Exploring Generative Models with Middle School Students

Safinah Ali*
safinah@media.mit.edu
MIT Media Lab
Cambridge, Massachusetts

Daniella DiPaola*
MIT Media Lab
Cambridge, Massachusetts
dipaola@mit.edu

Irene Lee
MIT STEP Lab
Cambridge, Massachusetts
ialee@mit.edu

Jenna Hong
MIT Media Lab
Cambridge, Massachusetts
jhong19@mit.edu

Cynthia Breazeal
MIT Media Lab
Cambridge, Massachusetts
cynthiab@media.mit.edu

Figure 1: A picture of the Jonas Brothers (left) was altered by FaceApp (right)

ABSTRACT

Applications of generative models such as Generative Adversarial Networks (GANs) have made their way to social media platforms that children frequently interact with. While GANs are associated with ethical implications pertaining to children, such as the generation of Deepfakes, there are negligible efforts to educate middle school children about generative AI. In this work, we present a generative models learning trajectory (LT), educational materials, and interactive activities for young learners with a focus on GANs, creation and application of machine-generated media, and its ethical implications. The activities were deployed in four online workshops with 72 students (grades 5-9). We found that these materials enabled children to gain an understanding of what generative models are, their technical components and potential applications, and benefits and harms, while reflecting on their ethical implications. Learning from our findings, we propose an improved learning trajectory for complex socio-technical systems.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning**; *Learning paradigms*; *Learning settings*; *Machine learning approaches*; *Machine learning algorithms*; • **Social and professional topics** → **Computing education**; *Model curricula*.

KEYWORDS

Generative Machine Learning, AI Education, Generative Adversarial Networks, Artificial Intelligence

*Both authors contributed equally to this research.

1 INTRODUCTION

In July of 2019, it seemed as if everyone's social media timelines were fast forwarded 50 years in the future. Celebrities, from Drake to the Jonas Brothers, were using FaceApp [15], a photo filtering app, to make themselves look older (Figure 1). These filtered images, often realistic to the point of uneasiness, began spreading through timelines, branded as the #AgeChallenge. The widespread use of this FaceApp, especially by children on social media, ignited discussions about the implications of high fidelity fake images and data collection by the companies that create these technologies [51].

To create these realistic photos, FaceApp used a generative machine learning model to style one's face in the style of an older adult. Applications of popular generative modeling techniques such as Generative Adversarial Networks (or GANs) have become commonplace on social media. Introduced in 2014, GANs are a novel application of machine learning that create new data instances that resemble the training data [21]. While largely used for experimental generation of art and entertainment [18], GANs have applications in science such as improving astronomical images [47] or detecting glaucomatous images that help with early diagnosis of blindness [8]. Concerns have also been raised about the use of GANs to generate hyper-realistic images and videos of humans, also known as Deepfakes, for sinister purposes, such as to produce fake, possibly incriminating, photographs and videos [40].

Middle school students are both creators and consumers of generative machine learning. Due to the speed and accuracy of generative AI, doctored images can be created with little to no effort. Students are using face filtering apps such as FaceApp and FaceTune to alter their own photos [15, 16]. Commercial applications that make use of Deepfakes, such as "Reface: Swap your faces now", use GANs

to swap faces in popular media with users' faces, are also being advertised on social networking applications such as TikTok, which is frequented by middle schoolers[35, 55]. This media gets posted on students' feeds and timelines, making it impossible to know which images are real and which have been doctored.

At the same time, generative machine learning poses exciting advancements for the future. For example, GANs can be used to create new types of art [59] as well as advancements in medical imaging [30]. If GANs are to enter these various fields, it is important that students can recognize and collaborate with these tools. By designing activities where students can practice creating with generative AI models, we can make collaboration with AI more accessible to the future workforce. Right now, there is little awareness about the existence and functioning of these networks outside of computer science and artificial intelligence professionals, and there exist no published frameworks, in our knowledge, to teach children about how GANs work and how they are used in the real world. If we make these tools accessible to students, we can prepare them for opportunities in the future workforce that involve collaborating with AI.

Questions arise as to what elements to teach for different age groups and how to teach it. In order to make these concepts accessible, we must consider how much we can teach about generative models before we introduce the underlying mathematics behind them, or how far do analogies or related context go in teaching the composition and applications of generative modeling techniques, and what learning goals are applicable for different age groups.

In this work we developed a generative models LT for middle school students with no technical knowledge of machine learning. We developed interactive learning activities to introduce to children: (1) the concept of generation, (2) how GANs work, (3) applications of GANs, their benefits, and harms (4) creation tools that leverage generative models, and (5) ethical implications of GANs with a focus on Deepfakes.

This paper aims to answer the following research question: *To what extent does our design of a learning trajectory support middle school students' development of understandings of technical composition, applications and ethical implications of generative models?*

In order to explore this question, the paper presents the following contributions:

- The design of a generative models learning trajectory (LT) with a focus on GANs, with learning goals accessible for middle school students.
- Descriptions of interactive learning activities that explore the technical composition, applications and ethical implications generative modeling techniques, with a focus on GANs, aimed at broad non-technical audiences, to be used in middle school classrooms.
- Findings on students' learning gains, artifacts generated, and perspectives relevant to generative modeling.

This paper is organized as follows. We present the theoretical frameworks that guided the developments of our learning trajectory and learning activities, followed by the background in Section 3, materials design in Section 4, followed by the methods and results. We conclude in Section 9.

2 THEORETICAL FRAMING

The Learning Trajectories (LTs) presented in the paper were developed after considering the theoretical background of LTs [48], Battista (2011) defines a Learning Trajectory (LT) as "a detailed description of the sequence of thoughts, ways of reasoning, and strategies that a student employs while involved in learning a topic" [7]. LTs consist of nodes representing learning goals connected by arrows that depict potential orderings between these learning goals. Learning trajectories must also be cognizant of students' cognitive development. Earlier nodes in an LT are ideas that are easier to comprehend for students, and they process to more complex novel concepts. In terms of thinking and learning, middle school students are able to think more abstractly than younger students [2, 12] and they can incorporate new knowledge into existing schemas [27, 28]. We made use of game-based learning approaches to make abstract concepts concrete since games have shown to enhance content engagement [19, 20, 31]. LTs have been extensively used for building curricula in mathematics [11], and more recently in Computer Science [17]. Learning trajectories have proven helpful when taking a constructivist approach to curriculum development wherein students' existing knowledge supplements learning new concepts [17].

Hypothetical LTs, developed prior to the learning process, consist of learning goals, learning activities, and predict a hypothetical learning process of how the students' thinking and understanding will evolve in the context of the learning activities [48]. HLT was suitable in this research because it serves as a link between instruction theory and a concrete teaching experiment [24], and was originally conceived as part of a model of teacher decision making [48]. HLTs were useful in articulating how learning goals, cognitive tasks, and instructional tasks fit together and formed a theory of action in design research [46]. In contrast, *actual* LTs are developed when the student has actually progressed through the learning path. In this work, we develop a hypothetical LT for generative models aimed at middle school students.

A Hypothetical LT (HLT) approach was followed in our work to develop the generative modeling LT in Section 4.1 (Design of the LT). We posited a conjecture regarding student learning based on students' presumed current understanding, then developed activities to support students in constructing more sophisticated ways of reasoning about generative AI. We followed the hierarchical ordering of cognitive skills in the (revised) Bloom's taxonomy: remember, understand, apply, analyze, evaluate, and create [5, 9]. Our conjecture was that students could progress from exposure and awareness of generative AI, to recall of its definition, to an explanation of the processes involved, to exploring the application of generative AI in a variety of media, and to the analysis and evaluation of the potential benefits and harms of applications of generative AI, before reaching the upper anchor of responsibly creating with generative AI.

