Engaging Teachers to Co-Design Integrated AI Curriculum for K-12 Classrooms

Phoebe Lin*
Harvard Graduate School of Design
Cambridge, MA, USA
phoebe.lin93@gmail.com

Jessica Van Brummelen*
Massachusetts Institute of Technology
Cambridge, MA, USA
jess@csail.mit.edu

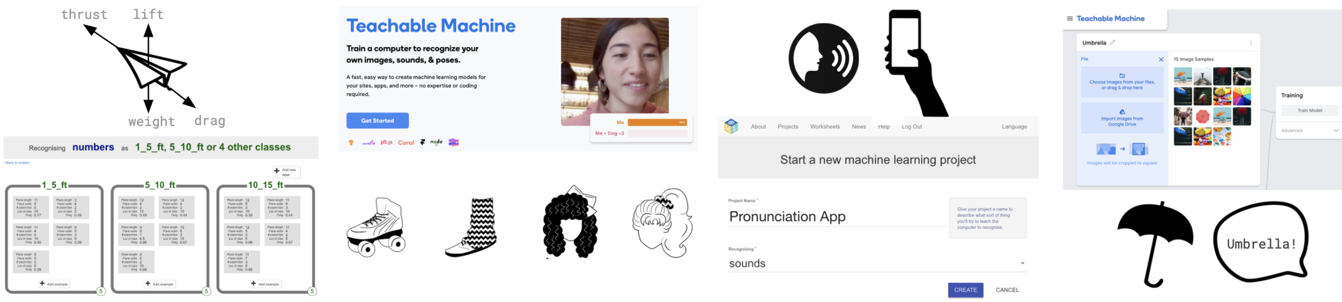

Figure 1: Representations of the four integrated curricula [81]. ***Left***: **The "exemplar" physics and AI curriculum [43].** ***Mid-left***: **Social studies and AI curriculum [13].** ***Mid-right***: **ESL and AI curriculum [43, 47].** ***Right***: **Literacy and AI curriculum for students with learning disabilities [13].**

ABSTRACT

Artificial Intelligence (AI) education is an increasingly popular topic area for K-12 teachers. However, little research has investigated how AI curriculum and tools can be designed to be more accessible to all teachers and learners. In this study, we take a Value-Sensitive Design approach to understanding the role of teacher values in the design of AI curriculum and tools, and identifying opportunities to integrate AI into core curriculum to leverage learners' interests. We organized co-design workshops with 15 K-12 teachers, where teachers and researchers co-created lesson plans using AI tools and embedding AI concepts into various core subjects. We found that K-12 teachers need additional scaffolding in AI tools and curriculum to facilitate ethics and data discussions, and value supports for learner evaluation and engagement, peer-to-peer collaboration, and critical reflection. We present an exemplar lesson plan that shows entry points for teaching AI in non-computing subjects and reflect on co-designing with K-12 teachers in a remote setting.

CCS CONCEPTS

• **Human-centered computing → Participatory design**; *User centered design.*

KEYWORDS

Artificial Intelligence, K-12 Education, Co-Design

ACM Reference Format:
Phoebe Lin and Jessica Van Brummelen. 2021. Engaging Teachers to Co-Design Integrated AI Curriculum for K-12 Classrooms. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3411764.3445377

*Both authors contributed equally to this research.

ACM ISBN 978-1-4503-8096-6/21/05…$15.00
https://doi.org/10.1145/3411764.3445377

1 INTRODUCTION

Artificial intelligence (AI) education is becoming an increasingly popular subject in the eyes of educators due to the rapid integration of AI technologies in user-facing services and products [31, 72, 90]. Researchers have called for formal K-12 education to prioritize AI literacy and teach children to interact with AI using a critical lens [96]. The AI4K12 research community has also published guidelines for what AI concepts K-12 curriculum should cover, known as the *Big AI Ideas*, and calls for AI researchers to help teachers and students understand AI [74]. As children interact more with AI technologies, it is critical that they are able to recognize AI, understand how AI algorithms work, use those algorithms to solve problems meaningful to them, and evaluate the impact of AI technologies on society [7].

However, introducing AI curriculum in K-12 classrooms can be challenging when available tools and curriculum are incompatible with the values and contexts of the people who teach and learn from them. For example, teachers of all subjects should feel empowered to teach AI, yet teachers often feel they lack sufficient understanding to teach AI and the capacity to include more curriculum on top of their existing curriculum [84]. Despite the proliferation of tools and AI curriculum in response to the recent calls to action, few are widely implemented due to constraints in the classroom that prevent curriculum from being usable. [57].

Similarly, AI as a discipline can span many other topics, such as government, journalism, and art [14, 36, 68], yet AI concepts are often confined to computing subjects such as computer science or data science. Tools and curriculum today often teach AI as an extension of computer science curricula or as standalone curricula that is difficult to adjust to other contexts [11, 50, 69]. Adapting those tools and curriculum then becomes especially difficult for teachers who teach core subjects, including English, math, social studies, and science, and may not have any AI experience. The lack of integrated AI curricula in core subjects has become one of the barriers to exposing AI to students with little access to computing disciplines.

These two challenges show that exposing AI education to a wider range of students involves more than creating useful tools, and requires involving relevant stakeholders to make them usable in classroom contexts. To address these challenges, we employ the Value-Sensitive Design approach [29, 30] to create teacher-centered design recommendations for AI curriculum and tools because teachers play a key role in the implementation of AI education. Value-Sensitive Design is defined as a "theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process" [30]. Value-Sensitive Design starts with the belief that technologies embody values and offers a proactive element to influence the design of tools early on in the design process.

In this paper, we partner with K-12 teachers to design AI curriculum that is integrated with core subjects. We set out to understand what is necessary and valued by K-12 teachers to effectively implement integrated AI curricula, and co-create lesson plans that address those needs and values. Specifically, our research questions are:

RQ1: How might we address K-12 teachers' values and considerations when designing AI curriculum and tools? (Teaching needs)

RQ2: How might AI tools and curriculum be integrated into core subject curriculum? (Integrated curriculum design)

To answer these research questions, we organized a multi-session workshop that spanned two days with fifteen teachers who teach various subjects at different schools. The first day of the workshop involved presentations and group discussions to level set everyone's basic understanding of AI. Between the first and second day of the workshop, participants were asked to complete a brainstorming assignment where they identified curriculum of their own to use as a potential base for an integrated AI curriculum. During the second day of the workshop, we split participants into three small groups to work together and design a lesson plan that integrates AI into a non-computing subject curriculum. The co-design process revealed that when the teachers designed curriculum, they considered four practical needs: evaluation, engagement, logistics, and collaboration. Furthermore, our analysis of the co-designed lesson plans showed opportunities for connections between AI and a core subject, with three points of integration: data, reflection, and scaffolding for ethics. Each of these needs and opportunities can be considered when developing AI education tools.

The contributions of this work are (1) identifying the values and needs of K-12 teachers teaching AI in the classroom and opportunities to address them with respect to AI tools and curricula, (2) showing ways in which teachers might incorporate AI into their classrooms and presenting integrated AI curriculum as outputs of the co-design sessions [81], and (3) reflecting on co-design sessions involving K-12 teachers in a remote setting to solicit design recommendations for AI curriculum and tools.

2 RELATED WORK

To the authors' knowledge, there are no papers describing a co-design process with teachers to integrate AI concepts into core curriculum, and few papers that purposefully integrate AI concepts into core curriculum. This is despite a recent sharp increase in education and learning-related human-computer interaction (HCI) research [51, 53, 89], as well as K-12 AI education research [10, 18, 19, 23, 44, 46, 67, 73, 78, 85]. Thus, we include discussion about broader topics influencing our study, including co-design with teachers of other course materials, co-design with relevant populations (e.g., children or parents), and the development of AI education tools and curricula.

