

Exploring AI Problem Formulation with Children via Teachable Machines

Utkarsh Dwivedi
utkarshdvd@gmail.com
University of Maryland
College Park, Maryland
USA

Salma Elsayed-Ali
sea@umd.edu
University of Maryland
College Park, Maryland
USA

Elizabeth Bonsignore
ebonsign@umd.edu
University of Maryland
College Park, Maryland
USA

Hernisa Kacorri
hernisa@umd.edu
University of Maryland
College Park, Maryland
USA

Input (s) or Action (s)	Your Teachable Technology		Output (s) or Action (s)
	What kind of device is it?		
	What do you call it?		
Storyboard: How does it work in real life?			
How would you train it?	How would you check that it works?	What kind of mistakes will it make?	How would you fix those mistakes?



Figure 1: Our co-design study comprises: a) modified *Big Paper* with a structured storyboard that b) child-adult pairs use to frame and discuss their ideas and then c) present their teachable machines while a researcher summarizes their input.

ABSTRACT

Emphasizing problem formulation in AI literacy activities with children is vital, yet we lack empirical studies on their structure and affordances. We propose that participatory design involving teachable machines facilitates problem formulation activities. To test this, we integrated problem reduction heuristics into storyboarding and invited a university-based intergenerational design team of 10 children (ages 8-13) and 9 adults to co-design a teachable machine. We find that children draw from personal experiences when formulating AI problems; they assume voice and video capabilities, explore diverse machine learning approaches, and plan for error handling. Their ideas promote human involvement in AI, though some are drawn to more autonomous systems. Their designs prioritize values like capability, logic, helpfulness, responsibility, and obedience, and a preference for a comfortable life, family security, inner harmony, and excitement as end-states. We conclude by discussing how these results can inform the design of future participatory AI activities.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0330-0/24/05

<https://doi.org/10.1145/3613904.3642692>

CCS CONCEPTS

• **Human-centered computing** → User studies; • **Computing methodologies** → Supervised learning by classification; • **Social and professional topics** → Children.

KEYWORDS

participatory machine learning, machine teaching, values, design metaphors, cooperative inquiry

ACM Reference Format:

Utkarsh Dwivedi, Salma Elsayed-Ali, Elizabeth Bonsignore, and Hernisa Kacorri. 2024. Exploring AI Problem Formulation with Children via Teachable Machines. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3613904.3642692>

1 INTRODUCTION

Many recognize that “the formulation of a problem is often more essential than its solution” [36]. In artificial intelligence (AI), problem formulation is a crucial step for developing technologies that maximize potential benefits and mitigate potential risks. Typically, problem formulation in this domain “involves determining the strategic goals driving the interventions and translating those strategic goals into tractable machine learning problems” [10]. Several practical resources exist to aid developers in AI problem formulation (e.g., [6, 48]). However, recent research suggests that UX practitioners, including designers and product managers, are seeking additional support during the early stages of ideation and problem

formulation in order to preempt AI product failures [109]. Given the rise of generative AI, problem formulation skills are being foregrounded as foundational, enduring, and adaptable competencies that “might prove to be as pivotal as learning programming languages was during the early days of computing” [4]. Thus, it is important that we center problem formulation in activities that aim to help children develop AI literacies [31, 35]. Yet, to date there are few empirical studies of what these activities look like and how they can be structured for meaningful engagement.

In this paper, ***we explore the affordances of participatory design activities with teachable machines¹ for engaging children in AI problem formulation.*** Specifically, we hypothesize that teachable machines, where “the user is a willing participant in the adaptation process and actively provides feedback to the machine to guide its learning” [82], can play an essential role in facilitating such activities. We already see recommendations for incorporating teachable machines in formal educational activities that aim to spark curiosity and promote children’s understanding of how AI works (e.g., [96]). Yet, in these activities children tend to have *little control* over the problem formulation [34, 54, 100]; children tinker with the input or output of *predefined machine learning model* (e.g., a classifier) in a *predefined context* (e.g., recognize objects [34, 99] or gestures [54]). These approaches are critical first steps. However, we believe that by expanding the level of control and creativity children can exercise with teachable machines, activities can go beyond enabling children to uncover AI black boxes or change their perception of AI abilities [30]—their current promise.

To illustrate our approach, we pair problem reduction heuristics [72] with a structured “Big Paper” storyboarding activity [40]. Our modified Big Paper is shown in Figure 1a. We then invite a university-based intergenerational design team of 10 children (aged 8–13 years) and 9 adult partners to co-design their own teachable machines for a problem of their choice, as shown in Figure 1b. This storyboarding activity is situated within a larger effort to involve children in participatory machine learning activities that enable them to practice and build upon AI literacies. Our partnership with children involves first focusing their attention on specific aspects of the machine teaching process (e.g., similar to Dwivedi *et al.* [34] and Vartiainen *et al.* [100]), then gradually removing such scaffolding so that they are equipped to tackle the broader landscape of problem formulation. Within the context of this study, we ask:

RQ1: What are the key aspects characterizing children’s formulated AI problems?

RQ2: What are the most prevalent design metaphors for AI in children’s ideas?

RQ3: Which personal values do children incorporate into their designs?

Our study, exploratory in nature, provides empirical evidence on the affordances of participatory design activities with teachable machines for engaging children, as young as 8 years old, in AI problem formulation. We find that children’s formulated AI problems draw from their life experiences, addressing needs at home like security, automation, and familial support, as well as challenges at school such as math, safety, and social interactions. They envision

AI with voice capabilities and often expect constant video monitoring. While some stick to supervised classification, others explore various machine learning approaches including learning by demonstration or unsupervised learning. Most anticipate errors and plan to address them through methods like adding more training data, retraining or, in some cases, imagining self-debugging capabilities.

We also find that most of the children’s ideas encompass a recent call on human-centered AI for a shift in the design metaphors around AI development to move away from “autonomous AI systems” towards systems that “center human capabilities and involvement” [20]. Yet, the concept of intelligent agents and social robots who operate as teammates or with assured autonomy [90] remained enticing for a few of the children; this was especially the case for those who described their AI-infused technologies as something that they would not have to train.

Last, we used the Rokeach Value Survey [85] as an analytical framework to examine the values reflected in children’s designs. The Rokeach framework includes terminal values, which represent the ultimate end goals people strive for, while instrumental values reflect the desirable modes of conduct that people exercise to reach those end goals. We found that instrumental values such as “capability,” “logic,” “helpfulness,” and “responsibility,” were shared among all children as preferred modes of behavior for their machines. “Obedience” was also common. Less frequent were “self-control,” “courage,” and “cheerfulness.” Terminal values were also less present with “a comfortable life” being most prevalent followed by “family security,” “inner harmony,” and “an exciting life.”

Our intertwined, two-fold aim of co-design with children illuminates their conceptualization of, and interest in AI-infused systems and also promotes designers’ efforts to enhance future AI designs. Our work thus aims to make the following contributions: (1) a structured storyboarding activity for AI problem formulation; (2) empirical results on how children (as young as 8 years old) formulate AI problems; (3) new insights on prevalent AI design metaphors among children that engage with teachable machines; (4) empirical results on shared instrumental and terminal values reflected in children’s designs of AI-infused technology; and (5) novel analytical lenses in AI problem formulation based on Shneiderman’s design metaphors [90] and Rokeach’s value survey [85].

2 RELATED WORK

In this section, we survey recent efforts in participatory AI and AI literacy learning. We explore how teachable machines are a means for children to practice AI literacies and to participate meaningfully in AI problem formulation. We close by discussing values’ frameworks and their utility as an analytical tool for revealing values reflected in children’s designs.

2.1 Problem Formulation via Participatory Machine Learning

There is an increasing literature in participatory machine learning that engages adults, those who typically have domain knowledge but may lack AI expertise, in the problem formulation stages e.g., [12, 26, 83, 88]. Fewer studies focus on children. The modified “Big Paper” in our work is informed by one of them, Woodward *et al.* [106], who used storyboarding with children across multiple sessions

¹A term first coined by Andreae and Andreae [8] in 1978 and resurfaced by Google Teachable Machine [69] in 2017.

and emphasized anticipating and responding to errors. A critical difference lie in the fact that children’s designs were constrained to a predefined problem space (a touchscreen handwriting game) and discussions around errors centered on existing technologies (voice assistants). While none of Woodward *et al.*’s [106] sessions included teachable machines, their findings strengthen our belief in the unique opportunity for teachable machines to involve children in AI problem formulation. Specifically, they found that children think of interactive systems in terms of “both user input and behavior and system output and behavior.” The input/output structure in teachable machines aligns well with this mental model.

A second example comes from Vartiainen *et al.* [100]. Their study, which ran in parallel to ours, also engaged children in machine learning problem formulation via teachable machines. Similar to our work, children interacted in multiple sessions with a teachable machine (Google Teachable Machine [49]), then used a design template to imagine their own teachable applications. Small differences between our studies lie in children’s ages (8-13 in ours vs. 12-13 years old in [100]) and group formation (child-adult pair in ours vs. teams of children in [100]). More critical differences that highlight the complementary nature of these studies relate to design constraints and analysis approaches. For example, children’s designs in Vartiainen *et al.* [100] strictly conformed to a multi-class single-label classification problem with input being image, sound, or pose in a given form factor (a laptop) and interface (Google Teachable Machine); in our study there were no such constraints. In contrast to Vartiainen *et al.* [100], we use a design-metaphor and human-values lens to analyze children’s engagement in problem formulation.

A more recent work from Druga *et al.* [29] strengthens our hypothesis and study design. They invited parent-child pairs to complete a series of activities along four sessions. Activities corresponded to the following AI literacies dimensions: multimodal situated practice, embodied situated practice, AI conceptual learning, critical framing of AI, and design future meaningful use. Similar to our study, activities for *situated practice* involved teachable machines. More so, they preceded activities corresponding to *AI conceptual learning* (e.g., “Draw What is Inside”) and *design future meaningful use* (e.g., “Design AI”) that best align with AI problem formulation. Even though these last activities were constrained in the context of smart assistants and the focus was on parents’ roles in helping their children develop AI literacies, the results indicate that they were engaging and effective.

2.2 Engaging Children with AI via Teachable Machines

Teachable machines are gaining traction in specific contexts such as accessibility, and aging, where users have domain knowledge (e.g. access needs) but may lack expertise in AI. For instance, they enable users to personalize object [61, 93], sound [16, 60], speech [47]), and activity [65] recognition models built in their smartphones or smartwatches by fine-tuning them with their own data. On the other hand, systems like Wekinator [43] and Google’s Teachable Machine [69] are more open-ended and encourage anyone to explore and fine-tune a machine learning system to their own use-cases or creative applications like creating a musical instrument [43].

Upon analyzing use cases of Google’s Teachable Machine, Carney *et al.* found that it empowered educators and hobbyists, those who don’t have expertise in AI, by making it easier to get started with machine learning [19]. Thus, it is not a surprise to see researchers envisioning opportunities for teachable machines to increase children’s familiarity and creativity with AI. One of the earliest studies was that of Dwivedi *et al.* [34], where children between the ages of 7-13 years old, used Google’s Teachable Machine [69] to explore machine learning concepts. Many studies followed (see [34] for a comprehensive analysis). Some of them e.g., Vartiainen *et al.* [99], involved even younger children, aged 3-9 years old.

Learning activities with teachable machines are also recommended by the K-12 guidelines for AI education [96]. Yet, we see few instances of teachable machines in repositories of formal learning activities designed for use in K-12 and undergraduate courses. For instance, our analysis of the Model AI Assignments repository² [3], hosted by the Association for the Advancement of Artificial Intelligence, indicates that only 2 out of the 79 assignments incorporate teachable machines. This could be partially explained by the overall low number (9 out of 79) of assignments involving *younger children* or *non computer science majors*, where teachable machines could be most helpful. The rest require learners to read or edit code in e.g., Python or Java. Many are skewed toward undergraduate students in computer science with a focus on concepts like metrics or path-finding algorithms [2]. While these higher-level programming activities promote learning goals for computational thinking, they do not support much young, non-programmers’ (8-13 years old) efforts to explore AI problem formulation, the goal of our paper.

The *Introducing AI* assignment [1], one of the two that included teachable machines, is the most similar to our work. After being introduced to computational thinking and AI literacy concepts like “algorithm” and “data,” children 6-9 years old construct their own imaginary AI device and investigate its weaknesses. Similar to our work, this activity empowers children to explore AI concepts and apply their knowledge to design their own applications. This activity was implemented with several constraints imposed by the researchers: a set of predefined algorithms, sensors, and data. Moreover, children’s creations were not analyzed in terms of reflected values or design metaphors. In contrast, we do not constrain the children’s ideas by system capabilities, instead we ask them to consider a problem first, identify the training data needed, and establish a test and error-correcting process. In addition, we gain deeper insights into children’s ideas using Rokeach’s value survey [85] and Shneiderman’s design metaphors [90] as an analytical tool.