Hypothetical Learning Trajectories (HLTs) were useful in articulating how learning goals, cognitive tasks; and instructional tasks fit together and formed a theory of action in design research [46]. However, in the case of knowledge about socio-technical systems (systems that constitute interactions between society's complex infrastructures and human behaviour technical components), applications of the technology, and ethical implications are interconnected

and inform each other [58]. Thus, the knowledge space cuts across social, technological, and ethical domains and the learning is not expected to be linear. In particular, the design of a learning trajectory for generative models is predicated on the need for making these connections between social, technical, and ethical domains. While learning trajectories are not the only learning framework for curriculum development, we chose to use this framework since the technical components of generative modeling are novel to young children, though the concept of generation and the applications of generative modeling are already accessible to them.

3 BACKGROUND

In this work, we develop a middle-school AI Education curriculum focusing on the theory, practice and ethical implications of generative models with a focus on GANs.

3.1 Generative models

Generative models describe a class of statistical models that *create* new data instances that resemble an existing set of training data. They are often seen in contrast with discriminative models, that tell apart two different data instances. A generative model could generate new photos of cats that look like real cats, while a discriminative model could tell a dog from a cat. There are several different kinds of generative models such as Generative Adversarial Networks (GANs) [21], Variational Autoencoders (VAEs) [13], and Autoregressive models such as PixelRNN [41]. Amongst other generative models, recent developments in GANs have demonstrated especially high accuracy in synthetic media generation [https://openai.com/blog/generative-models/], and have been prevalent in popular media. In this work, we explore the applications of different kinds of generative models, with a special focus on GANs.

3.2 Generative Adversarial Networks (or GANs)

First introduced by Ian Goodfellow in 2014, Generative adversarial networks (or GANs) are a new kind of generative machine learning model [21] that are used to generate different kinds of media, such as visual art, photographs, music, videos, and text. GANs are able to produce realistic new data by pairing a generator, that learns to produce the target output, with a discriminator, that learns to distinguish existing training data from the output of the generator. The generator tries to fool the discriminator that it's image is "real", while the discriminator tries to get better at distinguishing the "real" vs the "fake" image. The generator and the discriminator work together to help the generate new data that replicates patterns in the training data. The existence and use of GANs raises some ethical concerns such as "Should it be legal to produce and circulate fake media?", "Who owns GANs generated works of art?", "Who is responsible when fake media produces, whether directly or indirectly, harm to individuals?" and "Are the environmental impacts of training big AI models used in GANs proportional to the benefits produced?"

3.3 GANs and middle school AI

The impetus for educating middle school students about GANS is driven by their increasing contact with GANs in daily life. First, since many students acquire their first mobile device during middle

school [45], they start consuming data on social media websites such as Instagram where they are exposed to GANs-generated content [6]. Second, middle school students are already creators of media generated with GANS. Through social media apps such as Snapchat and Instagram, they view and create content with tools such as photo filters that integrate generative modeling techniques. Thus middle school students may be using GANs-enabled technology without realizing it. Third, children upload personal data, such as images, videos and text, on social media sites and may unwittingly be contributing data to datasets used to train GANs models. Finally, students are witness to and could be targeted by fake media that are generated by applications of GANs such as Deepfakes, like in the case of FaceApp.

This exposure to GANs, 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 GANS. Their awareness of manipulated media has ramifications for democracy, trust, security, and privacy. While some synthetic media are difficult to discern even for adults, the knowledge of the existence of GAN-generated media would empower children to take information that they witness online with essential skepticism. Thus, AI literacy focusing on generative machine learning techniques such as GANs and media generated by GANs is imperative for students to be informed citizens and critical consumers of online media. This curriculum addresses the need for students to be critical consumers, creators and detectors of GANs created media by introducing students to GANs, how they work, what kind of data they use, and applications of GANs. Through the curriculum, students are engaged in detecting and thinking critically about the benefits and harms of these applications of GANs.

As applications of AI become more common in children's lives, it becomes imperative that they learn how these technologies work and be able to responsibly create with them. In 2018, the AI4K12 initiative defined the "Five Big Ideas of AI" that K12 students must know as Perception, Representation and Reasoning, Learning, Natural Interaction, and Societal Impact. Long and Magerko (2020) described AI literacy as "AI literacy as a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace" [36]. Unsurprisingly, several new AI curricula and frameworks have been developed for k12 learners [56]. Topics of learning included basics of machine learning [39], audio and image classification through creative programming tools [10, 26, 52], data science [49], clustering algorithms [57], and AI ethics [4, 42]. Previous applications of generative modeling have been used to enhance children's creativity in collaborative interactions with intelligent agents [3], but, no K-12 curriculum were found that address the development or applications of generative models. While educators have developed courses focusing on generative machine learning aimed at post-secondary students and adult learners (especially aimed at artists) exist, including a few that focus on generative art [22, 33, 37, 44], most of these courses and activities require prior programming experience, mathematical knowledge or a sophisticated computing set-up, often requiring

cloud computing or GPU capabilities. To our knowledge, there are no Generative Machine Learning curricula or tools for K-12 students, that are suited for low technical requirements.

4 DESIGN CONSIDERATIONS

In this section we describe the design considerations that informed the development of the learning trajectory (Figure 2) and the activities aligned to the progression.
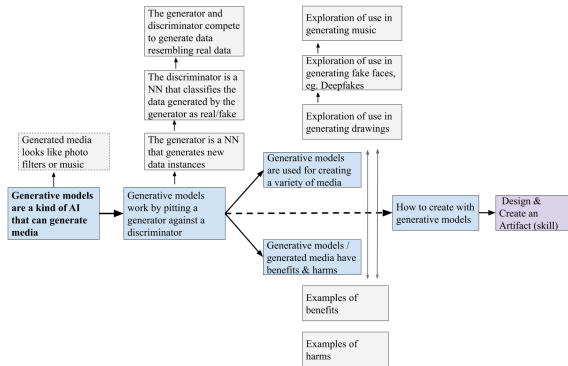

Figure 2: Generative models Learning Trajectory
.

4.1 Design of the Learning Trajectory

The design of a learning trajectory for generative models is predicated on the need for making connections between social, technical, and ethical domains evident. Importantly, instead of treating ethical implications as an independent learning goal, we take an integrated ethics approach, where we explicitly discuss ethical implications connected with applications of generative models and with creating with generative models. The resulting LT is multi-layered instead of linear. While existing approaches to generative models take either a mathematical or a programming-driven approach, our challenge was to create a learning trajectory for learners with no prerequisite knowledge in these fields. In designing the overall progression, we order our learning goals depending on the perceived accessibility to the concept, as well as our prior work in the field. In order to begin with lower "anchor points", or ideas that students already have experience with, we begin by introducing generative models through examples with which students are already familiar. Once students were aware of the existence of these models, we delved deeper into what GANs are, and how they work. Students applied their procedural knowledge about how GANs work to applications of GANs that they interacted with. Students discussed the possible benefits and harms of applications of GANs. We also discussed strategies to recognize Deepfakes. Finally, students used two generative tools to create generative stories. In this section, we define the four learning goals of this LT, followed by the learning activities that address these goals.

4.2 Design of the activities

The design of activities was informed by the following design principles and considerations about the learners, their environments, and learning goals.