2.1 Co-Design in HCI

Researchers have implemented co-design with many different populations to develop systems that better meet end-users' needs [61, 62]. For example, educational and AI technologies have been co-designed with children, allowing researchers to better understand children's learning, experience and motivational needs [54, 71, 82, 88]. These studies provide insight into how to better engage with children and learn from their creative ideas; for example, by providing scaffolding for collaboration and development [83] or by using technology to connect with children worldwide [34]. Similarly, co-design activities have been implemented alongside parents, who provide different perspectives, goals, and behaviors than children when designing [93]. Additionally, parents take on various roles, including teachers, spectators and scaffolders, when interacting with their children and educational computer science tools [94].

Although studies with children and parents provide many insights into effective co-design and user needs for K-12 technology, K-12 teachers' goals and experience likely align better with our goal of facilitating K-12 AI learning. Furthermore, although K-12 *AI* education has not yet benefited from tools and curriculum co-designed with K-12 teachers, other areas of education have (albeit relatively few, despite the apparent benefits teachers bring to co-design [92]). For example, in one study, researchers collaborated with teachers to develop new science curriculum materials. Researchers recognized the value in teachers' K-12 expertise and in promoting their agency throughout the design process [64]. Another science curriculum co-design study argued that the process of working with teachers had substantial effects on adoption of the tools and curricula, in addition to bringing social value and innovative ideas [24]. Furthermore, according to Penuel et al., co-designing with teachers can encourage curriculum innovation, classroom technology implementation, teacher ownership and sustainability of materials, and teacher learning [55].

In order to catalyze such benefits, one paper presents key considerations to co-designing with teachers. These include addressing a "concrete, tangible innovation challenge", investigating "current practice and classroom contexts", and involving a "central accountability for the quality of the products of the co-design", among others [57]. Other co-design studies emphasize the importance of diverse participants, including diversity in teachers, which can foster creativity and encourage unique perspectives [23, 48, 62]. Additional teacher co-design recommendations include actively democratizing innovation and developing collective capacity; for example, through professional development activities followed by design activities [55]. In our study, we utilize these considerations as well as guidance from values-led co-design [39, 82] to co-create AI-integrated core curricula.

2.2 AI Education for K-12

Computational thinking has been referred to as a key 21st-century literacy [56, 94], and AI literacy is quickly becoming another candidate due to its large societal impact, recent major advancements, and many misconceptions [46, 75]. For instance, technology design students—from K-12 to university-level engineering, computer science and architecture students—have been shown to have a bias towards observing positive scenarios and effects of their AI systems rather than negative [48, 79], despite the potential for AI to have detrimental effects on society [38, 91]. This demonstrates a need for AI education—especially in ethical technology design. Other studies also demonstrate this need, including an AI co-design study, which indicated children believe AI to be more advanced than reality [88], and a social robot AI education study, which showed children scoring an average of 66.8% on AI understanding assessments after the intervention [86].

Recent works have developed AI education tools and curricula for K-12. For example, one study presents an ethical AI design activity for middle school students [19]. Another presents a card-based ethical machine learning (ML) co-design workshop, which includes an educational introduction to ML [10]. Other works present software or hardware to teach AI concepts, including a k-means clustering visualization tool [85], Scratch extensions for probabilistic modeling and ML [18], a platform and robot to teach various AI concepts [86], a hands-on data science exercise for 5-9th grade students [67], and a blocks-based coding conversational AI development interface [79]. None of these works, however, include teachers as co-designers and many of them are also not readily available for teachers (especially teachers without AI backgrounds) to use. Even fewer incorporate material from core subject curricula, which would likely ease integration into current classrooms.

Other K-12 AI tools and curricula exist as standalone products or extend computer science curriculum. Many of these tools have the potential to be integrated into core subject curricula already being taught in K-12. Two widely-used AI teaching tools are *Teachable Machine* [13] and *Machine Learning for Kids* [43], which empower learners to develop classification models without needing to program. Other standalone AI teaching tools include *Any-Cubes*, which are toys to teach ML concepts [63]; *Calypso for Cozmo*, which is AI curriculum for a toy robot [76]; and extensions for *MIT App Inventor*, which enable students to develop AI-powered mobile apps [47]. Each of these tools could be integrated and taught in core classes; however, are presented as standalone AI tools.

In terms of K-12 AI education research involving instructors, most involve researchers rather than K-12 teachers as the instructors, and likely miss valuable expertise and feedback from professionals who have worked in the classroom [12, 19, 22, 44]. Nevertheless, some works involving K-12 teachers include an AI summer program for high school girls [77], an AI engineering course for high school students [66], and a STEM workshop for middle school students [59]. Each of these studies saw value in engaging with K-12 teachers.

Other works have also integrated core curriculum content into AI tools and curricula; however, most of these involve researchers as instructors and are often not in regular classroom settings. For example, one physical education curriculum involves students developing sports gesture classification models with researchers as facilitators [97]. Another science-based curriculum involves students teaching a conversational agent about animals, and observing it classify the animals into ecosystems with researchers as facilitators [44]. Although these works are state-of-the-art in K-12 AI education, it is unknown whether they are suitable for K-12 classrooms, since they have not been tested in regular classrooms and teachers were not involved in the design process.

In our literature review, we found one example of AI curriculum that was both integrated into a core course and designed or taught by K-12 teachers alongside researchers. This curriculum involved AI and science concepts, and was taught in Australian K-6 classrooms [35]. Although this example is insightful, much further research is needed to integrate teacher expertise and address widespread, integrated AI curriculum in K-12 classrooms [84].

3 METHOD

We conducted a two-day co-design workshop with fifteen instructors, ranging from K-12 teachers to educational directors. Participants completed pre-work before each day's activities, as well as pre- and post-workshop surveys. The co-design activity was split into three smaller group sessions to enable us to better identify differences in value and process of different teachers. This study was approved by the university's Institutional Review Board (IRB).

3.1 Participants

Fifteen teachers participated in the study, whom we recruited via a mailing list and snowball sampling through our personal network. Seven participants identified as female, four participants identified as male, and the remaining did not say. Their age ranged from 25 to 50 (M = 40.6, SD = 11.6). The only inclusion criteria was that they teach or previously taught in a K-12 classroom, were able to commit to the time and pre-work for the two-day workshop, and had an interest in teaching AI.

Because there was an overwhelming response in our recruitment, we unfortunately had to limit participation to preserve the intimacy of small-group interaction. In our selection process, we first grouped respondents into the subjects they taught: Computer Science, Technology, Non-STEM, and General (this means they taught all subjects). As we wanted to prioritize participants who taught non-STEM classrooms, such as English language arts (ELA), we accepted almost all respondents from that group. Then we randomly selected the remaining participants equally from the other three groups. During the co-design workshop, participants were placed in small groups, and we spread teachers out by their respective disciplines so that groups could have diverse perspectives. The work background of the final participant list is detailed in Tab. 1. While our participants were geographically diverse, we did not analyze these differences due to our small sample size. All participants provided informed consent to participate in compliance with our institution's IRB. Participants also provided their availability, which we used to select the workshop dates and times. Each day of the workshop lasted 2.5 hours. Every workshop session involved two researchers.

3.2 Workshop Design

The entire co-design workshop spanned two days, Session 1 on the first day and Session 2 on the second day. Session 1 consisted of discussions and a "What is AI" presentation to level set all participants (see Tab. 2), and Session 2 consisted of the co-design activity and an ethics presentation (see Tab. 3). Here, we describe our rationale and the activities in detail.

