



Funds of Knowledge used by Adolescents of Color in Scaffolded Sensemaking around Algorithmic Fairness

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

Alannah Oleson
olesona@uw.edu
University of Washington
Seattle, WA, USA

Lena Armstrong
lena318@sas.upenn.edu
University of Pennsylvania
Philadelphia, PA, USA

Audrey Le Meur
lemeu001@morris.umn.edu
University of Minnesota, Morris
Morris, MN, USA

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

ABSTRACT

With the ubiquity of computing technologies, adolescents are increasingly affected by algorithmic biases. While previous work provides insight into adolescents' perceptions of algorithmic bias, few provide guidance on how to engage adolescents in discourse on algorithmic bias that prioritizes both their agency and safety. To address this, we developed and conducted group discussions and design activities based on three scenarios of algorithmic bias with 15 adolescents of color (ages 15-17) in a summer academic program in the United States targeted at students from families with low-income backgrounds or who would be the first in their family to pursue post-secondary education. When sensemaking, all participants considered factors beyond the scenarios, using their situated knowledge to contextualize perceptions of unfairness. They also considered sources of bias and impacts of unfairness at different levels of individuals, communities, and society. However, when designing solutions, they tended to design for hypothetical "average users" instead of considering nuances of user populations. We offer insights for algorithmic fairness learning experiences that support situated reasoning in adolescents.

CCS CONCEPTS

• **Social and professional topics** → **Computing education**; **Adolescents**; *Computing literacy*.

KEYWORDS

algorithmic fairness, adolescents, sensemaking, funds of knowledge

ACM Reference Format:

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 *Proceedings of the 2023 ACM Conference on International Computing Education Research V.1 (ICER '23 V1)*, August 07–11, 2023, Chicago, IL, USA. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3568813.3600110>



This work is licensed under a Creative Commons Attribution International 4.0 License.

ICER '23 V1, August 07–11, 2023, Chicago, IL, USA
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9976-0/23/08.
<https://doi.org/10.1145/3568813.3600110>

1 INTRODUCTION

While computing provides immense benefits, it can also amplify oppression [4, 48, 52]. Prior work has explored youth and adolescents' perceptions of algorithmic bias, largely in the context of artificial intelligence (AI). Researchers observed that children often overly trust AI agents, limiting their ability to critically analyze AI technologies [20, 30, 41, 60, 64, 71]. Others discovered that with some instruction, children were capable of identifying unfair treatment from AI [22]. In contrast, Lee et al. [38] found that adolescents had both positive and negative intuitions around biased algorithms, no longer overwhelmingly positive like in younger children. These distinctions between children and adolescents mirror distinctions in conceptions of fairness with age and development. Adolescents often examine contextual factors at various scales when evaluating fairness as opposed to children, who frequently equate fairness with equal distribution [5, 14]. This growth suggests increased capacity for critical examination of algorithmic bias in adolescence.

Much prior work on algorithmic bias has specifically focused on AI and agents, but far more than just AI impacts youth and adolescents. There is emerging work on youth and adolescents' conceptions of algorithmic bias *beyond* AI, with Coenraad et al. discovering that youth were aware of explicit negative effects of technology more broadly, not only AI, without direct instruction [11]. However, previous work provides limited guidance on how youth and adolescents may participate in discourse around algorithmic fairness¹ in a way that allows them to draw from their existing knowledge and experiences but also safeguards their well-being in discussing potentially difficult topics. Understanding how we might engage adolescents in learning experiences around algorithmic fairness that balances both agency and safety will support the growing movement to educate them in the social impacts and ethics of AI [1, 65, 66] and computing more generally [24, 25, 32, 56, 68, 76].

In this investigation, we take a *funds of knowledge* stance on education. The funds of knowledge approach asserts that learners already have sophisticated skills and knowledge based on their lived experiences, social circumstances, and privileges [27, 45, 47], and that learners from marginalized backgrounds can have different

¹With our participants, we chose to use the term "fairness", instead of "bias" because prior work [22, 38] showed that "bias" may not be in youth and adolescents' vocabulary and they have limited comprehension of the word itself. We use "fairness" when specifically discussing our participants, but use the terms interchangeably elsewhere in the paper.

knowledge, though their experiences are often delegitimized in academic systems. Under this perspective, a key aspect of improving adolescent education on algorithmic fairness is understanding *what* these skills and knowledge are, so that learners from marginalized backgrounds may be equitably incorporated in the growing movement for socially responsible computing education.

To address these gaps, we explored two research questions. When engaging in scaffolded sensemaking:

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

To understand adolescents' existing skills and knowledge, we also draw from *sensemaking theory*, which posits that individuals process information from various sources to achieve understanding, rather than achieving an arbitrarily sufficient "amount" of knowledge [17]. Drawing from the *slow reveal graphs* instructional routine from math and data science classroom practice [35], we developed group sensemaking discussions and activities based on three scenarios of algorithmic fairness as *design probes* [44], employing these discussions as both instruments for discovery and blueprints for future classroom practice. We conducted these discussions and activities with 15 low-income and first-generation adolescents (ages 15-17) with diverse ethnic and linguistic backgrounds at a summer program in the United States. Since our current understanding of adolescents' existing funds of knowledge around algorithmic biases is nascent, this study is formative: our goal is to characterize our study population's experiences and understand the funds of knowledge that might guide their sensemaking processes, not to compare populations, contribute generalizable insights, or make causal claims.