- **No CS and advanced math prerequisites.** The generative models curriculum and tools were intentionally designed for use in K-12 environments with middle school students. The design took into consideration students' lack of exposure to computer science and advanced mathematics. Our workshop and lessons were designed for middle school students with no prior knowledge of artificial intelligence, machine learning, or neural networks. Since middle school students do not have the advanced mathematical or statistical skills, we chose a conceptual approach to teaching GANs that drew inspiration from other middle school CS and AI curricula that utilize unplugged activities.
- **Make activities accessible to those with limited bandwidth and/or older devices.** Specifically, the GANs activity was designed as a low-tech web-based activity that minimized prerequisite knowledge and technical knowledge to set up. The activities were designed for synchronous remote learning making it accessible to most students through a browser.
- **Support students' abstraction capabilities.** We take students' cognitive development into account when introducing abstract concepts and processes. We provide concrete representations to teach the complex processes in GANs in a simpler understandable manner. For example, the process of generating synthetic data were simplified to picking tiles to compose an image.
- **Use a game based learning approach.** In the case of the generator-discriminator game, the learners take the role of generator or discriminator, and follow steps or rules that are analogous to processes that each neural network performs. Since the rules mimic the processes used by the discriminator and generator, the increased exposure to these rules may lead to better recall and understanding of the processes themselves. Further, competitive game play between players acting as the generator and the discriminator also accurately depicts the roles of a generator and a discriminator in GANs which work in opposition to produce synthetic data.
- **Reduce cognitive load.** Efforts were made to decrease cognitive load [50] and reduce the space of possibilities during game play. A small 8 by 8 tiled grid was chosen to represent the image being reproduced thereby reducing the complexity of the game by limiting the number of possible guesses. Additionally, rather than having the discriminator provide weights or probabilities as feedback, tiles were colored in green (indicating a correct guess) or red (indicating an incorrect guess) thus simplifying the feedback to a binary response and minimizing the interpretation of the feedback.

5 ACTIVITY DESCRIPTIONS

5.1 Activity 1. Introduction to generative models

We begin the course by introducing students to the concept of generation. Students learn that AI can be used to generate synthetic media such as images, text, music, colors, paintings, digits or videos. We begin the activity by playing a game called "Created by GANs or Not". In the game, they are given various works of art (i.e. visual,

Learning Cluster	Learning Goal	Targeted Cognitive Skill	Activity
Introduction to generative models	Be introduced to the concept of generative models through familiar media	Remember	1
	Understand the breadth of what can be created with generative models	Understand	1
	Learn that AI-generated media are sometimes difficult to discern from human-created media	Understand	1
Structural Components of GANs	Understand the relationship between the generator and the discriminator, the two neural networks that make up a GAN	Understand	2
Applications of GANs	Get acquainted with applications of GANs for creating different kinds of media: images, drawings, music, and videos	Apply	3
	For a specific GAN, identify what the generator is trying to create	Apply	3
	For a specific GAN, identify the dataset used by the discriminator to make its decisions	Apply	3
	For a specific GAN, identify what the potential benefits are	Analyze	3
	For a specific GAN, identify what the potential harms are	Analyze	3
	Get acquainted with the concept of Deepfakes	Apply	4
	Learn some strategies to identify Deepfakes	Evaluate	4
Create with generative models	Partner with generative models to create stories	Create	5
	Experience style transfer in text and images	Create	5

Table 1: The Learning Trajectory: Generative modeling learning goals with the targeted cognitive skill [9] and the corresponding activity number.

audio, text) and are given three options to choose from: "Created by AI", "Not created by AI", or "I am not sure". Examples of media can be found in Figure 3. After completing the activity, students discuss which artwork they thought was created by AI and gave their reasoning. Students are then told that all of the artwork, in fact, was created using a kind of AI model called generative models. We then discuss whether students found this surprising, and which artworks being machine-generated were especially surprising for students. Learning is facilitated by eliciting an emotional response from students through cognitive dissonance, or contradicting their concept of "what is real". Through these examples of media created by generative models, students are introduced to the concept of generative models with a range of instances that students could be familiar with, for example, a photograph altered using a popular filter application Prisma [43], or a joke generated in the knock-knock joke style, or a generative "Happy Birthday" song.

5.2 Activity 2. How do GANs Work?
Generator/Discriminator Simulation Game
After students understand that GANs are used to generate media and explore some examples of GANs, they are introduced to how GANs work. Students learn that a GAN is made up of two neural networks, one called the generator, and one called the discriminator. They are told that the generator and discriminator have two

Figure 3: Students see examples of media created by generative models and try to guess if they are created by AI or not. Left to right: a. Image style transfer photo filter. b. Generated colors. c. Generated digits

different goals that are in competition with one another (1) The goal of the generator is to create something new that the discriminator will classify as "real", and (2) The goal of the discriminator is to detect if what the generator creates is "fake".

Then, the students break up into a "generator" group and a "discriminator" group. The generators are given a 6x6 grid and told that they need to insert 7-9 squares into the grid to create an arrangement of blocks that passes by the discriminator. The discriminators are given a dataset of images (Figure 4), and told that they must accept images that look like the images they were given. For this dataset, we chose pixelated faces, because they could be configured in many different ways and were common enough that students would recognize them.
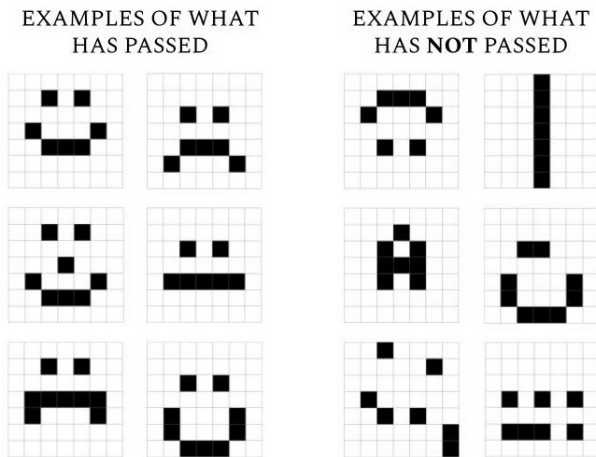
EXAMPLES OF WHAT
HAS PASSED

EXAMPLES OF WHAT
HAS **NOT** PASSED

Figure 4: Dataset presented to the discriminators

.

The game begins when the generators send over a configuration of 7-9 blocks to the discriminator. The discriminator team then has to decide whether or not the configuration should "pass". If it does not pass, then they must give the generator team feedback by giving them one block that is correct (marked in green) and one block that is incorrect (marked in red). The generator then gets another chance to produce a configuration of squares that will pass through the discriminator. The process continues until the generator produces a configuration that the discriminator determines fits with the rest of the data. An example of the back and forth from this game can be found in Figure 5. After the activity, students are told that a GAN goes through this process, but many more times to create much more detailed pieces of media.
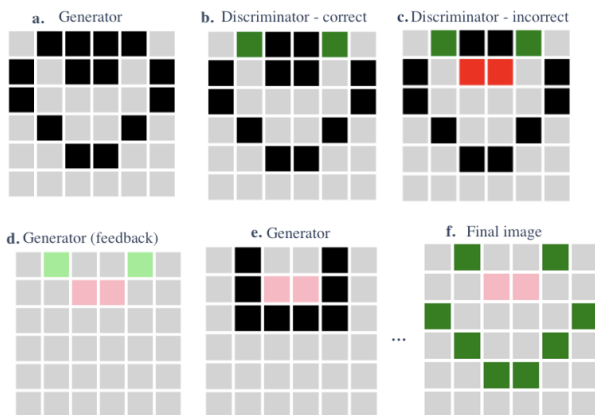
Figure 5: Activity 2. Gameplay of the generator/discriminator simulation game. a. Generator creates an initial guess; b. Discriminator shows 2 pixels that are correct; c. Discriminator shows 2 pixels that are incorrect; d. Generator receives feedback; e. Generator guesses again, (subsequent rounds occur but are not shown); f. Generator generates an image that the disciminator passes as "Correct"

5.3 Activity 3. Applications of GANs

After students understand how GANs work, they are asked to explore some interactive web-based tools that use GANs to create media.