3.2.1 Session 1. Before the first session, we asked participants to complete a pre-workshop questionnaire asking about participants' familiarity with AI, whether they have taught AI in the classroom before, and if so, what their experience was. This was to understand their backgrounds and enable us to tailor the content of Session 1 appropriately. Participants were also given detailed instructions on how to install and use Zoom [2], Slack [3], and Miro [1]—the tools used throughout the entire workshop. We started Session 1 with breaking participants into small groups on Zoom to discuss why participants thought AI is or is not important to teach their students. Having them describe what and why AI was important allowed us to understand their preconceptions about AI and their priorities as teachers. During the "Let's learn AI" presentation, participants learned the *Big AI Ideas* [75], categories of AI, and how to recognize what is and is not AI. During the "Let's learn AI tools" presentation, we demoed four distinct AI learning tools and provided participants with resources and links to explore further. We then used Miro for a card sorting activity [87], where we asked participants to generate categories for Google's A to Z of AI cards [4], where categories were limited to subjects taught in the classroom. The card sorting activity showed participants' enthusiasm for integrating AI topics into every classroom subject, including English language arts (ELA), writing and reading, social studies, math, science, economics, and social-emotional learning.

3.2.2 Session 2. Participants were asked to complete "pre-work" before Session 2. Participants had two days to complete their pre-work between Session 1 and Session 2. The pre-work asked participants to explore the rest of the AI learning tools, select one of the tools to go along with a curriculum they currently use or have used in their classrooms, and identify areas where they see potential to teach AI using the selected tool. Participants uploaded their submissions into a shared Google Drive folder. Participants had access to the workshop Google Drive folder, which contained all of the presentations and resources from Session 1, at all times, and could also post questions in the workshop Slack group, which was monitored closely by the researchers. From the pre-work submissions, we selected one idea to develop into an exemplar curriculum (see [81]).

For Session 2, participants were split into three groups of 4-5. Each group was asked to analyze the exemplar curriculum and discuss what they noticed. The co-design activity part 1 then began with each group responding to a prompt asking them to devise integrated AI curricula for specific subjects. We created the prompts from the pre-work submissions and organized the groups such that each would have a domain expert. For example, the group responding to the prompt asking participants to create a curriculum for students who are learning English as a Second Language (ESL) had an ESL teacher, who would be familiar with ESL students' needs. Each group was also paired with a researcher, who provided technical input and answered participants' questions about AI or learning tools. During the "Ethics & Diversity" presentation, we presented definitions of AI ethics, diversity statistics within the field, and resources for teaching and learning AI ethics. Participants then continued working in their groups on their integrated curriculum in co-design activity part 2.

Every group was successful in producing a first draft of an implementable AI curriculum that integrated with a core subject. The drafts can be found in the appendices [81]. Lastly, participants discussed why they thought AI was or was not important to teach for a second time, which acted as a reflection and a way to see if their mindset or preconceptions changed after the workshop. Participants were asked to complete a post-workshop questionnaire that asked how familiar they were with AI, how comfortable they felt teaching AI in their class, as well as feedback on the workshop itself and their demographics (i.e. age, gender, ethnicity).

3.3 Data Analysis

Our dataset consists of the audio recordings of the entire co-design workshop, participant questionnaires, and the deliverables of each participant, which include their pre-work submissions and their group work during the co-design activity. All audio recordings were transcribed to text and thematically coded by two researchers using open coding. We specifically examined their process, priorities, and challenges.

From the pre-work submissions, the most common tool used was Teachable Machine [13], with 4 usages, followed by MIT App Inventor [47], with 3 usages. ML4Kids [43], BERT Q&A [70] and Zhorai [44] each had 2 usages; and Pix2Pix Activity [6], Google Quick Draw [40], and AI for Oceans [16] each had 1 usage. By the end of the workshop, the 3 teacher groups co-created 3 different AI curriculum drafts. Two of them used Teachable Machine [13], and the other used ML4Kids [43]. More detailed information about these teacher work can be found in Tab. 4.

From the pre- and post-workshop questionnaires, we found that 9 out of 15 participants had never taught AI in the classroom. While some participants had experience teaching AI, they were interested in learning how to allow non-CS students experience AI and integrate AI into their teaching. Participants came into the workshop rating their own familiarity with AI an average of 4.8 out of 7, and

Table 1: Participants were selected to represent diverse profiles and/or subject areas.

ID	Grade taught	Subject taught	Location
P1	6th grade	English Language Arts (ELA)	North Carolina, USA
P2	5th grade	Science, General	Connecticut, USA
P3	6th-8th grade	Computer Science	Tunisia, North Africa
P4	9th-12th grade	Computer Science	Cuneo, Italy
P5	9th-12th grade	Chemistry and Math	British Columbia, Canada
P6	6th-8th grade	STEM, General	Florida, USA
P7	6th-8th grade	STEM	Florida, USA
P8	-	STEM	Pennsylvania, USA
P9	6th-8th grade	Computer Science	California, USA
P10	9th-12th grade	Career Exploration	Rhode Island, USA
P11	9th-12th grade	Computer Science	Massachusetts, USA
P12	9th-12th grade	Library Science	Rome, Italy
P13	6th grade	History	California, USA
P14	6th-9th grade	Computer Science	Turkey
P15	6th-12th grade	English as a Second Language (ESL)	Pennsylvania, USA

Table 2: Schedule for Session 1.

Time	Activity
15 min	Introduction
20 min	Why AI? (Discussion)
50 min	Let’s learn AI! (Presentation)
15 min	Break
25 min	Card sorting activity
25 min	Let’s learn AI tools! (Presentation)

Table 3: Schedule for Session 2.

Time	Activity
60 min	Co-design activity part 1
15 min	Break
40 min	Ethics & Diversity (Presentation)
20 min	Co-design activity part 2
15 min	Why AI? (Discussion)

finished the workshop with an average rating of 5.8 out of 7. Teachers also rated their confidence about integrating AI into their own curriculum with an average rating of 5.6 out of 7.

4 RESULTS AND DISCUSSION

Our teacher participants teach students with diverse needs. The co-design activity prompted rich discussion with three groups completing three curriculum drafts that integrated AI with a topic of their choice. The topics were: (1) “How Does Data Affect Government Policy?” (Social Studies Curriculum), (2) “Learn Vocabulary with an AI” (Literacy curriculum for students with learning disabilities), and (3) “Build an AI-powered Pronunciation Application” (ESL curriculum), as shown in the appendices [81]. During the co-designing process, all groups shared certain considerations for the curriculum, though each group addressed them differently. In the first section of results, we answer the first research question by outlining what the shared values and considerations were and showing how each group addressed them. We then answer the second research question by showing how each curriculum effectively integrated AI. Both research questions prompted design implications for AI tools for teachers, which are listed in Tab. 5 and labeled in the text with **DR** for “design recommendation”.

4.1 RQ1: How might we address the values and considerations of K-12 teachers when designing AI curriculum and tools?

We identified four categories of values that our teachers had while creating the curriculum drafts: *Evaluation*, *Engagement*, *Logistics*, and *Collaboration*, which can be considered when designing AI education tools.

4.1.1 Evaluation. All groups considered student evaluation to be critical to a curriculum. Teachers wanted to see evidence for learning and know their students understand relevant concepts correctly. To do so, teachers first considered their own objectives: “Do we have an end goal in mind, or like, what do we consider a success?” (P9). In the ESL curriculum, P12 referred to the *Big AI Ideas* to identify the what the group called the “AI objective”. P12 and P15 also frequently referred to the exemplar curriculum, suggesting that teachers require frameworks and scaffolding to devise the AI objective. To evaluate students, P5 and P12 both suggested non-traditional forms of evaluation, such as an “exit interview or on-the-fly assessments where students talk through all of the details, so we get a really good idea from a conversation with them whether they understand what they were doing” (P5) and “an engineer’s log where you’ve got their design and you’ve got to do it all official” (P12). In these drafts, teachers wanted to evaluate students on their conceptual knowledge, and not necessarily on their technical knowledge. This provides an opportunity for AI teaching tool developers to provide

[1] Some teachers submitted multiple ideas

Table 4: Project ideas developed in the teacher pre-work and co-created curricula activities organized by AI tool used.

