Scaffolding Children's Sensemaking around Algorithmic Fairness

Jean Salac
University of Washington
Seattle, WA, USA
salac@uw.edu

Rotem Landesman
University of Washington
Seattle, WA, USA
roteml@uw.edu

Stefania Druga
University of Washington
Seattle, WA, USA
st3f@uw.edu

Amy J. Ko
University of Washington
Seattle, WA, USA
ajko@uw.edu

ABSTRACT

Prior research has investigated children's perceptions of algorithmic bias, but provides little guidance on engaging children in conversations on algorithmic bias that center their agency and well-being. To address this, we developed discussions and design activities based on three scenarios of algorithmic (un)fairness. We conducted these discussions and activities with 16 children (ages 8-12) in the US, and examined our data using qualitative thematic analysis. Grounded in lived experiences and situated knowledge, participants were capable of reasoning around both explicit and implicit effects of algorithmic bias. Participants also expressed distrust of technology, doubting technology's abilities and preferring human approaches to resolve unfairness. This work contributes (1) a more nuanced understanding of children's situated reasoning of technology, suggesting their potential for critical engagement and (2) a blueprint for engaging children in scaffolded yet open-ended sensemaking around algorithmic fairness, informing the design of tools, curricula, and other learning experiences for children.

CCS CONCEPTS

• **Social and professional topics → Children**.

KEYWORDS

algorithmic fairness, children, sensemaking, funds of knowledge

ACM Reference Format:
Jean Salac, Rotem Landesman, Stefania Druga, and Amy J. Ko. 2023. Scaffolding Children's Sensemaking around Algorithmic Fairness. In *Interaction Design and Children (IDC '23), June 19–23, 2023, Chicago, IL, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3585088.3589379

1 INTRODUCTION

As computing becomes pervasive in children's lives, it brings extensive benefits, but also puts children's lives, families, communities, and futures at increased risk of harm. This has led to a flourishing body of work on youth's perceptions of algorithmic bias, focusing mainly on interactions with AI agents or AI learning experiences. Several scholars have found that children tend to overly trust AI agents, precluding them from critiquing AI technologies [20, 31, 42, 58, 62, 63]. In contrast, others observed that children could identify unfair treatment from AI with some instruction, suggesting some capacity to reason about algorithmic bias [22].

While most prior studies targeted algorithmic bias within AI systems, biases from algorithms in other computing technologies also affect the lives of children. Researchers have only recently extended beyond AI, with Coenraad et al. finding that even without instruction, youth were aware of visible negative effects of technology, such as non-consensual data collection and use [14]. Although prior work has characterized childrens' perceptions of algorithmic bias in AI, they do not provide much guidance on *how to engage* children in conversations around algorithmic biases that center their perspectives while safeguarding their well-being around a possibly difficult topic. This guidance is especially pertinent and timely with increasing calls to educate children on the social and ethical impacts of technology [11, 36, 64].

Based on the funds of knowledge position [28, 47] and sensemaking theory [17], in our prior work, we developed discussions and design activities to scaffold children in making sense of algorithmic fairness[1], and explored them with adolescents (ages 15-17) [55]. We found that adolescents introduced many factors that were not included in our prompts, such as lived experiences and power dynamics, into their sensemaking practices. Through this observed process, adolescents developed rich characterizations of algorithmic bias impacts by drawing from their funds of knowledge.

In this paper, we build upon prior work on youth perceptions of fairness in AI, as well as our own work [55] on scaffolding adolescent sensemaking around algorithmic fairness in both AI and other computing technologies, to examine youth sensemaking of algorithmic fairness more generally. Prior work in moral development indicates that children first reason about morality from a more egocentric stance and then learn to reason from others' perspectives as they grow older [9], suggesting interesting potential differences between youth and adolescents. Therefore, we pursued the following research questions:

[1]With our participants, we use the term "fairness" instead of "bias" because prior work [22] showed that "bias" might not be in youth's vocabulary. We use "fairness" when specifically discussing with our participants but use the terms interchangeably elsewhere in the paper.

IDC '23, June 19–23, 2023, Chicago, IL, USA

ACM ISBN 979-8-4007-0131-3/23/06.
https://doi.org/10.1145/3585088.3589379

(1) What funds of knowledge might children use to make sense of algorithmic fairness?
(2) How might the ages, identities, and backgrounds of children shape their sensemaking of algorithmic fairness?

After analyzing sensemaking discussions and design activities with 16 children (ages 8-12) in the US, we make two important contributions through this study. First, we contribute to a deeper understanding of children's situated knowledge around algorithmic bias in computing more broadly, not only AI. This suggests potential entryways, such as their own interpersonal relationships, for critical engagement with technology. Second, we contribute a blueprint for engaging children in scaffolded yet open-ended reasoning around algorithmic fairness, informing the design of tools, curricula, and other learning experiences in the growing movement to educate children on technology's social and ethical impacts.

2 BACKGROUND

To examine how children approach and view algorithmic fairness and its intricacies, it is important to first consider existing theories on children's moral development and its progression with children's age and experience. We then take a closer look at the gradual process of understanding the concept of fairness and the theoretical foundations behind our decision to consider children's context as a crucial factor in this study.

2.1 Children's Moral Development and Perceptions of Fairness

Chapman and Caperndale [12] proposed an interpretation of Kohlberg's general framework of moral development stages that emphasizes the process of opinion and stance formation to be rooted in action. According to them, children at different developmental stages construct moral values by internalizing their contextualized actions to develop moral structures. Therefore, children may develop differently across contexts based on their life experiences.

Furthermore, Piaget emphasizes the egocentric continuum in a child's moral development [9]. He posited that children start by reasoning from a more egocentric viewpoint, then by getting to know themselves in relation to the world around them, progress to a less egocentric view of values and morality. In Piaget's structural framework, this view would explain the tendency of younger children (up to the age of 11) to characterize morality based on outcomes rather than intentions, treating rules inflexibly and absolutely. Then, at an older age, children begin to see beyond themselves and may view situations from others' perspectives, using that context to decide if a rule is right or wrong.

Kohlberg [37] also worked to extend Dewey's work of cognitive moral development by differentiating between (1) the preconventional level of morality when a child's recognition of a good or bad action is directly tied to its consequences and their physical manifestations, such as punishments and rewards, and (2) the conventional level of morality, where older children nuance their moral reasoning while incorporating the moral codes prevalent in the adult society surrounding them. Therefore, it is important in our work to recognize our participants' ongoing development of moral standards, considering their stage based on their age and contextual backgrounds.

As a subset of moral development, the concept of fairness has been shown to mature alongside other values children acquire throughout their lives gradually. Children seem to have a deep aversion to inequality. They can identify unfair behavior from a young age [59] while tending to prefer equity over equality the older they become [4, 33, 51, 57]. While distinguishing between distributive fairness (favoring the results) and procedural fairness (recognizing the fairness in the decision-making process) occurs at a young age, research has shown that younger subsets of children favor fair processes over fair outcomes. In contrast, older children tend to favor distributive fairness, looking closely at the outcomes of an action to determine its fairness, even if the process leading to it was not 'fair' [23, 29]. Regarding actions, children older than eight were seen to be more likely to act upon perceived unfairness, while younger children are more likely only to recognize, but not act upon it [7, 8, 45].

2.2 Children's Perceptions of Algorithmic Fairness

With respect to children's understanding of algorithmic fairness, a wealth of literature identified children's perceptions almost exclusively within the field of AI and agent interactions. In this work, we focus both on children's direct interactions with computing and an 'ambient' computation that may affect children more indirectly in our discussions and activities. Reviews of the field [42] identified several studies that suggest that children often overestimate agent intelligence [21] and consequently overly trust agents [20, 31, 62, 63]. Skinner [58] similarly found that children equated kindness with fairness in AI agents, using kind communication with people to justify fairness. Further, Kim et al. observed that children expressed technosolutionistic and amoral preconceptions of AI [35]. Despite this, Druga et al. [22] have shown that after showing them videos of algorithmic bias examples, children could connect those examples to their daily lives, identifying situations of unfair treatment from AI based on race/ethnicity, age, and gender.