- **AI Duet:** Built by Yotam Mann and Google, this web tool utilizes generative piano music to let users play a duet with the computer [38]. Users press keys to play a music note, and AI Duet adds some notes to form a duet. The tool utilizes Tensorflow and Tone.js and has been trained on many MIDI examples and it learns about musical concepts, building a map of notes and timings.
- **Sketch RNN**: Built by Google Creative Lab, Sketch RNN is an interactive web experiment that lets you draw together with a recurrent neural network model [25]. The neural net has learned to draw by training it on millions of doodles collected from the Quick, Draw! Dataset [23]. Once the user starts drawing an object, Sketch-RNN will come up with many possible ways to continue drawing this object based on where they left off.
- **AI News Anchor**: Developed by Xinhua and the Chinese search engine, these AI-powered news anchors were developed through machine learning to simulate the voice, facial movements, and gestures of real-life broadcasters, to present "a lifelike image" of a human news anchor [32].
- **This Person Does Not Exist**: This tool utilizes StyleGan2 [29] and has been trained on human faces to generate fake human faces using GANs [53].

After exploration, they were asked the following questions:

- What do you think the generator in this GAN is trying to generate? What dataset is the discriminator basing its decisions on?
- How could this technology do the most good? How could this technology do the most harm?

5.4 Activity 4. Deepfakes

Students were already familiar with Deepfake-generated media from previous examples like AI News Anchor and This Person Does Not Exist. To make a connection with prior knowledge, we begin this lesson by redefining what Deepfakes are, by recollecting previous examples of Deepfakes that the students witnessed in Activity 3. Students are then presented a questionnaire with 10 videos (5 real and 5 Deepfakes) and are asked to identify the Deepfakes. The videos were taken from the public dataset released by the Kaggle Deepfake Detection Challenge [1]. This activity was followed by a classroom viewing of another series of video clips featuring stimulating examples of Deepfakes, during which students were encouraged to articulate why they thought a particular clip was a Deepfake, emphasizing what particular features of the video made it seem "less real" or "unconvincing". We showed a range of videos including Mona Lisa-styled Deepfakes and fake recordings of opinions voiced by political figures.

As a class, we then had an open discussion around how to spot Deepfakes on social media feeds or news sources. Having an open-ended conversation allowed us to present this topic as an ongoing discussion and a collaborative sharing of ideas - the structure with which we presented the material suggested to students that a list

for identifying GANs is not set in stone or comprehensive, and that their voices were just as important in forming the list.

As Deepfakes are synthetic media replacing an existing person with someone else's likeness, the cues that help distinguish Deepfakes from real media were often items that detracted from the naturalness of a person's face. We shared with students some tips for identifying Deepfakes such as blurry backgrounds or asymmetry in faces. Examples of other components that can help identify Deepfakes can be found in Figure 6. Given this list of 7 telltale signs of spotting Deepfakes, students were asked to complete a follow-up activity where they could put their new knowledge to the test; for a given video, students were asked to guess whether a video was real or fake and write down the reasoning behind their guess. In this way, we provided a framework with which they could think about the validity of media and gave them the opportunity to practice observing these components in real life examples.

Figure 6: Examples of how to recognize Deepfake photos.

This exercise both prepares them to spot synthetic media in the future as they continue to be exposed to Deepfakes online and helps them build general intuition around what GANs can change in a given image or video. By breaking down the flaws and gaps in GAN-produced media, students can gain a deeper understanding of why they should be questioning what they see on social media as well as what it is that they should look for when questioning that online content. Beyond developing strategies to recognize Deepfakes, an important learning goal of this lesson was also to understand that Deepfakes are very realistic and sometimes difficult to recognize even after knowing these strategies.

5.5 Activity 5. Create with generative models

Through the activities mentioned above, students experienced creating with many different types of generative models. Students were asked to create generative stories using the following two tools that leverage generative models:

- **Image Generation**: Students used previously introduced tools such as Sketch RNN (doodles completed by a recurrent neural network), GANimals (machine "breeded" animals created by GANs), This Person Does Not Exist (fake human faces using GANs), and pix2pix (translates doodles to pictures with style transfer). These images were used as machine-generated illustrations for the story.
- **Text Generation**: We introduced a Long Short Term Memory (LSTM) text-generation tool that outputs a string of text given human-inputted seed text and a selected training

corpus with a particular storytelling style. Students could choose from a variety of styles such as Dr. Seuss, Harry Potter, and Zora Neale Hurston. Upon selecting a style for the corpus, students could choose the seed text that started the generation of text, the length of the output text, and its temperature (randomness). Outputted sentences and phrases were used as machine-generated text for the story.

Students put together slides to showcase their stories, made up of a title slide and content slides that followed the trajectory of the story using GAN-created text and GAN-created illustrations to complement the text, much like a picture book. After writing their stories, students had the chance to reflect individually about the process of story creation and storytelling. In the reflections, we asked students to document observations on the generated text and evaluate the creativity demonstrated in their AI co-created stories. Some of the reflection questions include:

- What features did you notice from the writing style of text you chose?
- Could you have expected the ending you created with the text generator? If not, what was surprising about it?
- If you asked another person to create an alternate ending to your story intro, do you think their story would be more or less creative than the AI-generated story?

Students were also given the opportunity to share out their stories with the rest of the class, discussing the experience of storytelling as well as reading out the created narratives.

6 METHODS

6.1 Methods

This activity was piloted in four synchronous online summer workshops. All workshops were held virtually over Zoom, and the activities were made available to students on Google classroom. All courses were taught by a team of researchers and educators. The first three workshops were a part of a larger Introduction to AI curriculum, and the second 2 workshops were a part of a workshop just focused on GANs. Timing varied depending on the workshop, but most workshops met daily for 2-3 hours for anywhere between one and three weeks.

6.2 Participants

72 students (grade 5 – 9) participated in this IRB-approved study. The participants were spread across four different online summer programs for middle schoolers. A breakdown of student demographics can be found in Table 2. The workshops were led by two teachers, and were assisted by 2-3 teaching staff. All participants and their parents and teachers signed the assent and consent forms respectively to participate in these programs.

6.3 Assessment

Data were collected before and after the workshop as well as within each lesson.