Pre-work[1]		
Count	Tool	Project Idea(s)
4	Teachable Machine [13]	• Street object classification in a "smart car" • Image recognition and learning vocabulary • Build a taxonomy for recyclables • Recognizing poisonous mushrooms
3	MIT App Inventor [47]	• An app that can distinguish different painting styles • Mobile app for home automation • An app that recognizes texts from different Italian and English literature
2	ML4Kids [43]	• Text classification in a "smart home" • Predicting variables for flight
2	BERT QA [70]	• Writing a story with BERT • Understanding NLP
2	Zhorai [44]	• Chatbot in a "smart garden" • Interact with a chatbot
1	Pix2Pix Activity [6]	• Endangered plant identification
1	Google Quick Draw [40]	• Pattern recognition with quick draw
1	AI for Oceans [16]	• Understanding ethics
Co-created Curricula		
Count	Tool	Project Idea(s)
2	Teachable Machine [13]	• Understand the CROWN Act: professional vs. unprofessional dress codes and hairstyles • "Can you beat the machine" at recognizing new vocabulary words using images
1	ML4Kids [43]	• Correct pronunciation app (with MIT App Inventor [47])

scaffolding for evaluating student learning (if the tool is intended for classroom use)—whether it is directly embedded into the tool or presented as an auxiliary resource (**DR1: Design for Evaluation**).

4.1.2 Engagement. In a K-12 setting, engagement tends to be particularly challenging, which was a concern for our teachers. P8 and P10 grounded the Social Studies curriculum in law and government discourse by having students review an article around the Crown Act. Introducing context to the project gives students an "anchor" (P8) or hook to prompt further inquiry. Other anchors included asking students the "hard questions" about real-world applications of AI, such as "how do Siri and other personal assistants get to be at that point?" and "who used the machine learning and designed the app?" (P7). P5 and P15 both mentioned student-driven learning as a way to leverage students' interests. For example, "I can see a sixth grader coming in and going, I went to the baseball game and I couldn't say all these words. And they decide they're going to do baseball that day" (P12). Lastly, multiple groups brought up competition and gamification as effective methods of engagement: "the class creates a game that students use to quiz themselves on vocab by trying to be better than the system" and "module 1 can be a rock paper scissors game so that students get familiar with the interface"

Table 5: Design recommendations (DR) for AI Learning Tools & Curricula.

#	Design for...
DR1	Evaluation
DR2	Engagement
DR3	Easing Logistics
DR4	Collaborative Learning
DR5	Data Integration
DR6	Critical Reflection
DR7	Ethical Analysis

(P2). In the above examples, teachers were creative with incorporating engagement tactics into their curriculum, especially with attention being a limited resource and remote learning a new challenge. As a result, teachers need learning tools and curriculum that are similarly capable of engaging students effectively while students are learning (**DR2: Design for Engagement**). For example, tools that empower students to program their own AI projects [5, 80], engage students in embodied interactions [37, 98], and leverage learners' interests [41, 59], can be particularly engaging [46, 95].

4.1.3 Logistics. Logistics refers to the factors that enable the curriculum to be smoothly run in the classroom. Teachers tended to think about how the lesson itself would take shape before addressing which core standards the lesson intended to cover. For example, at the beginning of the co-design, P10 explained that what would be most beneficial was "thinking of how to structure the lesson and what resources we can use to pull in to have the engagement component". Most teachers struggled with identifying which technology resources and learning tools to use, for example, whether to use *Machine Learning for Kids* [43] or *Google Quick Draw* [40]. Our teacher participants generally looked to the researcher for guidance, and suggested that tools should be more explicit about when and how they can be applied in K-12 classrooms. Because the co-created curricula were created among teachers of various subjects (subjects taught in the classroom), teacher participants gravitated toward tools that provided flexibility. For example, teacher participants made key decisions to use *Teachable Machine* [13] and *ML4Kids* [43] because they could be more easily integrated into their curriculum ideas.

Teachers also paid close attention to grade-level considerations. They felt more comfortable having older students drive their own learning, but recognized even younger students are capable of deep reflection: "posing some challenging questions will vary a little depending on age, but you can get pretty deep with some—even fifth graders. They can get into this, and I think it's a good way of opening the door" (P15). AI teaching tool developers should effectively communicate how tools could be used in the classroom and what concepts students could learn that teachers can incorporate into their curriculum (**DR3: Design for Easing Logistics**).

4.1.4 Collaboration. All groups discussed the value of collaboration. In the ESL curriculum, teachers had their students collect data in groups and input the data into multiple models using Machine Learning for Kids. In the literacy curriculum, teachers had every student contribute 10 images to a class dataset to input into Google Teachable Machine. The presence of group work not only helps overcome the need to create many training examples for a machine learning model, but also provides students with opportunities to discuss design and ethics decisions with their peers and teacher. This also aligns with Long and Magerko's design consideration for *Social Interaction* [46].

Teachers also considered how collaboration can be implemented most effectively when designing curricula. For example, P8 described how "it's important to think about the group size because you want to make sure that students have a voice in the work. And when you start doing large group things those kids that process information internally never get to be heard." She went on to describe how, in her experience, "duos [of students] work really, really well" and how it is generally better to "go with smaller groups [of students in the classroom], but if you're using technology [...] you're bound by what you have." In order for curriculum to be unconstrained by tools, it is important to consider how AI learning tools may be used for, or even facilitate group interaction so that teachers can easily ensure all students have the opportunity to contribute and learn together (**DR4: Design for Collaborative Learning**).

4.2 RQ2: How might AI learning tools and curriculum be integrated into core curriculum to support teachers when they teach AI?

During the co-design, teachers made connections between the core subject material (e.g., social studies) and AI in three main ways: **(1) relating an AI tool or concept to the core subject**, **(2) relating content from the core subject to AI**, and **(3) noticing overlapping concepts in AI and the core subject**. For example, P14 related the AI tool, *Arbitrary Style Transfer* [49], to the core subject of history when he said, "If we give an image as input and try to modify [it] according to the old art [using] Style [Transfer][...] This can give us an idea about the history when we look at the picture, [...] but if you change the picture, the students may understand how people thought in the past". Other teachers related real-life applications of AI to core subjects, like how *YouTube* suggestion algorithms can be "tunnel visioned" in what they suggest, similar to how people can be "tunnel visioned" when considering politics or how recidivism risk analysis algorithms [9] can be related to social studies concepts (P10).

Teachers also often made connections by starting with a core subject concept and relating it to AI. For instance, one teacher connected physics data from one of their student's 3D printing projects to an AI flight prediction algorithm (P12). The same teacher also started with an English unit and asked, "What tools do we know that we [can] connect to language?", ultimately connecting English to a Shakespeare natural language processing algorithm. Another teacher began with the ELA concept of "argumentation" and connected it to the reflection and "data analysis" processes in AI (P8).

In terms of overlapping concepts between AI and core subjects, teachers often found connections using the *Big AI Ideas* [75]. For instance, the Big AI Ideas of *Societal Implications* and *Representation and Reasoning* are also core concepts in social studies. The AI concept of iterative development in ML was also directly connected to the social studies concept of iterative opinion making through "go[ing] back and forth" (P8) and adjusting beliefs.

Using these methods of connection, participants co-designed integrated curricula containing AI concepts and supports for teaching core subject requirements. The curricula contained three main points of integration: **(1) data**, **(2) reflection**, and **(3) ethics**. Creating AI tools with these points in mind could simplify integration of AI concepts into core curriculum for teachers.

4.2.1 Data. Educational activities often produce data, and AI systems often require data. This provides an obvious access point for AI systems to be integrated into ready-made educational activities. In our co-design workshops, participants used this fact to generate integrated curriculum. For instance, in the exemplar curriculum (see Fig. 1 and [81]) (which was based on a teacher's idea during the workshops) students would produce data as they construct airplanes for a physics activity. The paper airplane dimensions and time-of-flight data would then be used to train a ML model to predict the effectiveness of other potential paper airplanes, combining AI systems and physics concepts into a single curriculum.

