

# Designing with and for Youth: A Participatory Design Research Approach for Critical Machine Learning Education

**Golnaz Arastoopour Irgens\*, Ibrahim Adisa, Cinamon Bailey and Hazel Vega Quesada**

Clemson University, U.S.A // garasto@clemson.edu // iadisa@clemson.edu // cinamob@clemson.edu // hvegaqu@clemson.edu

\*Corresponding author

**ABSTRACT:** As big data algorithm usage becomes more ubiquitous, it will become critical for all young people, particularly those from historically marginalized populations, to have a deep understanding of data science that empowers them to enact change in their local communities and globally. In this study, we explore the concept of critical machine learning: integrating machine learning knowledge content with social, ethical, and political effects of algorithms. We modified an intergenerational participatory design approach known as cooperative inquiry to co-design a critical machine learning educational program with and for youth ages 9 - 13 in two after-school centers in the southern United States. Analyzing data from cognitive interviews, observations, and learner artifacts, we describe the roles of children and researchers as meta-design partners. Our findings suggest that cooperative inquiry and meta-design are suitable frameworks for designing critical machine learning educational environments that reflect children's interests and values. This approach may increase youth engagement around the social, ethical, and political implications of large-scale machine learning algorithm deployment.

**Keywords:** Critical data science, Machine learning, Algorithmic bias, Participatory design research, Community youth program

## 1. Introduction

The world is becoming increasingly reliant on digital technologies to navigate our lives. Such technologies often collect, store, and analyze data to improve efficiency and quality of life. Data is increasingly being utilized in the development of machine learning (ML) and artificial intelligence (AI) systems. These systems are often used to make critical decisions about people and communities. For example, ML algorithms are now used to predict cancer patient outcomes, measure natural disaster damage using social media postings, and quantify traffic dynamics to reduce air pollution (Data Science for Social Good, 2021). However, such reliance on ML algorithms can also be problematic. For example, widely used internet search algorithms that distribute knowledge to billions of people daily have been shown to reinforce racial and gender discrimination at large scales (Noble, 2018), and government software has been used to extrapolate data that preemptively criminalize the actions of low-income populations (Eubanks, 2018). In addition to having harmful effects on marginalized populations, the design and operation of ML algorithms are invisible to the public (O'Neil, 2016), and many young people are unaware of how such algorithms collect and use their personal data (Pangrazio & Selwyn, 2019). Thus, it is essential to engage youth early on in their education about ML concepts and how to think critically about the social, ethical, and political issues around modern large-scale algorithm deployment that can (re)enforce inequities and harm marginalized populations.

In this study, we explore how youth ages 9 - 13 and researchers co-designed a critical machine learning program implemented in their after-school program, discussing the triumphs and tensions that emerged from this process.

## 2. Background and theory

### 2.1. Critical machine learning education

Artificial intelligence is a broad field about how computers use data, symbolic rules, and numeric models to analyze their environment and behave in ways that generally require human intelligence (Boucher, 2020). One of the ways humans implement AI is through machine learning (ML), which employs algorithms and processes to allow computers to solve problems and make intelligent human decisions. Humans build ML models by training computers to understand and recognize patterns from large datasets. These models are, in turn, used to make predictions and automated decisions about people and systems (Samoili et al., 2021). ML applications rely heavily on data and are employed to make automated recommendations and decisions that, in some cases, can be harmful to those from non-dominant populations such as women, people of color, and those living in poverty.

For example, Buolamwini and Gebru (2018) evaluated three commercial image classification systems used for facial recognition technology. The study was spurred by Buolamwini's personal experiences of being misidentified when using facial recognition software.

The researchers found that darker-skinned females were the most misclassified group, with error rates up to 34%, while the maximum error rate for lighter-skinned males was 0.8%. These error rates become particularly concerning when facial recognition systems are being used by U.S. law enforcement and other government agencies to detect unlawful behaviors. Buolamwini and Gebru's (2018) study suggests that those who reside on the margins of society continue to be marginalized with biased datasets used for training ML algorithms. Thus, for youth to assess such technologies and advocate for themselves when models make unfair decisions against them, they need to understand the basics around how ML works and be able to reflect upon the risks and consequences. In our prior work, we have proposed a critical machine learning (CML) educational approach (Arastoopour Irgens et al., 2022) that integrates vital pedagogy (Freire, 1970; Giroux, 1985) into computer science and machine learning education. In this integrated approach, learners ask questions such as, who develops these technologies? What types of data are used to train machines? What is the history behind the data used? What decisions are made based on the outputs of the algorithms? These questions facilitate reflections, discussions, and actions around disrupting oppressive paradigms related to the deployment of ML-based technologies.

In recent years, researchers have explored approaches for engaging youth in learning about the social, ethical, and political implications of ML and AI. Williams and colleagues (2019) discovered that young learners' understanding of AI concepts was based on the degree to which they were able to actively participate in hands-on activities such as modifying a robot's input, teaching a robot how to play a game, or training it to tell the difference between objects. Studies with middle-school-aged children have shown their abilities to engage with ethical AI concepts and identify the societal impacts of racist and biased algorithms (Ali et al., 2019). Classroom intervention research suggests that upper elementary school-aged students can consider oppression from multiple perspectives, including the broader historical framework of how society is organized and how to create change (Fain, 2008), analyze and interrogate literature around societal issues such as immigration (Braden, 2019), and address and challenge social inequities in their curricula (Kersten, 2006). To prevent children from internalizing harmful messages about oppression, it is essential to proactively engage them in reflections early (Boutte & Muller, 2018), especially those who belong to groups that experience discrimination (Ayón, 2016).

Although there is evidence that middle school-aged children are capable of reflecting on how AI technologies can be unfair to themselves and others, discomfort and tensions may arise when discussing issues of AI oppression with youth. For example, Lee and colleagues (2021) worked with community partners and framed their activities with sensitivity when working with students from populations that are underrepresented in STEM and computing. Because many of the activities directly addressed discrimination against students from these groups, the researchers framed the discussion on how the AI community can solve such problems rather than pressuring the students themselves to solve AI bias problems. The researchers also noted a need to include more positive examples of AI technologies in healthcare, education, and art rather than exclusively discussing negative implications.

This empirical work demonstrates initial progress on critical ML education. It suggests that learners are likely to develop essential skills and an understanding of ML when engaged in hands-on activities that are personal, agentic, and sensitive to their identities. However, there is still a lack of consensus on how to design environments to support such learning (Wolff et al., 2019) that is meaningful for learners and their communities.

## **2.2. Participatory design research**

To engage youth in *culturally sustaining pedagogies* (Paris, 2012; Paris, 2021), in which we maintain the culture of their communities while simultaneously providing access to dominant ways of thinking around critical ML, it is necessary to include youth in the design of their learning. Traditionally, children do not play a significant role in the creation of their learning experiences. However, not including children's input in the design of their educational experiences leads to the possibility of a mismatch between designers' intentions and learners' interpretations. When youth are both users and designers, they engage in mutual learning and co-construction of knowledge with instructors, mentors, and other stakeholders (Robertson & Simonsen, 2013). Moreover, incorporating children's cultural values may increase the chances of sustained engagement and learning compared to only designing for fleeting youth interests (DiSalvo & DesPortes, 2017).

