice (MC) questions asking if students can distinguish between technologies that involve AI and one explanation item asking students what they think AI is.
2) Logic systems, including 4 MC questions to assess students' understanding of the processes decision trees utilize and their ability to apply the understanding to categorize things and make decisions.
3) Machine Learning (ML) general concepts, which includes 7 MC questions asking students to distinguish between example technologies of classifying and generative AI, and to discern examples of supervised and unsupervised learning.
4) Supervised Learning (8 MC questions), which tests student understanding of the processes of supervised learning and ability to apply their understanding to determine how an AI technology would classify things based on the labeled data on which it is trained.
5) Neural Network (4 MC questions) that examines student understanding of the processes of neural network.
6) Generative Adversarial Networks (GANs, 4 questions) that focuses on examining students' understanding of generator, discriminator, and processes of GANs through True or false and MC questions.

Attitudes toward AI Survey

The Attitudes toward AI Survey focuses on soliciting participants' interest in AI, anxiety about AI, and relevance of AI to their life. These survey questions are 5-Likert scale questions drawn and modified from validated instruments of the Science Motivation Questionnaire II (Glynn et al., 2011), the Modified Attitudes Towards Science Inventory (Weinburgh & Steele, 2000), and the AI anxiety scale (Wang & Wang, 2019). Each item presents a statement and asks students to indicate how strongly they agree/disagree with the statement (strongly agree to strongly disagree). The Interest in AI scale consists of 9 items (Cronbach's alpha = 0.94) and sample statements include "*I will take a class about AI if it is offered in my school*" and

Table 2 Example AI-CI questions

Scale	Sample Items
AI General Concepts	Review the following technologies and indicate whether or not they involve artificial intelligence (AI) A software that can classify different types of hats — AI ○　　Not AI ○
Logic Systems	Pick which question was most likely used at the "?" oval to classify this set of fruit: ○ Is it long?　○ Is it blue?　○ Is it small?　○ Is it sour?
Machine Learning (ML) General concepts	A doctor sends x-ray scans out to an AI to determine whether or not you have broken a bone. Is this an example of an AI classifying or a generating? ○ Classifying ○ Generating
Supervised Learning	An AI technology has been trained to classify images only using the following data: Banana　Broccoli　Banana　Broccoli Which of the following would you expect the technology to predict? ○ All things in multiples are broccoli ○ All crunchy things are broccoli ○ All things in multiples are bananas ○ All yellow things are bananas
Neural Network	Which of the following ways enables a neural network to learn? (check all that apply) ☐ By being told what to do by a supervisor ☐ By making a guess and finding out if the guess was right or wrong ☐ By rearranging the network ☐ By strengthening connections in the network ☐ By changing the algorithm it uses to make guesses
GANs	True　　False A generator and discriminator are both neural networks — ○　　○

"*I will talk to members of my household about what I know about AI.*" The relevance scale has 8 items (Cronbach's alpha of 0.71), with sample item statements like "*Learning about AI is relevant to my life.*" The anxiety scale includes 9 items with

Cronbach's alpha of 0.91 and sample item statements like "*I do not feel comfortable about using or building AI technologies*" and "*Working with AI would make me nervous.*" Each scale was examined for item-total correlations and Cronbach's alpha if deleted to determine that the items were measuring the same construct.

AI Career Futures Survey

The AI Career Futures survey examines students' awareness of AI impact on future jobs and career adaptability because the DAILy workshops engaged students in exploring careers that matched their interest, finding out how AI has been influencing the matched jobs, and becoming knowledgeable of preparations for entering their desired career fields. The career adaptability subscale was developed based on the revised Career Futures Inventory (Rottinghaus et al., 2012), which is built on a theoretically-derived, multidimensional definition of adaptive behavior for future careers such as identifying educational choices to join future jobs and investigating consequences of the actions. Similar to the Attitudes toward AI survey, the AI Career Futures Survey items are 5-Likert scale questions, each of which presents a statement and asks students to indicate how strongly they agree with the statement. Sample statements of the AI career awareness scale include "*I know about jobs that use AI*" and "*I know a role model of my background in fields related to AI.*" Sample statements of the career adaptability scale include "*I learned to prepare for the future,*" "*I learned about educational choices that I must make to get my dream job,*" and "*I became aware of my future job choices.*" In total, the AI careers awareness scale consists of 7 items with Cronbach's alpha of 0.822, and the career adaptability scale was composed of 10 items and had a reliability of alpha $= 0.934$.

Scoring and Data Analysis

All the multiple-choice questions in the AI Concept Inventory were coded as 0 (wrong) or 1 (correct). The 5-Likert scale questions in the Attitudes toward AI survey and the AI Career Futures survey were coded based on student rated agreement with the statements (1: strongly disagree, 5: strongly agree) and reversed coded if needed. Student responses to the items were scored and added to form an aggregate score of the instrument and of each subscale. We conducted statistical analysis such as paired t-test to compare student performance on the pre and posttest. We also examined student responses to each subscale to identify AI concepts and processes with which students had difficulties.

Besides the quantitative analysis, we also analyzed student responses to the open-ended question in the AI Concept Inventory that asked them to explain what they think AI is. We categorized the explanations using emerging categories. The categories reflect the ideas that students put forth in their definitions and are not meant to be treated as a "score." Categories were developed and iteratively revised by two authors of this paper who coded the explanations independently first and then discussed the coding with another author to solve conflicts. We compared the percentage of students with each category of ideas before and after the workshop to capture

a nuanced in-depth picture of what ideas students had about AI before the workshop and whether and/or how their ideas changed after the workshop.