Research into children's perceptions of algorithmic fairness that extends beyond AI to other computing technologies more broadly is nascent at best. Coenraad [14], for example, discovered that without instruction, youth demonstrated an awareness of visible negative impacts of technology more broadly, not only AI, and were able to provide examples of this bias within their lives. As educators and researchers increase efforts to educate children in critical computing literacies [5, 13, 48], this study offers a blueprint of how to leverage children's knowledge and backgrounds towards developing a more robust moral sensitivity to the complexities of algorithmic fairness.

2.3 Funds of Knowledge & Sensemaking Theory

We draw from funds of knowledge and sensemaking theories to support children in bringing their conceptions of fairness into computing. The funds of knowledge approach posits that learners already have various skills, knowledge, and competencies from their lives and their communities [28, 47]. This approach asserts that these assets are frequently invisible because of asymmetrical power relationships in education, and educators should identify and incorporate these skills when designing learning experiences. In K-12

STEM education, this approach has improved educational practices and outcomes [3, 16]. Our prior study exploring adolescents' sensemaking around algorithmic biases [55] showcased their ability to flesh out their reasoning with details from their own lives, drawing from their funds of knowledge.

Complementary to funds of knowledge is sensemaking theory, which postulates that knowledge is dynamic rather than static [17]. It proposes that individuals actively process information from various sources to achieve understanding rather than achieving an arbitrary pinnacle of knowledge. Through sensemaking, individuals can progressively develop new understandings by participating in complex activities where they may not always have prior knowledge, instead of simply receiving information through direct instruction. In computing, sensemaking practices allow children to play an active role in learning various concepts, such as AI [18, 19], and data literacy [53]. As our goal is to investigate how children may engage in conversations around algorithmic fairness that center their perspectives, we ground our methods in sensemaking theory to make space for different paths of achieving understanding.

3 METHODS

3.1 Study Context & Timeline

From July to November 2022, the first and second authors conducted three sensemaking discussions lasting a maximum of 45 minutes with 16 participants (ages 8-12) in the United States. As we had participants from three US states (North Carolina, Virginia, and Washington), we allowed local participants to participate in the study in-person or virtually, while non-local participants had to participate virtually. All participants chose to complete the study virtually. As the three discussions were independent, we offered participants the flexibility of scheduling their sessions 1-4 weeks apart to accommodate extracurricular activities, vacations, and other obligations.

3.2 Participant Demographics

Parents/guardians of the 16 participants in our study filled out a form with free-response questions to disclose their children's age, gender identity, ethnic identity, languages spoken at home, disabilities, and any other aspects of their identity they would like the research team to know. Parents/guardians also provided their child(ren)'s chosen pseudonym or 'superhero name.' If they did not choose a pseudonym, we use the last letters of their first and last names (Table 1). All participants had internet-connected devices at home, but we did not ask for any more information about their prior experiences with AI, data, or computing more broadly because (1) we did not require prior experience to participate in our study and (2) prior work has associated perceptions of having less prior experience with a lower sense of belonging and confidence in a computing context [44, 56], which may impact their participation in the study. Before each discussion, we asked each participant for their assent to research participation and session recordings.

3.3 Sensemaking Discussions

Table 2 shows the three sensemaking discussions participants engaged in, drawn from our prior study with adolescents [55]. Each discussion centered on a specific *scenario* designed to highlight different aspects of algorithmic unfairness. Each scenario started with seed text describing the *situation*. This was followed by the incremental reveal of different layers of algorithmic decision-making — whether a *computer* was used in decision-making, what *algorithm* was, what *data* was used, and what the composition of the *team* behind the algorithm was.

To facilitate participants' sensemaking and provide artifacts for us to analyze, each sensemaking discussion involved (1) a *warm-up question* before introducing the scenario, (2) *reflection questions* for each layer revealed, and (3) a semi-structured big paper *design activity* where participants brainstorm ideas on big paper to support the unconstrained generation of ideas [15, 25]. For virtual participation, we presented the warm-up and reflection questions on Google Slides and the big paper design activity on Google Jamboard. As the scenarios were adapted from a prior study with an older age group [55], we modified the discussion to be more suitable for this age group by (1) limiting the number of participants per discussion to 2, as opposed to 6-7 with the older participants, (2) having participants respond verbally to the reflection questions immediately, instead of having them write down their reflections first and then debating verbally, and (3) assisting the participants with typing for the big paper design activity as some of them had difficulty balancing between voicing their ideas and expressing those ideas on the big paper.

We made intentional choices to prioritize the safety and agency of participants, accounting for our participants' ages, backgrounds, and the power imbalance between the researchers and participants. As such, we did not directly ask participants about their harmful experiences with technology. Instead, we selected scenarios that might resonate with them based on prior literature. If they brought up their own experiences, we encouraged them to do so on their terms. If participants were siblings, they participated in discussions together because, in pilots of this method, we observed siblings being able to co-regulate and debate with each other. We wanted co-regulation because we were discussing topics that could be emotionally difficult, and siblings could perhaps support each other. We also wanted to encourage debate since all our scenarios do not have obvious conceptions of fairness. Lastly, we made deliberate terminology choices to minimize reliance on prior computing knowledge. As with 'fairness' (see the footnote in Section 1), we used the term 'rules' to describe the algorithm. We contextualized the data used in each scenario (e.g., voice recordings in the Speaker scenario) rather than simply using the term 'data'.

Regardless of the scenario, we adhered to the following protocol to encourage elaboration: (1) if the participant(s) mention another group of people not mentioned in the scenario, ask them how the rules would affect those people, and (2) if the participant(s) mention an interesting point, repeat their point back to them and ask them why.

3.3.1 Scenario Design. We created three scenarios of algorithmic decision-making that surfaced potential fairness issues to seed sensemaking discussions. These scenarios were selected because they do not have straightforward conceptions of fairness and, thus, may elicit interesting insights from participants.

(1) The Search Engine ('Search') scenario was based on biases in representation from search results [49, 61].

Pseudonym	Age	Gender	Ethnicity	Languages Spoken at Home	Disability
AN	12	Agender	Caucasian	English	No
Alex	11	Male	Asian	English, Chinese	No
Ashley*	9	Female	Middle Eastern	English, Arabic	n/a
Emily*	8	Female	Middle Eastern	English, Arabic	n/a
Blue Gamer†	8.5	Boy	Asian	English	No
Green Raven†	12	Boy	Asian	English	No
Ethan	9	Agender	Non-Hispanic, Ashkenazi Jewish	English	No
Kalex	11	Female	Caucasian	English	No
Kitkat Krystal‡	9	Female	Caucasian	English	No
Magentafied Moonstone‡	9	Female	Caucasian	English	No
Leroy	11	Male	White	English	No
Minecraft Coder	9	Male	European	English, Dutch	ADHD
Po	12	Female	Pakistani	Urdu, Punjabi, Potwari, English	n/a
Spider§	8	Girl	White	English	No
Squidney§	10	Nonbinary	White	English	No
StoofCorg	10	Female	Asian	English, Chinese	No

Table 1: Participant Demographics. "n/a" denotes parent/guardian declined to disclose. Matching symbols (∗, †, ‡, §) denote siblings.

Scenario	Seed Text	Sit	Comp	Algo	Data	Team
Search Engine ('Search')	*Ahmad is making a presentation for what he wants to major in college: nursing. When he searches online for images of nurses, he can barely find images of man nurses. Almost all the images are of women.*	✓		✓		✓
Smart Speaker ('Speaker')	*Alex and her friends are playing with her family's new smart speaker, Blurty. She notices Blurty responds to all her friends except Maximo, who just moved to the US from Mexico.*	✓		✓	✓	✓
School	*There are two schools, School A and School B, in the same city. There are the same number of kids who go to both schools. Here are some of the kids who go to School A (show a group of white children) and here are some of the kids who go to School B (show a group of Black children). In School A, every classroom has six boxes of school supplies, such as books, calculators, art supplies, and notebooks, to use when kids are learning. In School B, every classroom has one box of school supplies.*	✓	✓	✓	✓	✓

Table 2: Seed Text & Layers Discussed in Each Scenario (Situation, Computer, Algorithm, Data, and Team).

(2) The Smart Speaker ('Speaker') scenario was based on the failure of many voice recognition systems to recognize other languages or accents [39].
(3) The School scenario was adapted from the scenario used in [24] to understand youth's perceptions of social resource inequality to reflect algorithmic redlining [54] (see Table 3).