Druin (1999) argues that designing with and for children requires reimagining how to facilitate design processes to include children's voices effectively. Based on empirical work, Druin proposes *Cooperative Inquiry*, comprising three techniques: contextual inquiry, technology immersion, and participatory design. In contextual inquiry, adults and children collect data about how children and adults interact in a selected environment. These notes and data inform the creation of the technologies and programs. During technology immersion, children "tinker" with multiple, novel technologies in their own environment. Such technology-rich, time-intensive experiences allow researchers to observe multiple patterns of children's activity. The participatory design component involves adults and children creating low-tech prototypes of designs using materials such as sticky notes, clay, string, paper, or markers. One central assumption in participatory design and cooperative inquiry is that there are multiple and valued forms of expertise stakeholders bring to the design process. For example, in intergenerational design, children are experts in what it means to be a child today (Guha et al., 2013). Children can use their imaginations to propose creative design ideas that may inspire adults. Many of the child's ideas cannot be realized in the actual design process, but adult partners can help reformulate the ideas so that they are workable with existing technologies. Interactions of this form welcome children's ideas and give them agency in the design process.

However, one challenge that adult designers face is that no matter how much preparation goes into the design of open-ended technology-based learning activities, it is unclear how children will appropriate the technology. As Ehn (2008) puts it, a researcher's "envisioned use is hardly the same as actual use, no matter how much participation there has been in the design process" (p. 95). One approach to address this challenge is through meta-design, in which the tool is designed before users engage with it. Still, the design allows for flexibility such that users can act as co-designers and customize, extend, or redesign aspects of the tool (Fischer, 2021). Meta-design embraces a co-adaptive process between users and a system and provides opportunities, tools, and social reward structures to refine systems to fit users' needs. Through the lens of meta-design, adults and children do not necessarily design concurrently. Adults create a tool or learning activity that offers opportunities for creative production and modification of tools and procedures. In a learning context where adult designers are also educators, both educator and child learner play the role of meta-designer. This means all participants fluctuate between the roles of learners, designers, and contributors (Fischer et al., 2004). The activities are not finished products, and learners are informed participants who have the power to shift the learning goals and methods. The meta-design approach aligns well with interactive networked digital technologies, such as Scratch (<https://scratch.mit.edu/>) and ML-specific authoring tools, such as Google's Teachable Machine, that allow children to be consumers and producers of media (Jenkins, 2006). In particular, the sensitive and reflective nature of exploring ML tools through a critical lens with children also lends itself to the reflective, asynchronous design and implementation that occurs in meta-design. However, few studies have explored the roles of children and adults as meta-designers as they engage with ML tools through a critical lens, informing the design of learning environments.

In a previous study, we suggested that the youth who participated in this co-designed CML program made more sophisticated connections with socio-political orientations and ML content as they progressed through the program. They engaged in computational practices, such as experimenting and iterating, testing and debugging, reusing and remixing, and abstracting and modularizing (Arastoopour Irgens et al., 2022). In this current study, we focus on the program design that facilitated such learning. We rely on cooperative inquiry techniques and meta-design approaches to explore how children and adults partnered as meta-designers in a CML educational program co-design.

The research question in this study is: *How did children and adults engage as meta-design partners during the CML program?*

### 3. Methods

#### 3.1. Context, participants, and researcher positionalities

We implemented the project in after-school programs at two centers: Green Community Center and Sunshine Community Center (the names of the centers and children are pseudonyms). Both centers serve elementary schools in a Southern U.S. County that has a mix of urban and rural areas, a poverty rate of 13.4%, a household median income of \$56,609, and a population that is 67% White (non-Hispanic), 23% Black, 5% Hispanic, and 2% Asian. Participants included 44 youth ages 9 – 13, 3 staff counselors, and 4 researchers. Each youth participant used their school-assigned Chromebook. The youth population consisted of Black, Latino, and White children, with a mix of those who presented as girls and boys. Youth attendance was variable. The researchers

were university faculty and graduate students and consisted of a White/Middle Eastern woman, a White woman from the local region, a Nigerian Black man, and Costa Rican Latina woman. All researchers are actively opposed to big-data algorithms that convey and perpetuate historical and current racism and sexism at large scales. The lead author and director of the research project is a former computer science and mathematics instructor whose perspective has influenced the design of the current program.

### 3.2. Design of the CML educational program

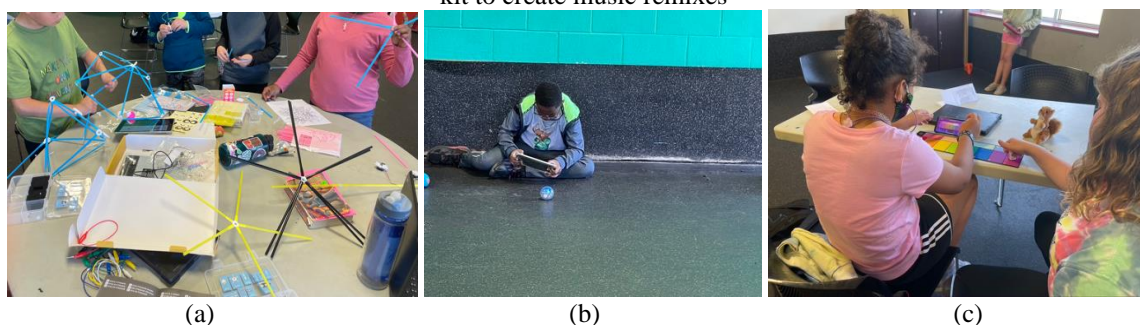
The program spanned 15 days, was implemented 2-3 days a week and occurred in three phases: *Initial Exploration, Activities and Discovery*, and *Youth Design of Machines*.

#### 3.2.1. Phase I: Initial exploration

During this phase, the objectives included: deciding where and how to implement the program initially; building relationships with youth and staff; determining youth values and interests; analyzing the program and adjusting as needed; and determining youth baseline knowledge about algorithms and ML.

Before any CML activities were fully designed and implemented, researchers volunteered at the after-school program. We planned to spend four days in this role observing children and staff in their day-to-day activities at the center, assisting children with their homework, and engaging in casual conversation to build relationships and get to know one another. According to our observations, the typical schedule of activities at the centers was staggered arrivals from different schools (10 minutes), homework time (45 minutes), and playing outside (until a parent arrived). At Green Community Center, two staff members sat at the front of the room watching children and occasionally shouting if they violated the rule of having three at a table. This rule was implemented because of the COVID-19 pandemic restrictions in the spring of 2021. Because of the global health crisis at the time, all students and adults wore masks during the implementations, and the after-school center shut down for one week in the middle of our schedule after a COVID-19 outbreak.

*Figure 1.* (a) children creating figures using Strawbees straws and connectors with optional robotics, (b) a child programming and moving a Sphero robot ball using a tablet, (c) children using the Specdrums app and physical kit to create music remixes



After we learned that most of the children did not have homework, we brought robotics toys during the next three volunteering as an alternative way to engage with the children and observe them with technologies. All children engaged with at least one of the toys. They built robotic sculptures with the Strawbees straws (<https://strawbees.com/>), created musical remixes using the Specdrums kit ([https://sphero.com/collections/all/family\\_specdrums](https://sphero.com/collections/all/family_specdrums)), and organized racing and bowling competitions with the Sphero (<https://sphero.com/>) robot balls (Figure 1).