Findings and Discussions

Students' Conceptions of AI Before and After the DAILy Workshop

Using emerging coding, we categorized student responses to the question "What do you think AI is?" on the pre and post-test. In total six categories of ideas emerged (see Table 3 for an explanation of the categories and sample student answers). In the "Incorrect" category, students explained wrong or naïve ideas of AI, e.g., "*AI is a human intelligence processed by machines.*" In the "Vague" category, students expressed some ideas of AI but the ideas were unclear, e.g., "*I think it's like a lot of computerized things and people make things or robots with AI. That's all I know.*" In the "General" category, students often referred to human intelligence and explained that AI is different from human intelligence. Students holding the "Societal" category idea of AI explained AI's impact on the society or its ethical implications. For instance, a student explained that "*I think it [AI] is technology that can change the world. Also [AI] can be an everyday technology. But it sometimes has malfunctions.*" Students with the "Technical" category of idea focused their explanation of AI on the technical components such as data, algorithm, and making predictions, e.g., "*A program that can process information. The information turns into an action that the AI was told to do.*" For students who had the "Complex" category of ideas, their explanation included two or more correct ideas about social, general, or technical aspects of AI and reflected a complicated understanding of AI, e.g., "*AI means intelligence made by humans in devices and technology. The intelligence is gained by training and testing a data set. It has a whole process which is called an algorithm. AI heavily affects us today and will even more in the future.*" Students who held the "Complex" ideas started to recognize both the technical and ethical/societal aspects of AI technology and probably consider AI as a sociotechnical system. Students with "Technical" or "Societal" ideas only considered one aspect of AI and did not have a sociotechnical perspective of AI.

We compared the percentages of students who demonstrated each type of ideas in their explanations on the pre and post-test (see Fig. 4). A majority of students started the workshop with a vague (43.5%) or incorrect (34.8%) description of AI. The most common incorrect answers were "a type of human intelligence" or "coding a robot" and vague answers typically defined AI as "intelligent technology" without a clear understanding of what made it intelligent (see Table 3 for examples). Fewer students defined AI in terms of the technical components (4.3%) or with some combination of ethical, general, and technical (4.3%). No students defined AI in terms of societal impact. On the post-test, students most commonly defined AI as incorrect (30.4%). However, several students described AI as a combination of ethical, general, and technical (or complex definitions = 21.7%), indicating that these students have started to develop a sociotechnical perspective of AI. Fewer students defined

Table 3 Categorization student ideas of what AI is

| Category	Description	Example student explanations
Incorrect	wrong ideas or naïve ideas around artificial intelligence	AI is a human intelligence processed by machines I think it means that the intelligence is artificial and needs to be grown before unlocked
Vague	vague ideas about AI	I think it's like a lot of computerized things and people make things or robots with AI. That's all I know Something artificial that can do things by itself and not have to have a human help it do a task
Societal	Focuses on the societal or ethical implications of AI	I think it is technology that can change the world. Also can be an everyday technology. But it sometimes has malfunctions What I think AI means is a program that has been made to help humans. it helps humans with their life because it helps in jobs. it can help you in so many things
General	Focuses on the general definition of AI, e.g., correct reference to human intelligence (viewing AI as explicitly different from human intelligence)	AI is intelligence demonstrated by machines unlike intelligence shown by humans and animals Artificial intelligence is about programming computers software that are able to replicate human actions
Technical	Focuses on the technical skills and knowledge around AI (e.g., data, algorithm, prediction)	A program that can process information. The information turns into an action that the AI was told to do
Complex	Includes two or more correct ideas about social, general, or technical aspects of AI	AI means intelligence made by humans in devices and technology. The intelligence is gained by training and testing a data set. It has a whole process which is called an algorithm. AI heavily affects us today and will even more in the future
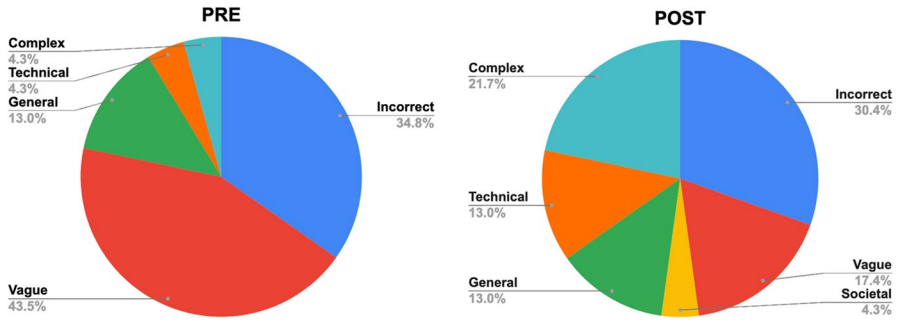
Fig. 4 Percentages of students who defined AI incorrectly, vaguely, in terms of societal impact, in relation to human intelligence, in terms of technical structure, and "complex," meaning two or more references to societal, general, and technical definitions on the pre and posttest

AI in vague terms (17.4%). The most common incorrect answer in the post test was describing artificial intelligence as "exactly like human intelligence" or "robots".

A closer examination revealed that of the students who started the workshop with an incorrect understanding of AI, 37.5% students developed a complex understanding of AI after the workshop, but 25% continued to stick to the incorrect understanding. For students who started the workshop with a vague understanding, some started to integrate ideas about AI's societal implications, underlying concepts and processes, and general definitions of AI (30%), although many students remained holding vague ideas (30%) or ended up adding non-normative ideas of AI (40%). With regard to students who started with a general understanding of AI, 66.7% incorporated new ideas (e.g., technical, ethical) and developed a complex understanding of AI after the workshop. Overall, we saw 21.7% of students who did not start with a complex understanding of AI develop it by the end of the workshop.

Comparing Students' Technical Knowledge and skills Before and After the Workshop

The paired t-test showed that students on average significantly improved their AI-CI scores after the workshop (pre: Mean = 24.37, SD = 4.83; post: Mean = 26.75, SD = 4.50; $p < 0.01$). On the post-test students on average achieved a gain of 2.36, which indicates that a majority of the students answered two or more questions correctly after the workshop. Students made the biggest learning gains on four sub-scales: (1) general concepts of AI, (2) logic systems, (3) general concepts of machine learning, and (4) supervised learning. Table 4 shows the paired t-test results of student performances on AI-CI. Next, we report student performances on each scale in more detail.