Concept Inventory Concept Inventories are assessment instruments that assess students' basic conceptual knowledge of a discipline, and their ability to discern common misconceptions in that

	Workshop 1	Workshop 2	Workshop 3	Workshop 4
n	22	12	16	22
Gender	F=13, M=9	F=5, M=7	F=5, M=11	F=13, M=9
Grade	5th(1), 6th(5), 7th(5), 8th(8), 9th(3)	5th(5), 6th(7)	7th(10), 8th(3), 9th(1), 10th(1), 11th(1)	6th(5), 7th(2), 8th(12), 9th(3)

Table 2: A breakdown of student demographics across the four online workshops

discipline [14]. CIs play an important part in relating teaching techniques to student learning. In our workshop, students completed a pre-test questionnaire before the workshop, and a post-test questionnaire again after the workshop as part of a larger CI developed around AI literacy [34]. For validating the content of the CI, we conducted panel reviews with psychometric measurement experts and AI education practitioners, assessed for face and content validity through cognitive interviews with 6 middle school students to revise the survey, and completed a pilot study with 20 participants from diverse backgrounds to determine the revised instrument's reliability. Two questions from the concept inventory specifically focused on how GANs work. One question focused on the roles of the generator and discriminator. This question was made up of the following three sub-statements and students were asked to mark each one independently as "True" or "False".

- Q1.1 A generator and discriminator are both Neural Networks (*True*)
- Q1.2 The generator and discriminator are working in competition with one another (*True*)
- Q1.3 The discriminator gives feedback to the generator (*True*)

Each sub-statement answered correctly was given 1 point, for a possible score of up to 4 points. Scores pre and post were averaged and a Wilcoxon Signed-Rank test was completed for statistical significance. Each sub-statement was marked as correct or not and individual tested for significance with a McNemar's Test.

The second question tested students' ability to understand the back and forth nature of a GAN:

Q2. A GAN is being trained to generate images of clouds. The generator creates an image and sends it over to the discriminator. The discriminator does not classify the image as a cloud. What happens next?

- The GAN stops running
- The generator and the discriminator switch roles
- The generator generates a new image (Correct Answer)
- The discriminator generates a new image

Answers for this question were scored for correctness and a McNemar's Test was performed to test for significance.

Embedded Assessment In the "GANs or Not" activity, students responses were recorded on Google slides. Results were then aggregated and average scores were calculated.

During the "How do GANs Work?" activity, data was collected during gameplay. Each round was documented and conversations

were transcribed. During the "Exploring GANs" activity, students recorded their responses in a Google form. The responses were coded by researchers for accuracy and theme with a percent agreement of 82.15%.

For the Deepfakes activity, each test was scored for correctly identifying if a video was a deepfake or not, as well as the correct reasoning. These results were compared pre to post activity through a t-test.

7 RESULTS

7.1 How do students understand the technical concepts of generative machine learning?

Students were asked two assessment questions before and after the workshop. The first question contained three statements about GANs that students had to mark independently as "True" or "False". If they marked all correctly, they obtained 4 points. Forty-three students completed this question pre-post. Individual statements within this item were analyzed during the pre and post test, and are shown in Table 3.

Question Statement	Correct answer	# correct (pre-test)	# correct (post-test)	p (McNemar test) (n=43)
Q1.1	True	21	31	0.024
Q1.2	True	16	12	0.502
Q1.3	True	30	36	0.181

Table 3: A McNemar test showed there was a statistically significant gain between pre- and post- for Q1.1. while there was no statistically significant change in Q1.2, & 1.3

The second question gave students an option of four answers and students had to pick the one that most accurately described how a GAN works. Thirty-two students answered this question at the beginning and end of the workshop. More students got the question correct at the end of the workshop (83.87%) versus the beginning of the workshop (45.16%). A McNemar's test shows that this was a significant change, $\chi2(1) = 7.56$, p=0.006.

7.2 How do students understand the applications of generative machine learning?

In Activity 1, students determined whether or not a piece of media was created by a generative model. The answer to all questions was "Yes, the media was made by a generative tool." A full breakdown of student responses can be found in Figure 7. More students thought that the Style Transfer image, generated colors and generated digits (Figure 3) were created by a GAN, relative to those who did not think it was created by a GAN.

In the Exploring GANs activity, students tried out up to four different web tools that utilize generative machine learning. After they explored these tools, they were asked to identify 1) what the generator in the GAN was trying to generate, and 2) what dataset the discriminator in the GAN was basing its decisions on. Students
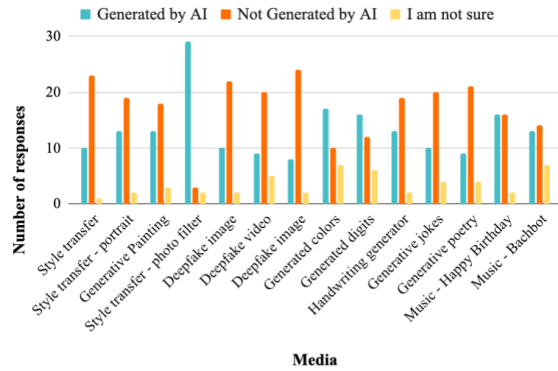
Figure 7: Results of the GANs or Not activity. All pieces of media were created by a GAN

were allowed to answer questions for one to four tools. We received 99 completed responses from 72 students. Overall, 88% of student responses were able to identify what the generator was trying to generate, and 60% of student responses were able to identify the dataset that the discriminator used.

Results were broken down further by media tool assessed. For the generator question, students were able to identify what "This Person Does Not Exist" and "AI News Anchor" were trying to generate with 100% accuracy, followed by Sketch-RNN with 87.2% accuracy, and AI Duet with 78.6% accuracy. Incorrect answers for Sketch-RNN were commonly associated with user input. For example, students thought that the generator was trying to copy the user input or trying to predict what the user would do next. A full breakdown of responses by tool can be found in Table 4.

Students were asked to identify whether something was a Deepfake or not, learned strategies for spotting Deepfakes, and then took the same Deepfake quiz. There was no significant change between the pre-quiz (M=51.25, sd=13.78) and post-quiz (M=54.89, sd=16.04), t(44)= -1.18, *p = 0.12*.

Moreover, we observed that students developed a new perception of AI as a potential artistic tool. In her post-workshop interview, when one participant was asked what she found surprising about the workshop, she said:

> That [AI] can make many things very similar to real life things, that they can make faces look very realistic.

7.3 How do students understand the ethics around generative machine learning?

Students reported the potential benefits and harms of the tools they interacted with using in descriptive answers in a form. Some students filled the form for multiple tools. We collected a total of 99 responses from 72 students. The responses were first coded for 1-2 word themes of benefits and harms independently by two coders. For instance, responses for potential benefits of Sketch-RNN such as "teach kids how to draw" and "It could help beginner drawers know how to draw" would both be coded as "learning". Some responses were coded with two themes, for example, the potential harm of "This Person Does Not Exist" "This technology could lead to identity theft and blaming incidents on a different person." was coded as

"Deception" and "Policing". We used the grounded theory of coding, where a coding scheme were formed from the data collected by two coders and validated using an inter-coder reliability (IR = 0.84) [54]. After the first round of coding, coders discussed dissimilarities in codes, created a new combined coding system, re-allotted codes using the new coding scheme. Students that responded that there were no potential benefits of harms were coded as "None" and responses that did not fall under any code categories were coded as "Other". The following codes were generated for the benefits and harms of GAN tools:

Students reported a total of more benefits (100) as compared to harms (85), however, this differed by the tools. Students were able to report harms and benefits of all the tools, and hence, we can infer that students could think about the potential uses and societal implications of these technologies.