#### 3.2.2. Phase II: Activities and discovery

During the second phase, the objectives included: continuing to grow relationships with youth and staff; continuing to discover youth and staff values and interests and implementing them into the activities; developing and supporting mutual learning through group activities and discussions; providing activities and tools that could assist youth in knowledge construction and flexible design.

Figure 2. (a) Example of one team’s pizza algorithm visualization, (b) example of another team’s everyday technology examples, and (c) example of another team’s helpful and harmful technologies list

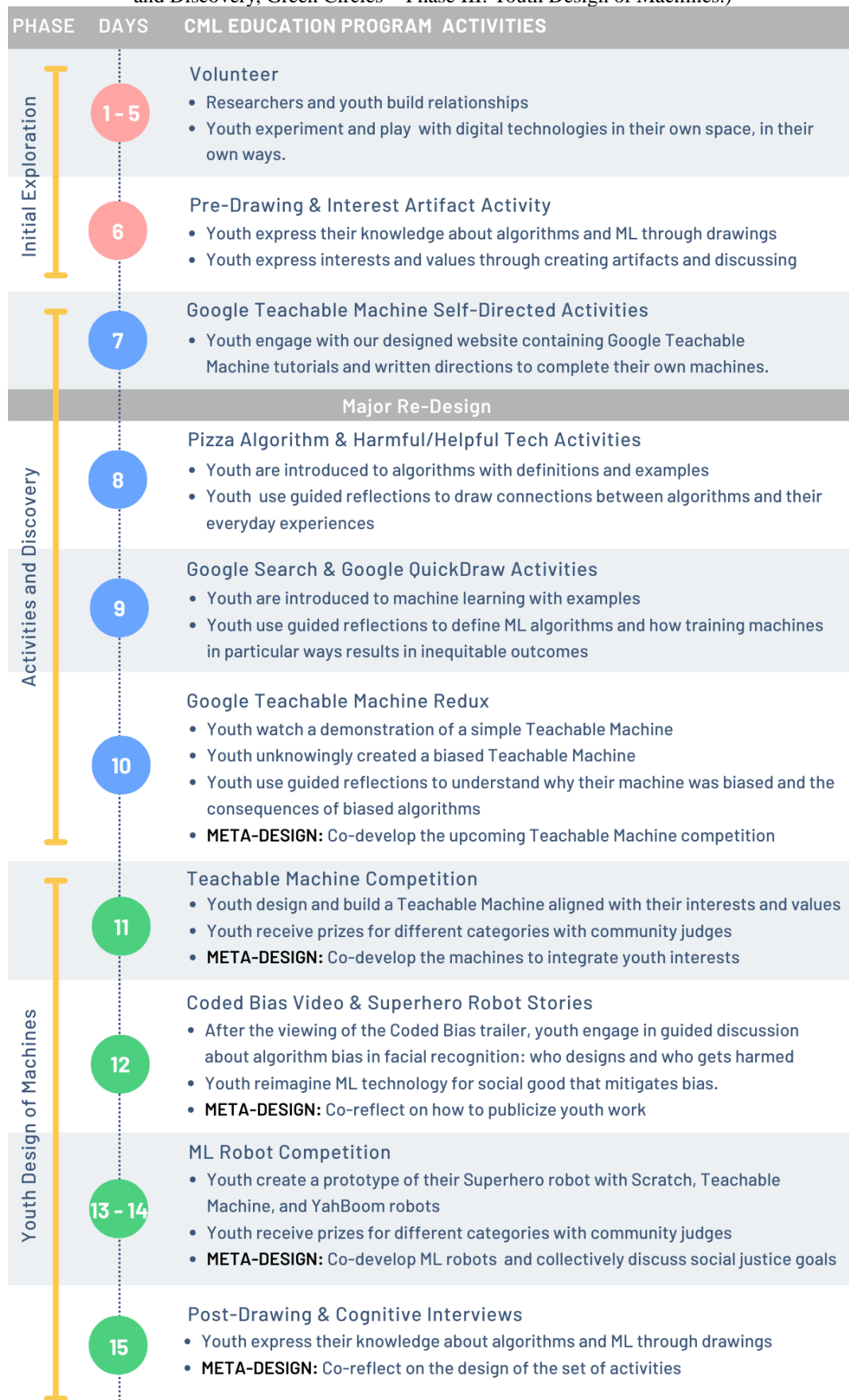


In the next activity, the youth used Google’s Teachable Machine (<https://teachablemachine.withgoogle.com/>) to classify images of cats versus dogs. Youth were told they were creating “a facial recognition software to determine whether a pet is a cat or dog” and given two envelopes containing printed images of cats and dogs. One envelope was labeled “training data,” and the other was labeled “test data.” Using the training dataset, youth used their Chromebook webcam to train the Teachable Machine to differentiate images of cats and dogs. However, the training dataset contained a more extensive variety and higher quantity of cats; thus, the resulting trained machine misclassified dogs more often than cats. Through this experience, youth engaged directly with training data, test data, and bias in training sets.

During this next phase, objectives included: providing youth with the tools they needed to design their own teachable machine; guiding youth in understanding algorithm bias; supporting youth as they designed and built (1) their own teachable machines and (2) trained their own robots; and supporting youth choice in their two designed products.

130

**Figure 3. Summary of CML education program activities and meta-design processes**  
 (Note. Numbers represent the days. Red Circles = Phase I: Initial Exploration, Blue Circles = Phase II: Activities and Discovery, Green Circles = Phase III: Youth Design of Machines.)

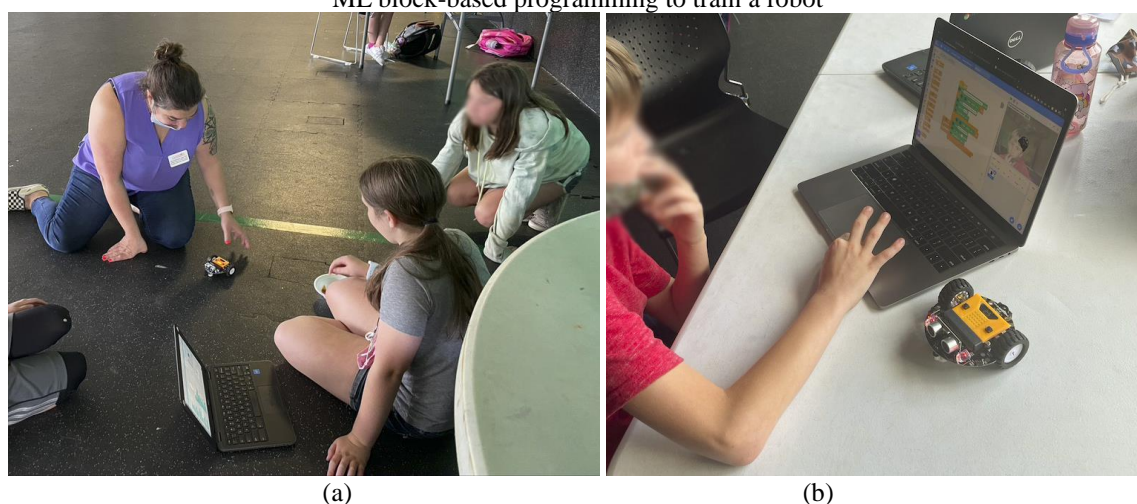




Before their next designed product, youth watched the *Coded Bias* film trailer (Kantayya, 2020), which featured Joy Buolamwini's realization of racist facial recognition technologies. Afterward, youth, staff, and researchers discussed racial discrimination embedded in ML technologies. After thinking critically about harmful technologies, youth shifted perspectives and were asked to create imaginative stories about helpful robots.