We presented scenarios in this order to highlight an increasing scope of harm (Table 2). In the Search Scenario, only a single individual is harmed. In the Speaker scenario, while only a single individual is harmed, the harm results in group exclusion. In the School scenario, a community is harmed. We also designed the scenarios to have varying technical focuses, with the Search scenario involving only software components, the Speaker scenario including hardware and software components, and the School scenario involving a covert, non-obvious technical component.

Table 3 gives a detailed overview of the School scenario. The Search and Speaker scenarios followed a similar structure, with some key differences. First, both scenarios had an apparent technical component that did not require uncovering. Second, they had different high-level abstractions of the algorithm. The Search Engine followed a naive search algorithm accounting for keyword presence in images' metadata and the Smart Speaker being activated by a specific phrase. Third, the Search scenario had no training data as it was not a machine learning-based algorithm, while the Speaker scenario had training data of voices from English-speaking countries. Lastly, the various teams in the Search scenario differed based on gender, while the Speaker scenario differed based on country of origin.

3.3.2 *Warm-up Questions.* Participants discussed a warm-up question at the beginning of each sensemaking discussion. These questions asked participants to share their experiences and were intended to help them get comfortable reflecting and voicing their perspectives. Table 3 shows the question used in the School scenario.

Stage	Phase	School: Questions & Revealed Layers
Warm-Up	—	*Ask:* Where do you go to school? When you walk into your school, what do you see? When you walk into your classroom, what do you see?
Worksheet		
Situation		*Reveal seed text (Table 2)*
	Understanding	*Ask:* Why do you think School A has more supplies than School B?
Computer		*Reveal:* A computer decided how much supplies each school should get.
	Evaluation	*Ask:* a. What do you think of a computer making that decision?
	Understanding	b. Why do you think a computer decided to give School A more supplies than School B?
Algorithm		*Reveal:* School A is in neighborhood A and School B is in neighborhood B. The computer made its decision using this rule: "For every $100 the neighborhood gives to the school, every classroom gets an extra box of school supplies."
	Evaluation	*Ask:* a. What do you think of the rules the computer used? [If participants don't mention fairness] How fair do you think the rules are? Why?
	Understanding	b. How do the rules impact different people?
	Evaluation	c. What are the pros and cons of using a computer to make that decision?
Data		*Reveal:* The computer used data about how much neighborhoods gave in the past to decide that each neighborhood should give $100 for each box of school supplies.
	Evaluation	*Ask:* a. What do you think of the data that the computer used?
	Evaluation	b. How fair is it that the computer used past data? Why?
Team		*Reveal:* The team who designed the rules and data the computer used was made up of all white people.
	Evaluation	*Ask:* a. What do you think of this team? [If participants do not mention fairness for questions a, b, and c] How fair do you think this team is? Why?
	Evaluation	b. What if the team was made up of all black people? What do you think of this team?
	Evaluation	c. What if the team was made up of people from different races? What do you think of this team?
	Evaluation	d. Which team is the most fair? Why? [If participants bring up other factors] If you don't think any of the teams are the most fair, what would be the most fair team? Why?
Design	Brainstorming	*Ask:* - Imagine you're the boss & you're in charge of the rules. What rules would you use to decide how much supplies each school should get?
Activity	Brainstorming	- Who will be applying the rules? Will it be a computer? A person? A team? Both?
	Brainstorming	- How do you make sure the rules are fair? [Follow-up questions if needed:]
	Brainstorming	- What kind of team would be the most fair in designing these rules?
	Brainstorming	- How would you and your team design the rules fairly?
	Brainstorming	- How would you and your team test the rules fairly?

Table 3: School Sensemaking Discussion Questions in full. *Italics* denote actions performed by facilitators.

3.3.3 Reflection Questions. After warm-ups, we revealed different layers of the algorithmic decision-making one at a time to scaffold the sensemaking process, inspired by sensemaking practices in math and data science education [38] (see Table 2). For each layer (e.g., algorithm, data) that was revealed, we prompted participants with reflection questions that were focused on either: (1) understanding or (2) evaluating the decisions in each scenario to encourage divergent or convergent thinking, respectively (Table 3).

3.3.4 Design Activities. After the reflection questions, participants brainstormed ideas to address the bias in the scenario. Consistent with the big paper method [25], they wrote their ideas either on the jamboard or on sticky notes (Figure 1).

For the design activities, we prompted participants to imagine that they were the boss and in charge of designing the algorithm and, if applicable, the data used in each scenario. We chose this framing because, in early trials of this method, pilot participants struggled with the agency they had in each scenario, getting preoccupied with whom they would answer to instead of the task at hand. Throughout the activity, we prompted them to consider the fairness of the different layers of decision-making they designed (see Table 3 for specific prompts). After the design activity, we debriefed participants, answering any questions they had about the scenarios or the study.

3.4 Data Collection & Analysis

We collected data over multiple sessions, with one scenario per session. We discussed the Search scenario with all 16 participants, the Speaker scenario with 14 participants (AN and Squidney dropped out), and the School scenario discussion with 12 participants (AN, Spider, Squidney, and Magentafied Moonstone dropped out). As siblings participated together, this resulted in 33 transcripts (32-45 minutes long) and big paper designs.

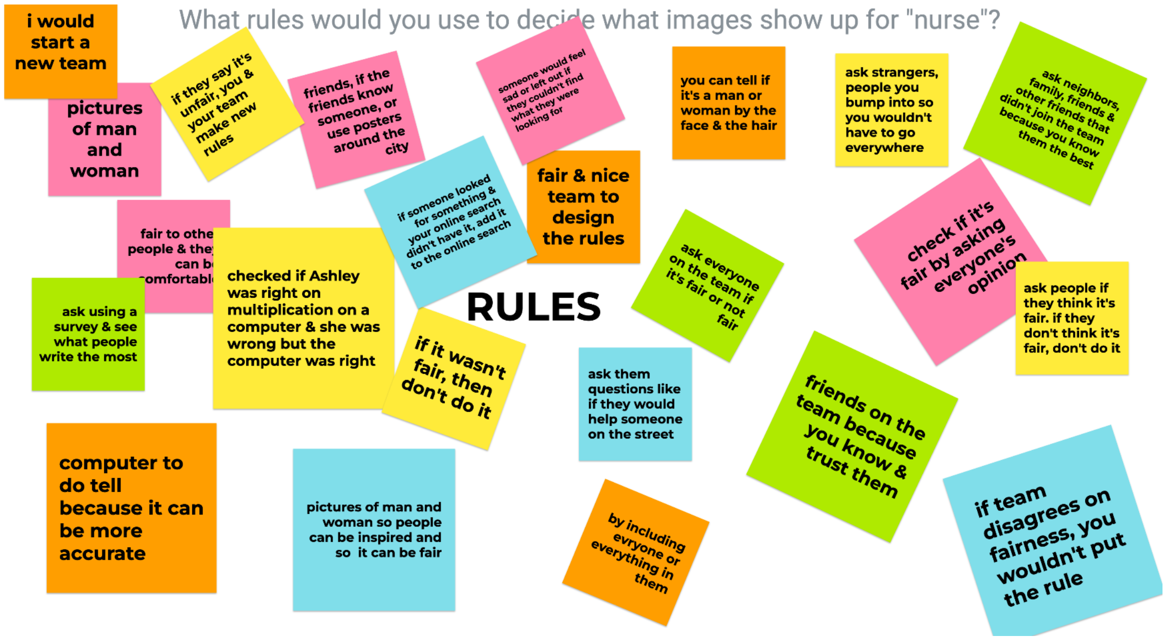

Figure 1: Example from a Big Paper Design Activity

To understand how participants engaged with the discussions and design activities, we used a deductive thematic analysis approach to analyze the transcripts and big paper designs using participants' responses as our data source, with at least two authors coding each transcript and resolving disagreements by consensus. We adopted the practice of taking participant quotes and big paper responses 'literally' to minimize inference. We did not capture agreement metrics such as inter-rater reliability throughout this process. Instead, we chose to resolve uncertainties through discussion and consensus-building, consistent with the position of Hammer and Berland [30] on qualitative coding that uses codes as an organizational aid for thematic claims about the data.