AI General Concepts: Recognizing AI and Identifying Features of AI Technology

Recognizing AI Students made statistically significant gains (Cohen's $d = 1.03$) on four items that describe different technologies and ask students whether the

Table 4 Paired t-test results of student performance on AI-CI before and after the DAILy workshop

| Scale	Pre		Post		Paired t-test results	Effect size (Cohen's d)
	Mean	SD	Mean	SD		
Aggregate (total score = 38)	23.69	4.61	26.05	4.34	t(24) = 3.37, p < .01	.53
AI general concepts (total score = 12)	7.96	1.73	8.56	1.55	t(24) = 1.56, p = .12	.37
Logic systems (total score = 4)	2.6	1.11	2.99	.76	t(24) = 2.27, p < .05	.41
ML general concepts (total score = 6)	4	1.04	4.44	1.53	t(24) = 1.74, p = .09	.34
Supervised learning (total score = 8)	4.04	1.43	4.6	1.5	t(24) = 1.51, p = .14	.38
NN (total score = 4)	2.89	1.33	2.87	1.14	t(24) = -.10, p = .93	.02
GANs (total score = 4)	2.38	.82	2.67	1.00	t(24) = 1.50, p = .07	.32
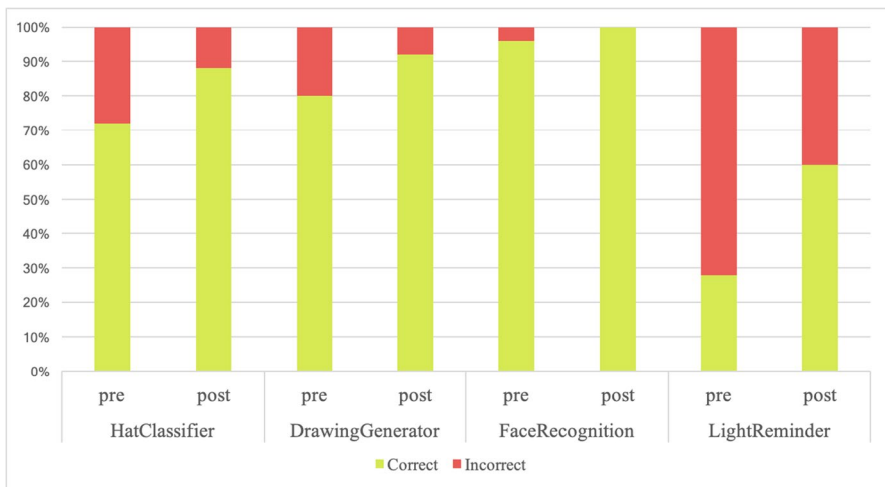
Fig. 5 Percentages of students who answered the Recognizing AI questions correctly on pre and posttest

technology utilizes AI (a software that classifies different types of hats, a drawing app that generates new paintings, a face recognition software, and a phone app that reminds people to turn off lights). On the pretest, most students were able to correctly determine the use of AI in the first three technology examples (average score = 2.76). On the posttest, over half of the students (n = 13, 52%) answered at least one more question correctly (4 students answered two or more questions correctly). Figure 5 shows the performance of students on these four questions.

One interesting finding is that students improved the most on the LightReminder question, which asks students whether AI is utilized in a smartphone app that sends a reminder for people to turn off light between 9am to 3 pm. On the pretest, only 18 students (72%) answered it incorrectly, i.e., they thought that AI was required in this app. On the posttest, 60% (n = 15) students realized that AI was not necessary for the app. This suggests that by exposing students to various example technologies that do

and do not use AI, the DAILy workshop helped improve students' ability to recognize AI, a critical competency for informed interactions with AI (Long & Magerko, 2020).

Identifying features of AI technology Four questions were designed to test whether students were able to identify key features of technologies that use machine learning. Each question describes a technology and asks students to determine if the technology uses datasets and makes predictions. We aggregated student responses to the dataset and predication questions respectively and compared the responses from pre to posttest. The results show that students achieved moderate gains in terms of identifying the use of dataset (pretest Mean=2.44; posttest: Mean=2.72, p=0.07, Cohen's d=0.40), but no gains on the prediction questions after the workshop (pretest: Mean=2.76, posttest: Mean=2.44, p=0.19). Nearly half of the students (n=10, 40%) still answered the prediction questions incorrectly in at least two scenarios on the posttest. Ten students' scores on these questions decreased on the posttest, suggesting a confusion about prediction among students.

Our observations of the "AI or Not" activity (Module 1) resonate with this finding. In this activity students were presented with a list of technologies and asked to distinguish whether it uses AI based on three criteria: algorithm, prediction, and dataset. We noted that students had a hard time determining what prediction in the context of AI means and whether/how it is different from human intelligence. For instance, in discussing whether automatic doors involve AI, some students thought it involves AI because "*it automatically opens*" and "*it makes a prediction to open when you step on the sensor*", whereas other students disagreed because automatic doors involve "*just a motion sensor that detects when you are in a close radius*". In another example of determining whether regular cars use AI, one student argued that "*it [regular car] uses AI because the brake predicts to stop when you press on it*". This shows that students tended to reason about prediction based on their everyday experience of human intelligence but may have had different interpretations of prediction. They encountered challenges when discerning technologies that make predictions. Incorporating more concrete examples of predictions and providing a definition of prediction in the context of AI can help contextualize this concept and improve student ability to analyze and determine if the technology involves AI.