For their final design, the youth experimented with Yahboom! Micro:bit Robots using block-based programming software from MIT's How to Train your Robot Curriculum. They were asked to use ML algorithms to train a robot to be helpful to society (Figure 4). The ML concepts embedded in the block-based language included incorporating a Google Teachable Machine and voice/text classification. Before designing and programming their robots, we provided them with a demonstration of a trained image classifier using the laptop's webcam that detected whether a human face was present or not. If the webcam detected a human face, the robot would print a smiley face on its display, shine magenta headlights, and spin around in circles. If the webcam did not detect a human face, the robot would print a sad face on its display, turn off the headlights, and move backward. Youth presented their robots to their peers, staff, and directors of the program and could earn prizes. Figure 3 summarizes the CML education activities and co-design processes by day.

Figure 4. (a) One researcher and two children working together to train and test a robot and (b) one child using ML block-based programming to train a robot



### 3.3. Data collection and analysis

During the implementation sessions, at least one researcher in the session kept a reflective research journal and took observational field notes. As Glesne (2014) suggested, the reflective research journal documented the researcher's reflections on tensions, successes, and other feelings as we interacted with the children and staff. In parallel, the observational notes documented the researcher's evidence from observations and corresponding analytical notes. The researchers interviewed participants who submitted signed consent forms from their parents. We followed a mixed cognitive clinical (Russ et al., 2012) and semi-structured interview protocol (Appendix A). In the first section of the interview, the researcher displayed the learner's teachable machines and asked learners to reflect on their work. In the last half of the interview, the children were asked about their thoughts and feelings toward their experiences in the program and what they would add or change in future iterations. We collected artifacts from learners, including their completed teachable machines, superhero robot block-based code, and photos of giant sticky note activities (pizza algorithm, types of technologies, harmful/helpful technologies, superhero robot stories). We also took pictures and videos during their work time.

These data were analyzed using a descriptive phenomenology methodology (Creswell & Poth, 2016), which describes the lived experiences of a group of people. We used thematic analysis (Braun & Clarke, 2006) to describe the meta-design phenomena in this specific CML education context. Specifically, the analytic phases included familiarizing with the data, generating codes, constructing themes, reviewing/reflecting on themes, defining themes, and producing the report (Terry et al., 2017). The findings describe what the children and adults experienced when engaged in the CML program and how they experienced it through the lens of meta-design.

## 4. Results

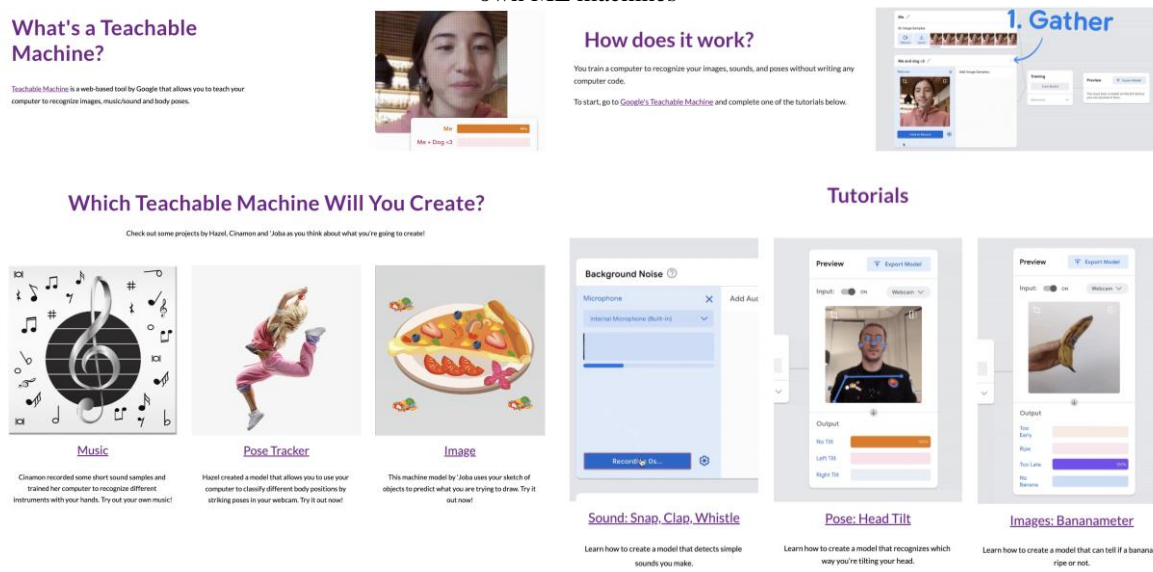
Through a meta-design conceptual lens, we analyzed the observational field notes, cognitive interviews, and youth artifacts. We identified 3 themes around adults and youth as meta-designers of a CML education program and what conditions facilitated co-design.

### 4.1. Researchers and youth persisting through failures and redesigning in real time

Following participatory design methods, we first volunteered at the after-school center to build relationships with the youth and understand how they interacted in their space. According to our observational notes, youth spent their time playing computer games, completing crossword puzzles, or conversing in small groups. When the weather was suitable, staff escorted the youth to play outside. On the second day at Green Community Center, the youth invited researchers to a fort they were building in a nearby wooded area. Youth included researchers in their imaginative play and used branches to continue building the fort.

Based on our observations and discussions with the youth, we designed self-directed and creative activities. We developed a website that linked to Google Teachable Machine, linked to tutorials, and linked to examples created by the research team (Figure 5). We added a discussion board to the site, which asked the youth to post links to their machines and discuss. On day 7 (see Figure 3), we implemented this activity at Green Community Center by taping printouts of a QR code on tables that led youth to the website. We planned to give youth access to the website and have them self-direct their experience, similar to what one would experience in other informal learning settings such as museums.

Figure 5. Two pages from our initial website for youth to explore Google Teachable Machine and design their own ML machines



Unfortunately, as one researcher put it afterward, “It was a complete failure.” According to researcher notes from day 7, “They [the youth] said they did not know anything about algorithms and machine learning... Once on the website, they just scrolled down without reading and didn’t know what to do. A few of them [8 out of 20 youth] got to try Teachable Machine.” We also observed that most youth who experimented with the Teachable Machine quickly abandoned the tool if they did not receive guidance, feedback, or intervention from an adult. However, one group of three girls sustained engagement with the device and used their toy figurines and webcam to create an image classification system that distinguished between their toys. They tested and modified their machine throughout the afternoon to improve its classification capabilities.