Our codebook was drawn from our prior study [55], using the metaphor of a camera as an analysis guide. Since the adolescent participants made sense of algorithmic fairness using many factors beyond those in the scenarios, in the post-hoc round of that study's thematic analysis, we organized those factors into a camera metaphor to describe how participants used different factors in their sensemaking and how these factors related to each other.

In this metaphor, we specifically used the lenses and filters of a camera. Photographers utilize lenses to modify the scale and resolution of a shot and affix different filters to a lens to photograph the same subject differently, resulting in varying final images. In this metaphor, the participants are photographers, viewing algorithmic fairness in different scales/resolutions and lights to make sense of them. Each participant has their camera with their own set of lenses and filters. In [55], we found that participants used two different *lenses* to make sense of algorithmic bias at different levels: (1) a human lens, which spanned individual to societal factors, and (2) a technical lens, which included technology creators and other technical factors. Along with adapting the scale and resolution with their lenses, we observed that participants employed different characteristics, such as gender and race, as *filters* to alter what was most relevant to their sensemaking in each scenario. This categorization scheme is reflected in our codebook (Table 4).

Categories in the codebook were not mutually exclusive and often overlapped. For example, if a participant used a programmer's gender bias to make sense of the Search scenario, their quote would be coded in both 'biases' under technical factors and 'gender' under characteristics. Factors that did not fit in any of the categories were coded as 'Other' under human factors, technical factors, or characteristics. While the first three authors reached a collective understanding of most codes, some were more often disagreed upon and required more discussion to build consensus: (1) *individual: principles* because for some participants, it was hard to determine if they espoused these principles themselves, or if they were simply naming a societal ideal, (2) *individual: stereotype* and *society: societal stereotypes* because for some participants, it was difficult to identify if they held the stereotype themselves or if they were naming a societal stereotype, (3) *community: membership* because it was sometimes difficult to identify the community that participants were referring to without inference. Upholding Hammer and Berland's stance on qualitative work [30], we report on these disagreements for transparency on which themes were more subject to different interpretations.

After all the transcripts were coded, the first and second authors collaboratively grouped the 'Other' factors into common categories using affinity diagramming [43]. The first author then coded the 'Other' factors into those common categories, with the second author verifying their codes and resolving uncertainties through consensus. Lastly, the first and second authors conducted a post-hoc analysis of all the codes to synthesize higher-level themes.

Factor	Explanation
Human Lens	
Individual: Stereotype	Participants drew from a stereotype they espoused themselves.
Individual: Lived Experience	Participants drew from their own lived experiences.
Individual: Principles	Participants drew from their views of right and wrong.
Community: Geographic Location	Participants drew from the location of a community.
Community: Membership	Participants drew from *who* is in a community.
Society: History	Participants drew from past events and phenomena.
Society: Societal Stereotype	Participants identified a problematic idea as a 'stereotype'.
Society: Power Distribution	Participants identified issues of power/agency or lack thereof.
Society: Systemic Marginalization	Participants problematized marginalization as a result of larger systems.
Technical Lens	
Creators: Qualifications	Participants expressed both technical and interpersonal qualitites that made tech creators (un)qualified for the job.
Creators: Biases	Participants cited biases held by the tech creators.
Creators: Power Dynamics	Participants accounted for power dynamics within teams of tech creators.
Characteristics as Filters	Characteristics participants used to describe the factors above: gender, country of origin, language, accent, (dis)ability, age, academic performance, economic status/class, race/ethnicity

Table 4: Codebook using the metaphor of a camera from [55] as an analysis guide. We coded for the presence/absence of these factors in participant transcripts and big paper designs.

3.5 Author Positionality

Positionality statements make explicit the relationship between the authors' identities and the research topic and the identities of the participants [32, 40, 52]. Each author wrote statements to describe experiences and perspectives that influenced their engagement with the research.

The first author identifies as a woman of color. Some participants used her visible identity facets in the sensemaking discussions. For instance, when making sense of why most nurses were women in the Search scenario, Blue Gamer hypothesized that it was because women "get less pay for some weird reason" and then asked the first author if she got enough pay. Her experiences with systemic marginalization in computing and society led to her interest in critical computing literacies for youth and adolescents. She led this project to understand the youth's perspectives and engagement with ideas around algorithmic bias.

The second author grew up and currently resides in a liberal and technology-centric city. Through their work in various communities throughout their career, they have come to recognize that their background may lead them to view technology as an integral and irresistible part of society - a reality that is not reflected by all participants in this study or otherwise. They are therefore motivated by their hope to challenge the culture of mystery, silence, and unchallenged acceptance of technological advancements. They seek to grant children the opportunity to see beyond the black box of technology and clearly understand its motives and inner workings.

The third author positions herself primarily as an activist for better and more inclusive technology education. Before deciding to embark on a Ph.D. journey, she worked for more than eight years on hands-on STEAM education in different communities worldwide as part of the organization she created called *HacKIDemia*. In the past three years, she has led multiple co-design sessions with families focused on AI literacy and created *Cognimates*, one of the first platforms for AI education, which is free and open-source.

The fourth author identifies as a mixed-race, queer, gender non-conforming parent and a computing researcher. She has had a lifelong interest in youth development and approached this project with a curiosity about youth capacity for moral reasoning about computation. Her role in the work was mentor, advisor, and facilitator.

4 RESULTS

Using the camera metaphor from our prior study with adolescents as an analysis guide(see Section 3.4), we characterize the two lenses, the *human* lens and *technical* lens, and the filters, or the *characteristics*, participants in this study used to make sense of algorithmic bias. Since participants often used the lenses and multiple filters in conjunction, featuring a quote for one particular lens or filter does not mean it did not contain others.

4.1 Human Lens: Different Scales of Human Factors

The human lens encompassed factors relevant to groups of people of different sizes: individual, community, and society.

4.1.1 Individual. The individual level included factors connected to an individual. One factor that participants accounted for in reasoning about the unfairness in a scenario was stereotypes (Search: 9/16; Speaker: 2/14; School: 1/12). When brainstorming how people designing a Search Engine would identify images of different gendered nurses, Emily suggested, *"You could tell by the face and hair?"*, reflecting physical gender stereotypes.

Participants also attributed the unfairness to an individual's pitfalls (Search: 6/16) or an individual's interaction with technology (Speaker: 5/12; School: 4/12). In reasoning about why the Smart Speaker did not respond to Maximo, Green Raven hypothesized, *"he might have said it wrong"*, faulting the individual in the scenario.

Across all scenarios, two factors were especially salient to participants. The first was their lived experiences (Search: 11/16, Speaker:

9/14, School: 8/12). Po drew from her school district when making sense of the unfairness in the School scenario, *"In the less wealthy side of <Po's hometown>, the district divides it purposely so the wealthier people go to one school, and then the less wealthy people go to another"*, using observations of wealth distribution in her neighborhood.

The other particularly salient factor was their principles or beliefs of right and wrong (Search: 16/16, Speaker: 14/14, School: 12/12). In critiquing the rules in the Search scenario, AN said, *"It'd be nice to have a search that would include some sort of diversity, but I don't know what kind of rule that would be."* Although they did not know how to accomplish it, AN believed that diversity was important in designing the rules for a search engine.

4.1.2 Community. The community level covered factors linked to a collective group of people. At this level, membership, or *who* was in a community, was especially relevant to participants in their sensemaking process (Search: 11/16; Speaker: 7/14; School: 8/12). When brainstorming who should design the rules in the School scenario, Alex wanted *"a team of parents who have children who go to all the schools in the city"*, conceptualizing potential stakeholders in that scenario.

Most participants also examined aspects of interpersonal relationships, such as bullying, empathy, and collaboration, when making sense of the unfairness in scenarios (Search: 10/16, Speaker: 12/14, School: 8/12). When making sense of Ahmad's feelings from seeing mostly women as nurses, Kalex compared it to an instance of bullying: *"I have a friend who is vegetarian [...] there was a kid in my class who would bully her for it."* Similarly, Minecraft Coder empathized with Maximo's struggle in the Speaker scenario: *"it would make you think you should speak another language, and you would ask every time Alexa can you repeat that again?"*. Participants also accounted for collaboration while brainstorming the design team. When asked if the team designing the Speaker should be from different countries, Magentafied Moonstone questioned, *"Can the team all speak one language? Because then it's not really a team if they can't collaborate"*, suggesting that good teamwork was a priority, perhaps even above diversity in countries or languages.