We observed that, amongst the students who interacted with the tools, a higher percentage could identify the benefits of purely artistic tools such as Sketch RNN and AI Duet and a higher percentage could identify the harms of GAN tools that generated human faces or videos such as AI News anchor or "This Person Does Not Exist". Students associate potential harms such as deception or policing only with tools that generated anthropomorphic media. Further, we also observed that there was wider distribution in potential harms of the artistic tools, however for the tools generating faces and news anchors, there was a greater consensus over what the potential harm can be. For instance, for "This Person Does Not Exist", 88% students reported deception as a potential harm, however, for AI duet, potential harms were spread across 10 different themes. We also observed that for more interactive tools (Sketch-RNN and AI Duet), where the output changes based on user interactions (drawings or music notes), students identified several learning benefits, such as "Create" or "Helping artists" or "Learning" or "AI innovation". However, for the less interactive tools which simply displayed a pre-generated output and did not take user input, students identified other benefits such as "Automation" or "War Decoy" or "Crime". A full breakdown of benefits and harms is shown in Table 4.

In the "Spotting Deepfakes" Activity, many students mentioned how difficult it was to tell if something was a Deepfake or not, even when they became aware of common techniques to detect them:

> Today, the main thing I learned is how to see if a video or picture is a deepfake/made by AI. I also learned how AI can generate things like pictures and stories using data... I found the Deepfake exercise really hard, this is because AI videos can be very convincing.

Students explain that difficulties came from the fact that they were required to "pay attention" to "small details":

> The activity if we had to identify if it was a Deepfake or not was hard. It was very hard to tell if it was. There are so many miniature details that you have to look at, and that makes it really tricky.

Students recognized how similar the Deepfake videos appeared to the real videos:

> That same activity about determining if the videos were real or fake was also challenging as some of the videos had really small differences.

Generative Model Tool	Responses	Generator Correct%	Discriminator Correct%	Reasons Discriminator Incorrect	Potential Benefits	Potential Harms
This Person Does Not Exist	23	100	78.3	Don't know(4), Incorrect(1)	Helping Artist(4), Protecting Identity(3), War Decoy(1), Crime(2), Entertainment(3), Other(6), None(3)	Deception(15), Policing(3), Other(4), None(2)
AI News Anchor	9	100	55.6	No dataset(3), Confused with input data(2), Don't know(1)	Automation(5), AI Innovation(1), Accessibility(2), Unbiased(1), Preservation(0), None(1)	Deception(1), Remove Jobs(4), Lying(3), Other(1)
Sketch RNN	39	87.2	64.1	No dataset(9), Confused with input data(7), Don't know(5)	Create(7), Helping Artist(12), Automation(1), AI Innovation(1), Crime(3), Learn(10), Accessibility(2), Entertainment(1), Other(7), None(4)	Deception(1), Policing(1), Inappropriate(10), Collecting Data(2), Copyright(3), Bias(3), Malfunction(4), Confusion(3), Miscommunication(1), Other(3), None(13)
AI Duet	28	78.6	42.9	No dataset(9), Confused with input data(8), Don't know(6)	Create(6), Helping Artist(6), Learn(3), Accessibility(3), Entertainment(4), Preservation(1), Other(5), None(1)	Copyright(5), Malfunction(3), Remove Jobs(3), Human Creativity(4), Bad Music(4), Other(4), None(6)

Table 4: Students report the role of the generator, the discriminator, the potential benefits and harms of the generative models in the Exploring GANs activity

8 DISCUSSION

We designed a learning trajectory for generative models then revised it based on testing it with students. Activities were aligned to the LT and sequenced to build off of one another, Below we discuss each learning goal and what we evidenced about middle school students' learning and perspectives on generative models.

8.1 Learning Goals

How, and to what extent, do students understand the technical concepts of generative machine learning? We found that students were able to identify the iterative process that happens between the generator and the discriminator in a GAN. Knowledge of the relationship between the generator and the discriminator was demonstrated (Concept Inventory Q2). However, students did not think that the generator and discriminator were working in opposition to one another (Concept inventory Q1.3). This confusion may arise because in the generator-discriminator game, the two teams were giving each other clues to get to one final answer, which could be viewed as collaborative instead of competitive. Students did not understand the role of the discriminator as well as they understood the role of the generator. The misunderstanding of the discriminator's role was evidenced in students' use of interactive tools that required user input (Sketch-RNN and AI Duet). In these activities, user input was conflated with the discriminator's training dataset. In the Exploring GANs activity, one student incorrectly perceives the generator and discriminator as working together and conflates the user input to the discriminator's dataset:

> The discriminator was playing more notes after the generator's notes to hint at a possibly continuation of a song. It did it successfully, since the generator played some more notes that smoothly made a song. Therefore, the generator and discriminator made a song together.

However, we found that students had an easier time identifying the parts of the GAN for tools that were most similar to the game simulation, like "This Person Does Not Exist":

> The generator is trying to generate images that look like real people... The discriminator based its decisions on a dataset of real people and their faces.

How do students understand the applications of generative machine learning? At the beginning of the workshop, students had a limited understanding of what might be generated by AI. Students tended to view the more artistic applications as GANs and the more photo-realistic applications as not computer generated. After students learned the breadth of what GANs can create, how they work, and student had the opportunity to use various generative AI tools, we found that more students showed interest in creative applications of GANs that allowed them to input their own information into the system to create something new. In the activities in which students applied generative modeling tools to their own projects such as generating stories and illustrations, we saw that students did not feel that the GAN produced "realistic enough" text, specifically calling out the differences in style versus semantics.

> I noticed that the words were kind of random and seemed like you put a bunch of random selections of words together.

> It didn't make sense sometimes. It just tries to put words together that are "Harry Potter-y" to make a story.

Most students noted that the generated text was "random" or "did not make sense", whereas only a couple of students said that the generated text resembled the style of the text in the dataset.

How do students understand the ethics around generative machine learning? Finally, we were interested in understanding students' perspectives on how these generative AI tools may be used in the future. We asked students to provide potential benefits and potential harms for each of the four GANs tools they played with. Overall, students were able to identify more benefits than harms for tools that gave them greater agency (Sketch RNN and AI Duet), and more harms than benefits for tools that did not give them agency ("AI News Anchor" and "This Person Does Not Exist"). In the tools that provided more agency, students were able to select and input media then "collaborate" with the GAN to finish creating an artifact. The most common benefit students described for these tools were "helping artists create new things."

> It could be used to give ideas to artists who started drawing but have no idea on what to draw next.

The most common harms were around "inappropriate drawings" or "copyright infringement."

> This technology could do the most harm by drawing something that is not accepted in a certain culture.

The tools that provided less agency ("AI News Anchor" and "This Person Does Not Exist") showed fake humans that were completely generated by AI. Students perceived "This Person Does Not Exist" as a way to create something new, yet deceptive. "AI News Anchor" was seen as automation that took away jobs from current news anchors. From these findings, we conjecture that students can understand ethical issues in generative models and see the benefits in the creative aspects of GANs, but they are simultaneously concerned about generative AI's ability to deceive users and take away jobs.

We saw that students were able to categorize Deepfakes as a type of generative AI and were aware of the consequences they may have on the spread of misinformation. Though students were exposed to indicators of manipulated media and Deepfakes, they did not improve in their ability to identify Deepfakes. This is not surprising as the technology has become so sophisticated that even experts can be fooled.

8.2   Lessons Learned

While students could identify how the generator and discriminator work in opposition to generate synthetic media, students had trouble identifying the role of input datasets, and conflated them with use input in collaborative applications of generative models. Students seemed to lack the fundamental background about how a network gets trained through data. We hypothesize that students' understanding of GANs can be strengthened by their knowledge of Neural Networks, which could preface the generative models learning trajectory.