The day after this implementation, the researchers decided to hold an emergency design meeting to redesign tools and activities for the youth that provided less information at one time, less text, more discussion, and direct interaction with adults. Based on our observations and discussions with the group of three girls who successfully created a machine and seemed to enjoy the process, we decided to directly connect to and actively support students’ everyday interests with ML. During this significant re-design between days 7 and 8 (see Figure 3), researchers created an outline of activities that gradually introduced youth to algorithms, ML, training datasets,



test datasets, classification systems, bias in datasets, and harmful technologies. We adapted activities from MIT's How to Train Your Robot curriculum to better fit with our population of children. For example, one researcher noted in their observations and interactions with youth that "these kids like to have objects and things they can touch and manipulate." As a result, we printed out pictures of cats and dogs for the biased Google Teachable Machine activity. This way, children could hold physical photos to their webcams to train their machines rather than download and upload digital images.

Researchers also scheduled reflection discussions with the youth at the start and end of the sessions. These reflection discussions functioned as a pedagogical tool where youth could reflect on their learning and co-construct understanding with others. These reflection discussions also served as a meta-design tool where youth could express their desires for changes or modifications to the learning activities. These changes in the program engaged youth more deeply in the CML content because they could see the benefits for them in terms of participation and feel that their voices were being heard. For example, during one discussion, a researcher asked if one of the children could summarize what they did the previous day.

Justin summarized the Teachable Machine activity in which he trained a facial recognition system for cats and dogs with biased training data. He explained, "We went on this website and programmed a computer to see if it knew if it was a cat or dog. And it would actually say cat because it had more cat pictures instead of dogs. And every time you got a cat, you see more pop up [referring to the increased percentage of confidence in the classifier]." The researcher added, "Okay, so it was better for cats than dogs because there were a lot more cats than dogs in the training set." Several children added how they tested the machine on their faces, and the machine confidently classified them as either dogs or cats. The group laughed together when these stories were shared. Joe insightfully added that he was likely identified as a cat because his face was in the webcam's line of sight while training his cat classifier. The researchers did not anticipate that youth would test their *own faces* when creating a facial recognition system for cats and dogs. However, this exploration of their faces led to joyful interactions with the ML tool and playful discoveries around human error when training datasets. This portion of a reflection discussion highlights how researchers and youth created understanding together, both around CML knowledge and the design of the activities, when reflecting on their open-ended learning activities.

#### **4.2. Technical and social conditions to facilitate researcher and youth participation in design**

Our observations inspired another critical part of the significant redesign during the first five days of volunteering. We observed that all children engaged with at least one of the robotic toys we brought to the centers. One researcher noted that on the second day of robotics play that "since they knew that we were coming with the toys, they almost jumped and barely allowed us to walk. They spent the whole time with the toys... They even wanted to continue playing when it was time to clean up." Based on these observations and interactions, the researchers changed direction, made physical computing a major part of the program, and offered opportunities for children to build ML robotic machines of their interests with guidance from adults. Specifically, we chose a Scratch programming interface with a Teachable Machine extension which was used to program robots that used a micro:bit processor. Youth were asked to create narrative stories about robots that can be helpful to people, which we called "superhero" robots. Although their stories were fantastical in nature, with researchers' help, youth could translate aspects of their stories into a workable, programmable robot that used ML algorithms.

The youth were aware of the technical flexibility of the program and tools. In his interview, Lucas said he had no prior experience with programming and had not heard anyone talk about ML before. Lucas added, "When you started telling us those things [referring to bias in training datasets], I was surprised. So, now that I know this, about technological machines, I really like them because it's like, it doesn't have to be like anything in the world. You can be creative on your own and you can actually make something new." Lucas was referring to his learning about biased datasets. He explained that he enjoyed the ability to create a machine and contribute something to the world that was not there before. In turn, Lucas acted as a meta-designer, designing his own learning experience and designing a new ML application, which could inspire and change future iterations of the program. He continued to explain the benefits of participating in this program and how creativity is essential for his future profession, adding, "Cuz the way you work or something, you have to be creative. I want to be successful in life." Similarly, Emma specifically enjoyed the creative aspect of designing and programming a superhero robot. She noted, "I really liked the robots. And I also really liked when we got to come up with our hypothetical robots." Emma's use of the phrase "got to come up with" indicates her perspective that she had the agency to design her robot story. She stated that she also enjoyed "the drawing part of it and, like, kind of creativeness." Like Lucas, Emma valued the creative and flexible nature of the design phase and the ability to implement her design into an ML robot. Emma felt that what she learned in the program was valuable to her

future education: “I feel as we go into middle school, or high school, college, and so on, we’re always going to have to know something like this.”

In addition to flexible technical conditions for participation, involvement social conditions encouraged youth to make sense of CML and reflect on their participation in the program with researchers. For example, researchers held discussions with youth about how to showcase their creations. Youth and researchers created a list of options for presenting their ML machines and robots: post on a website and share publicly, have a “science fair” style presentation, invite friends and family, or hold a competition with prizes. Ultimately, the youth chose to hold a contest in which staff from the after-school center were the judges. These conversations were one example of creating comfortable social conditions for broad participation in the design of activities.

Although many conversations among youth and adults were benevolent, some discussions around discrimination brought discomfort and distrust for some youth. The most notable examples come from discussions following the Coded Bias viewing of how ML algorithms harm marginalized communities. At Sunshine Community Center, when the researcher asked what was interesting to the youth about the film trailer, Carter, a White girl, responded, “This is talking about... racism, so it said that, like, it can change whether people get property or have to pay the same prices for things as others.” Here, Carter noticed “racism” and inequities regarding how “people” obtain housing or purchase goods when ML algorithms are involved. Kendall, an African American girl, responded to Carter, “What I thought was interesting was the same reason, just because of the software, the person wasn’t recognized... they could get locked out of their house, or they could be denied for housing.” Sitting close to Kendall, Justin, an African American boy, nodded and affirmed her statement.

Carter then clarified who the “people” were being discriminated against, and she used an outdated term for African Americans, which Justin found offensive. This caused Justin to exclaim, “What?” and put his head down on the table. The researcher did not engage further with this language and continued the conversation. Kendall also ignored Justin and continued, “I thought what was interesting was the fact that the problem was so big that she [Joy Buolamwini] had to take it to court. The researcher asked, “Yes, who did she say it was mostly representing...” Justin popped his head up and shouted, “White Men!” The researcher responded, “That’s right, Justin.” Here, Justin rejoined the discussion by using his own racial vocabulary.

The researcher then asked a direct question about biased training datasets, “What do you think the training data looks like for the stuff that Joy was using?” Kendall answered and used the language that Justin introduced, “Umm mostly White men, because they didn’t have any other people in there to help them like, create their software besides White men.” Justin whispered, “Racist.” The researcher directly addressed Justin this time, “Right, yeah. I mean, it was biased towards white men because those are the people making a lot of the software.” Kendall then put the pieces together and said, “Well, bias is also like being racist.” The researcher extended Kendall’s comment, “Bias is a general term. Racist is a specific term for being discriminatory towards people who are a certain color or a certain race and that’s what’s happening here, right?” Justin, Kendall, and Carter nodded and affirmed. As the conversation continued towards defining bias, racism, and how ML algorithms can be programmed to be anti-Black, Justin fluctuated between engaging passionately in the discussion and disengaging.