4.1.3 Society. The society level encompassed factors attributed to larger structural issues. Participants accounted for history (Search: 10/16, Speaker: 1/14, School: 11/12), issues of power distribution (Search: 12/16, Speaker: 3/14, School: 11/12), systemic marginalization (Search: 9/16, Speaker: 2/14, School: 9/12), and societal stereotypes (Search: 11/16, Speaker: 7/14, School: 8/12). Interestingly, while participants considered systemic issues in their sensemaking, they tended to attribute these issues to individuals. This tendency was exemplified in the following exchange between Ethan and a researcher while discussing the Speaker scenario:

> *Researcher: Why do you think the speaker would respond differently to someone from a different place?*
> *Ethan: Because America is mostly a bad place, and so we build bad things.*
> *Researcher: Alright, why do you say it's a bad place?*
> *Ethan: It's getting better, but [...] it doesn't accept a lot of stuff.*
> *Researcher: What do you mean when you say it doesn't really accept that much?*
> *Ethan: Not the president now, but who used to be president like, a few years ago, Trump.*

In this exchange, Ethan voiced a larger structural issue, a country's unwelcoming environment, but blamed it on a specific person. Similarly, when making sense of why schools received different amounts of school supplies, Green Raven hypothesized, *"The white person's like white supremacy cuz it's what white people think"*. While Green Raven identified the systemic issue of white supremacy, he attributed it to individual white people. Both these instances reveal an awareness of systemic issues but only a vague, individualistic understanding of them.

Participants also considered societal ideals. These are distinct from the principles at the individual level because it was not clear if the children espoused these ideals themselves (Search: 16/16, Speaker: 12/14, School: 9/12). In his evaluation of an all-men team in the Search scenario, Minecraft Coder critiqued, *"the team would think the same"*. While he accounted for the societal ideal of diversity in his sensemaking, it was not definitive if he believed in this ideal.

4.2 Technical Lens: Different Resolutions of Technology Factors

Participants used the technical lens to account for technology-related factors to varying degrees of specificity. This lens was not as well-formed as the human lens, which was expected because we did not require prior computing experience.

4.2.1 Technology Creators. We grouped engineers, programmers, designers, and others involved in developing technology under the umbrella of 'creators', as participants did not meaningfully distinguish between them. When sensemaking, participants accounted for the creators' qualifications (Search: 16/16, Speaker: 14/14, School: 12/12) and biases (Search: 15/16, Speaker: 7/14, School: 8/12), as well as power dynamics within teams of tech creators (Search: 11/16, Speaker: 5/14, School: 6/12). For instance, Alex contemplated the skill of navigating team disagreements in deciding who would be qualified to design rules for the School scenario: *"Some people might disagree on how much you should price it [...] but as long as they could always find some way to agree they'll be a very good team."*

4.2.2 Users. Participants also considered technology users (Search: 7/16; Speaker: 10/14; School: 7/12). This largely came in the form of testing the technology with users conceptualized by the participants, both real and hypothetical. In brainstorming how she would test the Speaker, Ashley said, *"they could have many people who speak different languages"*. She also included people in her life in testing, *"Your family? Maybe your neighbor? Maybe other people you see?"*

In conceptualizing hypothetical users, some participants made assumptions about them, such as Squidney in making sense of the rules in the Search scenario: *"They probably would know that having a label would increase the chance of having it in the list or photos that come up when you search nurse."* Squidney indicated an assumption of some technical knowledge from a user, which may not necessarily be the case.

4.2.3 Attitudes towards Technology. Lastly, participants reflected on their attitudes toward technology in their sensemaking. Most participants voiced doubts over a computer's abilities (Search: 11/16;

Speaker: 11/14; School: 12/12). These doubts were often cited as reasons to absolve the computer, such as StoofCorg in the Search scenario: *"the computer is just following whatever the people say"*. These doubts might have also shaped the computers' role in the solutions participants designed. For instance, an idea from Ethan's design board was "`Use the computer for the emails and the phone`", relegating technology to a communication role.

Participants also demonstrated a preference for humans over computers in addressing the unfairness in the scenarios (Search: 6/16; Speaker: 1/14; School: 7/12). In making sense of the computer's decision-making role in the School scenario, Po critiqued, *"Instead of making a computer do it and being extremely specific, [...] then I think it's just be a better idea for a human to do it. Because there's less room for error if a human just did it."*

4.3 Filters: Characteristics Salient in Sensemaking

After participants determined the scale/resolution of their sensemaking within the human or technical lens, they decided which filter to attach to the lens to make different *characteristics* more relevant. Participants tended to use characteristics from both the prompts and their conceptions to reason about the algorithmic unfairness in the scenarios. Participants sometimes used multiple characteristics simultaneously, so using a quote for a specific characteristic does not mean it did not contain other characteristics. We begin with the characteristics used only in the scenario where they were prompted, followed by characteristics that participants introduced into only one scenario, and end with characteristics used by participants in all scenarios regardless of mention in the prompts.

4.3.1 Characteristics Used only in the Prompted Scenario. All participants (14/14) used both language and a closely related though not explicitly stated factor, accent, to make sense of the fairness in the Speaker scenario. When evaluating a multilingual design team for the Speaker, Blue Gamer prioritized language inclusivity, *"Yeah, it would include everyone. Even if they made bad decisions, it's still better than everyone knowing only one language."*. As for accents, when making sense of the rule used to trigger the Speaker, Kitkat Krystal hypothesized: *"If you think about people with accents, like Maximo might have, maybe when they say 'Hey, Blurty', it may sound a little different that Blurty may not understand."*

4.3.2 Characteristics Introduced by Participants. Some participants (4/16) used economic status to make sense of the unfairness in the Search scenario, although it was not mentioned at all. Participants often reasoned about economic status through gender, a characteristic included in the prompt. When assessing the fairness of an all-male design team, Spider hypothesized, *"this team probably gets paid a lot of money [...], and so the woman is out lots of money"*.

Another characteristic that participants introduced was (dis)ability into the Speaker scenario (3/12). In critiquing the rule used to trigger the Speaker, Alex described his own pronunciation difficulties: *"If people speak with accents or speech impediments, like lisps, then perhaps it will be harder to understand. For example, I have to wear these retainers every night. When I wear them, it's hard for me to pronounce my Ls and Rs. If I tried to say, 'Hey, Blurty', it might sound a little weird."* Participants also considered different traits of speech (10/14). For example, Ethan brainstormed testing with voices of different pitches, *"If it's high pitch, accent, or low, or deep, we can find out if it can still recognize it as Hey, Blurty."*

In the School scenario, participants introduced the school or neighborhood population (5/15) to make sense of the unfairness. When it was revealed that a computer decided the distribution of school supplies, Kitkat Krystal rationalized its decision based on class size, *"It can also just be based on how many kids are in your class. In school A, say there were 32 students in a class. But then for school B, there were 22. So it's okay, they can get a smaller amount of boxes."* On a related note, participants also introduced the distribution of resources (4/15). In brainstorming solutions, Ashley ideated: *"I think they should split the money to both schools. They can get the supplies they need"*.

4.3.3 Characteristics Salient across all Scenarios. Four characteristics were used by participants to make sense of algorithmic fairness in all the scenarios: gender, country of origin, race/ethnicity, and age.

Gender was only mentioned in the Search scenario, but was relevant in all scenarios (Search: 16/16; Speaker: 7/14; School: 3/12). For instance, when deciding who he would want in his design team for the School scenario, Leroy suggested, *"a mix of races and genders"*, layering gender in addition to race (which was included in the scenario).

Similarly, country of origin was only brought up in the Speaker scenario but was added by participants in all scenarios (Search: 3/16; Speaker: 14/14; School: 1/12). As an example, in designing a solution to the Search scenario, Green Raven said, *"I guess every country has a say. So it's a worldwide thing."*

Race/ethnicity was also only included in the School scenario, but became pertinent to participants across all scenarios (Search: 9/16; Speaker: 4/14; School: 12/12). For example, when choosing which accents and dialects to include in the training data for the Speaker, StoofCorg brainstormed, *"definitely Mexican because you get a lot of Mexican immigrants, Chinese because there's a lot of Chinese immigrants, African Americans, maybe British."*

Interestingly, age was not prompted in any of the scenarios and yet was utilized by all participants in their sensemaking (Search: 9/16; Speaker: 4/14; School: 1/12). Age was often used as a gateway for participants to introduce their own conceptions of fairness into their sensemaking process, such Magentafied Moonstone (age 12) who was participating with her sibling Kitkat Krystal (age 9): *"When you have someone that's 9 versus someone that's 12, clothes might cost more. So you might have to give us different amounts of money, but it's still fair."*

5 DISCUSSION

5.1 RQ1: What funds of knowledge might children use to make sense of algorithmic fairness?

Through the scaffolded sensemaking in these scenario discussions and design activities, all participants reasoned around algorithmic fairness using factors both explicitly and not explicitly mentioned in the scenarios. Participants used two lenses to adjust the scale and resolution of their sensemaking: (1) the human lens and (2) the technical lens.