Students also had trouble with understanding the role of the Discriminator network in the generative applications they interacted with. In future work, we could benefit from making interactive applications of the technology more transparent - for instance, discussing the dataset each application uses, and uncover a progression of media generated by the generator that the discriminator rejects. Finally, we observed that having students reflect on the uses and ethical implications of their creations, while interacting with them, helped them bridge their technical, applied, and ethical understandings of generative AI.

8.3   Implications

We designed a generative models LT for middle school students, followed by learning activities that address these learning goals. The LT was successful in helping students understand the technical components of generative models, their applications and ethical implications, and we have no evidence to indicate that the ordering of learning goals was erroneous. However, results highlight that this LT misses out on some key learning goals associated with generative models.

8.3.1   Neural Networks. While discussing the technical components of GANs, we do not discuss neural networks beyond mentioning that they are a "kind of AI algorithm." While students understand that the role of the generator and the discriminator, we do not focus enough on the fact that they are both neural networks. In Activity 3, we found that students have difficulty in recognizing what dataset discriminators are trained on. Further, students have very little understanding on what mechanisms the discriminator uses in providing its feedback to the generator. Gaining an understanding of the feed-forward, evaluation, and backpropagation processes of a Neural Network will help students gain a better understanding of how the discriminator uses training dataset to classify the samples generated by the generator, and penalizes the generator for producing fake samples.

8.3.2   Dataset. While creating generative stories using the text generator tool in Activity 5, several students were confused about why the network was generating a certain kind of text. Students also have difficulty in recognizing what input datasets these models are trained on. While we focus on how the generator and discriminator work, we do not lay enough emphasis on the training datasets used to train these GANs. An essential learning goal, datasets should be a part of the learning trajectory while discussing the technical components of GANs.

8.3.3   Style Transfer. In our learning activities, we introduce the concept of GANs by simulating a simple GAN attempting to create a graphic after being given feedback (Activity 2). However, when we have students explore GANs, several applications such as sketch-RNN, or AI Duet, involve an essential component of *style transfer* involved. These applications learn the style of an input dataset and not only generate a synthetic instance, but also transfer this generated style on a target media. We learned that, for these applications, it was difficult for students to comprehend the role of the discriminator. Also, applications of GANs that students interact with in their daily lives such as photo filters, or Deepfakes, all make use of style transfer. Hence, we recommend that in the technical components learning goals, we must also add "How Style Transfer GANs work" as an addition to "How GANs work."

8.3.4   Transfer of Learning from Technical Components to Applications. In the Exploring GANs Activity, we asked students to apply what they learned about how GANs work to real generative AI tools. Students who explored "This Person Does Not Exist" were able to most accurately answer the questions about both the generator and discriminator. "This Person Does Not Exist" was the tool that most closely resembled the generator/discriminator game, by using a visual generation as well as no user input. The generative tool that was the least similar to the game, AI duet, had both audio generation and user input. This was the tool that students were least able to answer, often confusing the audio generation with the audio that they inputted. Future tools would expand from visual generation to other types of media and teach students about the role of user input in GANs.

We designed a modified LT (Figure 8) to begin with Neural Networks as a learning goal preceding generative models. While understanding the technical components of a GAN, we added Style Transfer models as an additional learning goal. In order to make clear the role of training datasets, when students learn about the applications of generative models, we lay an extra emphasis to expose
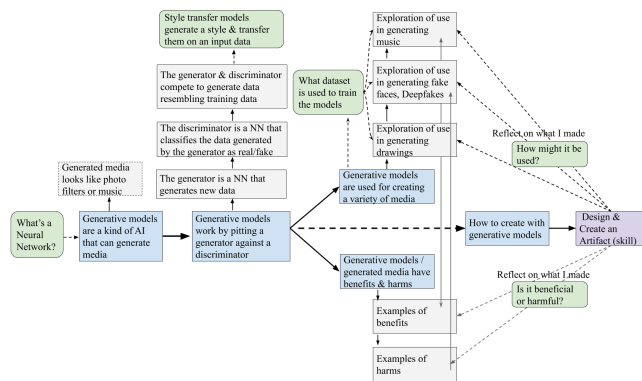
Figure 8: Modified generative models Learning Trajectory

.

to students that datasets were used to train these generative models. Finally, when students create with generative models, they are asked to reflect upon how their creations can be used, and identify the potential benefits and harms of their creation. This allows students to bridge their technical, applied, and ethical understandings of generative AI into practice.

9 CONCLUSION

In this work, we use a learning trajectory approach to teach a highly technical topic, generative models, to middle school students. We acknowledge that to teach students about generative models is to teach them about the entire socio-technical system: a combination of technical, applied, and ethical concepts. We integrate these three different domains into our learning trajectory, guiding students along a path of learning technical concepts, understanding the applications of them, and then understanding how these applications are integrated into society in good and bad ways. One limitation of our assessment methods is that the True/False questions in the assessment (Section 7.3) have *True* as their correct response. In future work, we should balance this out with some statements that are *False.*

The main contribution of this work is that through a case study of generative modeling concepts, we demonstrate how we adapted an HLT to a socio-technical system. Our LT ties in the technical concepts, practical applications and ethical implications of generative modeling. Results indicate learning gains across all categories. While existing works outline learning trajectories for technical systems, this is the first work, to our knowledge, that tackles teaching a socio-technical system to young learners. While advances in technology have been leading to serious social and ethical implications, this approach can be adopted by other educators that aim to teach concepts pertaining to other socio-technical systems. As technology, such as generative models, becomes more complex, its applications more accessible, and its ethical implications more widespread, we hope that future AI learning designers can benefit from the same integrated approach.

REFERENCES

[1] [n.d.]. Deepfake Detection Challenge. https://www.kaggle.com/c/deepfake-detection-challenge

[2] 2020. Young Teens (12-14 years of age). *Center for Disease Control and Prevention* (2020). https://www.cdc.gov/ncbddd/childdevelopment/positiveparenting/adolescence.html

[3] Safinah Ali, Hae Won Park, and Cynthia Breazeal. 2020. A social robot's influence on children's figural creativity during gameplay. *International Journal of Child-Computer Interaction* (2020), 100234.

[4] Safinah Ali, Blakeley H Payne, Randi Williams, Hae Won Park, and Cynthia Breazeal. [n.d.]. Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education.

[5] Lorin W Anderson, Benjamin Samuel Bloom, et al. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Longman,.

[6] Monica Anderson, Jingjing Jiang, et al. 2018. Teens, social media & technology 2018. *Pew Research Center* 31 (2018), 2018.

[7] Michael T Battista. 2011. Conceptualizations and issues related to learning progressions, learning trajectories, and levels of sophistication. *The Mathematics Enthusiast* 8, 3 (2011), 507–570.

[8] Tomaz Ribeiro Viana Bisneto, Antonio Oseas de Carvalho Filho, and Deborah Maria Vieira Magalhães. 2020. Generative adversarial network and texture features applied to automatic glaucoma detection. *Applied Soft Computing* 90 (2020), 106165.

[9] Benjamin Samuel Bloom. 1956. Taxonomy of educational objectives: The classification of educational goals. *Cognitive domain* (1956).

[10] 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.

[11] Jere Confrey, Alan P Maloney, and Andrew K Corley. 2014. Learning trajectories: A framework for connecting standards with curriculum. *ZDM* 46, 5 (2014), 719–733.

[12] Susanne A Denham, TM Wyatt, Hideko Hamada Bassett, D Echeverria, and SS Knox. 2009. Assessing social-emotional development in children from a longitudinal perspective. *Journal of Epidemiology & Community Health* 63, Suppl 1 (2009), i37–i52.