There was a similar dynamic at Green Community Center, in which some youth engaged in the discussion and others, both African American and White, withdrew from the conversation. Afterward, we realized the youth were leading the direction of the discussion based on what they noticed, and the researcher mainly was following. This youth-led discussion facilitated an open system in which youth could shape what they wanted to discuss around algorithm bias and racism and how they wanted to continue with the activities. On the other hand, adults’ lack of structure or setting of ground rules and common language before the discussion led to some discomfort and insensitive language. The open nature of the discussion led to situations in which the researcher was unprepared. In turn, after the Coded Bias discussions, the researchers did not mention discrimination, racism, or other sociopolitical contexts around ML unless an individual child said it first and wanted to discuss it further. We let the youth’s level of experience and interests around institutional racism direct how we pursued the implementation of the remaining activities. Some children, such as Kendall, had experience discussing issues of racism and discrimination, were passionate about the issues, and could apply their knowledge to ML models. We provided additional content knowledge with these youth and helped them negotiate and build upon their prior experiences.

### 4.3. Negotiations and tensions between researcher and youth goals

According to our research notes, our goals were for youth to program robots for social good that incorporated some form of ML and to explain how training datasets were biased. However, we wanted to provide a flexible, modifiable environment such that learners could accomplish the CML goals in creative ways that aligned with their interests and values.

The youth took advantage of the flexible learning environment by designing their own experiences. For example, before the program began, Bianca, Ian, and Eric founded a gaming club called the Super Phenomenal Gamers (pseudonym). They explained to researchers that they formed this club based on their common interests in watching competitive gamers stream videos and wanting to explore esports as a career option. All three children were inconsistently involved in the CML activities; some days, they would participate, and other days they would play games on their mobile devices or update their YouTube channel. However, at the end of the program, they decided to develop an ML robot that relied on speech recognition. The purpose of their robot was to promote the Super Phenomenal Gamers YouTube channel. When the machine was turned on, the robot asked, “Are you subscribed to the Super Phenomenal Gamers?” If a person responded “yes,” the robot changed its headlight color to green, spun around several times, and replied, “Thanks for subscribing.” If a person responded “no,” then the robot remained still, changed its headlight color to red, answered, “Subscribe for more great content,” and displayed “Why not?” on its scrolling marquee.

Because of his interest in streaming videos, Eric used his mobile device to record the team working on their robot. He also recorded the group receiving feedback from one researcher and presenting their robot to the judges during the competition. The children posted these videos on the Super Phenomenal Gamers’ YouTube channel, which became an unexpected source of data, detailing how the children worked together when adults were not interacting or recording them. The videos provided Eric and his team with a method for documenting their work and formulating a narrative from their point of view. For example, when a researcher, Bianca, and Ian were debugging their code and ignoring the camera, Eric turned the camera to the researcher and said, “Say hi to the channel! People who go to Clemson you might know her, so please subscribe.” Most importantly, the children made an unprompted design choice to record and post videos of their work because it directly connected to their goals of being competitive gamers. The videos and robot design promoted their YouTube channel that contained their gaming videos. Bianca, Eric, and Ian’s choices illustrated the flexible nature of the ML tools and activities, which allowed them to show researchers alternative ways for the tools to be used in conjunction with other popular media in ways that the children valued. Although the Super Phenomenal Gamers enthusiastically used ML algorithms to program a robot and made creative choices for marketing purposes, the *critical* lens was missing from their project. These children did not design a robot for broader social good, nor did they reflect on algorithm bias and how to mitigate bias in their design. In this case, the researchers compromised their goal of having youth design for social good.

In contrast, other children designed robots for social good and were able to incorporate their interests and values. For example, Kendall decided to integrate her work using Google’s Teachable Machine with her superhero robot project. Using the webcam on her laptop, she trained her machine to classify objects by color. During the cognitive interview, she explained that she built this machine to help children learn their colors, specifically reflecting on her younger cousin’s lack of resources for learning. She explained, “My cousin... when she was growing up, she didn’t have the opportunity to, like sit down every day and like watch, like TV shows that teach her colors and stuff. And so, the only time she had stuff to learn is when I came down with my books and stuff and like taught her. And so, I thought to myself, that could be happening to multiple other kids all over America. And so, I thought, well, maybe I could make a machine that can help kids with that.” In that moment, it also occurred to her that her machine could be adapted to assist those with vision impairments. She imagined recreating her machine as a phone app that could help blind children identify colors: “They could download the app on to their phones... And if they’re like doing something, and they need to know what color it is, they turn on the camera, and it sends a link to their phone. And for each color has a different frequency. For orange or for red, it would be high pitch. But for black, it would be really low to be able to tell which one it is.”

When asked how she trained her machine, Kendall said she uploaded images with various colors from Google and tested the machine on “sticky notes, an orange, my backpack, and outfit.” However, she noticed that her machine did not accurately classify objects with different textures. She explained, “at first, I just put in solid colors, but then I realized when I put my mask up to the camera that there’s multiple different textures and stuff of different colors. So, for each color, I inserted a different texture.” During the interview, she demonstrated the updated functionality of the machine by holding her hand up to the webcam: “Yeah, my hands are really wrinkly. But if I put my hand into the camera, I’d be able to recognize that my hand is brown.” Kendall said she also walked around the center to test her friends’ machines. She pointed out that other youth created training datasets

that were biased and, in turn, were not as functional as she had expected. She referred to one person who “only used a picture of himself” to classify faces and concluded that “his data was biased.” In contrast, Kendall said her machine worked well and noted that she won the most creative teachable machine award at her center. Kendall’s work is an example of youth’s ability to design machines for social good and understand bias in training datasets and how to retrain data to minimize bias, all while aligning with their interests and values.

## 5. Discussion

In this study, we integrated cooperative inquiry (Druin, 1999; Guha et al., 2013) techniques with meta-design (Fischer, 2021) to explore how researchers and children interact as meta-design partners in the context of CML education. The findings in this study described the conditions that allowed adults and youth to be meta-designers of a CML education program, as well as the tensions and negotiations that emerge from an intergenerational design process involving sociopolitical contexts.

In our initial role as volunteers, we built reciprocal trusting relationships by being curious and valuing children’s ideas, ingenuity, and practices in their after-school space. Being a volunteer was similar to the *least-adult role* in Cumbo and colleagues’ (2019) study in which research is situated in the children’s familiar play environment, child-led interactions with adults shape the activities, and adults are reflexive about their changing relationship with the children. This phase of the relational building laid the foundation for engaging as meta-designer partners with children.

When the adult’s role changed to meta-designer, we developed tools and activities that we thought would meet our goal of youth creating robots for social good that incorporated ML and explaining the consequences of biased training datasets. However, our initial design assumptions led to a failed implementation in which youth were not interested in creating ML machines or exploring biased datasets. After discussing further with the youth, researchers redesigned the implementation plan. Like Williams and colleagues (2019), we discovered that young learners preferred physical computing and relied on adult-guided, hands-on activities to develop their understanding of ML. Aligning with the meta-design concept around developing open systems in which users can use products but also design them as they use them (Fischer & Scharff, 2000), we implemented technical tools and pedagogical activities such that children could modify their artifacts and learning experiences.