For the ESL integrated curriculum, students produced data as they were practicing word pronunciation, which was then utilized in a pronunciation teaching app. For example, students would create data by recording saying a word correctly (as guided by a teacher) and incorrectly, which would then train a classification model for an app developed in *MIT App Inventor* [47]. This app would then be used to further help students learn correct pronunciation. Future AI-integrated curricula might consider utilizing the data inherent in core curricula activities, such as speech pronunciation data, to teach AI. This may be through teaching data-related AI competencies, like *Data Literacy*, *Learning from Data*, and *Critically Interpreting Data* [46], or through using data to train ML models, which can teach other AI competencies.

For the literacy curriculum for students with learning disabilities, participants also used data from core curriculum—vocabulary words—to integrate AI concepts. From the vocabulary words, students would find relevant images, generating further data, and use this to train a classification model. This addressed the aforementioned data-related AI competencies, as well as other competencies, including the *ML Steps* and *Human Role in AI*, in addition to relevant English literacy concepts.

In these curriculum drafts, as well as the pre-work, teachers leveraged tools like *Teachable Machine* [13] and *ML4Kids* [43] for their transparent data training capabilities. We believe that tools that allowed students to go through the steps of the ML cycle, such as data collection, model training, and model evaluation, and in some way understand why an AI system behaves the way it does, were some of the reasons why teachers gravitated toward them. Therefore, AI tool developers should consider integrating "how" and "why" concepts, specifically surrounding data, directly into the tool to make it easier for teachers (and thus, students) to make connections between data and core curricula (**DR5: Design for Data Integration**).

4.2.2 Reflection. Another point of AI integration was student reflection on core curriculum content and AI methods. Many common core standards as well as AI competencies can be addressed through student reflection. For example, the common core standard, 1-ESS1-1: "Use observations of the sun, moon, and stars to describe patterns that can be predicted." [52], and the AI concept, "Learning from Data", could be addressed by reflecting on patterns in a constellation classification model's input and output. In the exemplar curricula, students were asked to reflect on what did and did not work and why, and on the real-world implications of a biased dataset in airplane development. This reflection addressed both a standard from the common core, 3-5-ETS1-3: "Plans and carries out fair tests in which variables are controlled and failure points are considered to identify aspects of a model or prototype that can be improved" as well as a number of AI literacy competencies, including *AI Strengths & Weaknesses*, *Critically Interpreting Data*, and *Ethics* [46].

Teachers also used this method to integrate AI concepts into the social studies curriculum. For example, students were asked to reflect on the amount of data in each image category, social norms and peer opinion, people's ability to access resources, and consensus agreement in this curriculum. These reflection questions address a number of the AI competencies, including *Data Literacy*, *Critically Interpreting Data*, and *Ethics* [46], as well as core social studies and English language arts standards, including NSS-EC.5-8.1: *Scarcity*, NSS-C.5-8.3: *Principles of Democracy*, NL-ENG.K-12.4: *Communication Skills*, and NL-ENG.K-12.7: *Evaluating Data* [25]. Since many core curricula standards involve reflecting on learning, AI teaching tool developers may consider embedding opportunities for reflection directly into tools (**DR6: Design for Critical Reflection**). This could empower teachers to easily and effectively integrate AI learning into core curricula.

4.2.3 Ethics. The final point of integration we present is through ethics, which is one of the AI literacy competencies [46]. Technology ethics is increasingly being called for in education and HCI [28, 33, 42]. Researchers and educators are investigating how to best implement ethics in their course and technology designs [32, 33, 45, 60]. For instance, in computer science education, researchers hypothesize that integrating ethics and social good into curricula motivates students through contextualization [32]. Similarly, ethics is being brought into ML and HCI courses [60, 65]. It is also being brought into AI and HCI technology; for example, through explainable AI [17, 27] and human-centered design [8]. In our study, we found teachers were very interested in incorporating ethics into their classes, especially since ethics is a common topic in core curricula standards [25, 52]. This provided an opportunity for integration of *AI* ethics topics into core classes.

One example of core curricula ethics includes environmental ethics, which can be found in life science standards (K-ESS3-3: "Communicate solutions that will reduce the impact of humans on the land, water, air, and/or other living things in the local environment."), and design ethics, which can be found in engineering standards (MS-ETS1-1: "Define the criteria and constraints of a design problem [...] taking into account [...] potential impacts on people and the natural environment") [52]. Furthermore, social justice principles, which are highly related to AI ethics, are commonly advocated for within standards-based K-12 education [20, 21]. By teaching ethical principles with respect to AI, teachers can also address standards related to the common core.

Each curricula designed in the co-design sessions had an ethics component. In the exemplar, students would engage in a brainstorming session about how AI bias affected the accuracy of ML models and relevant implications in the real world. Similarly, the ESL curriculum addressed ethics through discussing AI bias, socioeconomic norms for "correct" pronunciations, and the implications of an AI system judging people's pronunciations in the real-world. The social studies curriculum was developed around the ethics of the "CROWN Act" [26], what it means for students to design AI algorithms to classify outfits and hairstyles as "professional" or "unprofessional", and how this might affect different people groups. The literacy curriculum for students with learning disabilities addressed ethics through discussion about the accuracy of the image classification system and reasons for any bias observed. Each of these curricula touched on environmental, social justice or other ethical issues, addressing both AI and common core ethics standards.

Though we found that teachers were highly interested in teaching ethics (e.g., the social studies curricula was entirely focused on ethics), they also seemed apprehensive about actually implementing ethics activities in the classroom. For example, P5 described how there is a "barrier that comes up for teachers" when "kids often bring ethics up with questions and sometimes teachers will avoid it because they're afraid to say something wrong [...] even though those discussions would be so rich." Nevertheless, P5 also mentioned how if it was in a "planned lesson", it would be "less scary because you know what you're going to say". These opportunities and challenges in teaching ethics are also reflected in the literature [15, 58]. Designing scaffolding for AI ethics within teaching tools (**DR7: Design for Ethical Analysis**), for example, how the Zhorai K-12 teaching tool engages students in ethical dialog activity [44], would not only enable core curricula integration, but would also empower teachers to more confidently teach students about ethics.

4.3 Reflecting on Remote Co-design

Due to the COVID-19 pandemic, we organized and ran this co-design workshop completely remotely. Among our activities, the perceived helpfulness from most to least helpful was: Presentations (11 votes), Co-design activity (9 votes), Why AI? and Ethics discussions (both 7 votes), and the Card sorting activity (4 votes). Participants also indicated meeting like-minded educators from around the world and having access to the list of tools and links to be particularly rewarding takeaways. Overall, we noticed a slight increase in familiarity with AI after the workshop and a high level of confidence for integrating AI into their classrooms, though we did not establish that baseline. When asked if the workshop changed their opinion about teaching AI, teachers cited "introducing AI is the gateway to so much learning...now I am seeing and starting to understand the vast world of opportunities that exist for coding beyond being video game designers" (anonymous), as well as seeing the necessity of teaching AI and understanding that AI can be accessible to not "just the computery people" (anonymous).

4.3.1 Factors to a Successful Co-Design. At the beginning of the workshop, we established norms as an entire group to make facilitating easier. For example, setting expectations for "warm" calling to ensure equal representation of voices in the room meant participants expected to be called on to share their thoughts. Other norms included being present, having discussions in breakout rooms, and keeping cameras on. These norms helped establish general expectations for what it meant for teachers to actively participate during the workshop.

We received feedback from teachers following the co-design that they learned a lot from the workshop, and felt the co-design activity was particularly well-run. We believe this was due to the fact that we shared our goals with them at the beginning of the activity. Our goals were to imagine how AI might be integrated to core curricula, and more importantly, what their approach as teachers would be. We also shared an exemplar of a quality deliverable, which is a common teacher practice. By taking these steps, it seemed to ease participants' concerns about "being helpful" and knowing what to do. If we contrast the co-design activity to the card sorting activity, which was rated as least helpful, we realized that stating explicit goals and providing specific instruction were key factors to the activity's effectiveness and teacher perceptions of learning.