[13] Carl Doersch. 2016. Tutorial on variational autoencoders. *arXiv preprint arXiv:1606.05908* (2016).

[14] DL Evans, Gary L Gray, Stephen Krause, Jay Martin, Clark Midkiff, Branislav M Notaros, Michael Pavelich, David Rancour, Teri Reed-Rhoads, Paul Steif, et al. 2003. Progress on concept inventory assessment tools. In *33rd Annual Frontiers in Education, 2003. FIE 2003.,* Vol. 1. IEEE, T4G–1.

[15] FaceApp [n.d.].

[16] FaceTune [n.d.]. FaceTune. https://www.facetuneapp.com/

[17] Diana Franklin, Jen Palmer, Woorin Jang, Elizabeth M Lehman, Jasmine Marckwordt, Randall I Landsberg, Alexandria Muller, and Danielle Harlow. 2020. Exploring Quantum Reversibility with Young Learners. In *Proceedings of the 2020 ACM Conference on International Computing Education Research.* 147–157.

[18] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2015. A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576* (2015).

[19] James Paul Gee. 2003. What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE)* 1, 1 (2003), 20–20.

[20] James Paul Gee. 2009. Video games, learning, and "content". In *Games: Purpose and potential in education.* Springer, 43–53.

[21] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems.* 2672–2680.

[22] Google. 2016. Artists and Machine Learning. https://ami.withgoogle.com/.

[23] Google. 2017. Quick, Draw! https://quickdraw.withgoogle.com/.

[24] Koeno P Gravemeijer. 1994. *Developing realistic Mathematics Education (Ontwikkelen van realistisch reken/wiskundeonderwijs).*

[25] David Ha and Douglas Eck. 2017. A neural representation of sketch drawings. *arXiv preprint arXiv:1704.03477* (2017).

[26] Ken Kahn and Niall Winters. 2017. Child-friendly programming interfaces to AI cloud services. In *European Conference on Technology Enhanced Learning*. Springer, 566–570.

[27] Neil K. Kaneshiro. 2018. Adolescent development. *MedlinePlus* (2018). https://medlineplus.gov/ency/article/002003.htm

[28] Neil K. Kaneshiro. 2018. School-age children development. *MedlinePlus* (2018). https://medlineplus.gov/ency/article/002017.htm

[29] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2019. Analyzing and Improving the Image Quality of StyleGAN. arXiv:1912.04958 [cs.CV]

[30] Salome Kazeminia, Christoph Baur, Arjan Kuijper, Bram van Ginneken, Nassir Navab, Shadi Albarqouni, and Anirban Mukhopadhyay. 2020. GANs for medical image analysis. *Artificial Intelligence in Medicine* (2020), 101938.

[31] Eric Klopfer, Scot Osterweil, Katie Salen, et al. 2009. Moving learning games forward. *Cambridge, MA: The Education Arcade* (2009).

[32] Lily Kuo. 2018. World's first AI news anchor unveiled in China. *The Guardian* 8 (2018).

[33] MIT Media Lab. 2019. Computer Visions: Generative Machine Learning Tools for Creative Applications. http://visions.media.mit.edu/.

[34] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing Middle School Students' AI Literacy. *52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21)* (March 2021). "In press".

[35] Natasha Lomas. 2020. Deepfake video app Reface is just getting started on shapeshifting selfie culture. https://techcrunch.com/2020/08/17/deepfake-video-app-reface-is-just-getting-started-on-shapeshifting-selfie-culture/?guccounter=1&guce_eferrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS88&guce_eferrer_sig=AQAAAN1tfqNlH9AaIPEYrK6121tJ7pQV1QMu3prDqVEsaI8DrS67JBiyRNUk09CUyWaq1hlqQj-XXlvO7gXrwlSMj8H0ACpyfUtJ9kOVsaAK9WeXG4Gpt2d2_Sb9I7o5XfYSStx0CdaVKdcbol0rld_wrTtLom85Eq9M-a6KifbsoE-R

[36] 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*. 1–16.

[37] Machine Learning for Artists 2016. Machine Learning for Artists. https://ml4a.github.io/classes/.

[38] Yotam Mann. 2016. AI duet. *Experiments with Google. See, https://experiments.withgoogle.com/ai/ai-duet* (2016).

[39] Lívia S Marques, Christiane Gresse von Wangenheim, and Jean CR HAUCK. 2020. Teaching Machine Learning in School: A Systematic Mapping of the State of the Art. *Informatics in Education* 19, 2 (2020), 283–321.

[40] Thanh Thi Nguyen, Cuong M Nguyen, Dung Tien Nguyen, Duc Thanh Nguyen, and Saeid Nahavandi. 2019. Deep learning for deepfakes creation and detection. *arXiv preprint arXiv:1909.11573* (2019).

[41] Aaron van den Oord, Nal Kalchbrenner, and Koray Kavukcuoglu. 2016. Pixel recurrent neural networks. *arXiv preprint arXiv:1601.06759* (2016).

[42] B.H. Payne. 2020. *Can my algorithm be my opinion?: An AI + Ethics Curriculum for Middle School Students*. Master's thesis. Massachusetts Institute of Technology, Media Lab, Cambridge, MA, USA.

[43] Prisma [n.d.]. Prisma. prisma-ai.com

[44] MIT Architecture Renaud Danhaive. 2019. Creative Machine Learning for Design. https://architecture.mit.edu/subject/spring-2020-4453.

[45] Victoria J Rideout and Michael B Robb. 2019. *The common sense census: Media use by tweens and teens*. Common Sense Media.

[46] Julie Sarama, Douglas H Clements, Jeffrey Barrett, Douglas W Van Dine, and Jennifer S McDonel. 2011. Evaluation of a learning trajectory for length in the early years. *ZDM* 43, 5 (2011), 667.

[47] Kevin Schawinski, Ce Zhang, Hantian Zhang, Lucas Fowler, and Gokula Krishnan Santhanam. 2017. Generative adversarial networks recover features in astrophysical images of galaxies beyond the deconvolution limit. *Monthly Notices of the Royal Astronomical Society: Letters* 467, 1 (2017), L110–L114.

[48] Martin A Simon. 1995. Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for research in mathematics education* (1995), 114–145.

[49] 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.

[50] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science* 12, 2 (1988), 257–285.

[51] Kara Swisher. 2019. Does Russia Want More Than Your Old Face? https://www.nytimes.com/2019/07/19/opinion/faceapp-privacy-russian.html

[52] Danny Tang et al. 2019. *Empowering novices to understand and use machine learning with personalized image classification models, intuitive analysis tools, and MIT App Inventor*. Ph.D. Dissertation. Massachusetts Institute of Technology.

[53] This Person Does Not Exist 2019. This Person Does Not Exist. https://www.thispersondoesnotexist.com/.

[54] Robert Thornberg and Kathy Charmaz. 2014. Grounded theory and theoretical coding. *The SAGE handbook of qualitative data analysis* 5 (2014), 153–69.

[55] TikTok [n.d.]. TikTok. https://tiktok.com

[56] 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.

[57] 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.

[58] Langdon Winner. 1980. Do artifacts have politics? *Daedalus* (1980), 121–136.

[59] Zheng Xu, Michael Wilber, Chen Fang, Aaron Hertzmann, and Hailin Jin. 2018. Learning from multi-domain artistic images for arbitrary style transfer. *arXiv preprint arXiv:1805.09987* (2018).