Moreover, we adapted the cooperative inquiry techniques of observing children use technologies in their own space (Druin, 1999) and supporting children in developing “low-tech prototypes” (Yip et al., 2013) of their designs by creating a visual story of a superhero robot using giant sticky-notes and markers. Researchers also supported children in realizing their fantastical designs (Guha et al., 2013) with the available robotics ML tools. Thus, through our iterative design process with children, we learned how to design mostly open systems in terms of allowing for daily reflective discussions, flexible ML technical tools, and flexible social conditions to reflect on social, ethical, and political ideas around ML but also around the design of the activities. However, these open systems were unsuccessful if the young learner had to fully control their own learning, likely overwhelmed with the amount of information and choice offered (Kirschner & van Merriënboer, 2013). Providing children autonomy in developmentally appropriate ways with adult guidance and support provided the conditions needed for meta-design.

Although we provided a particular set of tools, we discovered that children took advantage of the flexible technical conditions to create artifacts that went beyond the adult designer’s expectations. For example, Kendall created an ML robot that could teach underprivileged children, mitigated bias in her machine, and aligned her project to her interests and personal values. In another example, the Super Phenomenal Gamers team created a robot that promoted their YouTube channel, recorded their robot creation process, and posted these videos to promote their channel further. Before Bianca, Eric, and Ian found a way to integrate competitive gaming into their robot construction, they marginally participated in the program. Once they saw the program’s benefits and that they could augment their everyday tasks (Fischer, 2011), they became fully engaged in designing a ML robot. However, this success story had its limitations. The Super Phenomenal Gamers did not engage critically with ML, which was not aligned with the researchers’ goals. These findings speak to the multiplicity of voices that emerge during open systems in the meta-design of CML educational activities, which introduces complexities and tensions. In this case, if the researchers pushed their critical agenda with the Super Phenomenal Gamers, would the children have entirely disengaged from their ML project? These questions and the balancing of interests between children and researchers are essential to anticipate and engage with during meta-design around CML education.

In another form of meta-design interaction, researchers created flexible social conditions and engaged with youth in daily discussions about tools and the pedagogical activities. These discussions created an open culture of reflection that lowered the barriers to sharing design suggestions and allowed children to see that changes to the program and tools were indeed possible while the program was occurring (Fischer, 2011). However, when discussing the sociopolitical aspects of ML, the open culture of reflection became uncomfortable and unsafe for some of the youth who were part of the group being discriminated against by ML. Some youth felt passionate about discussing discrimination, but for some, the conditions did not facilitate this form of discussion. For example, Justin withdrew from the discussion when insensitive language was used and never addressed, and other children withdrew from the content around anti-Black racism as well. These examples suggest that although it is essential to engage youth in sociopolitical reflections at early ages (Ayón, 2016; Boute & Muller, 2018), much preparation and training must go into developing such open systems for critical discussions around ML technologies. This preparation must go beyond consulting with community partners (Lee et al., 2021) to include how to respond compassionately and sensitively to children, how to allow them to opt out of discussions safely, and, more generally how to build trust among partners in politicized contexts in technology education (Vakil et al., 2016).

Although asking children to co-design the CML learning environment while simultaneously participating in the learning environment was fundamentally messy, conceptualizing the roles of children and adults as meta-designers and incorporating cooperative inquiry techniques benefited the design of the program and engaged more children in critical thinking around ML as the program continued and was re-designed. A cooperative inquiry approach provided appropriate tools for children to participate as design partners where their expertise and curiosity were valued. All in all, the children's projects surprised and inspired the researchers to redesign the activities during the program and for the future. In subsequent implementations, we will encourage children to incorporate digital media of their own choice and promote the idea of videoing and narrating their work. We will also encourage youth to think more critically about their designs. In addition to redesigning activities, we will also update the robot's functionality based on the children's desires, such as adding music or adding more crafting opportunities.

## 6. Conclusion

This study integrated cooperative inquiry techniques with meta-design approaches to describe how adults and children collaborated as meta-design partners in a CML program and simultaneously engaged in the program as learners. We argue that conceptualizing adults and children as meta-designers is a practical approach for collaborative design and ML applications for social good. The exploration presented in this study is just one example of the multiple possible approaches towards engaging youth early on in their education about machine learning concepts and how to think critically about the social, ethical, and political issues around modern large-scale algorithm deployment. Educational research explorations are crucial for breaking the harmful tradition of technology development and consumption without a critical lens.

## Acknowledgment

Thank you to the lovely children at the local after-school centers for being our meta-design partners and developing this CML program. This work is funded in part by the National Science Foundation (ECR-2024965). The opinions, findings, and conclusions do not reflect the views of the funding agencies, cooperating institutions, or other individuals.

## References

- Ali, S., Payne, B. H., Williams, R., Park, H. W., & Breazeal, C. (2019). Constructionism, ethics, and creativity: Developing primary and middle school artificial intelligence education. *International Workshop on Education in Artificial Intelligence K-12 (EDUAI '19)* (pp. 1-4).
- Arastoopour Irgens, G., Vega, H., Adisa, I., & Bailey, C. (2022). Characterizing children's conceptual knowledge and computational practices in a critical machine learning educational program. *International Journal of Child-Computer Interaction*, 34, 100541. <https://doi.org/10.1016/j.ijcci.2022.100541>
- Ayón, C. (2016). Talking to Latino children about race, inequality, and discrimination: Raising families in an anti-immigrant political environment. *Journal of the Society for Social Work and Research*, 7(3), 449–477. <https://doi.org/10.1086/686929>