4.3.2 Lessons for Future Co-Designs. Since most teachers were unfamiliar with teaching AI, we grouped them into smaller groups of 4-5 so that each group could have a dedicated researcher co-designing with them. Grouping teachers who taught different subjects and had different perspectives this way was helpful with idea generation. For example, each pre-work that teachers submitted correlated strongly with the class they taught (e.g., science teachers generated ideas connecting AI to science). When teachers co-designed together, the ideas were more general. While this meant some teachers worked on curriculum that was unrelated to their discipline, we believed this trade-off was necessary given the complexity of the task and the benefits of collaboration. This setup may have worked better if teachers from the same school joined, and groups could be formed by school. Similarly, if there were more teachers from the same subjects, groups could also be formed by subject and thus leverage more specific AI tools.

Several teachers also requested more time to play with the AI learning tools and digest the presentations. This could have been addressed by scheduling more time between Session 1 and 2, so teachers would have more time to complete the pre-work for Session 2. One suggestion from a participant was to introduce the AI tools using a jigsaw game where every teacher explores an assigned AI tool and presents it back to the group.

4.4 Limitations and Future Work

The above findings contribute to the under-explored need to collaborate with teachers when designing AI curriculum, as well as the potential for AI to be integrated into K-12 core curriculum. However, there were limitations in this research that should be considered when interpreting the findings. First, while the teachers had full control over design decisions in the curriculum, the co-design process was not fully teacher-driven, as the workshop structure was decided by the researchers. Strategies used to address the power dynamics of this study include group facilitation techniques and activities designed to promote equal participant contributions. Future studies could involve teacher stakeholders earlier in the workshop planning process. Second, the curricula that teachers co-created were not evaluated on their efficacy with students. While we are aware of many teacher participants later reporting using these curricula in their classrooms post-workshop, a follow-up evaluative study would be necessary to close the loop on how successful integrated AI curricula can be in front of students. Lastly, our findings could have been impacted by our workshop logistics. Future studies could organize larger-scale workshops to address our small sample size, as well as provide a longer workshop time frame, which our teachers identified as a limitation.

5 CONCLUSION

In this paper, we stressed the importance of teacher perspectives in the adoption of AI curricula and learning tools, and explored teacher values in K-12 classrooms and how AI education and tools can be integrated with existing core curriculum. We engaged K-12 teachers and researchers in a two-day co-design workshop, where we co-created lesson plans that embedded AI concepts into curricula for social studies, ESL, and literacy for students with learning disabilities. We found that teachers value curriculum that address evaluation and engagement of students, which could be built into the learning tool or curriculum. Teachers also successfully connected AI with their subject by having students examine subject-related datasets, as well as reflect on real-world implications and AI ethics. Our work highlights an opportunity to increase accessibility of K-12 AI education by embedding AI into core subjects (e.g., English, social studies), and reaching students outside of CS and technology classrooms.

ACKNOWLEDGMENTS

We thank our teachers who made this work possible. We also thank our anonymous reviewers, Loren Terveen, Randi Williams, and Hal Abelson for their feedback and guidance.

REFERENCES

[1] 2011. Miro. https://miro.com/. Accessed: 2020-08-13.
[2] 2011. Zoom. https://zoom.us/. Accessed: 2020-08-13.
[3] 2013. Slack. https://slack.com/. Accessed: 2020-08-13.
[4] 2020. A to Z of AI with Google. https://atozofai.withgoogle.com/intl/en-US/about/. Accessed: 2020-08-13.
[5] Adam Agassi, Hadas Erel, Iddo Yehoshua Wald, and Oren Zuckerman. 2019. Scratch nodes ML: A playful system for children to create gesture recognition classifiers. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.
[6] Safinah Ali and Brian Jordan. 2020. Doodles to Pictures! https://mitmedialab.github.io/GAN-play/. Accessed: 2020-09-10.
[7] Safinah Ali, Blakeley H Payne, Randi Williams, Hae Won Park, and Cynthia Breazeal. 2019. Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education. In *Proceedings of the International Workshop on Education in Artificial Intelligence in K-12*. Springer, Cyberspace.
[8] Jan Auernhammer. 2020. Human-centered AI: The role of Human-centered Design Research in the development of AI. In *Synergy - DRS International Conference 2020*. Design Research Society. https://doi.org/10.21606/drs.2020.282
[9] Richard Berk. 2017. An impact assessment of machine learning risk forecasts on parole board decisions and recidivism. *Journal of Experimental Criminology* 13, 2 (2017), 193–216.
[10] Karl-Emil Kjær Bilstrup, Magnus H Kaspersen, and Marianne Graves Petersen. 2020. Staging Reflections on Ethical Dilemmas in Machine Learning: A Card-Based Design Workshop for High School Students. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 1211–1222.
[11] Harald Burgsteiner, Martin Kandlhofer, and Gerald Steinbauer. 2016. IRobot: Teaching the basics of artificial intelligence in high schools. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. 4126–4127.
[12] Harald Burgsteiner, Martin Kandlhofer, and Gerald Steinbauer. 2016. Irobot: Teaching the basics of artificial intelligence in high schools. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 30.
[13] Michelle Carney, Barron Webster, Irene Alvarado, Kyle Phillips, Noura Howell, Jordan Griffith, Jonas Jongejan, Amit Pitaru, and Alexander Chen. 2020. Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
[14] H. Chen. 2009. AI, E-government, and Politics 2.0. *IEEE Intelligent Systems* 24, 5 (2009), 64–86.
[15] Licui Chen and Sihan Xiao. 2020. Perceptions, challenges and coping strategies of science teachers in teaching socioscientific issues: A systematic review. *Educational Research Review* (2020), 100377.
[16] Code.org. 2019. AI for Oceans. https://code.org/oceans. Accessed: 2020-09-10.
[17] Mathieu d'Aquin. 2020. The Data Ethics Challenges of Explainable AI and Their Knowledge-Based Solutions. In *Knowledge Graphs for eXplainable Artificial Intelligence: Foundations, Applications and Challenges*, Ilaria Tiddi, Freddy Lécué, and Pascal Hitzler (Eds.). Studies on the Semantic Web, Vol. 47. IOS Press, Berlin, 262–275.
[18] Manuj Dhariwal and Shruti Dhariwal. 2020. Let's Chance: Playful Probabilistic Programming for Children. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–7.
[19] Daniella DiPaola, Blakeley H Payne, and Cynthia Breazeal. 2020. Decoding design agendas: an ethical design activity for middle school students. In *Proceedings of the Interaction Design and Children Conference*. 1–10.
[20] Andrea L Dixon, Catherine Tucker, and Mary Ann Clark. 2010. Integrating social justice advocacy with national standards of practice: Implications for school counselor education. *Counselor Education and Supervision* 50, 2 (2010), 103–115.
[21] Alison G Dover. 2013. Getting "up to code": Preparing for and confronting challenges when teaching for social justice in standards-based classrooms. *Action in Teacher Education* 35, 2 (2013), 89–102.
[22] Stefania Druga. 2018. *Growing up with AI: Cognimates: from coding to teaching machines*. Master's thesis. Massachusetts Institute of Technology.
[23] Stefania Druga, Sarah T Vu, Eesh Likhith, and Tammy Qiu. 2019. Inclusive AI literacy for kids around the world. In *Proceedings of FabLearn 2019*. 104–111.
[24] Eva Durall, Merja Bauters, Iida Hietala, Teemu Leinonen, and Evangelos Kapros. 2020. Co-creation and co-design in technology-enhanced learning: Innovating science learning outside the classroom. *Interaction Design and Architecture (s)* 42 (2020), 202–226.
[25] Education World Inc. 2020. U.S. Education Standards. https://www.educationworld.com/standards/national/. Accessed: 2020-09-15.
[26] Torrie K Edwards. 2020. From the Editorial Board: Tangled Discrimination in Schools: Binding Hair to Control Student Identity. *The High School Journal* 103, 2 (2020), 53–56.
[27] Juliana J Ferreira and Mateus S Monteiro. 2020. What Are People Doing About XAI User Experience? A Survey on AI Explainability Research and Practice. In *International Conference on Human-Computer Interaction*. Springer, 56–73.
[28] Casey Fiesler, Natalie Garrett, and Nathan Beard. 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*. 289–295.
[29] Batya Friedman, David G Hendry, and Alan Borning. 2017. A survey of value sensitive design methods. *Foundations and Trends in Human-Computer Interaction* 11, 2 (2017), 63–125.
[30] Batya Friedman and Peter H Kahn Jr. 2003. Human values, ethics, and design. *The human-computer interaction handbook* (2003), 1177–1201.
[31] Ashok Goel. 2017. AI Education for the World. *AI Magazine* 38, 2 (2017), 3–4.
[32] Michael Goldweber, John Barr, Tony Clear, Renzo Davoli, Samuel Mann, Elizabeth Patitsas, and Scott Portnoff. 2013. A framework for enhancing the social good in computing education: a values approach. *ACM Inroads* 4, 1 (2013), 58–79.
[33] Barbara J Grosz, David Gray Grant, Kate Vredenburgh, Jeff Behrends, Lily Hu, Alison Simmons, and Jim Waldo. 2019. Embedded EthiCS: integrating ethics across CS education. *Commun. ACM* 62, 8 (2019), 54–61.
[34] Mona Leigh Guha, Allison Druin, and Jerry Alan Fails. 2013. Cooperative Inquiry revisited: Reflections of the past and guidelines for the future of intergenerational co-design. *International Journal of Child-Computer Interaction* 1, 1 (2013), 14–23.
[35] Clint Heinze, Janet Haase, and Helen Higgins. 2010. An action research report from a multiyear approach to teaching artificial intelligence at the K-6 level. In *1st symposium on educational advances in artificial intelligence*.
[36] Aaron Hertzmann. 2020. Visual Indeterminacy in GAN Art. *Leonardo* 53, 4 (2020), 424–428.
[37] Tom Hitron, Yoav Orlev, Iddo Wald, Ariel Shamir, Hadas Erel, and Oren Zuckerman. 2019. Can Children Understand Machine Learning Concepts? The Effect of Uncovering Black Boxes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.
[38] Ayanna Howard and Jason Borenstein. 2018. The ugly truth about ourselves and our robot creations: the problem of bias and social inequity. *Science and engineering ethics* 24, 5 (2018), 1521–1536.
[39] Ole Sejer Iversen, Kim Halskov, and Tuck W Leong. 2012. Values-led participatory design. *CoDesign* 8, 2-3 (2012), 87–103.