Logic Systems

Students achieved significant gains on items that assessed student understanding of logic systems (pretest Mean=2.6, SD=1.11, posttest Mean=2.99, SD=0.76, p<0.05). They started with high pretest scores on this scale, with fifteen students (60%) correctly answering the three questions that ask to categorize things using a decision tree. This suggests that students may have acquired knowledge of how computers categorize things from previous experiences. They, however, had little or no prior knowledge with regard to how to build a decision tree (15 students incorrectly answered the question of choosing steps and sequence involved in building decision
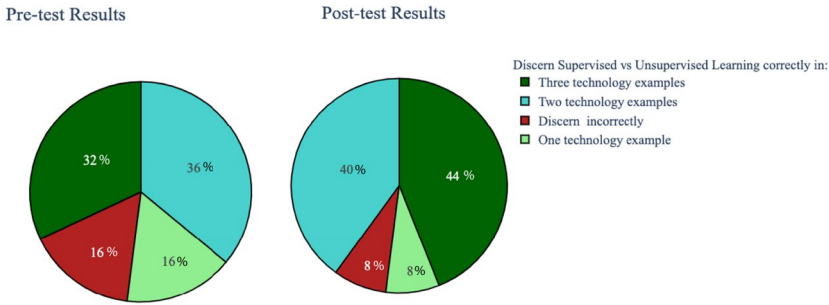
Fig. 6 Percentages of students who answered the supervised and unsupervised learning questions correctly in one, two, or three technology examples on the pre and post-tests

trees on the pretest). On the posttest all students answered at least one question correctly and 80% of the students correctly answered three or all four questions. Nine students chose the correct steps and sequences for making decision trees and five students chose the correct steps but put them in a wrong order.

This finding is promising given the short duration of the intervention (the decision tree/Pastaland activity lasted approximately one hour). It was also noteworthy that the implementation of the Pastaland activity did not explicitly teach students the steps or sequence. By building and using decision trees to categorize different types of pasta, students developed ideas about the processes involved and were able to answer the questions correctly.

Machine Learning General Concepts

Supervised vs. unsupervised learning Students achieved a significant improvement on the set of supervised vs. unsupervised learning questions after the workshop (Pretest: Mean = 1.84, SD = 1.07; Posttest: Mean = 2.2, SD = 0.91; $p < 0.05$). After the workshop, ten students (40%) improved their scores and correctly determined the use of supervised or unsupervised learning in more technology examples. Nearly half of the students ($n = 11$, 44%) correctly identified the use of supervised or unsupervised learning in all the three scenarios on the posttest. This shows that after the workshop, students were able to discern supervised and unsupervised learning by examining whether the technology uses labeled data. Figure 6 shows the percentages of students who correctly identified supervised or unsupervised learning in none, one, two, or three technology scenarios.

Classifying vs. generative AI We did not find significant gains in students' performance on the set of classifying or generative AI questions (Pretest: Mean = 2.16, SD = 0.55; Posttest: Mean = 2.24, SD = 0.93, $p = 0.70$). One reason for this was the ceiling effect: on the pretest 23 students (92%) were able to distinguish between classifying and generative AI correctly in at least two technology examples. This indicates that students may have developed normative ideas about classifying and generative technologies from previous experiences.

Supervised Learning

Students improved their understanding of supervised learning after the workshop: the average score on this scale increased from 4 on the pretest to 4.6 on the posttest (total score of 8). In particular students achieved significant gains on four items asking about the processes of supervised learning (e.g., *"What is the difference between the training and testing phases in Supervised Learning?"*, *"What is the purpose of a label on an image in Supervised Learning?"*). They on average had little or no prior knowledge (pretest Mean = 1.38, SD = 0.97) before the workshop. On the posttest, over half of the students (*n* = 14) were able to tell the purpose of the labelled data *in supervised learning*. Sixteen students were able to apply their understanding of supervised learning to identify one reason for a flawed machine learning system that can recognize men's shoes accurately but not women's shoes. These findings suggest that the hands-on activities in Module 3 of the DAILy curriculum was effective in promoting students' understanding of supervised learning. The Teachable Machine activity engaged students in experiencing how supervised learning takes place by training and testing their models. Students also experimented how using an unbalanced dataset can result in a biased model. These activities helped demystify supervised learning processes and also reinforced the idea that AI technology can be biased due to the dataset it was trained on.

Neural Networks (NN)

Students made no gains on instruments that examine student understanding of the processes of neural networks (Pretest Mean = 2.89, SD = 1.33; Posttest Mean = 2.87, SD = 1.14). After the workshop, 40% of the students (*n* = 10) did not recall any processes involved in neural networks. Over half of the students (*n* = 13) could not tell whether learning takes place during the testing or training phase in neural networks. Three reasons may account for this. First, the intervention was too short. Students in total spent approximately 40 min on the Neural Network activity where they played a participatory simulation game to make sense of how a neural network works (Module 4). The processes (e.g., feeding forward, evaluation, and back propagation) were only mentioned once on a set of 3 slides. Students may need more time and scaffolding to understand the steps of NN. Therefore, although almost all the students we interviewed expressed that they liked "the telephone game" (as we called the NN game), they struggled to make sense of the NN processes. Second, the online learning format limited student interactions and interpretations. Unlike the face-to-face sessions where students can act out and observe how others respond and how the system as a whole behaves, playing the simulation game via videoconferencing greatly limited the interactions between students and made the observations much more challenging, and thereby hindered student interpretation of NN processes. Third, the NN activities in Module 4 did not include real-world applications of NN to foster student understanding of this concept. (Neural network was the most abstract concept discussed in the DAILy curriculum.) Without concrete examples, it may be difficult for young adolescents to grasp the concept of NN. Future revisions of the curricular activity need to provide more scaffolding (e.g., whole group

Table 5 Student performance on the Attitudes toward AI and Career Futures surveys before and after the DAILy workshop (higher score indicates more agreement with the problem statement, total score = 5)

Scale	Pre		Post		Paired t-test results
	Mean	SD	Mean	SD	
Interest in AI	3.56	.48	3.56	.73	$t(24) = .04, p = .97$
Relevance of AI to their life	3.84	.47	3.85	.40	$t(24) = .06, p = .95$
*Anxiety about AI	2.48	.68	2.53	.74	$t(24) = .32, p = .75$
AI career awareness	3.15	.68	3.48	.60	$t(24) = 3.81, p < .001$
Career adaptability	3.82	.55	3.85	.61	$t(24) = .19, p = .85$

*A higher score (maximum = 5) in Anxiety about AI means higher anxiety

discussions and explicit instructions on what to observe) and utilize examples of the use of NN in our life to help students connect what they did in the game to the NN processes.