2.3 Children’s Values in Design

Technology is inherently value-laden [39, 58, 98, 101]. Computing research on values has largely focused on values of ethical import [44, 77] such as privacy and security, as well as values of a deontological or moral import [52]. While these studies are critical, research on individual, material values remains sparse [42, 52, 76, 102]. Those that do focus on personal values largely emphasize the values of adults and researchers rather than children [63, 91, 98, 108],

²The repository is populated annually for the past 12 years by a workshop that collects and structures various assignments to help instructors [79].

let alone children who are from historically marginalized communities. Skovbjerg *et al.* [91] recommend that Child-Computer Interaction (CCI) researchers acknowledge and discuss the embedded values in systems. This is especially important for AI problem formulation. Some of the most important implications of a system can emerge during this early stage and raise profoundly different ethical concerns, such as possible threats to fairness and civil rights [81]. Thus, it is critical to examine and explicitly foreground children’s personal values in the design of AI-infused technology, and technology more broadly [33, 37, 91, 94, 108]. This is increasingly important as personal values play a significant role in the design process [91] and children’s value trajectories are changing with societal changes [108]. This may also help us attain a clearer understanding of what it means to align AI with human values and ways to go about this.

Social psychologist Milton Rokeach defines values as enduring beliefs and personal standards that guide and determine actions, attitudes, ideologies, judgments, justifications, and presentations of self [85]. In his established value survey (RVS), Rokeach identified 36 values: 18 instrumental values reflecting preferred behaviors and 18 terminal values reflecting preferred end-states of existence [37, 58, 85, 102]. Despite its rigorous validation and importance in value research—having laid the foundation for other frameworks such as Hofstede’s Cultural Dimensions and Schwartz’s Value Survey (SVS)—the RVS has been seldom used in computing research: only in five papers across 40 years [7, 53, 64, 78, 102] and once in CCI research [37]. He *et al.* [53] used the RVS as an analytical tool in broader human-centric computing literature analyzing persuasive energy feedback technologies. Elsayed-Ali *et al.* [37] used the RVS in CCI literature to scaffold children’s designs of novel technologies. Until now, however, the RVS has not been used as an analytical lens in CCI research; other value frameworks have. For example, one study examined human values in adopting ubiquitous technology to support attendance control service in a primary school using the SVS [57, 86]. Values including Benevolence, Achievement, Power, Conformity, and Self-Direction were identified as the values exhibited by children’s behaviors in adopting the service [57]. However, the SVS is more concerned with motivational goals [86] and Hofstede’s model is more concerned with cultural values within organizational settings [56], as opposed to Rokeach’s focus on individual values and guiding principles in life. In this study, we use the RVS as an analytical tool to examine the perceived values reflected in children’s designs. We aim to uncover the values and interests children embed in their everyday AI systems, especially during the problem formulation phase. As values give expression to human needs [85], we must strive to understand the values children embed in the AI technologies they design to better meet their needs [37, 39, 91, 98].

3 METHODS

This work is part of a larger series of co-design sessions in which we explored children’s conceptualizations of AI with machine teaching (see Figure 2 for an overview). In our first two sessions, we constrained the children’s interactions and exposure to teachable machines. Children tinkered with ML inputs in the first session (using Google’s Teachable Machine [34]) and created personally

meaningful outputs in the second. Specifically, for the second session, children visited a local museum exhibit where they trained their teachable machine within an augmented reality app, which would then display their own 3D artworks upon encountering specific artifacts from the larger art installation. These experiences with similar machine learning models (3-way image classifiers) sensitized children to specific input/output aspects of the machine teaching process. Building from this grounding in machine learning models, we then invited children to imagine how they might incorporate teachable machines and machine teaching use cases into their every day lives. Our approach is aligned with established life-relevant learning approaches [23] that help children see the importance of scientific and AI literacy practices such as problem formulation in daily life.

3.1 Child and Adult Co-designers

Our work is part of a multi-part project exploring machine teaching and AI literacy activities with children (see Figure 2). In this study, 10 children aged 8–13 years old (4 girls, 6 boys) and 9 adult co-design partners (7 women, 2 men) participated. The first two sessions (Figure 2) afforded all the children an opportunity to gain hands-on familiarity with specific aspects of the machine teaching process; however, none of the children had coding experience beyond a handful of Hour of Code activities [24] at their schools. All personally identifiable data has been removed to protect the children’s anonymity. Children’s pseudonyms, gender, age, and race/ethnicity are shown in Table 1. Overall, the children’s demographics reflect our effort to include more non-dominant youth (i.e., non-White, immigrant backgrounds [73]), historically underrepresented in STEM learning, as co-design partners in the design of new technologies. With regards to adult co-designers, two adults had a machine learning background (e.g., coursework and research experience) while others had backgrounds in human-computer interaction and education. All adult co-designers have higher education qualifications.

3.2 Selection and Participation of Children

This study’s child participants are part of a university-based participatory design team. The children are recruited by word of mouth, from neighborhoods and municipalities local to the University of Maryland, based on family interest. Children are selected from a wait list, which is always open to new and prospective child members, with a goal of balancing the children’s age ranges (typically

Pseudonym	Gender	Age	Race/Ethnicity
Brian	M	8	Black/African American
Ed	M	8	Asian/White biracial
Luke	M	8	Asian American
Nancy	F	8	White/Caucasian
Ollie	F	8	Asian American
Adrian	M	11	Asian/White biracial
Alan	M	11	Black/African American
Denny	F	11	Black/African American
Penny	F	11	Black/African American
Kevin	M	13	Black/African American

Table 1: Pseudonym and demographics of the children.

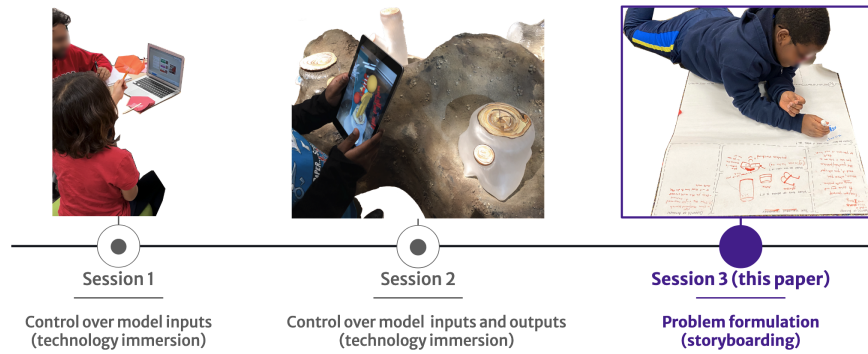


Figure 2: Our work is part of multiple co-design sessions. In session 1, technology immersion [32], children engaged with Google’s Teachable Machines, trained the classifier by exerting control over the input (training examples of the objects to be recognized), and tested each other’s training efforts. In session 2, also a technology immersion, children engaged with a teachable augmented reality application in a museum. The application, developed by our team, enabled children to control both the input (training examples of objects in a museum exhibit) and the output (their own 3D designs that were triggered upon successful recognition of the object). In this study, we present results from session 3, that engaged children in AI problem formulation via a modified “Big Paper” storyboarding activity. In session 3, children explored how they might formulate and approach a machine teaching problem of their own design, influenced by their every day life experience.

7-12 years old) and gender. In the case of the 13-year old for this particular study, the child started the academic year as a 12-year old and turned 13 during his year of co-designing with the team. The design team obtains parental consent and child assent at the start of each academic year, and all child participants are protected under university Institutional Review Board approval (#357390). We obtained signed parental consent and child assent, including audio/video recording consent. The adult researchers review participatory design and study goals with parents and remind the child participants that they can stop participation at any time during the co-design sessions. All personally identifiable data was removed to protect the children’s anonymity.

3.3 Study Design & Rationale

All co-design sessions across our study followed a Cooperative Inquiry-based [32] approach to understand children’s experiences with machine teaching and problem formulation, and each adhered to a similar 3-act structure [50]:

Circle Time: A warm-up to establish context and guide discussion.

Main Design Activity: The larger team forms smaller groups of child-adult co-designers (typically in adult-child pairs or 2-3 children per adult) in a focused design activity that aims to uncover how children imagine emerging technologies or learn, perceive, appropriate, and evaluate existing designs.

“Big Ideas”: Each group shares their design ideas in a whole group presentation to surface, summarize, and synthesize common themes and potential design requirements [41, 50].

Circle Time: Specifically for the storyboarding session, all co-designers (children and adults) answered the question, *Imagine you are helping a friend learn to do something. As they are learning, how would you help them fix their mistakes?* This prompt enabled the team to consider challenges and successes related to how people make mistakes and how we can help fix them. For example, Alan shared that he would “probably tell them what’s wrong and what’s

right” and Penny said that she would “*explain what they did wrong.*” The team also discussed the potential for a variety of everyday applications of machine teaching by exploring examples from prior image recognition work [62].

Main Design Activity: The main co-design activity consisted primarily of a modified “Big Paper” technique. We chose to employ storyboarding over other co-design techniques (such as prototyping, sketches only, or brainstorming with post-it notes) because our goal was to support children’s freedom to brainstorm in problem formulation, and storyboarding is a graphic visualization that can depict imagined user scenarios or sequences early in the design process [40, 51].

Typically, co-designers use large, blank sheets of paper for their “Big Paper” storyboarding. Often, however, more structured variations of storyboarding have been found to support younger children’s ideation process and design volume [40]. For example, Comicboarding uses a “familiar construct, the comic” to capture their ideas within comic strip panels, which has been found to increase the volume of children’s ideas during brainstorming [75].

Similarly, to scaffold and surface the children’s ideas regarding potentially complex problem formulation, we structured the “Big Paper” technique to employ problem reduction heuristics [103] by including sections that listed key questions as explicit design considerations (see Figure 1a). Specifically, we included sections for (1) breaking down the machine learning process into functional units (e.g., inputs and outputs); (2) specifying the technology used; (3) imagining how it works in real life; (4) explaining how to train it; (5) anticipating how and when their machines might make a “mistake”; and, (6) ensuring their systems could recover from mistakes.

Children constructed their storyboards in collaboration with adult co-designers; each child worked with an adult co-designer. In this study, both the children and adults took on the roles of *design partners*, meaning they were both equal stakeholders and active participants throughout the design of the teachable machines [15, 110].

Additionally, there were multiple advantages to pairing children with adult co-designers. First, this pairing supported a more balanced partnership through elaboration [15, 110], meaning children and adults worked collaboratively to generate and mix ideas together. Prior CCI research with children has found retrospectively that children appreciate and benefit from the complementary support roles that adults take on in co-design [110] to make space for and elaborate upon children’s ideas [74] and that the resulting design ideas may be more inclusive, open-ended, and nuanced than expected with complex topics such as privacy [67]. Second, adults provided children individualized attention and help with time management, structuring questions and children’s responses via storyboarding. Third, the adult-child pairing promoted children’s critical reflection. Adults encouraged children to engage in reflection-in-action [87] by asking them probing questions like *How would you test that it works?* and *How does it fix its mistakes?*. Questions like *Would violence or stealing be nice or good things to do?*, afforded the group with opportunities to discuss and reflect upon potential ethical shortcomings in the children’s designs.

Big Ideas: During their “Big Ideas” presentation, (see Figure 1c), children elaborated upon their overarching idea and explained how to train, test, and tackle problems that they anticipated. One adult captured design ideas from each adult-child dyad on a whiteboard for everyone to see [41, 51]. For most co-design approaches with children, particularly with the cooperative inquiry method, the smaller sub-groups assemble together again after the main co-design activity to share and compare their respective creations (e.g., sketches, stories, or prototypes). This “Big Ideas” process can uncover emerging common themes, overlapping ideas, unique perspectives, help developers prioritize or rank requirements [41], and enable both children and adults to engage in critical dialogue about their designs [50]. For example, during “Big Ideas”, many teams discuss what is technically possible, what design metaphors might underlie a seemingly whimsical idea, or what might be feasible alternatives and design requirements [51].

3.4 Data Collection and Analysis

We used multiple cameras to record the sessions and two researchers kept field notes. Overall, our data consisted of 10 “Big Paper” storyboards, more than 3.5 hours of video recordings, more than 1.5 hours of audio recordings and 3 sets of field notes. Via a thematic analysis [22] we aim to (RQ1) understand how children conceptualize, experience, and reflect on their engagement with AI problem formulation via teachable machines; (RQ2) uncover design metaphors for AI that are most prevalent in children’s ideas; and, (RQ3) examine and explicitly foreground children’s personal values in their AI designs. We selected verbatim quotes from children to support our findings. Below, we outline our steps.

Problem Formulation: Our structured storyboard was a silhouette of the design decisions that the children made, with a focus on the machine teaching process. The structured activity employs a series of decision strategies [103] to reduce complexity in the problem formulation phase such as factoring into subproblems (input and outputs); determining problem boundaries (device, use cases, training, testing); and examining changes (anticipating and fixing

errors). One researcher used this lens initially to code problem formulation themes in the data; then a second researcher, in discussion with the first, iteratively refined the themes and sub-themes.