[40] Jonas Jongejan, Henry Rowley, Takashi Kawashima, Jongmin Kim, and Nick Fox-Gieg. 2017. Google Quick, Draw! https://quickdraw.withgoogle.com/. Accessed: 2020-09-16.
[41] K Megasari Kahn, Rani Megasari, Erna Piantari, and Enjun Junaeti. 2018. AI programming by children using Snap! block programming in a developing country. In *Thirteenth European Conference on Technology Enhanced Learning 2018*. Springer, Leeds, UK, 1–14.
[42] Lorraine Kisselburgh, Michel Beaudouin-Lafon, Lorrie Cranor, Jonathan Lazar, and Vicki L Hanson. 2020. HCI Ethics, Privacy, Accessibility, and the Environment: A Town Hall Forum on Global Policy Issues. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–6.
[43] Dale Lane. 2020. Machine Learning for Kids. https://machinelearningforkids.co.uk/. Accessed: 2020-09-05.
[44] Phoebe Lin, Jessica Van Brummelen, Galit Lukin, Randi Williams, and Cynthia Breazeal. 2020. Zhorai: Designing a Conversational Agent for Children to Explore Machine Learning Concepts.. In *AAAI*. AAAI, New York, NY, USA, 13381–13388.
[45] Daria Loi, Christine T Wolf, Jeanette L Blomberg, Raphael Arar, and Margot Brereton. 2019. Co-designing AI Futures: Integrating AI Ethics, Social Computing, and Design. In *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*. 381–384.
[46] Duri Long and Brian Magerko. 2020. What is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376727
[47] MIT App Inventor. 2020. Artificial Intelligence with MIT App Inventor. https://appinventor.mit.edu/explore/ai-with-mit-app-inventor. Accessed: 2020-09-10.
[48] Henrik Mucha, Dennis Mevißen, Sebastian Robert, Ricarda Jacobi, Kirsten Meyer, Winfried Heusler, and Daniel Arztmann. 2020. Co-Design Futures for AI and Space: A Workbook Sprint. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.
[49] Reiichiro Nakano. 2019. Arbitrary Style Transfer in the Browser. https://reiinakano.com/arbitrary-image-stylization-tfjs/. Accessed: 2020-09-16.
[50] Taro Narahara and Yoshihiro Kobayashi. 2018. Personalizing homemade bots with plug & play AI for STEAM education. In *SIGGRAPH Asia 2018 Technical Briefs*. 1–4.
[51] Norzila Ngadiman, Shahida Sulaiman, Norsham Idris, and Hasnah Mohamed. 2020. A Comparative Study on Quality Characteristics in Designing Educational Applications. *International Journal of Information and Education Technology* 10, 8 (2020).
[52] NSTA. 2014. Access the Next Generation Science Standards by Topic. https://ngss.nsta.org/AccessStandardsByTopic.aspx. Accessed: 2020-09-14.
[53] Viktoria Pammer-Schindler, Erik Harpstead, Benjamin Xie, Betsy DiSalvo, Ahmed Kharrufa, Petr Slovak, Amy Ogan, Joseph Jay Williams, and Michael J Lee. 2020. Learning and education in HCI: A reflection on the SIG at CHI 2019. *Interactions* 27, 5 (2020), 6–7.
[54] Samiullah Paracha, Lynne Hall, Kathy Clawson, Nicole Mitsche, and Fatima Jamil. 2019. Co-design with Children: Using Participatory Design for Design Thinking and Social and Emotional Learning. *Open Education Studies* 1, 1 (2019), 267–280.
[55] William R Penuel. 2019. Co-design as infrastructuring with attention to power: Building collective capacity for equitable teaching and learning through design-based implementation research. In *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning*. Springer, Cham, 387–401.
[56] Marc Prensky. 2008. Programming is the new literacy. *Edutopia magazine* (2008).
[57] Jeremy Roschelle, William Penuel, and Nicole Shechtman. 2006. Co-design of innovations with teachers: Definition and dynamics. In *Proceedings of the 7th international conference on Learning sciences*. International Society of the Learning Sciences, 606–612.
[58] Troy D Sadler, Aidin Amirshokoohi, Mahsa Kazempour, and Kathleen M Allspaw. 2006. Socioscience and ethics in science classrooms: Teacher perspectives and strategies. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching* 43, 4 (2006), 353–376.
[59] Bawornsak Sakulkueakulsuk, Siyada Witoon, Potiwat Ngarmkajornwiwat, Pornpen Pataranutaporn, Werasak Surareungchai, Pat Pataranutaporn, and Pakpoom Subsoontorn. 2018. Kids making AI: Integrating Machine Learning, Gamification, and Social Context in STEM Education. In *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. IEEE, 1005–1010.
[60] Jeffrey Saltz, Michael Skirpan, Casey Fiesler, Micha Gorelick, Tom Yeh, Robert Heckman, Neil Dewar, and Nathan Beard. 2019. Integrating ethics within machine learning courses. *ACM Transactions on Computing Education (TOCE)* 19, 4 (2019), 1–26.
[61] Elizabeth B-N Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *Co-design* 4, 1 (2008), 5–18.
[62] Michael Scaife, Yvonne Rogers, Frances Aldrich, and Matt Davies. 1997. Designing for or designing with? Informant design for interactive learning environments. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. 343–350.
[63] Alexander Scheidt and Tim Pulver. 2019. Any-Cubes: A Children's Toy for Learning AI: Enhanced Play with Deep Learning and MQTT. In *Proceedings of Mensch und Computer 2019*. 893–895.
[64] Samuel Severance, William R Penuel, Tamara Sumner, and Heather Leary. 2016. Organizing for teacher agency in curricular co-design. *Journal of the Learning Sciences* 25, 4 (2016), 531–564.
[65] Michael Skirpan, Nathan Beard, Srinjita Bhaduri, Casey Fiesler, and Tom Yeh. 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*. 940–945.
[66] Ahuva Sperling and Dorit Lickerman. 2012. Integrating AI and machine learning in software engineering course for high school students. In *Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education*. 244–249.
[67] Shashank Srikant and Varun Aggarwal. 2017. Introducing data science to school kids. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. 561–566.
[68] Jonathan Stray. 2019. Making artificial intelligence work for investigative journalism. *Digital Journalism* 7, 8 (2019), 1076–1097.
[69] Danny Tang, Yuria Utsumi, and Natalie Lao. 2019. PIC: A Personal Image Classification Webtool for High School Students. In *Proceedings of the 2019 IJCAI EduAI Workshop*. IJCAI.
[70] TensorFlow. 2020. Mobile BERT Q & A. https://storage.googleapis.com/tfjs-models/demos/mobilebert-qna/index.html. Accessed: 2020-09-10.
[71] T. Toivonen, I. Jormanainen, J. Kahila, M. Tedre, T. Valtonen, and H. Vartiainen. 2020. Co-Designing Machine Learning Apps in K–12 With Primary School Children. In *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*. 308–310.
[72] David Touretzky, Christina Gardner-McCune, Cynthia Breazeal, Fred Martin, and Deborah Seehorn. 2019. A year in K-12 AI education. *AI Magazine* 40, 4 (2019), 88–90.
[73] David Touretzky, Christina Gardner-McCune, Cynthia Breazeal, Fred Martin, and Deborah Seehorn. 2019. A year in K-12 AI education. *AI Magazine* 40, 4 (2019), 88–90.
[74] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12: What should every child know about AI?. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 9795–9799.
[75] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12: What should every child know about AI?. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. AAAI, Honolulu, Hawaii, USA, 9795–9799.
[76] David S Touretzky and Christina Gardner-McCune. 2018. Calypso for Cozmo: Robotic AI for everyone. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*. ACM, Baltimore, MD, USA, 1110–1110.
[77] Marie E Vachovsky, Grace Wu, Sorathan Chaturapruek, Olga Russakovsky, Richard Sommer, and Li Fei-Fei. 2016. Toward more gender diversity in CS through an artificial intelligence summer program for high school girls. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. 303–308.
[78] Jessica Van Brummelen. 2019. Conversational Agents to Democratize Artificial Intelligence. In *2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, 239–240.
[79] Jessica Van Brummelen, Tommy Heng, and Viktoriya Tabunshchyk. 2021. Teaching Tech to Talk: K-12 Conversational Artificial Intelligence Literacy Curriculum and Development Tools. In *EAAI at AAAI*. AAAI, New York, NY, USA, 13381–13388.
[80] Jessica Van Brummelen, Tommy Heng, and Viktoriya Tabunshchyk. 2021. Teaching Tech to Talk: K-12 Conversational Artificial Intelligence Literacy Curriculum and Development Tools. In *2021 AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI)*.
[81] Jessica Van Brummelen and Phoebe Lin. 2020. Appendix. https://gist.github.com/jessvb/f415307d4696e6754e2f2e39a0a897c0. Accessed: 2020-12-23.
[82] Maarten Van Mechelen, Jan Derboven, Ann Laenen, Bert Willems, David Geerts, and Vero Vanden Abeele. 2017. The GLID method: Moving from design features to underlying values in co-design. *International Journal of Human-Computer Studies* 97 (2017), 116–128.
[83] Maarten Van Mechelen, Ann Laenen, Bieke Zaman, Bert Willems, and Vero Vanden Abeele. 2019. Collaborative Design Thinking (CoDeT): A co-design approach for high child-to-adult ratios. *International Journal of Human-Computer Studies* 130 (2019), 179–195.
[84] A. Vazhayil, R. Shetty, R. R. Bhavani, and N. Akshay. 2019. Focusing on Teacher Education to Introduce AI in Schools: Perspectives and Illustrative Findings. In *2019 IEEE Tenth International Conference on Technology for Education (T4E)*. 71–77.
[85] Xiaoyu Wan, Xiaofei Zhou, Zaiqiao Ye, Chase K Mortensen, and Zhen Bai. 2020. SmileyCluster: supporting accessible machine learning in K-12 scientific discovery. In *Proceedings of the Interaction Design and Children Conference*. 23–35.
[86] Randi Williams, Hae Won Park, and Cynthia Breazeal. 2019. A is for artificial intelligence: the impact of artificial intelligence activities on young children's