In the human lens, participants used factors at increasing group sizes ranging from individual to society, which reflects ecological system theory that views an individual relative to their communities and larger society [10]. At the individual level, participants often grounded their sensemaking in their lived experiences and principles. Similarly, at the community level, participants tended to base their reasoning in both real and hypothetical interpersonal relationships, which often arose when hypothesizing the impacts of algorithmic bias. However, while participants also reasoned at the societal level, this sensemaking was more vague — participants tended to attribute structural issues to individual bad actors, not fully comprehending the large, systemic scale. Participants also expressed societal ideals when making sense of the unfairness, but it was not always clear if they believed in themselves. This vagueness may be because these issues are more abstract, coming from surrounding adult society, and less grounded in their own lived experiences [37].

As for the technical lens, participants often developed specific conceptualizations of users, drawing from both real people in their lives and hypothetical users. This mirrors the relevance of lived experiences and interpersonal relationships observed through the human lens. In contrast with prior work [20, 31, 42, 62, 63], participants also exhibited a distrust towards technology, doubting computers' abilities and displayed an inclination towards a human approach to address unfairness. This may be due to various reasons, including but not limited to (1) the lack of personification of the technology in the scenarios [26], (2) the child characters in the scenarios were easier to empathize with, (3) lived experiences with or exposure to adult tech use [50], and (4) a broader attitude change towards technology in society. Given the salience of our participants' lived experiences regardless of the lens, we encourage designers, educators, and other stakeholders to consider centering children's lived experiences, and their resultant funds of knowledge, in discussions of algorithmic fairness.

5.2 RQ2: How might the ages, identities, and backgrounds of children shape their sensemaking of algorithmic fairness?

Compared with the adolescent participants in our prior study [55], participants in this study tended to draw more heavily from their own experiences, perhaps indicative of a more egocentric viewpoint of morality characteristic in children [9]. Interpersonal relationships were very salient to participants in this study but not as salient for adolescents. This observation may potentially be because, at this age group, most children are learning to navigate interpersonal relationships and refining those skills [46]. In the human lens, the differences between the child and adolescent participants seemed to be reflected through the larger relevance of the individual and community levels and the relative vagueness of the societal level (Figure 2). In the technical lens, this study's participants often expressed detailed characterization of users in their testing, as opposed to the previous study's adolescent participants often designing for the 'average user', possibly indicating the salience of interpersonal relationships. This study's participants also drew from their own skeptical attitudes around technology, potentially reflecting a more egocentric viewpoint.

In addition to reasoning fairness at different scales and levels of specificity, participants often introduced characteristics beyond the prompt in their sensemaking. Participants introduced economic status into the Search scenario, (dis)ability into the Speaker scenario, and population and resource distribution into the School scenario, none of which were prompted. Participants may have added economic status into the Search scenario because it involved genders and occupations; some may have been sensitive to occupational stereotypes [27] or have existing knowledge of the gender pay gap [60]. Participants may include (dis)ability in the form of speech impediments because of their own experiences [6]. Similarly, participants may have accounted for school or neighborhood population, as well as resource distribution because schools, classrooms, and neighborhoods are contexts they are familiar with.

Participants seemed to be especially attuned to gender, race/ethnicity, country of origin, and age, as they used them to make sense of the unfairness in all scenarios regardless of whether they were prompted. In contrast, race/ethnicity and economic status/class were particularly salient for the adolescent participants in [55]. Children develop identities around gender and race from a young age as part of learning social competence [34], which may explain the salience of gender and race. While we did not specifically ask in our demographics form, some participants brought up their immigrant backgrounds, which may account for the relevance of country of origin across the scenarios. Age may be particularly salient to participants because many developmental milestones in childhood are tied to age [1]. The seemingly lower relevance of economic status/class for participants in this study compared with the adolescent participants in [55] may be grounded in lived experience. We did not recruit based on economic background for this study, but we recruited the adolescent participants from a program targeting students from low-income backgrounds in [55]. With this understanding of how learners' funds of knowledge may change based on ages, stages of development, identities, and backgrounds, we encourage designers, educators, and other stakeholders to consider scaffolding discussions of algorithmic fairness accordingly, thus allowing learners to leverage their funds of knowledge for deeper reasoning.

5.3 Limitations, Contributions, & Future Work

Although our study provides valuable insights, its design has some limitations. While we prioritized participants' well-being and safety throughout the study, elements of the inherent power dynamic between an adult researcher and child participants still persist. For example, by following up on a participant's idea to encourage their thinking, a researcher signals importance to the participant, which can influence them to emphasize ideas differently. Participants may also have perceived us as authority figures, possibly affecting their behavior in sessions. They may have acted differently to impress us or may have been more engaged because perceived authority figures were listening to them. Also, while our participants were from different US states, they were all from locales that had a technology industry presence, which may have shaped their attitudes toward technology. Lastly, our participants only represent their own

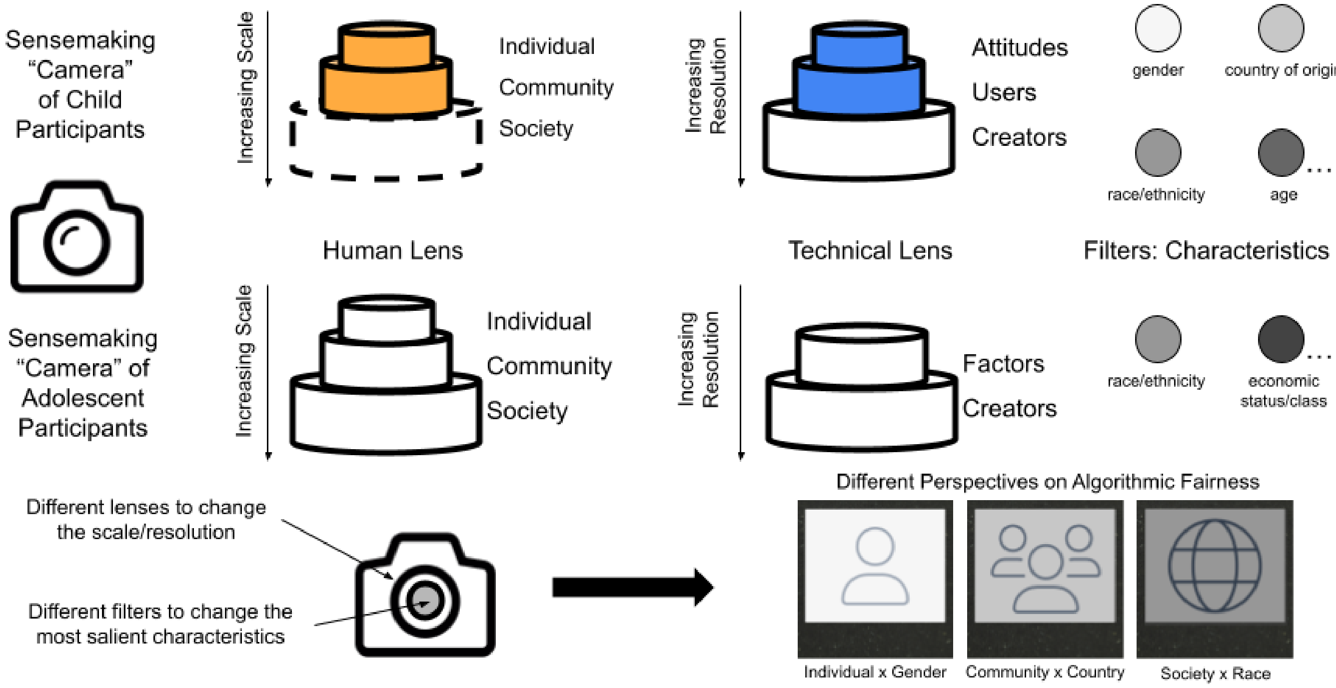

Figure 2: Comparison of the Sensemaking 'Cameras' of this study's participants & adolescents from [55]. The individual and community levels of the human lens were more salient (orange), while the societal level was more vague (dashed line). The technical lens had new aspects of 'users' and 'attitudes' (blue).

unique views and experiences, which do not generalize (neither was generalizability an objective of this study).