Design Metaphors: Shneiderman’s [90] recent overview of emerging theories on human-centered AI highlights systems’ design approaches that “*support human self-efficacy, promote creativity, clarify responsibility, and facilitate social participation*” [97]. We situate our co-design efforts with children within this shift in perspective. Shneiderman describes how distinct design goals (science and innovation) “*have value in thinking about design metaphors for future technologies*” [90], with science goals favoring “*more automation*” and innovation goals favoring “*greater human control*.” To explore how children relate to their machines, we mapped children’s ideas into Shneiderman’s four pairs of design metaphors: *Intelligent Agents* vs. *Supertools*, *Teammates* vs. *Tele-bots*, *Assured Autonomy* vs. *Control Centers*, and *Social Robots* vs. *Active Appliances* [90]. Similar to our iterative approach to unpack how children approached problem formulation, one researcher initially coded the data for evidence across the spectrum of “*more automation*” in contrast with “*human control*” and then iteratively refined the themes in discussion with a second researcher.

Human Values: We used the RVS [85] as an analytical framework to examine the values reflected in children’s designs. We developed an understanding of these values from past work by He *et al.* [53] who used the RVS to classify persuasive energy feedback technologies. Similarly, we sought to employ the RVS to analyze children’s values in AI problem formulation. Using Rokeach’s 36 values (both instrumental and terminal) and their definitions, two researchers found patterns in the data that supported a smaller subset of these values. They individually coded the data, discussed classifications, resolved disagreements, and reached consensus.

4 RESULTS

To answer RQ1, we focus on the problem formulation aspect of children’s designs, mainly how they describe purpose of their teachable machines, context, machine learning model, inputs and outputs, anticipated errors, and recovery from errors. For RQ2, we then interpret children’s ideas from the point of view of AI design metaphors. Last, to answer RQ3, we examine the designs and explicitly foreground the conversations for the human values they appeal to. We present the conversations between adults and children, inline quotes from children, and the corresponding storyboards (see Figure 3) to support our findings. The block quotes specify child or adult, where children’s physical or non-verbal responses are within curly brackets and italicized, e.g., {*she nods*} or e.g., {*points to the output section on the storyboard*}. Any addition to complete missing words within the rest of the quote is denoted with square brackets (e.g., [the]). Apostrophes are used for inline quotes, which are also italicized. We highlight any part of a quote in **bold**.

4.1 Children’s Problem Formulation

Children designed the following storyboards, as presented and ordered in Figure 3, indicating the context that they would operate in as well as their input (for recognition) and output (for response).

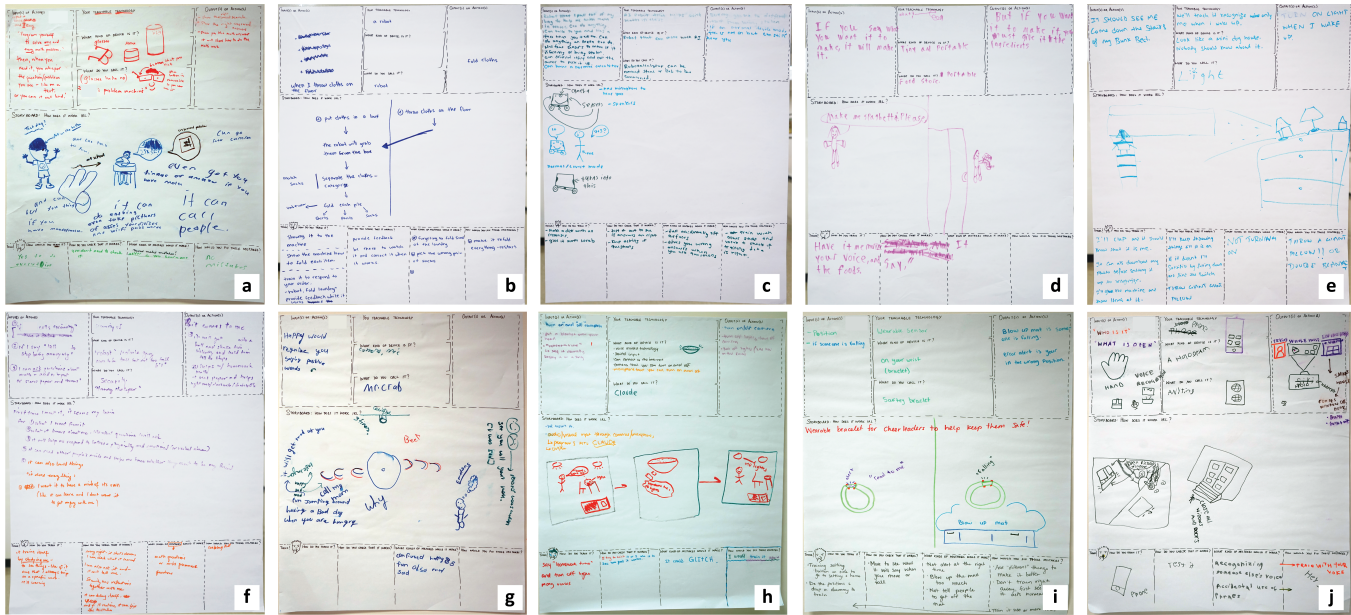


Figure 3: Children’s storyboards: a) Brian’s smart-glasses based “*Problem Machine*” helps solve any math problem just by looking at it, b) Ed’s robot helps you with chores such as folding clothes, c) Luke’s “*Robocalculator*” can travel to you and solve math problems, d) Nancy’s portable food store makes food for you if you get the ingredients and tell it the recipe, e) Ollie’s smart light controller, “*Light*”, recognizes her and turns on a light when she wakes up, f) Adrian’s robot security guard, “*Handy Helper*”, keeps his younger brother in check and protects him from bullies, g) Alan’s wall-mounted emotion detection machine, “*Mr. Crab*”, responds to your emotions, h) Denny’s smart light controller, “*Claude*”, creates your ideal homework environment, i) Penny’s safety bracelet for cheerleaders alerts you if you are about to fall and blows up a mat to cushion any fall, and j) Kevin’s holographic smart home security system, “*Holo*,” alerts you of possible intruders and open windows/doors.

- a) **Brian’s smart glasses-based Problem Machine** is a camera mounted on smart glasses that solves any math or crossword puzzle; “*you whisper it onto the glasses and it tells you or pops up [the answer] on your glasses.*”
- b) **Ed’s robot that folds clothes** helps Ed to do house chores like folding clothes that are on the floor based on the type of garment like shirts, pants, and socks. It is called “*robot*”. If it makes a mistake Ed will “*make it refold everything.*”
- c) **Luke’s “Robocalculator” for solving math problems** can move around, it will come to you if you name it and call for it. The Robocalculator can be snuck in your schoolbag and you can “*tell it to whisper*” answers to math problems.
- d) **Nancy’s portable food store** can ingest ingredients and recipes for so you can order it to make food with commands like “*Make me spaghetti please.*”
- e) **Ollie’s smart light controller called “Light”** can respond to claps close or open far away lights to help you avoid bumping into things at night. But, Ollie wanted “*it to make it difficult for other people*” to use it. It only works when she claps, which “*Light*” confirms by recognizing her face.
- f) **Adrian’s robot security guard called “Handy Helper”** is a personal security guard that can bellow abuses to those who bully Adrian and even discipline his younger brother if he’s being annoying. “*Handy Helper*” is constantly monitoring

Adrian and “*it can train itself by studying*” Adrian. It also can help with math and homework.

- g) **Alan’s wall-mounted emotion detector, “Mr. Crab,”** can learn your emotions. It uses a camera and mic as inputs and can make video calls. If you are sad, it “*calls [your] mom.*”
- h) **Denny’s smart light controller called “Claude”** controls the lights either by recognizing your activities (e.g., lights ON when you do homework) or by responding to your voice commands (e.g., “*homework time*”).
- i) **Penny’s safety bracelet** alerts cheerleaders before they fall and inflates a mattress to cushion their fall. It also keeps track of who is on the mat and “*tells people to get off the mat so that the faller can fall on the mat.*”
- j) **Kevin’s home-security hologram called “Holo”** is a full home-automation that responds to voice commands, acts upon alert sounds like fire alarms, and displays camera feeds of “*who’s outside*”. For example, it recognizes a specific user’s voice and closes or opens doors and windows in the house.

4.1.1 Defining Model, Inputs, and Outputs. In teachable machines, which fall within the broader machine teaching paradigm [112], problems are often formulated as supervised machine learning tasks; specifically, multi-class single-label classification. Thus, the teaching signal (i.e., input) typically refers to a labeled dataset of examples that the teacher (e.g., a child) provides the learner (e.g., a classifier). Similar to Dwivedi *et al.* [34] and Vartiainen *et al.* [100],

the children in this study had previously been exposed to teachable machines that employed a single supervised classification model that mapped one type of input (e.g., images) to an output (e.g., GIFs or sounds). Yet, many of the children’s designs deviate from this machine learning example in their problem formulation. Even when children’s designs employed supervised classification, each of their designs often included multiple such models capturing different functionalities. For example, Penny’s machine is capable of both (i) early fall detection (via the sensor in the bracelet) and (ii) collision detection (via weight sensors in the mat detecting presence of others).

Still, the concept of supervised learning appears even if the problem is not formulated as classification. Some children suggested that they would teach their teachable machines by “*showing it*” the correct way to do a task, a common AI problem called *learning by demonstration* [9]. Specifically, these children’s designs match problems formulated as *behavior cloning* [46], where given limited data over a short time an agent can generate similar behavior. Ed’s robot would fold the clothes, and he would teach it how to fold a shirt or pants differently from folding a sock.

Adult: How do you train the machine to do all of this?

Ed: By showing it.

Adult: By showing it. So what do you need to show it?

Ed: How to fold the clothes.

Adult: Do you need to show it, would you show it like how to fold a shirt or would you show it how to fold a sock? What would you show it?

Ed: How to fold all of them.

The promise of teachable machines is that end-users can fine-tune them by training them. However, children often assume existing capabilities in their machines without having to explicitly train them. This could be either because others trained them (pre-trained) or simply because they are “*smart*”; surprisingly, these assumptions don’t fall far away from recent discourse on foundation models even though the study was completed prior to the rise of ChatGPT [5] in the public discourse. All, except Penny and Ollie, assumed that their system comes equipped with built-in voice recognition, which may be related to their awareness of so-called intelligent agents [11] like “*Siri*” and “*Alexa*”. Specifically, Kevin’s Holo could automatically detect people without Kevin’s instruction, Adrian’s security robot would “*know*” if he’s being bullied, Nancy’s portable food robot did not need a recipe just a command because it “*would know how to make it*”, and Brian’s problem machine could solve any math or crossword puzzle without explicit training from him (and this was before ChatGPT’s [5] rise in popularity.)

We were also surprised to see a high variation among input types, with video being the most common. More importantly, children often envisioned a state of constant data collection/monitoring, providing a ripe opportunity for discussions around security and privacy in AI. For example, Adrian’s Handy Helper will always be monitoring Adrian and learning (unsupervised) from him and Kevin’s Holo will monitor his house with cameras,

Adult: So it works with cameras in your house too?

Kevin: Yeah. Cameras like in the kitchen and living room.

Adult: So cameras throughout the house that it talks to.

Kevin: Yeah.

4.1.2 Conceptualizing Teaching Strategies. For the machine capabilities that children decided to provide training examples, their teaching strategies diverged. Penny had a clear approach from the start. Each cheerleader would record their correct motion (“*like 200 or more*”) by using a button on the bracelet,

Adult: How do you train it to know the proper positions?

Penny: You do the proper positions and like {*Penny places her right hand on her left wrist as if she is touching a button on the safety bracelet*} you tell it to train.

Apart from correct motions, Penny suggested in her storyboard that “*you could drop a dummy to train it for falls*”, closely related to the *synthetic data* approach in machine learning often deployed when there are ethical concerns from using real data (e.g., [66]).

The others vacillated on their machine’s specific functions as well as what and how they needed to teach it. This is not a surprise, as problem formulation is often an iterative process. We observed that during this process, children would typically add to the teaching signal to provide more *context* or *discriminatory features*. For example, Alan initially thought that Mr. Crab would recognize (“*hunger*”, “*happy*”, “*fun*”, “*sad*”, and “*mad*”) based on his voice and his training labels like “*I am hungry*”. Then, he added that physical signals like “*jumping*” could also be indicators:

Alan: Yeah and just tell when you’re hungry by saying “*I’m hungry!*” You say it out loud ... And it would order you ... anything you want, or get you something from your fridge.

At this point, he adds “*saying positive words*” to the storyboard and proceeds to explain how he would teach it,

Alan: It can tell when you’re happy and having a good day or when you’re jumping and you’re excited or something like that...it recognizes by you saying positive words. You have to say it 6 times before it recognizes.

While Penny and Ollie wanted a button or an app to label start and stop for recording the training moves or actions, other children used speech for labeling. For example, Alan prompted “*I’m hungry now*”, Denny wanted to say “*Homework time!*” and to start doing homework on the table, and Kevin wanted, “*you go to each door and like tell it what door it is.*” We suspect children’s labeling approaches are reflective of their interactions with existing voice technologies.

4.1.3 Anticipating Failure. All children presumed that there would be some pre-training done for their system which they could train upon with new data or scenarios (e.g., fine-tuning). Hence the most common error they foresaw would be the likely scenario of insufficient pre-training. For example, Alan anticipated that Mr. Crab would not be “*properly trained*”; thus, it would not be able to differentiate between *similar* emotions like “*happy*” and “*fun*”,

Adult: So what mistakes would it make? Do you think it will confuse sad with happy or?