GANs

Students on average improved their understanding of how the GANs work, however the learning gains were not statistically significant (pretest: Mean = 2.36 (SD = 0.82), posttest: Mean = 2.64 (SD = 0.99), $p = .07$). The majority of students understood that the generator and discriminator were neural networks and that they followed a back-and-forth process to eventually create new media. However, students did not see the generator and discriminator as working against one another. This could be due to the fact that the short intervention was not sufficient in terms of helping students develop a solid understanding of the complex mechanisms of GANs (students spent 30 min playing the roles of generators and/or discriminators to see how a generator and discriminator network in a GAN works). More scaffoldings are necessary to explain how GANs work. Many of our lessons focused more on applications of GANs rather than reflecting and discussing the underlying mechanisms of GANs.

Comparing Students' Attitudes toward AI and Career Futures Before and After the Workshop

Attitudes toward AI

Students started the workshop with a positive attitude toward AI, which was not surprising given that this was a self-selected program. We did not find significant differences between the performances on the pre and posttest (pretest: Mean = 3.53, SD = 0.46; posttest: Mean = 3.63, SD = 0.29; $p = 0.32$). Table 5 shows the paired t-test results of each scale of the Attitudes toward AI pre- and post-survey. The largest increase was observed in an item that asked for students' interest in taking class in AI if it is offered at their school. On the pretest, 60% of the students ($n = 15$)

indicated vague or no interest (selected "disagree" or "not sure" to the statement), whereas on the posttest, 15 students agreed or strongly agreed with the statement. This shows that the workshop sustained students' interest in AI and students would like to continue with their AI education after this experience.

Students also continued to find relevance of AI to their lives after the DAILy workshop (see Table 5). Most students agreed or strongly agreed with each item statement and felt AI to be highly relevant to their life. The highest increase in scores was observed in an item asking for students' agreement with the statement "*It is useful to me right now to know something about AI*". This suggests that the DAILy workshop enabled more young adolescents to realize the importance of learning AI at their age. The biggest drop in mean scores was in the item of "*I can use what I learn about AI to help my community*". While students see AI as relevant, they may not know how to use it in their capacity to benefit others in their community. One reason for this may be that the DAILy workshop emphasized the bias of AI would have negative impacts on people from underrepresented groups but did not provide any real-life examples of how AI can help people in students' communities.

With regard to students' anxiety about AI, we found a minor increase from pre to posttest. More students agreed to the statements "*I think AI can be dangerous or harmful*", "*The use of AI in everyday life scares me*", and "*I am afraid that AI will make us lazy*" after the workshop. This indicates that more students left the workshop with an increased sense of the negative consequences of AI. Despite this increased anxiety, we also found decreases in agreement with the statements "*I think AI can be threatening to society*" and "*Working with AI would make me nervous*". Altogether this suggests that while students recognized the potential harms of AI, they felt using AI in a correct way would mitigate the damage. Meanwhile the experience of working with AI technologies during the workshop helped reduce their anxiety about AI.

AI Career Futures

Overall, students showed a significant increase in the awareness of AI related careers from the beginning of the program (Mean = 3.15, SD = 0.68) to the end (Mean = 3.48, SD = 0.60; $p < 0.001$; Cohen's $d = 0.52$). The biggest positive differences were related to (1) knowing about jobs that use AI, (2) knowing role models in the AI field, and (3) discussing AI with friends and family. The first two increases were likely due to the fact that the workshop exposed students to various jobs that may include the use of AI and people who work in AI related field. The third increase, however, indicates that the workshop has sparked an early interest in AI careers and students tended to discuss with their family and friends to find more resources and opinions for the topic (Maltese & Tai, 2010).

We also found that students reported developing some aspects of career adaptability skills (see Table 5). The biggest increases were observed in students' agreement with statements of "*I learned to think about what my future will be like*", "*I became aware of my future job choices", "I learned to recognize resources available to me and use them*", and "*I learned about educational choices that I must make*

to get my dream job". The career training sessions of the workshop engaged students in exploring their matched future job choices and creating a roadmap to enter those fields. Through these activities, students developed not only more concrete ideas of their future jobs but also skills of how to become more adaptive in the AI era. One student spoke about how knowing about AI might be helpful for them in the future, specifically their ability to provide clarity about AI to others who may be confused about its implications, *"They keep saying that the more and more in the future we go, the more we use AI. So, I think that if that's going to happen, the fact that I know a lot about AI, and if somebody has questions, or is worrying, or anything, I can reassure them that nothing bad is going to happen. Or if something goes wrong, I can fix it. I have more knowledge of what's going on, so I don't have a bunch of questions. I have a bunch of answers."*

Students' Ideas of Ethical and Societal Implications of AI

To explore what kinds of ideas students developed about AI's ethical and societal implications through the DAILy workshop, we examined the interviews conducted with the 19 students after the workshop. The students were selected to represent a wide range of gender, age/grade, and ethnicity. As noted above our analysis focused on student ideas about the benefits and harms of AI and how to make AI technology less biased.

After the workshop almost all the interviewees (except one) articulated both the benefits and harms of AI technologies. They believed that AI can make life easier, help complete dangerous work ("*AI is doing some, maybe, risky tasks that humans now don't have to do*"), generate more objective conclusions ("*it (AI) wouldn't be based on one person's thoughts, because it will collect data from different sources. And this could help the news reporters. I mean the journalists.*"), and accomplish tasks more efficiently ("*A great consequence about it, about AI is that we can shorten, or we can expedite a bunch of jobs or tasks that we do everyday that we can do over 50% faster.*").