In spite of its limitations, this work reveals crucial insights into children's reasoning around algorithmic fairness. When provided opportunities to use their situated knowledge, such as in these discussions and design activities, children were not only aware of algorithmic bias but also capable of sensemaking around both its explicit and implicit negative impacts. This characterization of participants' funds of knowledge can help designers uphold design principles for children, by gathering and respecting children's unique perspectives [41] and by supporting proactive measures to protect them from algorithmic bias [2], such as education. Understanding potential funds of knowledge also inform the design of learning experiences on algorithmic fairness by understanding potential paths for children to engage in a way that centers their agency and well-being. Since learners' funds of knowledge evolve with age, experiences, and social circumstances [28, 47], it is important that such experiences designed for children consider both their lived experiences as well as their moral and interpersonal development so that their existing experiences and knowledge may serve as a bridge into a new domain.

While this study offers one blueprint for incorporating children's funds of knowledge through open-ended sensemaking, future work could explore other tools and techniques to do so. As our participants only represent their own knowledge and experiences, it is also crucial that future studies replicate this study with participants of different identities and backgrounds. Future studies may also investigate how different stakeholders in different contexts, such as families, teachers, designers, and policymakers, may be responsive to children's funds of knowledge when engaging them in learning experiences on algorithmic bias. Learning experiences, tools, and other interactions with technology that meaningfully integrate children's situated knowledge can better enable them to take advantage of their unique perspectives in navigating this increasingly technological world.

6 SELECTION AND PARTICIPATION OF CHILDREN

We recruited children through our networks, social media, and local parent groups. We selected children based on their ages (7-12) and attempted to get a mix of genders, ethnicities, languages spoken at home, and disability status. Before any discussions with children, both the children and their parents read and signed a consent form, describing the study purpose, procedures, potential risks, stress, or discomfort, confidentiality, and the de-identified public dissemination of research results. At the start of each discussion, the researcher first told the children that it would take up to 45 minutes, and that we would start by discussing a scenario with computers, followed by an activity where they would get to be the boss and design some rules. We then asked the children if they assented to the discussion being recorded and that they can stop the recording at any time. Lastly, we reinforced to the children that they should only share what they are comfortable sharing and that if they no longer want to participate, they can tell us without getting in trouble with us or their parents. Throughout the discussion, we also looked out for signs of discomfort from the children, so that we could skip questions or shorten/end discussions if they were uncomfortable.

ACKNOWLEDGMENTS

We would like to thank Carla Agard-Strickland and her daughter Keira Strickland, Jen Palmer and her daughter Ida Hellige, Dr. Yuliana Zamora and her son, and the other parents and children who preferred to not be acknowledged by name for their help in piloting the methods. This material is based upon work supported by the National Science Foundation under grants 2127309, 1539179, 1703304, 1836813, 2031265, 2100296, 2122950, 2137834, 2137312, and unrestricted gifts from Adobe, Google, and Microsoft.

REFERENCES

[1] 2019. Brain architecture. https://developingchild.harvard.edu/science/key-concepts/brain-architecture/
[2] 2022. Blueprint for an AI bill of rights. https://www.whitehouse.gov/ostp/ai-bill-of-rights/#discrimination
[3] Jane Andrews and Wan Ching Yee. 2006. Children's 'funds of knowledge'and their real life activities: Two minority ethnic children learning in out-of-school contexts in the UK. *Educational review* 58, 4 (2006), 435–449.
[4] Nicolas Baumard, Olivier Mascaro, and Coralie Chevallier. 2012. Preschoolers are able to take merit into account when distributing goods. *Developmental psychology* 48, 2 (2012), 492.
[5] Karl-Emil Kjær Bilstrup, Magnus Høholt Kaspersen, Mille Skovhus Lunding, Marie-Monique Schaper, Maarten Van Mechelen, Mariana Aki Tamashiro, Rachel Charlotte Smith, Ole Sejer Iversen, and Marianne Graves Petersen. 2022. Supporting critical data literacy in K-9 education: three principles for enriching pupils' relationship to data. In *Interaction Design and Children*. 225–236.
[6] Dorothy VM Bishop and Laurence B Leonard. 2000. *Speech and language impairments in children: Causes, characteristics, intervention and outcome.* Psychology press.
[7] Peter R Blake and Katherine McAuliffe. 2011. "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition* 120, 2 (2011), 215–224.
[8] Peter R Blake, Katherine McAuliffe, and Felix Warneken. 2014. The developmental origins of fairness: The knowledge–behavior gap. *Trends in cognitive sciences* 18, 11 (2014), 559–561.
[9] Jan Boom. 2011. Egocentrism in moral development: Gibbs, Piaget, Kohlberg. *New Ideas in Psychology* 29, 3 (2011), 355–363.
[10] Urie Bronfenbrenner. 1992. *Ecological systems theory.* Jessica Kingsley Publishers.
[11] Christina Caron. 2023. The A.I. Chatbots have arrived. time to talk to your kids. https://www.nytimes.com/2023/03/22/well/family/ai-chatgpt-parents-children.html
[12] Jeremy IM Carpendale. 2000. Kohlberg and Piaget on stages and moral reasoning. *Developmental Review* 20, 2 (2000), 181–205.
[13] Vicky Charisi, Laura Malinverni, Marie-Monique Schaper, and Elisa Rubegni. 2020. Creating opportunities for children's critical reflections on AI, robotics and other intelligent technologies. In *Proceedings of the 2020 ACM interaction design and children conference: extended abstracts*. 89–95.
[14] Merijke Coenraad. 2022. "That's what techquity is": youth perceptions of technological and algorithmic bias. *Information and Learning Sciences* ahead-of-print (2022).
[15] Merijke Coenraad, Jen Palmer, Diana Franklin, and David Weintrop. 2019. Enacting identities: Participatory design as a context for youth to reflect, project, and apply their emerging identities. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. 185–196.
[16] Maya Denton and Maura Borrego. 2021. Funds of knowledge in STEM education: A scoping review. *Studies in engineering education* 1, 2 (2021).
[17] Brenda Dervin. 1998. Sense-making theory and practice: An overview of user interests in knowledge seeking and use. *Journal of knowledge management* (1998).
[18] 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*. 1–17.
[19] Stefania Druga and Amy J Ko. 2021. How do children's perceptions of machine intelligence change when training and coding smart programs?. In *Interaction design and children*. 49–61.
[20] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. 2017. " Hey Google is it ok if I eat you?" Initial explorations in child-agent interaction. In *Proceedings of the 2017 conference on interaction design and children*. 595–600.
[21] Stefania Druga, Randi Williams, Hae Won Park, and Cynthia Breazeal. 2018. How smart are the smart toys? Children and parents' agent interaction and intelligence attribution. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*. 231–240.
[22] Stefania Druga, Jason Yip, Michael Preston, and Devin Dillon. 2021. The 4As: Ask, Adapt, Author, Analyze-AI Literacy Framework for Families. In *Algorithmic Rights and Protections for Children*. PubPub.
[23] Yarrow Dunham, Allison Durkin, and Tom R Tyler. 2018. The development of a preference for procedural justice for self and others. *Scientific reports* 8, 1 (2018), 1–8.
[24] Laura Elenbaas and Melanie Killen. 2017. Children's perceptions of social resource inequality. *Journal of Applied Developmental Psychology* 48 (2017), 49–58.
[25] Jerry Alan Fails, Mona Leigh Guha, Allison Druin, et al. 2013. Methods and techniques for involving children in the design of new technology for children. *Foundations and Trends® in Human–Computer Interaction* 6, 2 (2013), 85–166.
[26] Janik Festerling and Iram Siraj. 2022. Anthropomorphizing technology: a conceptual review of anthropomorphism research and how it relates to children's engagements with digital voice assistants. *Integrative Psychological and Behavioral Science* 56, 3 (2022), 709–738.
[27] Linda D Gettys and Arnie Cann. 1981. Children's perceptions of occupational sex stereotypes. *Sex Roles* 7, 3 (1981), 301–308.
[28] Norma González, Luis C. Moll, and Cathy Amanti. 2006. *Funds of knowledge: Theorizing practices in households, communities, and classrooms.* Routledge.
[29] Patricia Grocke, Federico Rossano, and Michael Tomasello. 2015. Procedural justice in children: Preschoolers accept unequal resource distributions if the procedure provides equal opportunities. *Journal of experimental child psychology* 140 (2015), 197–210.
[30] David Hammer and Leema K. Berland. 2014. Confusing claims for data: A critique of common practices for presenting qualitative research on learning. *Journal of the Learning Sciences* 23, 1 (2014), 37–46.
[31] Clint Andrew Heinze, Janet Haase, and Helen Higgins. 2010. An action research report from a multi-year approach to teaching artificial intelligence at the k-6 level. In *First AAAI Symposium on Educational Advances in Artificial Intelligence*.
[32] Andrew Gary Darwin Holmes. 2020. Researcher Positionality–A Consideration of Its Influence and Place in Qualitative Research–A New Researcher Guide. *Shanlax International Journal of Education* 8, 4 (2020), 1–10.
[33] Elizabeth Huppert, Jason M Cowell, Yawei Cheng, Carlos Contreras-Ibáñez, Natalia Gomez-Sicard, Maria Luz Gonzalez-Gadea, David Huepe, Agustin Ibanez, Kang Lee, Randa Mahasneh, et al. 2019. The development of children's preferences for equality and equity across 13 individualistic and collectivist cultures. *Developmental science* 22, 2 (2019), e12729.
[34] Phyllis A Katz and Jennifer A Kofkin. 1997. Race, gender, and young children. *Developmental psychopathology: Perspectives on adjustment, risk, and disorder* 21 (1997), 51–74.
[35] Keunjae Kim, Kyungbin Kwon, Haesol Bae, Ottenbreit-Leftwich Anne, and Krista Glazewski. 2023. Exploring the middle school students' common preconceptions of AI concepts:"AI is a cure-all solution!". In *Society for Information Technology & Teacher Education International Conference*. Association for the Advancement of Computing in Education (AACE), 2124–2129.
[36] Amy J. Ko, Alannah Oleson, Mara Kirdani-Ryan, Yim Register, Benjamin Xie, Mina Tari, Matthew Davidson, Stefania Druga, and Dastyni Loksa. 2020. It is time for more critical CS education. *Commun. ACM* 63, 11 (2020), 31–33.
[37] Lawrence Kohlberg. 1975. The cognitive-developmental approach to moral education. *The Phi Delta Kappan* 56, 10 (1975), 670–677.
[38] Jenna Laib. 2022. https://slowrevealgraphs.com/
[39] Halcyon M. Lawrence. 2013. *Speech intelligibility and accents in speech-mediated interfaces: Results and recommendations.* Illinois Institute of Technology.
[40] Calvin A. Liang, Sean A. Munson, and Julie A. Kientz. 2021. Embracing four tensions in human-computer interaction research with marginalized people. *ACM Transactions on Computer-Human Interaction (TOCHI)* 28, 2 (2021), 1–47.
[41] Chris Lindgren, Jonna Tötterman, and Petter Karlsson. 2022. The Designing for Children's Rights Guide. https://joanganzcooneycenter.org/2022/09/26/the-designing-for-childrens-rights-guide/
[42] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–16.
[43] Andrés Lucero. 2015. Using affinity diagrams to evaluate interactive prototypes. In *IFIP conference on human-computer interaction*. Springer, 231–248.
[44] Jane Margolis, Allan Fisher, and Faye Miller. 2000. The anatomy of interest: Women in undergraduate computer science. *Women's Studies Quarterly* 28, 1/2 (2000), 104–127.
[45] Katherine McAuliffe, Peter R Blake, Grace Kim, Richard W Wrangham, and Felix Warneken. 2013. Social influences on inequity aversion in children. *PloS one* 8, 12 (2013), e80966.
[46] R Steve McCallum and Bruce A Bracken. 1993. Interpersonal relations between school children and their peers, parents, and teachers. *Educational Psychology Review* 5, 2 (1993), 155–176.