- Bailey, C. S., Adisa, I. O., Vega, H., & Arastoopour Irgens, G. (2021). Cognitive, affective, and politicized trust in a community youth program: A Participatory design research project. In *2021 Conference on Research in Equitable and Sustained Participation in Engineering, Computing, and Technology (RESPECT)* (pp. 1-2). IEEE. <https://doi.org/10.1109/RESPECT51740.2021.9620658>
- Boucher, P. (2020). *Artificial intelligence: How does it work, why does it matter, and what can we do about it?* EPRS: European Parliamentary Research Service.
- Boutte, G., & Muller, M. (2018). Engaging children in conversations about oppression using children's literature. *Talking Points*, 30(1), 2–9.
- Braden, E. G. (2019). Do you have papers?: Latinx third graders analyze immigration policy through critical multicultural literature. *Children's Literature in Education*, 50(4), 464–480. <https://doi.org/10.1007/s10583-018-9359-5>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, in *Proceedings of Machine Learning Research 81*, 77-91. <https://proceedings.mlr.press/v81/buolamwini18a.html>.
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
- Cumbo, B. J., Eriksson, E., & Iversen, O. S. (2019). The “Least-adult” role in participatory design with children. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 73–84. <https://doi.org/10.1145/3369457.3369464>
- Data Science for Social Good. (2021). *Data science for social good*. <https://www.datascienceforsocialgood.org/>
- DiSalvo, B., & DesPortes, K. (2017). Participatory design for value-driven learning. In B. DiSalvo, J. Yip, E. Bonsignore, & C. DiSalvo (Eds.), *Participatory Design for Learning* (1st ed., pp. 175–188). Routledge. <https://doi.org/10.4324/9781315630830-17>
- Druin, A. (1999). Cooperative inquiry: New technologies for children. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 592–599). <https://doi.org/10.1145/302979.303166>
- Ehn, P. (2008). Participation in design things. *Proceedings Participatory Design Conference (PDC)* (pp. 92–101). ACM Digital Library.
- Eubanks, V. (2018). *Automating inequality. How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Fain, J. G. (2008). “Um, They weren't thinking about their thinking”: Children's talk about issues of oppression. *Multicultural Perspectives*, 10(4), 201–208. <https://doi.org/10.1080/15210960802526102>
- Fischer, G. (2011). Beyond interaction: Meta-design and cultures of participation. *Proceedings of the 23rd Australian Computer-Human Interaction Conference* (pp. 112–121). <https://doi.org/10.1145/2071536.2071553>
- Fischer, G. (2021). End-user development: Empowering stakeholders with artificial intelligence, meta-design, and cultures of participation. In D. Fogli, D. Tetteroo, B. R. Barricelli, S. Borsci, P. Markopoulos, & G. A. Papadopoulos (Eds.), *End-User Development* (Vol. 12724, pp. 3–16). Springer International Publishing. [https://doi.org/10.1007/978-3-030-79840-6\\_1](https://doi.org/10.1007/978-3-030-79840-6_1)
- Fischer, G., Giaccardi, E., Ye, Y., Sutcliffe, A. G., & Mehandjiev, N. (2004). Meta-design: A Manifesto for end-user development. *Communications of the ACM*, 47(9), 33–37. <https://doi.org/10.1145/1015864.1015884>
- Fischer, G., & Scharff, E. (2000). Meta-design: design for designers. *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 396-405). <https://doi.org/10.1145/347642.347798>
- Freire, P. (1970). *Pedagogy of the oppressed* (30th anniversary ed). Continuum.
- Giroux, H. A. (1985). Critical pedagogy, cultural politics and the discourse of experience. *Journal of Education*, 167(2), 22–41. <https://doi.org/10.1177/002205748516700204>
- Glesne, C. (2014). *Becoming qualitative researchers: An Introduction* (5th ed.). Pearson.
- Guha, M. L., Druin, A., & Fails, J. A. (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), 14–23. <https://doi.org/10.1016/j.ijcci.2012.08.003>
- Jenkins, H. (2006). *Fans, bloggers, and gamers: Exploring participatory culture*. NYU Press.
- Kantayya, S. (Director). (2020). *Coded bias* [Documentary]. 7th Empire Media. <https://www.codedbias.com/>
- Kersten, J. M. (2006). Why's everyone white? Moving toward critical pedagogy in an elementary classroom. *Journal of Urban Learning, Teaching, and Research*, 2, 35–44.

- Kirschner, P. A., & van Merriënboer, J. J. G. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist*, 48(3), 169–183. <https://doi.org/10.1080/00461520.2013.804395>
- Lee, I., Ali, S., Zhang, H., DiPaola, D., & Breazeal, C. (2021). Developing middle school students' AI literacy. *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 191–197). <https://doi.org/10.1145/3408877.3432513>
- Noble, S. (2018). *Algorithms of oppression: How search engines reinforce racism* (1st ed.). NYU Press.
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy* (1st ed.). Crown.
- Pangrazio, L., & Selwyn, N. (2019). "Personal data literacies": A Critical literacies approach to enhancing understandings of personal digital data. *New Media & Society*, 21(2), 419–437. <https://doi.org/10.1177/1461444818799523>
- Paris, D. (2012). Culturally sustaining pedagogy: A Needed change in stance, terminology, and practice. *Educational Researcher*, 41(3), 93–97. <https://doi.org/10.3102/0013189X12441244>
- Paris, D. (2021). Culturally sustaining pedagogies and our futures. *The Educational Forum*, 85(4), 364–376. <https://doi.org/10.1080/00131725.2021.1957634>
- Robertson, T., & Simonsen, J. (2013). Participatory design: An Introduction. In *Routledge International Handbook of Participatory Design* (pp. 1–17). Routledge.
- Russ, R. S., Lee, V. R., & Sherin, B. L. (2012). Framing in cognitive clinical interviews about intuitive science knowledge: Dynamic student understandings of the discourse interaction. *Science Education*, 96(4), 573–599. <https://doi.org/10.1002/sce.21014>
- Samoili, S., Cobo, M. L., Delipetrev, B., Martinez-Plumed, F., Gomez, E., & Prato, G. D. (2021). AI watch. Defining Artificial Intelligence 2.0. Towards an operational definition and taxonomy of AI for the AI landscape. In *JRC Working Papers (No. JRC126426; JRC Working Papers). Joint Research Centre (Seville site)*. <https://ideas.repec.org/p/ipt/iptwpa/jrc126426.html>
- Terry, G., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic analysis. In *The SAGE Handbook of Qualitative Research in Psychology* (Vol. 2, pp. 17–37). Sage.
- Vakil, S., McKinney de Royston, M., Suad Nasir, N., & Kirshner, B. (2016). Rethinking race and power in design-based research: reflections from the field. *Cognition and Instruction*, 34(3), 194–209. <https://doi.org/10.1080/07370008.2016.1169817>
- Williams, R., Park, H. W., & Breazeal, C. (2019). A is for artificial intelligence: The Impact of artificial intelligence activities on young children's perceptions of robots. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–11). <https://doi.org/10.1145/3290605.3300677>
- Wolff, A., Wermelinger, M., & Petre, M. (2019). Exploring design principles for data literacy activities to support children's inquiries from complex data. *International Journal of Human-Computer Studies*, 129, 41–54. <https://doi.org/10.1016/j.ijhcs.2019.03.006>
- Yip, J., Clegg, T., Bonsignore, E., Gelderblom, H., Rhodes, E., & Druin, A. (2013). Brownies or bags-of-stuff?: Domain expertise in cooperative inquiry with children. *Proceedings of the 12th International Conference on Interaction Design and Children*, 201–210. <https://doi.org/10.1145/2485760.2485763>

## Appendix A

### Critical Data Literacies Project

#### Semi-structured Cognitive Interview Protocol

##### Introduction:

- A. Obtain consent and answer questions about the study (if this is the first meeting and consent form has not been signed)
- B. Review the study procedures with the participant
- C. Verify permission to record the interview using a digital voice recorder/camera (are you okay with recording the interview?)
- D. Say the participant's name and the date at the beginning of the recording
- E. Ask the following questions:

##### Questions:

- 1. Can you show me your Teachable Machine?
  - 1. Walk me through your TM. How does it work?
- 2. How did you come up with ideas for your TM?
  - 1. Did you include any of your interests?
  - 2. How did you start your TM?
  - 3. What steps did you take?
  - 4. What was important for you in this stage? How about in the next stages?
- 3. Once you decided \_\_\_\_\_, how did you move on? What did you do?
  - 1. Why did you \_\_\_\_\_?
  - 2. How did you \_\_\_\_\_? / How did you learn to \_\_\_\_\_?
- 4. I noticed that you \_\_\_\_\_, tell me more about it.
- 5. Were there any challenges in \_\_\_\_\_?
  - 1. How did you work them out?
- 6. How do you feel about your process of creating your TM?
  - 1. What was something you enjoyed about it?
  - 2. What was something that you did not like about it?
- 7. How did you feel about the results of your machine?
  - 1. Did you share it with anyone else and if so, what did they think?
  - 2. Do you think anyone else would like to use your machine? Who?
  - 3. Is there anything you would change about it if you were to do this again?
- 8. If we were to come back in the summer or next year, what sort of things would you like to work on with us?