perceptions of robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.

[87] Jed R. Wood and Larry E. Wood. 2008. Card Sorting: Current Practices and Beyond. 4, 1 (Nov. 2008), 1–6.

[88] Julia Woodward, Zari McFadden, Nicole Shiver, Amir Ben-hayon, Jason C Yip, and Lisa Anthony. 2018. Using co-design to examine how children conceptualize intelligent interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[89] Benjamin Xie, Erik Harpstead, Betsy DiSalvo, Petr Slovak, Ahmed Kharrufa, Michael J. Lee, Viktoria Pammer-Schindler, Amy Ogan, and Joseph Jay Williams. 2019. Learning, Education, and HCI. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/3290607.3311761

[90] Xiaozhe Yang. 2019. Accelerated Move for AI Education in China. *ECNU Review of Education* 2, 3 (2019), 347–352.

[91] Adrienne Yapo and Joseph Weiss. 2018. Ethical implications of bias in machine learning. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

[92] Svetlana Yarosh, Iulian Radu, Seth Hunter, and Eric Rosenbaum. 2011. Examining values: an analysis of nine years of IDC research. In *Proceedings of the 10th International Conference on Interaction Design and Children*. 136–144.

[93] Jason C Yip, Tamara Clegg, June Ahn, Judith Odili Uchidiuno, Elizabeth Bonsignore, Austin Beck, Daniel Pauw, and Kelly Mills. 2016. The evolution of engagements and social bonds during child-parent co-design. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3607–3619.

[94] Junnan Yu, Chenke Bai, and Ricarose Roque. 2020. Considering Parents in Coding Kit Design: Understanding Parents' Perspectives and Roles. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[95] Xiaofei Zhou, Jessica Van Brummelen, and Phoebe Lin. 2020. Designing AI Learning Experiences for K-12: Emerging Works, Future Opportunities and a Design Framework. *arXiv preprint arXiv:2009.10228* (2020).

[96] Michelle Zimmerman. 2018. Teaching AI: Exploring New Frontiers for Learning. *International Society for Technology in Education* (12 2018).

[97] Abigail Zimmermann-Niefield, R Benjamin Shapiro, and Shaun Kane. 2019. Sports and machine learning: How young people can use data from their own bodies to learn about machine learning. *XRDS: Crossroads, The ACM Magazine for Students* 25, 4 (2019), 44–49.

[98] Abigail Zimmermann-Niefield, R Benjamin Shapiro, and Shaun Kane. 2019. Sports and machine learning: How young people can use data from their own bodies to learn about machine learning. *XRDS: Crossroads, The ACM Magazine for Students* 25, 4 (2019), 44–49.