Alan: Probably happy and fun. They’re kind of the same.

Adult: What else do you think would be challenging?

Alan: Mad and sad.

However, we observed that children anticipated errors in their designs that go beyond the components that they taught. For example, when Penny discussed testing her safety bracelet with acrobatic moves, she anticipated that it might not alert at the right time, a *latency* challenge, in contrast to classification challenges such as

blowing up the mat too often (*false positives*) and not telling people to get off the mat (*false negatives*).

Those who used voice to interact with their system (most of the children) brought up pairs of phrases with potential for confusion like “*Ok cool*” instead of “*Ok Google*” similar to existing listing of misrecognized wake words [25]. Indeed, one child, Kevin, explicitly demonstrated a wake word mistake by pulling out his phone and saying a wrong phrase. Kevin’s Holo, supporting personalized phrases and commands by the home owner, could not only misrecognize some words but also mistake the identity of the person who uttered them,

Kevin: Some mistakes that it can make is if it recognizes someone else’s voice, and to prevent that you would have to train it with your voice. And if you accidentally use a phrase like “*Alarm*” and let’s say your friend said “*My alarm just randomly turned on*” and after that it would start an alarm randomly calling 9-1-1. You don’t want that to happen! You would have to say like “*Hey whatever the name is, do whatever.*”

4.1.4 Proposing Recovery. While children’s storyboards diverged among design ideas, teaching strategies and anticipated errors, they tend to converge when fixing errors: the majority (5) of those who opted to train their machines also opted to add more examples for any misrecognized scenario. For instance, in the case of a misrecognized home owner, Kevin suggested adding more examples from that person’s voice much like the initial training that is done by existing systems (e.g., saying “*Hey Google!*” many times).

Penny also said that she would add more examples but they would be “*different*” ones hinting at variation in training,

Adult: So would you just keep training it [and] training it?

Penny: No. Maybe we could add different things.

Perhaps it is not surprising to see this as the most common strategy e.g., many children have used it with Google Teachable Machine [69]. Even adult non-experts opt for *more data* as a fix [107].

Out of all the children, only Denny and Luke would help their machines recover from errors by starting the training from scratch; in contrast, Brian expected his problem machine to make no mistakes at all as “*it knows everything,*”

Adult: How would you check if the answer [to the math problem] was correct?

Brian: You don’t need to check. There’s a built in calculator, just in case you need to check.

Adult: You don’t think a calculator does not make a mistake? What if someone plugged in the wrong number?

Brian: It’s got a built in calculator, just in case. It never makes a mistake, the glasses never make a mistake {*for emphasis, Brian underlined “never” on his storyboard*}.

Like Brian, Luke and Adrian wanted their machines to solve math problems and never make mistakes. When we concluded this study in October 2022, we had yet to anticipate that ChatGPT would exhibit this behavior of “solving” math problems when released in November 2022. Children might be too impressed by ChatGPT, “never failing” to solve Grade 2-7 math problems, which could overshadow the limitations of this and other AI-infused technologies [21, 84, 104].

In this section, we illustrated how children can construct different scenarios where a machine learning system can fail. The following section discusses the design metaphors best describe children’s AI-infused technologies and whether they retain oversight over their error-prone or never-failing machines.

4.2 Design Metaphors Reflected in Children’s Designs

4.2.1 “Intelligent Agents” and “Supertools”. We consider the first pair of Shneiderman’s design metaphors [90]: “*Intelligent Agents*” and “*Supertools*” (or “*AI-infused Tools*” [20]) and how they relate to children’s ideas. “*Intelligent Agents*” are capable of thinking and making decisions with little or no oversight while “*Supertools*” offers a high degree of human control and automate the more repetitive aspects of a task. The two are often thought of in contrast as there is a trade-off between more automation and greater human control. Surprisingly, we did not see such contrast in many children’s designs; 7 of 10 children’s ideas had characteristics of “*Supertools*”. Only Adrian had aspects of “*Intelligent Agents*” in his idea, while Brian and Alan had characteristics of both “*Intelligent Agents*” and “*Supertools*”, particularly the automatic solving of *all* math problems and a *human-level* detection of emotions.

During circle time, we showed children examples of people training a machine learning system to respond with a specific action (e.g., a person hiding under bed covers to trigger an image detector to turn off a light). They asked insightful questions; for example, Brian asked “*where does the camera feed go?*” However, when discussing what their teachable machines “*know*” beforehand and what they are capable of, children’s responses reflected a technology utopian perspective [28, 45]. Their system had some *magic*, or *know it all* before any training was needed. For example, Brian noted that his machine would know all the math answers and not need training. Instead, it would train and help the human (“*it’s a computer! it already knows!*”). Or, when Adrian, who’s idea only had the aspects of “*Intelligent Agent*” described how “*Handy Helper*” would be independent to take its own decisions in a particular scenario based on human-like emotion recognition and mind reading capabilities,

Adrian: It can tell sort of ... tell what I am thinking so then I can do ... really cool stuff. If somebody bullies me on the street and I’m like, “I may be small BUT” then Security can read my mind and become un-invisible and be like “SURPRISE!”, slap.

...

Adult: And then it will help me respond to bullying. Only physically or?

Adrian: Physically and also I can give it emotional damage such as “smackdown” or something.

Adult: So verbal abuse.

Adrian: Verbal abuse, [because] words hurt more than weapons.

Adult: So does it just reads “a bully’s” mind or other people you are talking to.

Adrian: It can read any person’s mind. So if I meet anybody I can tell what they are thinking.

As “*Supertools*”, children’s machines would control lights, call parents, inflate mats, and open windows, all on their own, by making appropriate inferences. All their machines (except Penny’s

bracelet) would also do these actions upon explicit user commands. The following quote from Kevin shows how he wanted “Holo” to automate opening doors and windows both on command and in response to a kitchen accident,

Kevin: Let’s say you were cooking some eggs. And after that it just starts getting set on fire and smoke is going everywhere. And you want the window to be open. So all that bad smoke could go out. What you would do is that you would say “Open kitchen windows” or those carbon monoxide things would detect the smoke, like always, and it would start beeping and also it would open the window near it. Probably every window sometimes.

Only Brian’s problem machine had characteristics of both “*Intelligent Agents*” and “*Supertools*” because he had a superhuman expectation that it would never make any mistake on a math problem. Independently of how the “intelligence” came to be, Brian’s problem machine was there to help, a concept well aligned with the goals of Supertools to “amplify, augment, empower, and enhance humans” [90]. Like Brian’s idea, Vartiainen *et al.* [100] also noted that more than a few of the children’s ideas related to homework automation. They wanted a teachable machine that would “show mistakes or correct answers when taking a picture of homework.”

4.2.2 “Teammates” or “Tele-bots”. When looking at children’s ideas under the lens of the “*Teammates*” and “*Tele-bots*” pair of design metaphors, we observed a shift towards the second goal. We found that only 1 child assumed human-human interactions in their designs, where machines would be perceived as teammates, partners, or collaborators. For example, all children leveraging voice interactions imagined them being one-sided commands rather than conversations, such as Denny’s smart light controller that would respond to her command “*Homework time!*” but only Adrian wanted his “*Handy Helper*” robot to tell him if the person talking to Adrian wants to be his friend. As he continues to describe how the robot can scan anyone’s mind he says,

Adrian: It can read any person’s mind. So if I meet anybody I can tell what they are thinking. So if I am trying to make a friend.

Adult: But what if?

Adrian: So if they don’t want to be my friend then they just don’t want to be my friend because I have a super awesome robot that is currently reading their mind.

“*Handy Helper*” ensures that Adrian’s younger brother does not annoy him. Much like a helping hand or a teammate, the “*Handy Helper*” warns him when “*it sees me about to make a mistake it can tell me “Adrian don’t chase that dog around,” or like “Adrian don’t grab that pointy stick.”*” These are all responses or alerts to a situation. Even when Adrian asks his robot to reflect on its mistakes and fix them, he does not imagine much of a conversation but just something it does automatically.

As Tele-bots, children’s machines often provided “*superhuman perceptual and motor support while allowing human-human teamwork to succeed*” [90]. For example, Kevin’s Holo could signal “*who is outside*” his home and Penny’s bracelet would alert cheerleaders if they were about to fall to them and their teammates,

Penny: For example you’re holding someone or someone is in the air and you can’t read it yourself you could tell it to read to you and it will tell you everything you are doing wrong, and it will also alert when someone is falling ...

Another example is Ed’s robot, that is assumed to have motor actuators and grasping capabilities that let it find and grab the clothes lying around, perceive what kind of clothes are folded in what way and then fold the clothes.

4.2.3 “Assured Autonomy” and “Control Centers”. When looking at children’s designs under the lens of the “*Assured Autonomy*” and “*Control Centers*” pair of design metaphors, we observe that children’s ideas aligned more closely to the second goal. While systems with “*Assured Autonomy*” act independently, many children’s machines instead depend on and support human control and oversight (supervised autonomy) through “*Control Centers*” or Control Panels [90]. Only 2 of 10 children had aspects of “*Assured Autonomy*” and the rest had aspects of “*Control Centers*.” Children could predict the response of their machine learning system, maintain situation awareness, and take control of their machines at any time. For example, Kevin’s Holo supported supervisory control and situation awareness,

Kevin: What it does is that if you wave your hand toward it, it pops up a hologram showing everything about your house. Over here, it would probably be like security cameras of the front door, even though there are more security cameras around the house, on the outside and the inside. It’s just showing you the front door because that’s the main place. There would also be a picture of the house—like [the] holographic picture of your house that you could rotate around.

In contrast, Adrian gave complete discretion to his robot, and Nancy expected her food machine to make any food without supervision. Just put in the ingredients, and it does the rest. Adrian’s “*Handy Helper*” learn autonomously, “it can train itself by studying me and seeing if I make any mistakes, like continuously.” When the adult asked Adrian about the kind of mistakes his robot would make or how it could fix them, Adrian provided evidence of “*Assured Autonomy*” by highlighting an inherent issue with the concept of complete autonomy.

Adrian: ... I want it have a mind of it own. Right. So do you know rainbow friends?

Adult: No.

Adrian: Okay. Rainbow friends is this roblox game. I guess you could just look it up. But, um, it’s like rainbow friends have minds of their own. I want my robot to have a mind of its own so I don’t have to be like “why do you learn to [do something]”. Because it might turn against me and I don’t like it. And I don’t like a 6 foot 4 like super robot shooting rocket launchers and stuff to hit me.

In illustrating the tension between control and autonomy, Shneiderman provides the critical reason: that “operators [should] have a clear mental model of what will happen next,” *i.e.*, *predictable behavior*. Adrian could use the storyboard’s structure, the questions on mistakes, and fixing errors to critique his idea and push back on any unpredictable behavior.

4.2.4 “Social Robots” or “Active Appliances”. When viewing children’s designs under the lens of the “Social Robots” and “Active Appliances” pair of design metaphors, we find more evidence of the second design goal being prominent with some caveat. Children envisioned their machines as physical devices attached to a ceiling, placed somewhere as speakers, put on as glasses, worn as bracelets, or as holograms (*i.e.*, of a house, not a humanoid), or humanoid like in Adrian’s case. Many could listen and respond to commands. Alan, Kevin, and Denny named their machines. However, even those with human-like names like “Claude” and “Mr. Crab” did not have human-like forms. While having a human-like name and being able to listen/speak is at some level anthropomorphism [71], these features are standard even in everyday devices and appliances that are not considered social robots (*e.g.* our smartphones).

Emotion, often associated with Social Robots, is prevalent in children’s designs. For example, in Vartiainen *et al.* [100], one of the apps would detect your mood using facial expressions and “*if you are feeling sad, the app will comfort you.*” Woodward *et al.* [106] made a similar observation, that children wanted their intelligent user interfaces to recognize their emotions and respond with appropriate emotion, stating that “*this type of intelligence was a fundamental part of many of the children’s designs*” [106]. We found that only two children, Alan and Adrian, have characteristics of social robots, particularly the ability to recognize emotions or read minds. However, Alan’s “Mr. Crab” or Adrian’s “Handy Helper” could only recognize emotions rather than express them. When Alan needed comfort, he wanted it to come from his mother, not “Mr. Crab”,

Alan: ...and you can add contacts to it

Adult: Oh cool

Alan: So then it can call my mom

When presenting his idea to the larger group, he said that “Mr. Crab” would call his mom when he’s sad because, “*when you’re sad, I just call my mom, basically. I just call her.*” In contrast to Social Robots mentioned by Shneiderman, “Mr. Crab” plays the intermediary role and recognizes Alan’s emotional state and then automatically calls his mom rather than engaging Alan in a conversation or having synthetic fur for him to touch.

The rest of the ideas, by 8 of 10 children, fall under “Active Appliances” with robotic characteristics like actuators in Ed’s clothes folding robot or treaded robots like Luke’s “Robo-calculator” would come to you when you called it. The most interesting characteristic that made children’s ideas an “Active Appliance” was how they would fix its mistakes. As Shneiderman points out, there is “much room for improvement in the frustrating designs [of Active Appliances], which are often internally inconsistent, difficult to learn, and vary greatly across devices.” Much like our response to frustrating appliances, when some people “hit the TV” or “just restart it” to fix it, Ollie would be annoyed if her “Light” didn’t work causing her to stay stuck in the dark on her bed,

Adult: What will you do if it doesn’t turn on?