With regard to the bad consequences of AI, students expressed three major concerns:

- potential for laziness or loss of autonomy: 8 students mentioned that they were afraid that human beings would become lazy and do not work, e.g., "*People will definitely get lazy and God forbid if anything were to happen and they would like stop working*"; "*It could make people not want to do as much work as before, maybe.*"
- discrimination against people of color due to the use of biased AI technology such as facial recognition systems: 4 students explained how the bias in AI would lead to bad consequences. For instance, one student explained this in detail, "*A bad use of AI would be how the government or the police, a police station using face recognition to recognize where you are from. Because black people, since you're a darker color and they think you're dangerous or something. Those are things that have happened before. They think that just because you live where*

lots of darker people live, there's going to be a higher crime rate, even if you don't do anything. Because many conditions can force a person to do something. We just never know about it. Which is why I think it's unfair for people to be labeled based on where they live. Because if you have the money to live in a rich neighborhood, you'll most likely take that chance to…"

- harms of the spread of deepfakes: 4 students mentioned the potential harms of generative AI. They worried that this may lead to faking peoples' words, pictures, videos, and identity, e.g., *"That sometimes it (AI) can affect the people in the real world. So hackers can do something, or maybe they can make a video with AI of someone's saying something that they never actually said, because of editing and people saying certain words, and then they just put all the words together. So they make it seem like someone actually said something. So can it be used for bad stuff."*

Further, 16 students provided constructive solutions when asked how to build an unbiased AI system. They emphasized the importance of using a dataset that includes diverse and unbiased data. For instance, one student talked about setting up a good dataset for the training purpose, *"I would use like different types of data sets to program my AI. Like if I was doing on facial recognition system, I would use people of like different ages, genders, ethnicities, and stuff like that to make it a little less biased and have like an equal amount of said people."* Another student explained using a diverse dataset for the testing phase, *"I would get people of color to come and also different types of people, different cultures too, to test it out and see."*

Overall, the interview results demonstrate that most students incorporated ethical and societal implications into their views of AI technology after the workshop. Analysis of students' final presentations suggested similar themes—students frequently referenced anecdotes of bias as memorable learning moments of the workshop, *"I learned from this program that artificial intelligence can be biased and that we put information like racism and sexism in our AI without realizing which can make the people using the AI not want to use it."* Further examining student explanations of what AI is yielded further evidence, as more students added ethical and societal impact of AI in their definition of AI on the posttest.

Conclusions

The rise of AI has attracted many researchers' attention and led to calls for creating opportunities to engage young learners in AI education. While the content and pedagogy of AI education at the K-12 level have not been established, recent review studies (Marques et al., 2020; Zhou et al., 2020) pointed out a dearth of AI curriculum and programs that incorporate AI ethics, a topic that is critical to prepare young students to become informed users and developers of AI technology. This paper aims to help fill in this gap by reporting the design and implementation of the DAILy curriculum that interweaves student learning of the technical, ethical or societal, and career aspects of AI. Our results demonstrate that the DAILy workshop

was engaging and productive in supporting student learning of general concepts of AI concepts, logic systems, and machine learning. Students enjoyed ethics-related activities embedded in the DAILy curriculum. They were able to internalize what they had experienced, connected to ethical implications of technology design, and developed positive ideas about their future selves with AI after the workshop. Overall, the approach of combining ethics and technical learning of AI is age appropriate and promising in terms of developing AI literacy among middle schoolers. Our work also contributes to the AI education field by providing a working definition of AI literacy grounded in findings of what middle school students are capable of learning and doing with AI.

Further our approach of positioning AI as closely related to society successfully engaged students from groups underrepresented in STEM and computing. We found that female students of color were particularly active in the investigation of bias and ensuing discussion. For instance, upon viewing the disproportionate representation of women of color in commercial datasets used for facial recognition, a female student stated "*What I also noticed is that you can see that the women either way … This is another thing about sexism. Women are also either way still have a lower percentage than the men in their class, I guess you could say.*" She further noted how AI systems are biased against people like her while she positions herself as a woman and part of a group who should be represented.

We also found that the students talked with their family members and friends about AI's impact on current and future jobs. A student in his interview described that *"I asked her [my mom], "Are you scared if AI takes your job?" She's like, "Yes, but I know I'll have another job." So, it's really cool what AI could do and make new things for other people."* This provides further evidence of student engagement as he was active in leveraging existing familial and aspirational capital (Yosso, 2005) to make sense of AI and refine his existing perceptions of AI. Overall, the emphasis on AI's ethical and societal implications offers an effective way to excite underrepresented students, which resonates with previous research that minority students are often more drawn to STEM/CS programs that teaches the content via the context of solving social justice issues (Mark et al., 2013; Vakil, 2018). This offers valuable insights into broadening participation in Computer Science and AI education.

Limitations and Next Steps

There are a few limitations of this study. First, the students were recruited to participate instead of being randomly selected. The students were from a very specific population, low-income families in urban areas and may become the first generation of college students in their family. Further, the workshop was taught by a team of researchers and educators who developed the DAILy curriculum. The findings may differ from situations involving participants, treatments, and settings different from those in the study. Another limitation is that this study measured immediate effects of the intervention using a pre-posttest design. Complex AI processes such as neural networks may require longer exposures and more time for students to reflect and internalize. Conducting a delayed posttest

and interviews could clarify the long-term effects of the DAILy experience on student ideas about AI.

As our next step, we plan to revise the DAILy curriculum based on the findings from this implementation. For instance, many students were found to remain confused about what prediction means in the context of AI after the workshop. In Module 1 (introduction of AI), we plan to add more concrete examples of prediction in AI (based on what is learned to make a prediction/inference) and engage students in discussions or debates on whether and why technologies make predictions or not. These activities would help clarify students' misconceptions and enhance their understanding of prediction in the context of AI. Another challenge we found was that activities of NN were engaging yet difficult for students. Students had a hard time connecting their virtual actions (playing of the role of the nodes) with the abstract processes of NN. We plan to include scaffolds to help students make the connections, e.g., asking students to record their actions, make analogies to the NN processes, and discuss with peers. In addition, some students provided the feedback that they hope to apply their knowledge of AI and ethics to help their community. Informed by DiPaola et al. (2020) work, we plan to include a capstone project for students to explain how their knowledge of AI ethics can help others. The capstone project could engage students in developing technologies with ethical considerations such as developing a fair local news feed and redesigning the recommendation system of social media. Such learning experience will not only engage students in applying their AI knowledge and skills, but also enable them to recognize their own AI capability to help others and become more confident in working with AI.