[47] Luis C. Moll, Cathy Amanti, Deborah Neff, and Norma Gonzalez. 1992. Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into practice* 31, 2 (1992), 132–141.

[48] Luis Morales-Navarro, Naomi Thompson, Yasmin Kafai, Mia Shaw, and Nichole Pinkard. 2022. Reimagining and Co-designing with Youth an Hour of Code Activity for Critical Engagement with Computing. In *Interaction Design and Children*. 288–296.

[49] Safiya Umoja Noble. 2018. Algorithms of oppression. In *Algorithms of Oppression*. New York University Press.

[50] Lydia Plowman, Joanna McPake, and Christine Stephen. 2010. The technologisation of childhood? Young children and technology in the home. *Children & Society* 24, 1 (2010), 63–74.

[51] Michael T Rizzo, Laura Elenbaas, Shelby Cooley, and Melanie Killen. 2016. Children's recognition of fairness and others' welfare in a resource allocation task: Age related changes. *Developmental psychology* 52, 8 (2016), 1307.

[52] Steven O. Roberts, Carmelle Bareket-Shavit, Forrest A. Dollins, Peter D. Goldie, and Elizabeth Mortenson. 2020. Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on psychological science* 15, 6 (2020), 1295–1309.

[53] Andee Rubin. 2020. Learning to reason with data: How did we get here and what do we know? *Journal of the Learning Sciences* 29, 1 (2020), 154–164.

[54] Sara Safransky. 2020. Geographies of algorithmic violence: Redlining the smart city. *International Journal of Urban and Regional Research* 44, 2 (2020), 200–218.

[55] Jean Salac, Alannah Oleson, Lena Armstrong, Audrey Le Meur, and Amy J Ko. 2023. Funds of Knowledge used by Adolescents of Color in Scaffolded Sensemaking around Algorithmic Fairness. *In Review* (2023).

[56] Linda J Sax, Jennifer M Blaney, Kathleen J Lehman, Sarah L Rodriguez, Kari L George, and Christina Zavala. 2018. Sense of belonging in computing: The role of introductory courses for women and underrepresented minority students. *Social Sciences* 7, 8 (2018), 122.

[57] Marco FH Schmidt, Margarita Svetlova, Jana Johe, and Michael Tomasello. 2016. Children's developing understanding of legitimate reasons for allocating resources unequally. *Cognitive Development* 37 (2016), 42–52.

[58] Zoe Skinner, Stacey Brown, and Greg Walsh. 2020. Children of Color's Perceptions of Fairness in AI: An Exploration of Equitable and Inclusive Co-Design. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[59] Jessica A Sommerville. 2018. Infants' understanding of distributive fairness as a test case for identifying the extents and limits of infants' sociomoral cognition and behavior. *Child Development Perspectives* 12, 3 (2018), 141–145.

[60] Christia Spears Brown and Rebecca S Bigler. 2005. Children's perceptions of discrimination: A developmental model. *Child development* 76, 3 (2005), 533–553.

[61] Carla Strickland. 2019. We needed a wizard Sprite for a #scratch project. Stock wizard sprite is old bearded white man. Wasn't feelin it. Was excited to see a partner T made a custom wizard sprite BUT they only changed the adjectives "old" & "bearded. Time for some #BlackGirlWizardMagic. https://twitter.com/CisforCarla/status/1128844596421779456

[62] Fumihide Tanaka, Aaron Cicourel, and Javier R. Movellan. 2007. Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences* 104, 46 (2007), 17954–17958.

[63] Randi Williams, Christian Vázquez Machado, Stefania Druga, Cynthia Breazeal, and Pattie Maes. 2018. " My doll says it's ok" a study of children's conformity to a talking doll. In *Proceedings of the 17th ACM Conference on Interaction Design and Children*. 625–631.

[64] Aman Yadav, Marie Heath, and Anne Drew Hu. 2022. Toward justice in computer science through community, criticality, and citizenship. *Commun. ACM* 65, 5 (2022), 42–44.