Ollie: Don’t worry even if it takes an hour to do, I’ll still figure it out. I’ll be like here’s a pillow and keep on throwing things for an hour and sacrifice myself and go switch on the lights. And if it doesn’t work **I’ll get so mad I’ll scream into my pillow.**

Adult: What do you think? What kinds of mistakes will it

make?

Ollie: Umm, probably not turning on the lights. That’s the only mistake it can do. And maybe I can just, I have a giant giant pillow that I’ll throw at this so this turns on. Or just a giant blanket. Like a double blanket.

4.3 Values Reflected in Children’s Designs

When looking at the values represented in children’s designs, we group our observations following Rokeach’s framework [85], distinguishing between instrumental and terminal values.

4.3.1 Instrumental Values in Children’s Designs. Rokeach describes instrumental values as preferred modes of behavior. For example, the value of “capability” entails competency and efficiency. Of the 18 instrumental values, the two coders agreed upon 10 instrumental values that were apparent in children’s designs, with the following 4 values represented in all their designs: “capability,” “logic,” “helpfulness,” and “responsibility.” We also found the following values apparent in children’s designs: “obedience” in 6 of the designs (including all of the younger children’s designs); “cleanliness” in Kevin and Ed’s designs; “self-control” and “courage” in Adrian’s design; and “cheerfulness” in Luke’s design. Below, we provide examples of the 4 most frequent instrumental values in children’s designs.

All of the children’s designs appealed to the value “capability,” making explicit reference to how their system might increase their efficiency or competency, or instances where the system itself exhibits efficiency and competency. For example, Kevin’s Holo would help him close the windows and doors of his house when he is going to sleep or outdoors, thus helping him complete the task more efficiently,

Kevin: Let’s say you’re about to go to bed or go somewhere and you have no time to go around the house and close every single window because you live in a mansion. You could say “What is open?” Let’s say there was a window all the way upstairs you didn’t know about. What you would do is say “What is open?” and it would tell you every single door that is open and every single window that is open.

All of the children’s designs reflected the value “logic,” which Rokeach defines as being consistent and rational. Children primarily touched on how they would program the logic of their classifiers including the context, input, and outputs. For example, Penny would train her safety bracelet using different positions,

Penny: I can train it by going to the settings. It has a setting where you can go to...train it that way.

Adult: How many times did you [want] to train it to make sure the mat blew up if you were the one falling?

Penny: 200 or more, I’ll train it 200 or more.

All of the children’s designs appealed to the value “helpfulness,” which means working for the welfare of others. Children described how their teachable machines would assist them in accomplishing a specific task, oftentimes explicitly using the word “help” to describe their machines. Adrian described his robot security guard as “*a handy helper*”. He’s a handy and he’s like a helper and he helps me with everything.” Similarly, Penny stated “*My machine is a wearable sensor, like a bracelet, and it’s a bracelet for cheerleaders to help them stay safe.*”

Rokeach's value "responsibility" was reflected in all of the children's designs and means dependability and reliability. Children trusted their machines would work reliably and provide the intended outcomes. For example, Alan relied on his emotion detection machine to detect the correct emotion and respond to it accordingly. Luke also relied on his Robocalculator to not make mistakes, *"I would make this special disk and this has everything. So when somebody ask[s] some question the disk will automatically go to its brain. And it's always right and never wrong."* We see this value also in ability to recover from errors. For example, every night Adrians' robot *"will do his reflection"* and could learn from mistakes.

Adrian: So I'll explain to him patiently that it did wrong and it will be like "Oh okay." And then we can, it can ask me to help it but if it doesn't then it can figure it out himself. Then it can debug it.

Adult: It can debug itself?

Adrian: Yes. But if it needs help from me then I can help him.

The value "obedience," was also reflected in 6 of the children's designs, and entails dutifulness and respect as well as user control over the AI-infused system. Adrian used the analogy of a dog to describe how his robot security guard obeys his orders: *"when I call 'Security' it will come towards me. It's like dogs when you call its name it will come to you."*

4.3.2 Terminal Values in Children's Designs. Rokeach describes terminal values as preferred end-states. For example, the value "family security" entails taking care of loved ones. Of the 18 terminal values, the two coders agreed upon 4 present in children's designs. We classify children's designs as appealing to the following terminal values: "family security," "a comfortable life," "inner harmony," and "an exciting life." While values relating to family and comfort are of highest priority to children [37], these values often go unacknowledged in existing AI discourse and frameworks. We call explicit attention to these values in our analysis. Below are examples of the 4 terminal values in children's designs.

The value "family security" was reflected in 3 designs belonging to the older children. Alan's machine called his mother when he felt sad; Penny's protected her teammates; and Kevin's safeguarded his home,

Kevin: Let's say someone was knocking on your door, right, and you said "Who is it?" to them but they didn't hear you or you kept saying it and they're trying to get in for some reason. You use the hologram so you could see who is outside using the cameras, and if there's somebody there you could speak to them using a mic or you could see what they're doing, or you could also contact 9-1-1.

The value "a comfortable life" means a prosperous life and was reflected in 7 designs. This value is represented in both Ollie and Denny's smart light controllers intended to help them create a comfortable homework environment,

Adult: For you, Claude turns off the lights. If Adult owned Claude, could Adult train Claude to turn on music when it's homework time, or is Claude just a Denny thing that always turns off the lights?

Denny: Adult could train him.

Adult: Okay, so it could help anyone to make their homework

environment no matter what, this is just your specific one?
{Denny nods}

Alan's teachable machine reflected Rokeach's value "inner harmony," which means to have freedom from inner conflict. Alan designed an emotion detection machine in order to identify and alleviate times when someone may be experiencing negative emotions like sadness or anger, as well as detect and reward times when someone is happy or excited.

Last, Penny's safety bracelet appealed to Rokeach's value, "an exciting life," alternatively defined as a stimulating, active life. Penny drew inspiration from her extracurricular activity cheerleading involving difficult stunts and physical exertion. Penny describes how her teachable machine would detect cheerleader's quick actions and body movements,

Penny: It's gonna also alert when someone is falling or like some quick action or movement. When someone's falling it blows up the mat and tells people to get off the mat so that the faller can fall on the mat.

5 DISCUSSION

Our user study, exploratory in nature, shows both promising results and future research directions for teachable machines in problem formulation activities that aim to help children develop AI literacies. In this section, we first reflect on lessons learned: some offer new insights, others strengthen prior empirical and anecdotal evidence. We also discuss implications for designing AI problem formulation activities and broader use of design metaphors and human values as analytical lenses in AI problem formulation. We then discuss limitations in our study that may affect both the applicability of our approach and the generalizability of our findings.

Key aspects characterizing children's formulated AI problems (RQ1). We find that children draw from their life experiences when identifying problems. In the home, their AI designs can provide security, control lights, automate cooking, fold clothes, discipline siblings, and call one's mom for comfort. At school they solve math and puzzles, ensure safety in cheerleading, and bellow abuses to bullies. While the form factor of their designs varied, almost all assumed voice capabilities and many envisioned a state of constant video monitoring. Few restricted their designs to supervised classification, the machine learning paradigm in the teachable machines they were exposed to. Some moved away from classification to other supervised machine learning approaches such as learning by demonstration or imagined that their machines would learn in an unsupervised fashion, by just following them around and observing them. Others presumed no learning at all; their machines are "smart" and they just "know". The majority of those who chose to explicitly train their machines opted to provide labels via voice. Interesting concepts like synthetic data, discriminatory features, similarity, and fine-grained classification emerged in few of the designs. Most of the children anticipated errors. They were related to false positives, false negatives, grasping and object handling, force and torque control, and latency. Many of them would fix the errors by providing more training data, two opted to restart the training, and one imagined that the machine *"can debug itself so there isn't anything to worry about."*

Prevalent design metaphors for AI in children’s ideas (RQ2). We find that when children are first asked to train a machine and then build their own to solve a problem of their choice, they mostly do not anthropomorphize their machines beyond naming them or giving them pronouns. Instead, children’s designs tend to align with proposed changes in design metaphors for AI [20, 90]; they moved away from *intelligent agents*, *teammates*, *assured autonomy*, and *social robots*. The most prevalent design metaphors reflected in children’s ideas were *supertools*, *tele-bots*, *control centers*, and *active appliances*. Shneiderman suggests these same metaphors for designers to pursue, i.e., “successful designers avoid mimicking human models and pursue supertools, tele-bots, active appliances, and control centers that support human control over technology while ensuring high levels of automation.” Such an alignment concurs with what Shneiderman calls the “*Innovation Goal*” that “drives researchers to develop widely used products and services by applying HCAI methods.”

Personal values incorporated into children’s designs (RQ3). Ongoing discourse on human-centered AI routinely advocates for considering values such as reliability, safety, and trustworthiness in future AI technology. We found that children’s designs for teachable machines often reflected similar values including “Responsibility,” “Capability,” and “Family Security.” Furthermore, many of the children’s designs reflected the value “Obedience,” which draws analogy to Shneiderman’s [90] design metaphors that stress extending human abilities such as human control. However, we also found more nuance in terminal values related to security and safety. Current discourse around safety and security in AI tends to emphasize misuse, potential weaponization or ethically questionable goals e.g. hacking or gaming a system loophole to complete a task [17]. While the children’s designs mapping to the terminal goal of “family security” echoed these discussions, they also expanded the value space to consider comfort, familial connection and harmony, along with excitement e.g. keeping acrobatic cheerleaders safe. These values for family and comfort were high priorities for children; yet nuances around comfort and harmony are often overlooked in existing AI design considerations, ethical debates, or frameworks.

5.1 Implications

5.1.1 Youth & AI Literacies. We provide empirical evidence on the affordances of participatory design activities with teachable machines for engaging children, as young as 8 years old, in AI problem formulation. We demonstrate that established co-design approaches [32, 40, 51] with problem reduction heuristics [72] in the context of teachable machines can promote children’s AI literacies and support their efforts to formulate problems for AI-infused technologies with personal relevance in their daily lives. Despite their inexperience with machine learning, the children worked closely with adults to explore what was feasible and what was not, iteratively reformulating the problem space as they considered their machine’s specific failure and recovery conditions. Their designs showcase their creativity not only in the personally meaningful problems they articulated, but also in how they tried to construct their training strategies.

We situate our work in Long and Magerko’s [70] definition of AI literacy and connect our study with several of their design considerations, along with competency goals that our sessions fulfill. Since we based our sessions on teachable machines that don’t require coding and used pen and paper-based storyboarding, our sessions exemplify Long and Magerko’s design considerations for *Low Barrier to Entry*. By structuring our storyboard and asking children questions that promote reflection on their systems, we promote the design considerations on *Critical Thinking*. Finally, children were free to solve a problem of their choice. We found personally meaningful ideas like Penny’s safety bracelet for her cheerleader team, which reflected their design consideration of *Identity, Values, and Backgrounds*.

As AI and machine learning technology proliferate in children’s everyday lives, we gain more opportunities to equip them better to tackle overhyped claims about the infallibility of AI. To this end, our AI problem formulation activity fulfill many competency goals from Long and Magerko’s [70] work 5 that intersect with the learning goals laid out by Touretsky [96]. Empowering children to make a teachable machine ensures that they are not reduced to mere “users” of AI. Our work shows how children ground their ideas in the “data” used to train a machine. Hence, our approach to using teachable machines, similar to Dwivedi *et al.* [34] and Vartiainen *et al.* [100] followed by a structured storyboarding session can be beneficial in fulfilling the competency goal of *Learning from Data*. We saw that the storyboard’s scaffolded structure helped children reflect on different aspects of their designs. In combination with recent work by Vartiainen *et al.* [100] and Hitron *et al.* [55], our storyboarding activity, which focuses on potential errors and options for error recovery, demonstrates how we can encourage children to reflect critically on the utility and limits of AI in their everyday lives and fulfill the competency goal of *AI’s Strengths and Weaknesses*. Finally, the children design their machines and formulate the machine learning problems they want to solve by imagining their own AI and fulfilling the competency goal to *Imagine Future AI*.