Materials Availability The materials presented in this work can be found on our website https://raise.mit. edu/daily/index.html.

Declarations

Ethics Approval The work presented in this paper was approved by MIT's Internal review Board, Committee on the Use of Humans as Experimental Subjects (COUHES) under approval no. 1912000072 and Boston College's Internal Review Board, under approval no. 20.207.01e.

Informed Consent to Participate and Publish All participants signed an assent form to participate in the study, which indicated that we will be collecting their in-classroom audio, video, chat and activity data. Since participants were minors, all participants' parents provided written consent prior to enrollment in the study. Students were told that they can withdraw from the study at any point, and they may choose to turn off their audio and video if they preferred.

References

Ali, S., Payne, B. H., Williams, R., Park, H. W., & Breazeal, C. (2019). Constructionism, ethics, and creativity: Developing primary and middle school artificial intelligence education. In *International workshop on education in artificial intelligence K-12 (EDUAI'19)* (pp. 1–4).

Ali, S., DiPaola, D., Lee, I., Hong, J., & Breazeal, C. (2021). Exploring Generative Models with Middle School Students. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–13).

Almstrum, V. L., Henderson, P. B., Harvey, V., Heeren, C., Marion, W., Riedesel, C., Soh, L.-K., & Tew, A. E. (2006). Concept inventories in computer science for the topic discrete mathematics. In *Working group reports on ITiCSE on Innovation and technology in computer science education* (pp. 132–145).

Badham, R., Clegg, C., & Wall, T. (2000). Socio-technical theory. *Handbook of Ergonomics.* John Wiley.

Boss, J. A. (1994). The effect of community service work on the moral development of college ethics students. *Journal of Moral Education, 23*(2), 183–198.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77–91). PMLR.

Coeckelbergh, M. (2020). *AI ethics*. MIT Press.

Crain, W. C. (1985). *Theories of development*. Prentice-Hall, Inc.

Creswell, J. W., & Plano Clark, V. (2018). *Designing and conducting mixed methods research*. Sage Publications.

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. L., Sonnert, G., Sadler, P. M., & Hazari, Z. (2012). Out-of-school time science activities and their association with career interest in STEM. *International Journal of Science Education, Part B, 2*(1), 63–79. https://doi.org/10.1080/21548455.2011.629455

Danyluk, A., Howe, A., Gini, M., & Anderson, M. (2014). Broadening Participation in AI. *Computing Research News, 26*(10), Article 10.

DiPaola, D., Payne, B. H., & Breazeal, C. (2020). Decoding design agendas: an ethical design activity for middle school students. In Proceedings of *the Interaction Design and Children Conference* (pp. 1–10).

Druga, S., Williams, R., Breazeal, C., & Resnick, M. (2017). " Hey Google is it ok if I eat you?" Initial explorations in child-agent interaction. In *Proceedings of the 2017 conference on interaction design and children* (pp. 595–600).

DuBois, J. M., & Burkemper, J. (2002). Ethics education in US medical schools: A study of syllabi. *Academic Medicine, 77*(5), 432–437.

Duffy, R. D., Blustein, D. L., Diemer, M. A., & Autin, K. L. (2016). The psychology of working theory. *Journal of Counseling Psychology, 63*(2), 127.

Emery, F. E., & Trist, E. L. (1960). Socio-technical systems. In C. W. Churchman & M. Verhulst (Eds.), *Management science models and techniques.* (Vol. 2). Pergamon Press.

Fiesler, C., Garrett, N., & Beard, N. (2020). What do we teach when we teach tech ethics? A syllabi analysis. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (pp. 289–295).

Gardner, H. (1991). The tensions between education and development. *Journal of Moral Education, 20*(2), 113–125.

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching, 48*(10), 1159–1176. https://doi.org/10.1002/tea.20442

Gordon, J., & Wrenn, C. B. (2020). Ethics of artificial intelligence. In *The Internet Encyclopedia of Philosophy.* https://iep.utm.edu/ethics-of-artificial-intelligence/. Accessed 5 May 2022.

Grosz, B. J., Grant, D. G., Vredenburgh, K., Behrends, J., Hu, L., Simmons, A., & Waldo, J. (2019). Embedded ethics: Integrating ethics across CS education. *Communications of the ACM, 62*(8), 54–61.

Guerreiro-Santalla, S., Bellas, F., & Duro, R. J. (2020). Artificial intelligence in pre-university education: what and how to teach. In *Multidisciplinary Digital Publishing Institute Proceedings* (Vol. 54, No. 1, p. 48).

Gunkel, D. J. (2012). *The machine question: Critical perspectives on AI, robots, and ethics*. MIT Press.

Hasse, A., Cortesi, S., Lombana, A., & Gasser, U. (2019). Youth and artificial intelligence: Where we stand. In *Berkman Klein Center Research Publication.* Article 2019–3.

Ho, J. W., Scadding, M., Kong, S. C., Andone, D., Biswas, G., Hoppe, H. U., & Hsu, T. C. (2019). Classroom activities for teaching artificial intelligence to primary school students. In Proceedings of International Conference on Computational Thinking Education (pp. 157–159).

Jing, M. (2018). China looks to school kids to win the global AI race. In *South China Morning Post.* https://www.scmp.com/tech/china-tech/article/2144396/china-looks-school-kids-win-global-ai-race. Accessed 5 May 2022.

Kenny, M., Bower, M., Perry, J., Blustein, D., & Amtzis, A. (2004). *The Tools for Tomorrow Program: Integrating school-to-career psychoeducation into high school curriculum*. Chestnut Hill, MA: Boston College.

Kirkpatrick, K. (2016). Battling algorithmic bias: How do we ensure algorithms treat us fairly? *Communications of the ACM, 59*(10), 16–17.

Klopfer, E., & Yoon, S. (2005). Developing games and simulations for today and tomorrow's tech savvy youth. *TechTrends, 49*(3), 33–41.