5.1.2 Youth & AI Ethics. Gaining insight into children’s values has been a longstanding challenge in Child-Computer Interaction (CCI) [33, 37, 58, 59, 91, 94, 98], and attending to children’s values can promote critical computational action practices [95] and positive dispositions toward STEM. Our study demonstrates that when children and adults co-design together, they engage in and enjoy critical reflection, playfully questioning the boundaries of privacy and AI ethics issues. For example, children asked their machines to cheat on math quizzes to help reduce math anxiety, to read others’ minds as a means for nurturing friendships and mitigating social anxiety, and to serve as powerful physical and verbal allies against bullying. These discussions indicate that storyboarding approaches like ours may promote reflexive conversations around the ethics of AI model development and deployment. Similarly, Williams *et al.* [105] and Zhang *et al.* [111] showed that middle schoolers brought up ethics when designing their own AI systems that included image recognition and generation. Another recent example comes from Kusuma *et al.* [68] involving adults who explored ethical considerations in facial recognition for finding relatives in Civil War databases. Future researchers could extend the approaches from this prior work with a storyboarding session like

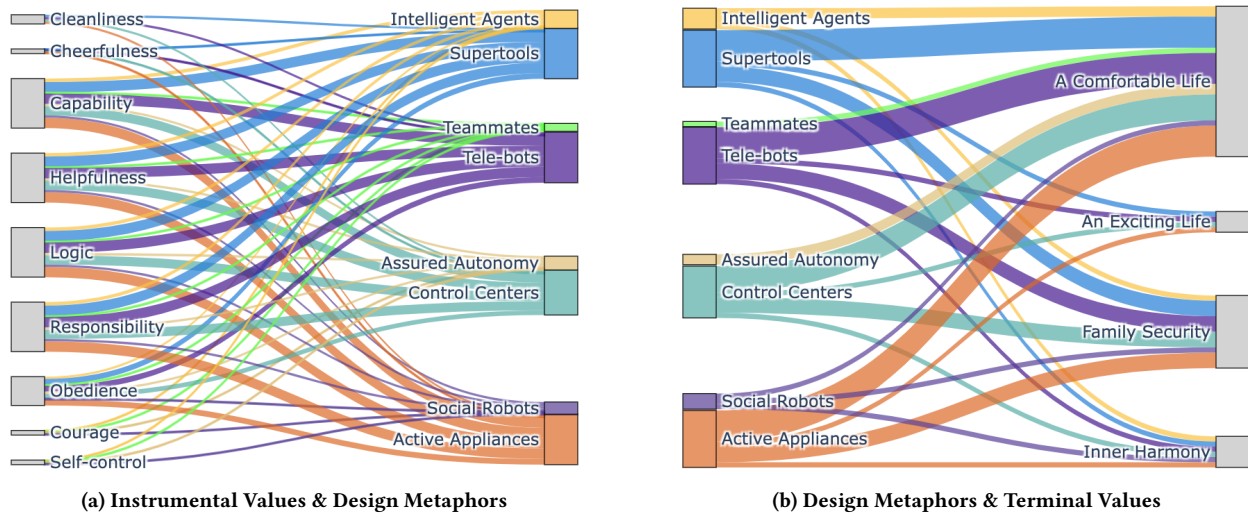


Figure 4: These Sankey diagrams illustrate the relations between the design metaphors and human values in the children’s designs. Each flow line represent the strength of a link (or affinity) between the two e.g., “Active Appliances” are closer to Instrumental Values like “Helpfulness,” “Capability,” “Logic” and “Responsibility”; and the Terminal Values of “A Comfortable Life” and “Family Security.”

ours that includes prompts for participants to fix their systems’ ethical issues. We show that when grappling with technical AI issues related to personally meaningful topics, children can be creatively involved and immersed in problem formulation and critical of potential ethical limits of AI. We hope that our choice of widely available resources and materials could encourage others to replicate and expand on this work to facilitate reflections on AI ethics and engage children from other developed or developing regions and across cultures. For instance, one could combine the idea of a pre-defined toolkit from Druga *et al.* [29] and our structured storyboarding activity to further ground the session within machine learning concepts.

5.1.3 UX & AI Practitioners. There is an increasing need among practitioners for support during the early stages of ideation and problem formulation in order to preempt AI product failures [109] with relevant skills being foregrounded as foundational, enduring, and adaptable competencies [4]. We see how insights from this study could contribute to future research with this broader population. For example, we observed that after engaging with a teachable machine, children would be critical of AI if asked to find out how their machines would fail, did little anthropomorphizing of their machines, and aligned more closely with Shneiderman’s call for a shift in design metaphors for AI [20, 90] despite influence by existing technologies and science fiction. This observation strengthens findings in prior work by Druga *et al.* [30] of children’s increased awareness of the potential fallibility of AI systems, after they experimented with teachable machines. However, the children’s shifts toward more nuanced, less techno-utopian views also lead us to reflect whether activities with teachable machines, further enriched with our analytical lenses, could also be structured for meaningful engagement of UX practitioners, who may not have machine learning expertise, in AI problem formulation. For instance, after

interacting with teachable machines, UX practitioners could storyboard with other stakeholders imagined AI systems, map the ideas to Shneiderman’s design metaphors and Rokeach’s value framework, reflect, and iterate. Such activities would build upon prior value-sensitive design by Shilton [89] that established “practices that open new conversations about social values and encourage consensus around those values as design criteria” [89, p. 374].

Another example could be future research focusing on AI practitioners. Machine learning researchers are taking AI-centric approaches to label human values, borrowing from deontological or moral values instead of material values [38, 80]. However, emerging work in value alignment of AI [52] (such that it benefits human society) suggests using material values to guide the “actions of AI agents that are preferable to other actions” because material values are realized through established human value frameworks like the Rokeach Value Survey [85]. We can understand this tension by highlighting the relationship between design metaphors and human values: the same tension exists between “Innovation Goals” and “Science Goals.” We intentionally use the word “lens” when describing problem formulation. By combining design metaphors, or value lenses, we gain different perspectives of the design space. We can see Figure 4 as a reflection of this design space within children’s ideas. When using design metaphors to guide the design robots [27], designers could use Rokeach Values to understand what human values it reflects. For example, we see that children’s ideas with “Control Centers” and “Tele-bots” are related to human values of “Responsibility” and “Obedience.” And so if a robot is designed in contravention of this design space, say, as an “autonomous teammate” but one which fails to communicate that it is “obedient” to the human, we can expect such a system to fail because a user would not feel in control of the robot. Our work may inspire future research into how to weave human values into design metaphors

and how AI practitioners can leverage human value frameworks such as Rokeach's in their optimization objectives.

5.2 Limitations

Our co-design study is exploratory and subjective in nature, thus it is not conclusive. It provides a rich set of observations and insights, generating hypotheses that need further investigation. For instance, to demonstrate the effectiveness of our storyboarding activity over other approaches with teachable machines, such as sketching, prototyping, and brainstorming, a mixed-methods approach is needed. Similarly, demonstrating effectiveness on learning outcomes would require pre-post measurements.

5.2.1 One-to-one child-adult ratio. First, while all participating children had experience training and testing a teachable machine, we had a one-to-one ratio of adult to children co-designers. Such a ratio may not be possible in other AI literacies spaces, such as in a formal classroom where it is likely that only one adult can facilitate. In addition, future studies should explore how problem formulation discussions unfold when there are more children than adults and in a formal learning context.

5.2.2 Acknowledging child development differences. Second, the children participating in this study's problem formulation activity ranged in age from 8-13 years old, and we acknowledge child development differences across this six-year time span. The small number of child partners ($n=10$), while representative of sample sizes in user studies in human-computer interaction [18], did not allow us to substantiate or unpack potential effects of child development. However, as noted above, all the children had experience testing and training teachable machines (figure 2, session 1), so they all had similarly grounded exposure to machine teaching concepts. In addition, all the older children (11-13) and only one 8-year old (Brian) took part in session 2 (the museum-based machine-teaching experience). Anecdotally, we also observed age-related patterns in the children's creations: despite working more closely with the older children, 8-year old Brian's final machine revealed similar values, design metaphors, and overall functionality as the other 8-year old boys.

5.2.3 Limited accessibility. Third, we constructed our printed storyboards and selected other session materials like videos that were included in presentation slides for this activity with children that we knew apriori would be sighted. While both the storyboard and the videos could be made accessible via a screen reader (e.g., convert the storyboard to an accessible slide and provide audio description for the videos), more work is needed to explore equally effective co-design sessions especially for mixed-abilities teams and groups.

6 CONCLUSION

Recent efforts in participatory machine learning [13, 14, 92] call for designers and researchers to carefully consider what we intend by "meaningful participation" and who we involve as partners in our design work. Our work responds to these calls by preparing children – those who may be the most affected by increasingly pervasive AI technologies – as partners in this process. The goal of this study was to expand the level of control and creativity children can exercise with teachable machines by engaging them in a specific problem

formulation exercise. To pursue this goal, we conducted a modified storyboarding session with youth 8-13 years old. Our findings and exploratory insights contribute to the design of learning activities that use teachable machines. Particularly, they could benefit from allowing children to formulate their machine learning problems, using children's values to be usable and enjoyable, and showcasing their utility to support children's goals. We call on future designers and researchers to conduct more studies that involve children as active agents in the design of everyday AI systems imbued with their values.

ACKNOWLEDGMENTS

We thank the KidsTeam children for their ideas, input, and contributions. We also thank graduate students Merijke Coenraad, Nitzan Koren, Zahra Zuzar Dhuliawala, Pooja Bipin Gajera, Naishi Jain, Dinesh Kumar Nanduri, and undergraduate students Anumta Ali and Subhata Sivam for their contributions in conducting the study sessions as adult co-designers of KidsTeam. Utkarsh Dwivedi was partially supported by the Ann G. Wylie Dissertation Fellowship and NSF grant #1816380. Hernisa Kacorri was partially supported by NSF grants #1816380, #1955568, and #2229885.

REFERENCES

- [1] AAAI. 2021. Introducing AI. <http://modelai.gettyburg.edu/2021/intro/>.
- [2] AAAI. 2021. Rushhour: Designing and comparing A* heuristics for a children's puzzle. <http://modelai.gettyburg.edu/2021/rushhour>.
- [3] AAAI. 2022. Model AI Assignments. <http://modelai.gettyburg.edu/>.
- [4] Oguz A. Acar. 2023. AI Prompt Engineering Isn't the Future. <https://hbr.org/2023/06/ai-prompt-engineering-isnt-the-future>
- [5] Open AI. 2022. Introducing ChatGPT. <https://openai.com/blog/chatgpt>.
- [6] Amazon. 2022. Formulating the problem. <https://docs.aws.amazon.com/machine-learning/latest/dg/formulating-the-problem.html>.
- [7] Ronald E. Anderson. 1978. Value orientation of computer science students. *Commun. ACM* 21, 3 (March 1978), 219–225. <https://doi.org/10.1145/359361.359365>
- [8] Peter M. Andreae and John H. Andreae. 1978. A teachable machine in the real world. *International Journal of Man-Machine Studies* 10, 3 (1978), 301 – 312. [https://doi.org/10.1016/S0020-7373\(78\)80048-0](https://doi.org/10.1016/S0020-7373(78)80048-0)
- [9] Brenna Argall, S. Chernova, Manuela M. Veloso, and Brett Browning. 2009. A survey of robot learning from demonstration. *Robotics Auton. Syst.* 57 (2009), 469–483. <https://api.semanticscholar.org/CorpusID:1045325>
- [10] Donald Martin Jr. au2, Vinodkumar Prabhakaran, Jill Kuhlberg, Andrew Smart, and William S. Isaac. 2020. Participatory Problem Formulation for Fairer Machine Learning Through Community Based System Dynamics. *arXiv:2005.07572 [cs.CY]*
- [11] Erin Beneteau, Ashley Boone, Yuxing Wu, Julie A. Kientz, Jason Yip, and Alexis Hiniker. 2020. *Parenting with Alexa: Exploring the Introduction of Smart Speakers on Family Dynamics*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi-org.proxy-um.researchport.umd.edu/10.1145/3313831.3376344>
- [12] Abeba Birhane, William Isaac, Vinodkumar Prabhakaran, Mark Diaz, Madeleine Clare Elish, Jason Gabriel, and Shakir Mohamed. 2022. Power to the People? Opportunities and Challenges for Participatory AI. In *Equity and Access in Algorithms, Mechanisms, and Optimization* (Arlington, VA, USA) (EAAMO '22). Association for Computing Machinery, New York, NY, USA, Article 6, 8 pages. <https://doi.org/10.1145/3551624.3555290>
- [13] Abeba Birhane, William Samuel Isaac, Vinodkumar Prabhakaran, Mark Diaz, Madeleine Clare Elish, Jason Gabriel, and Shakir Mohamed. 2022. Frameworks and Challenges to Participatory AI. <https://arxiv.org/pdf/2209.07572.pdf>
- [14] Elizabeth Bondi, Lily Xu, Diana Acosta-Navas, and Jackson A. Killian. 2021. Envisioning Communities: A Participatory Approach Towards AI for Social Good. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) (AI/ES '21). Association for Computing Machinery, New York, NY, USA, 425–436. <https://doi.org/10.1145/3461702.3462612>
- [15] Elizabeth Bonsignore, June Ahn, Tamara Clegg, Mona Leigh Guha, Jason C. Yip, and Allison Druin. 2013. Embedding Participatory Design into Designs for Learning: An Untapped Interdisciplinary Resource?. In *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013*

- Conference Proceedings Volume 1 — Full Papers & Symposia.*, Vol. 1. International Society of the Learning Sciences, Madison, WI, 549–556.
- [16] Danielle Bragg, Nicholas Huynh, and Richard E. Ladner. 2016. A Personalizable Mobile Sound Detector App Design for Deaf and Hard-of-Hearing Users. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) (ASSETS '16). Association for Computing Machinery (ACM), New York, NY, USA, 3–13. <https://doi.org/10.1145/2982142.2982171>
 - [17] CSET Policy Brief. 2019. *AI Safety, Security, and Stability Among Great Powers: Options, Challenges, and Lessons Learned for Pragmatic Engagement*. Technical Report. CSET Policy, Georgetown University. <https://doi.org/10.51593/20190051>
 - [18] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 981–992. <https://doi.org/10.1145/2858036.2858498>
 - [19] 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 Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI EA '20). Association for Computing Machinery (ACM), New York, NY, USA, 1–8. <https://doi.org/10.1145/3334480.3382839>
 - [20] Joel Chan, Hal Daumé III, John P. Dickerson, Hernisa Kacorri, and Ben Shneiderman. 2021. Supporting human flourishing by ensuring human involvement in AI-infused systems. In *Human Centered AI Workshop, 35th Conference on Neural Information Processing Systems (NeurIPS 2021)*. NeurIPS, Virtual, 5 pages. https://drive.google.com/file/d/1o2Rk8qrciy2_vfuopEoiyVwvK6LK9Cf/view
 - [21] Vincent Cheng and Yu Zhang. 2023. Analyzing ChatGPT's Mathematical Deficiencies: Insights and Contributions. In *Proceedings of the 35th Conference on Computational Linguistics and Speech Processing (ROCLING 2023)*. The Association for Computational Linguistics and Chinese Language Processing (ACLCLP), Taipei City, Taiwan, 188–193.
 - [22] Victoria Clarke, Virginia Braun, and Nikki Hayfield. 2015. Thematic analysis. *Qualitative psychology: A practical guide to research methods* 222 (2015), 248.
 - [23] Tamara Clegg, Elizabeth Bonsignore, Jason Yip, Helene Gelderblom, Alex Kuhn, Tobin Valenstein, Becky Lewittes, and Allison Druin. 2012. Technology for promoting scientific practice and personal meaning in life-relevant learning. In *Proceedings of the 11th International Conference on Interaction Design and Children*. ACM, Bremen, Germany, 152–161.
 - [24] Code.org. 2023. Hour of Code. <https://hourofcode.com/us/learn>.
 - [25] Fast Company. 2020. Tired of saying 'Hey Google' and 'Alexa'? Change it up with these unintentional wake words. <https://www.fastcompany.com/90524070/tired-of-saying-hey-google-and-alexa-change-it-up-with-these-unintentional-wake-words>.
 - [26] Fernando Delgado, Solon Barocas, and Karen Levy. 2022. An Uncommon Task: Participatory Design in Legal AI. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 51 (apr 2022), 23 pages. <https://doi.org/10.1145/3512898>
 - [27] Nathaniel Dennler, Changxiao Ruan, Jessica Hadiwijoyo, Brenna Chen, Stefanos Nikolaidis, and Maja Mataric. 2022. Using Design Metaphors to Understand User Expectations of Socially Interactive Robot Embodiments.
 - [28] Sascha Dickel and Jan-Felix Schrape. 2017. The Logic of Digital Utopianism. *NanoEthics* 11, 1 (April 2017), 47–58. <https://doi.org/10.1007/s11569-017-0285-6>
 - [29] Stefania Druga, Fee Lia Christoph, and Amy J Ko. 2022. Family as a Third Space for AI Literacies: How do children and parents learn about AI together?. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans, Louisiana, USA, 1–17.
 - [30] Stefania Druga and Amy J Ko. 2021. How Do Children's Perceptions of Machine Intelligence Change When Training and Coding Smart Programs?. In *Proceedings of the 20th Annual ACM Interaction Design and Children Conference* (Athens, Greece) (IDC '21). Association for Computing Machinery, New York, NY, USA, 49–61. <https://doi.org/10.1145/3459990.3460712>
 - [31] Stefania Druga, Jason Yip, Michael Preston, and Devin Dillon. 2021. *The 4As: Ask, Adapt, Author, Analyze - AI Literacy Framework for Families*. MIT Press, Boston, Massachusetts, USA, Chapter 10, 193–231. <https://doi.org/10.7551/mitpress/13654.003.0014> <https://wip.mitpress.mit.edu/pub/the-4as>.
 - [32] Allison Druin. 1999. Cooperative inquiry: developing new technologies for children with children. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*. ACM, Pittsburgh, Pennsylvania, United States, 592–599. <https://doi.org/10.1145/302979.303166>
 - [33] Allison Druin. 2002. The role of children in the design of new technology. *Behaviour & Information Technology* 21, 1 (2002), 1–25. <https://doi.org/10.1080/01442900110108659>
 - [34] Utkarsh Dwivedi, Jaina Gandhi, Raj Parikh, Merijke Coenraad, Elizabeth Bonsignore, and Hernisa Kacorri. 2021. Exploring Machine Teaching with Children. In *2021 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*. IEEE, St. Louis, Missouri, USA, 1–11. <https://doi.org/10.1109/VL/HCC51201.2021.9576171>
 - [35] Catherine D'Ignazio and Rahul Bhargava. 2015. Approaches to building big data literacy. , 6 pages. <https://api.semanticscholar.org/CorpusID:13689606>
 - [36] Albert Einstein and Leopold Infeld. 1938. *Evolution of physics*. Vol. 10. Simon and Schuster, New York, USA.
 - [37] Salma Elsayed-Ali, Elizabeth Bonsignore, Hernisa Kacorri, and Mega Subramaniam. 2020. Designing for Children's Values: Conceptualizing Value-Sensitive Technologies with Children. In *ACM Interaction Design and Children Conference: Extended Abstracts* (London, United Kingdom) (IDC 2020). Association for Computing Machinery (ACM), New York, NY, USA, 296–301. <https://doi.org/10.1145/3397617.3397826>
 - [38] Denis Emelin, Ronan Le Bras, Jena D Hwang, Maxwell Forbes, and Yejin Choi. 2020. Moral stories: Situated reasoning about norms, intents, actions, and their consequences.
 - [39] Eva Eriksson, Elisabet M. Nilsson, Anne-Marie Hansen, and Tilde Bekker. 2022. Teaching for Values in Human-Computer Interaction. *Frontiers in Computer Science* 4 (Feb. 2022), 830736. <https://doi.org/10.3389/fcomp.2022.830736>
 - [40] Jerry Alan Fails, Mona Leigh Guha, and Allison Druin. 2012. Methods and Techniques for Involving Children in the Design of New Technology for Children. *Foundations and Trends® in Human-Computer Interaction* 6, 2 (2012), 85–166. <https://doi.org/10.1561/11000000018>
 - [41] Jerry Alan Fails, Mona Leigh Guha, and Allison Druin. 2012. Methods and Techniques for Involving Children in the Design of New Technology for Children. *Foundations and Trends® in Human-Computer Interaction* 6, 2 (2012), 85–166. <https://doi.org/10.1561/11000000018>
 - [42] Maria Angela Ferrario, Will Simm, Stephen Forshaw, Adrian Gradinar, Marcia Tavares Smith, and Ian Smith. 2016. Values-first SE: research principles in practice. In *Proceedings of the 38th International Conference on Software Engineering Companion - ICSE '16*. ACM Press, Austin, Texas, 553–562. <https://doi.org/10.1145/2889160.2889219>
 - [43] Rebecca Fiebrink. 2009. Wekinator | Software for real-time, interactive machine learning. <http://www.wekinator.org/>
 - [44] Batya Friedman. 1996. Value-sensitive Design.
 - [45] Mahtab Ghazizadeh, John D. Lee, and Linda Ng Boyle. 2012. Extending the Technology Acceptance Model to assess automation. *Cognition, Technology & Work* 14, 1 (March 2012), 39–49. <https://doi.org/10.1007/s10111-011-0194-3>
 - [46] Vinicius G. Goecks, Gregory M. Gremillion, Vernon J. Lawhern, John Valasek, and Nicholas R. Waytowich. 2020. Integrating Behavior Cloning and Reinforcement Learning for Improved Performance in Dense and Sparse Reward Environments. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems* (Auckland, New Zealand) (AAMAS '20). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 465–473.
 - [47] Google. 2021. Project Euphonia: Google Research Initiative. <https://sites.research.google/euphonia/about/>.
 - [48] Google. 2022. Formulate Your Problem as an ML Problem. <https://developers.google.com/machine-learning/problem-framing/formulate>.
 - [49] Google. 2022. Teachable Machine. <https://teachablemachine.withgoogle.com/>.
 - [50] Mona Leigh Guha, Allison Druin, and Jerry Alan Fails. 2013. Cooperative Inquiry revisited: Reflections of the past and guidelines for the future of intergenerational co-design. *International Journal of Child-Computer Interaction* 1, 1 (Jan. 2013), 14–23. <https://doi.org/10.1016/j.ijcci.2012.08.003>
 - [51] Mona Leigh Guha, Allison Druin, and Jerry Alan Fails. 2013. Cooperative Inquiry revisited: Reflections of the past and guidelines for the future of intergenerational co-design. *International Journal of Child-Computer Interaction* 1, 1 (Jan. 2013), 14–23. <https://doi.org/10.1016/j.ijcci.2012.08.003>
 - [52] Shengnan Han, Eugene Kelly, Shahrokh Nikou, and Eric-Oluf Svec. 2022. Aligning artificial intelligence with human values: reflections from a phenomenological perspective. *AI & SOCIETY* 37, 4 (Dec. 2022), 1383–1395. <https://doi.org/10.1007/s00146-021-01247-4>
 - [53] Helen Ai He, Saul Greenberg, and Elaine M. Huang. 2010. One size does not fit all: applying the transtheoretical model to energy feedback technology design. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 927. <https://doi.org/10.1145/1753326.1753464>
 - [54] Tom Hitron, Yoav Orlev, Iddo Wald, Ariel Shamir, Hadas Erel, and Oren Zuckerman. 2019. Can Children Understand Machine Learning Concepts? The Effect of Uncovering Black Boxes. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery (ACM), New York, NY, USA, 1–11. <https://doi.org/10.1145/3290605.3300645>
 - [55] Tom Hitron, Iddo Wald, Hadas Erel, and Oren Zuckerman. 2018. Introducing children to machine learning concepts through hands-on experience. In *Proceedings of the ACM Conference on Interaction Design and Children*. Association for Computing Machinery (ACM), Trondheim, Norway, 563–568. <https://doi.org/10.1145/3202185.3210776>
 - [56] Geert Hofstede. 2001. *Culture's consequences: comparing values, behaviors, institutions, and organizations across nations* (2. ed., [nachdr.] ed.). Sage Publ, Thousand Oaks, Calif. OCLC: 711858494.