Loescher, K. J., Hughes, R. W., Cavico, F., Mirabella, J., & Pellet, P. F. (2005). The impact of an "Ethics across the curriculum" initiative on the cognitive moral development of business school undergraduates. *Teaching Ethics, 5*(2), 31–72.

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

Maltese, A. V., & Tai, R. H. (2010). Eyeballs in the Fridge: Sources of early interest in science. *International Journal of Science Education, 32*(5), 669–685. https://doi.org/10.1080/09500690902792385

Manyika, J., Lund, S., Chui, M., Bughin, J., Woetzel, J., Batra, P., Ko, R., & Sanghvi, S. (2017). *Jobs lost, jobs gained: Workforce transitions in a time of automation*. McKinsey Global Institute. https://www.mckinsey.com/~/media/BAB489A30B724BECB5DEDC41E9BB9FAC.ashx

Mark, S., DeBay, D., Zhang, L., Haley, J., Patchen, A., Wong, C., & Barnett, M. (2013). Coupling social justice and out-of-school time learning to provide opportunities to motivate, engage, and interest under-represented populations in STEM fields. *Career Planning and Adult Development, 29*(2), 93–104.

Marques, L. S., Gresse von Wangenheim, C., & Hauck, J. C. (2020). Teaching machine learning in school: A systematic mapping of the state of the art. *Informatics in Education, 19*(2), 283–321.

McGovern, A., Tidwell, Z., & Rushing, D. (2011). *Teaching introductory artificial intelligence through java-based games*. AAAI Symposium on Educational Advances in Artificial Intelligence, North America.

Müller, V. C. (2020). Ethics of Artificial Intelligence and Robotics. In *Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/ethics-ai/. Accessed 5 May 2022.

Narahara, T., & Kobayashi, Y. (2018). Personalizing homemade bots with plug & play AI for STEAM education. In SIGGRAPH Asia 2018 Technical Briefs (pp. 1–4).

Narayanan, A., & Vallor, S. (2014). Why software engineering courses should include ethics coverage. *Communications of the ACM, 57*(3), 23–25.

National Science Board (2020). *Vision 2030.* https://www.nsf.gov/nsb/publications/2020/nsb202015.pdf. Accessed 5 May 2022.

Payne, B. H. (2020). *Can my algorithm be my opinion?: An AI+ ethics curriculum for middle school students* (Unpublished master's thesis). Massachusetts Institute of Technology, Cambridge.

Pew Research Center (2018). *Teens, Social Media and Technology 2018*. Retrieved September 23, 2021, from https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/

Porter, L., Taylor, C., & Webb, K. C. (2014). Leveraging open source principles for flexible concept inventory development. In *Proceedings of the 2014 conference on Innovation & technology in computer science education* (pp. 243–248).

Rottinghaus, P. J., Buelow, K. L., Matyja, A., & Schneider, M. R. (2012). The career futures inventory–revised: Measuring dimensions of career adaptability. *Journal of Career Assessment, 20*(2), 123–139.

Selbst, A. D. (2017). Disparate impact in big data policing. *Ga. l. Rev., 52*, 109.

Skirpan, M., Beard, N., Bhaduri, S., Fiesler, C., & Yeh, T. (2018). Ethics education in context: A case study of novel ethics activities for the CS classroom. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (pp. 940–945).

Squire, K., & Klopfer, E. (2007). Augmented reality simulations on handheld computers. *The Journal of the Learning Sciences, 16*(3), 371–413.

Strauss, A., & Corbin, J. (1994). Grounded Theory Methodology: An Overview. *Handbook of Qualitative Research. 1st Ed.*, *17*, 273–285.

Tai, R. H., Qi Liu, C., Maltese, A. V., & Fan, X. (2006). Planning early for careers in science. *Science, 312*(5777), 1143–1144.

Taylor, C., Zingaro, D., Porter, L., Webb, K. C., Lee, C. B., & Clancy, M. (2014). Computer science concept inventories: Past and future. *Computer Science Education, 24*(4), 253–276.

Torrey, L. (2012). *Teaching problem-solving in algorithms and AI*. 3rd Symposium on Educational Advances in Artificial Intelligence.

Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). 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).

Vakil, S. (2018). Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education. *Harvard Educational Review, 88*(1), 26–52.

Van Brakel, R. (2016). Pre-emptive big data surveillance and its (dis) empowering consequences: The case of predictive policing. In B. van der Sloot, D. Broeders & E. Schrijvers (Eds.), Exploring the boundaries of big data (pp. 117–141). Amsterdam University Press.

Wang, Y.-Y., & Wang, Y.-S. (2019). Development and validation of an artificial intelligence anxiety scale: An initial application in predicting motivated learning behavior. *Interactive Learning Environments*, 1–16. https://doi.org/10.1080/10494820.2019.1674887

Weinburgh, M. H., & Steele, D. (2000). The modified attitudes toward Science Inventory: Developing an instrument to be used with fifth grade urban students. *Journal of Women and Minorities in Science and Engineering, 6*(1), 87–94. https://doi.org/10.1615/JWomenMinorScienEng.v6.i1.50

Wood, C. (1997). Yardsticks: Children in the Classroom Ages 4-14. A Resource for Parents and Teachers. Northeast Foundation for Children Publishing Division.

Yosso, T. J. (2005). Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education, 8*(1), 69–91.

Zhou, X., Van Brummelen, J., & Lin, P. (2020). Designing AI learning experiences for K-12: emerging works, future opportunities and a design framework. arXiv preprint arXiv:2009.10228.

Authors and Affiliations

Helen Zhang[1] · **Irene Lee**[2] · **Safinah Ali**[3] · **Daniella DiPaola**[3] · **Yihong Cheng**[1] · **Cynthia Breazeal**[3]

1   Lynch School of Education and Human Development, Boston College, Chestnut Hill, MA, USA

2   MIT Scheller Teacher Education Program, Massachusetts Institute of Technology, Cambridge, MA, USA

3   MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA, USA