- [57] Minna Isomursu, Mari Ervasti, Marianne Kinnula, and Pekka Isomursu. 2009. Examining human values in adopting ubiquitous technology in school. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '09*. ACM Press, Bonn, Germany, 1. <https://doi.org/10.1145/1613858.1613933>
- [58] Ole Sejer Iversen and Tuck W Leong. 2012. Values-led participatory design: mediating the emergence of values. *NordiCHI 7* (2012), 468–477. <https://doi.org/10.1145/2399016.2399087>
- [59] Ole Sejer Iversen, Tuck W. Leong, Peter Wright, Judith Gregory, and Geoff Bowker. 2012. Working with human values in design. In *Proceedings of the 12th Participatory Design Conference on Exploratory Papers Workshop Descriptions Industry Cases - Volume 2 - PDC '12*. ACM Press, Roskilde, Denmark, 143. <https://doi.org/10.1145/2348144.2348191>
- [60] Dhruv Jain, Khoa Huynh Anh Nguyen, Steven M. Goodman, Rachel Grossman-Kahn, Hung Ngo, Aditya Kusupati, Ruofei Du, Alex Olwal, Leah Findlater, and Jon E. Froehlich. 2022. ProtoSound: A Personalized and Scalable Sound Recognition System for Deaf and Hard-of-Hearing Users. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 305, 16 pages. <https://doi.org/10.1145/3491102.3502020>
- [61] Hernisa Kacorri. 2017. Teachable Machines for Accessibility. *SIGACCESS Access. Comput.* 19, 119 (Nov. 2017), 10–18. <https://doi.org/10.1145/3167902.3167904>
- [62] Bjørn Karmann. 2017. Objectifier - Spacial Programming (User testing). <https://youtu.be/3a825NjMLjk>
- [63] Saba Kawas, Ye Yuan, Akeiyah DeWitt, Qiao Jin, Susanne Kirchner, Abigail Bilger, Ethan Grantham, Julie A Kientz, Andrea Tartaro, and Svetlana Yarosh. 2020. Another decade of IDC research: examining and reflecting on values and ethics. In *Proceedings of the Interaction Design and Children Conference*. ACM, London United Kingdom, 205–215. <https://doi.org/10.1145/3392063.3394436>
- [64] Alex Kayal, Willem-Paul Brinkman, Mark A. Neerinx, and M. Birna Van Riemsdijk. 2018. Automatic Resolution of Normative Conflicts in Supportive Technology Based on User Values. *ACM Transactions on Internet Technology* 18, 4 (Nov. 2018), 1–21. <https://doi.org/10.1145/3158371>
- [65] Young-Ho Kim, Diana Chou, Bongshin Lee, Margaret Danilovich, Amanda Lazar, David E. Conroy, Hernisa Kacorri, and Eun Kyoung Choe. 2022. My-Move: Facilitating Older Adults to Collect In-Situ Activity Labels on a Smart-watch with Speech. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 416, 21 pages. <https://doi.org/10.1145/3491102.3517457>
- [66] Yo-whan Kim, Samarth Mishra, SouYoung Jin, Rameswar Panda, Hilde Kuehne, Leonid Karlinsky, Venkatesh Saligrama, Kate Saenko, Aude Oliva, and Rogerio Feris. 2022. How Transferable are Video Representations Based on Synthetic Data? *Advances in Neural Information Processing Systems* 35 (2022), 35710–35723.
- [67] Priya Kumar, Jessica Vitak, Marshini Chetty, Tamara L. Clegg, Jonathan Yang, Brenna McNally, and Elizabeth Bonsignore. 2018. Co-designing online privacy-related games and stories with children. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*. ACM, Trondheim Norway, 67–79. <https://doi.org/10.1145/3202185.3202735>
- [68] Manisha Kusuma, Vikram Mohanty, Marx Wang, and Kurt Luther. 2022. Civil War Twin: Exploring Ethical Challenges in Designing an Educational Face Recognition Application. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society* (Oxford, United Kingdom) (AI/ES '22). Association for Computing Machinery, New York, NY, USA, 369–384. <https://doi.org/10.1145/3514094.3534141>
- [69] Google Creative Lab. 2017. Teachable Machine. <https://teachablemachine.withgoogle.com/v1/>
- [70] Duri Long and Brian Magerko. 2020. What is AI Literacy? Competencies and Design Considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–16. <https://doi.org/10.1145/3313831.3376727>
- [71] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (CHI '16). Association for Computing Machinery, New York, NY, USA, 5286–5297. <https://doi.org/10.1145/2858036.2858288>
- [72] Kenneth R MacCrimmon and Ronald N Taylor. 1976. Decision making and problem solving. *Handbook of industrial and organizational psychology* 1, 976 (1976), 1397–1463.
- [73] Sheron L. Mark. 2016. Psychology of Working Narratives of STEM Career Exploration for Non-dominant Youth. *Journal of Science Education and Technology* 25, 6 (Dec. 2016), 976–993. <https://doi.org/10.1007/s10956-016-9646-0>
- [74] Brenna McNally, Mona Leigh Guha, Matthew Louis Mauriello, and Allison Druin. 2016. Children's Perspectives on Ethical Issues Surrounding Their Past Involvement on a Participatory Design Team. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 3595–3606. <https://doi.org/10.1145/2858036.2858338>
- [75] Neema Moraveji, Jason Li, Jiarong Ding, Patrick O'Kelley, and Suze Woolf. 2007. Comicboarding: using comics as proxies for participatory design with children. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 1371–1374. <https://doi.org/10.1145/1240624.1240832>
- [76] Davoud Mougouei, Harsha Perera, Waqar Hussain, Rifat Shams, and Jon Whittle. 2018. Operationalizing human values in software: a research roadmap. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering - ESEC/FSE 2018*. ACM Press, Lake Buena Vista, FL, USA, 780–784. <https://doi.org/10.1145/3236024.3264843>
- [77] Michael J. Muller, Cathleen Wharton, William J. McIver, and Lila Laux. 1997. Toward an HCI research and practice agenda based on human needs and social responsibility. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. ACM, Atlanta Georgia USA, 155–161. <https://doi.org/10.1145/258549.258640>
- [78] J. Michael Munson and Barry Z. Posner. 1979. The values of engineers and managing engineers. *IEEE Transactions on Engineering Management* EM-26, 4 (1979), 94–100. <https://doi.org/10.1109/TEM.1979.6447357>
- [79] Todd Neller, John DeNero, Dan Klein, Sven Koenig, William Yeoh, Xiaoming Zheng, Kenny Daniel, Alex Nash, Zachary Dodds, Giuseppe Carenini, et al. 2010. Model AI assignments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 24. AAAI, Georgia, USA, 1919–1921.
- [80] Alexander Pan, Jun Shern Chan, Andy Zou, Nathaniel Li, Steven Basart, Thomas Woodside, Hanlin Zhang, Scott Emmons, and Dan Hendrycks. 2023. Do the rewards justify the means? measuring trade-offs between rewards and ethical behavior in the MACHIAVELLI benchmark. In *Proceedings of the 40th International Conference on Machine Learning (ICML'23)*. JMLR.org, Honolulu, Hawaii, USA, Article 1117, 31 pages.
- [81] Samir Passi and Solon Barocas. 2019. Problem Formulation and Fairness. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, Atlanta GA USA, 39–48. <https://doi.org/10.1145/3287560.3287567>
- [82] Rupal Patel and Deb Roy. 1998. Teachable interfaces for individuals with dysarthric speech and severe physical disabilities. In *Proceedings of the AAAI Workshop on Integrating Artificial Intelligence and Assistive Technology*. AAAI, Wisconsin, USA, 40–47.
- [83] Vinodkumar Prabhakaran and Donald Jr Martin. 2020. Participatory Machine Learning Using Community-Based System Dynamics. *Health Hum Rights* 22, 2 (Dec 2020), 71–74.
- [84] Nitin Rane. 2023. Enhancing Mathematical Capabilities through ChatGPT and Similar Generative Artificial Intelligence: Roles and Challenges in Solving Mathematical Problems. *Available at SSRN 4603237* 0, 0 (2023), 1–9. <https://doi.org/10.2139/ssrn.4603237>
- [85] Milton Rokeach. 1973. *The nature of human values*. Free Press, New York, NY, US. Pages: x, 438.
- [86] Shalom H. Schwartz. 1992. Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. In *Advances in Experimental Social Psychology*, Vol. 25. Elsevier, Amsterdam, Netherlands, 1–65. [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6)
- [87] Donald A. Schön. 1983. *The reflective practitioner: How professionals think in action*. Basic books, New York, USA.
- [88] Mark P Sendak, William Ratliff, Dina Sarro, Elizabeth Alderton, Joseph Futoma, Michael Gao, Marshall Nichols, Mike Revoir, Faraz Yashar, Corinne Miller, et al. 2020. Real-world integration of a sepsis deep learning technology into routine clinical care: implementation study. *JMIR medical informatics* 8, 7 (2020), e15182.
- [89] Katie Shilton. 2013. Values Levers: Building Ethics into Design. *Science, Technology, & Human Values* 38, 3 (May 2013), 374–397. <https://doi.org/10.1177/0162243912436985>
- [90] Ben Shneiderman. 2022. *Human-Centered AI*. Oxford University Press, Oxford, UK.
- [91] Helle Marie Skovbjerg, Tilde Bekker, and Wolmet Barendregt. 2016. Being Explicit about Underlying Values, Assumptions and Views when Designing for Children in the IDC Community. In *Proceedings of the The 15th International Conference on Interaction Design and Children - IDC '16*. ACM Press, Manchester, United Kingdom, 713–719. <https://doi.org/10.1145/2930674.2932224>
- [92] Mona Sloane, Emanuel Moss, Olaitan Awomolo, and Laura Forlano. 2022. Participation Is Not a Design Fix for Machine Learning. In *Equity and Access in Algorithms, Mechanisms, and Optimization* (Arlington, VA, USA) (EAAMO '22). Association for Computing Machinery, New York, NY, USA, Article 1, 6 pages. <https://doi.org/10.1145/3551624.3555285>
- [93] Joan Sosa-García and Francesca Odono. 2017. "Hands On" Visual Recognition for Visually Impaired Users. *ACM Trans. Access. Comput.* 10, 3, Article 8 (aug 2017), 30 pages. <https://doi.org/10.1145/3060056>
- [94] Katta Spiel, Emeline Brulé, Christopher Frauenberger, Gilles Bailly, and Geraldine Fitzpatrick. 2018. Micro-ethics for participatory design with marginalised children. In *Proceedings of the 15th Participatory Design Conference: Full Papers - Volume 1*. ACM, Hasselt and Genk Belgium, 1–12. <https://doi.org/10.1145/3210586.3210603>

- [95] Mike Tissenbaum, Josh Sheldon, and Hal Abelson. 2019. From computational thinking to computational action. *Commun. ACM* 62, 3 (Feb. 2019), 34–36. <https://doi.org/10.1145/3265747>
- [96] David Touretzky. 2019. AI for K-12. https://github.com/touretzkyds/ai4k12/blob/master/documents/CSTA2020_Learning_Activities_K-5.pdf.
- [97] University of Maryland, College Park and Ben Shneiderman. 2020. Human-Centered Artificial Intelligence: Three Fresh Ideas. *AIS Transactions on Human-Computer Interaction* 12, 3 (2020), 109–124. <https://doi.org/10.17705/1thci.00131>
- [98] Maarten Van Mechelen, Gavin Sim, Bieke Zaman, Peggy Gregory, Karin Slegers, and Matthew Horton. 2014. Applying the CHECK tool to participatory design sessions with children. In *Proceedings of the 2014 conference on Interaction design and children*. ACM, Aarhus Denmark, 253–256. <https://doi.org/10.1145/2593968.2610465>
- [99] Henriikka Vartiainen, Matti Tedre, and Teemu Valtonen. 2020. Learning machine learning with very young children: Who is teaching whom? *International Journal of Child-Computer Interaction* 25 (Sept. 2020), 1–11. <https://doi.org/10.1016/j.ijcci.2020.100182>
- [100] Henriikka Vartiainen, Tapani Toivonen, Ilkka Jormanainen, Juho Kahila, Matti Tedre, and Teemu Valtonen. 2021. Machine learning for middle schoolers: Learning through data-driven design. *International Journal of Child-Computer Interaction* 29 (2021), 100281. <https://doi.org/10.1016/j.ijcci.2021.100281>
- [101] Peter-Paul Verbeek. 2011. *Moralizing technology: understanding and designing the morality of things*. The University of Chicago Press, Chicago ; London.
- [102] Amy Voids and Elizabeth D. Mynatt. 2005. Conveying user values between families and designers. In *CHI '05 extended abstracts on Human factors in computing systems - CHI '05*. ACM Press, Portland, OR, USA, 2013. <https://doi.org/10.1145/1056808.1057080>
- [103] Roger J. Volkema. 1983. Problem Formulation in Planning and Design. *Management Science* 29, 6 (1983), 639–652. <https://doi.org/10.1287/mnsc.29.6.639> arXiv:<https://doi.org/10.1287/mnsc.29.6.639>
- [104] Yousef Wardat, Mohammad A Tashtoush, Rommel AlAli, and Adeeb M Jarrah. 2023. ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education* 19, 7 (2023), em2286.
- [105] Randi Williams, Safinah Ali, Nisha Devasia, Daniella DiPaola, Jenna Hong, Stephen P Kaputsos, Brian Jordan, and Cynthia Breazeal. 2023. AI+ ethics curricula for middle school youth: Lessons learned from three project-based curricula. *International Journal of Artificial Intelligence in Education* 33, 2 (2023), 325–383.
- [106] Julia Woodward, Zari McFadden, Nicole Shiver, Amir Ben-hayon, Jason C. Yip, and Lisa Anthony. 2018. Using Co-Design to Examine How Children Conceptualize Intelligent Interfaces. In *Proceedings of ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery (ACM), Montreal QC, Canada, 1–14. <https://doi.org/10.1145/3173574.3174149>
- [107] Qian Yang, Jina Suh, Nan-Chen Chen, and Gonzalo Ramos. 2018. Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models. In *Proceedings of the ACM Conference on Designing Interactive Systems Conference*. Association for Computing Machinery (ACM), Hong Kong, China, 573–584. <https://doi.org/10.1145/3196709.3196729>
- [108] Svetlana Yarosh, Iulian Radu, Seth Hunter, and Eric Rosenbaum. 2011. Examining values: an analysis of nine years of IDC research. In *Proceedings of the 10th International Conference on Interaction Design and Children - IDC '11*. ACM Press, Ann Arbor, Michigan, 136–144. <https://doi.org/10.1145/1999030.1999046>
- [109] Nur Yildirim, Mahima Pushkarna, Nitesh Goyal, Martin Wattenberg, and Fernanda Viégas. 2023. Investigating How Practitioners Use Human-AI Guidelines: A Case Study on the People + AI Guidebook. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 356, 13 pages. <https://doi.org/10.1145/3544548.3580900>
- [110] Jason C. Yip, Kiley Sobel, Caroline Pitt, Kung Jin Lee, Sijin Chen, Kari Nasu, and Laura R. Pina. 2017. Examining Adult-Child Interactions in Intergenerational Participatory Design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 5742–5754. <https://doi.org/10.1145/3025453.3025787>
- [111] Helen Zhang, Irene Lee, Safinah Ali, Daniella DiPaola, Yihong Cheng, and Cynthia Breazeal. 2023. Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory study. *International Journal of Artificial Intelligence in Education* 33, 2 (2023), 290–324.
- [112] Xiaojin Zhu, Adish Singla, Sandra Zilles, and Anna N. Rafferty. 2018. An Overview of Machine Teaching. *arXiv preprint NA*, NA (2018), 1–18. <http://arxiv.org/abs/1801.05927>