This study makes two key contributions. First, we contribute to better understandings around the funds of knowledge that adolescents may bring into learning experiences on algorithmic bias. This study focused on adolescents of different ages, ethnic, linguistic, and socioeconomic backgrounds from prior work [11, 22, 38], whose perspectives are understudied yet critical to equitable education. Second, we contribute a blueprint for scaffolding discussions around algorithmic bias with adolescents that prioritizes both their agency in incorporating their existing skills and knowledge, as well as safety in engaging in a possibly challenging topic. This study provides insights that can inform the efforts of computing educators, curriculum designers, and other stakeholders working to advance critical computing education.

2 BACKGROUND

2.1 Youth & Adolescents' Perceptions of Fairness

Research from psychology and sociology indicates that adolescents have sophisticated conceptions of fairness, though definitions of fairness change with age and development. Abilities to reason around fairness grow from mainly revolving around equal distribution or treatment (in young children) to considering different contextual factors and learning that equality does not necessarily

mean fairness (in older children and adolescents) [5]. These abilities also shift over time from being more self-oriented to learning to perspective-take and consider impacts for others at expanding scales [14]. Adolescents are capable of resolving conflicts around fairness within familial and friend relationships [8, 40, 70]. Adolescents also exhibit nuanced understandings of fairness in group settings, such as coordinating moral and group concerns, recognizing how group dynamics contribute to unfairness, and considering the status of disadvantaged groups [55]. Scholars have also explored how adolescents negotiate unfairness at a societal level. Adolescents are keenly attuned to their own unfair social circumstances [2], group-level exclusion, inequality of opportunity, and unfairness in society more broadly [3, 9, 46]. Further, they reflect this awareness in their moral and non-moral (e.g. personal choice) reasoning and their judgments of others' emotions [2, 9]. Adolescence presents a promising developmental stage at which individuals can effectively reason about the intricate and highly contextual nature of algorithmic fairness.

2.2 Youth & Adolescents' Perceptions of Algorithmic Fairness

This study also builds upon prior work investigating perceptions of algorithmic bias. Most prior work investigates adults within a range of scales and contexts. For instance, in a study about algorithms making social division decisions (e.g. dividing chores among roommates), Lee et al. found that participants perceived algorithmic decisions as less fair because they did not account for multiple concepts of fairness [37]. As for bias at a larger scale, Woodruff et al. explored how participants from marginalized race or class backgrounds in the United States felt about algorithmic fairness, finding that it elicited negative feelings about racial injustice and economic inequality [74]. In contrast, a similar study in India found that AI evoked an unquestioning aspiration in their participants due to the lack of an ecosystem to interrogate high-stakes AI [58].

For youth and adolescents specifically, Long et al.'s literature synthesis [41] identified several studies suggesting that children often overestimate agent intelligence [21] and consequently overly trust agents [20, 30, 64, 71]. This attribution of socio-emotional characteristics to AI agents can inhibit critical examination of AI [41]. This was reinforced by a more recent study by Skinner et al. [60], who found that children equated kindness with fairness, using kind communication with people to justify fairness. However, Druga et al. discovered that after showing them videos of algorithmic bias examples, children were able to connect those examples to their daily lives, identifying situations of unfair treatment from AI based on race/ethnicity, age, and gender [22].

In contrast, the literature on adolescents (around ages 14-18) is relatively sparse. Early work from Lee et al. [38] found that teenagers had varied intuitions regarding biases in data-trained technologies, ranging from human biases manifesting in technology to technology expressing preferences for some content over others. They also wrestled with the costs, such as limiting access to resources, and benefits, such as a better user experience, of different algorithms. These differences between children and teenagers suggest a growing, yet under-explored, potential for critical examination with development.

Nonetheless, most prior work only examined perceptions of AI, but algorithms *beyond* AI also influence adolescents' lives. Some researchers have recently started investigating conceptions of algorithmic bias beyond AI in youth. Coenraad et al. discovered that without direct 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 [11]. While this prior work provides insights into youth and adolescents' initial conceptions of technological biases more generally, our understanding of adolescents' reasoning around such biases is still nascent. Therefore, our study is formative in that we aim to richly characterize adolescents' reasoning around algorithmic bias and their potential contributing factors, providing foundational understandings for future work to build upon, instead of making generalizable or inferential claims about the population. Further, this study presents a blueprint for incorporating adolescents' existing knowledge and backgrounds in learning experiences around algorithmic bias in the growing movement for critical computing education [32, 76].

2.3 Funds of Knowledge & Sensemaking Theory

In our study, we draw from funds of knowledge and sensemaking theories to support adolescents in bringing their conceptions of fairness into computing. The funds of knowledge approach posits that learners, especially from marginalized backgrounds, already have various skills, knowledge, and competencies from their lives and their communities [27, 47]. This approach asserts that these assets are frequently invisible or delegitimized because of asymmetrical power relationships in education, and educators should identify and incorporate these skills when designing learning experiences. Recently, this approach has been applied in STEM education, improving educational practices and student learning outcomes [16, 63]. Previous research showed that employing funds of knowledge can enable adolescent learners to align their everyday skills and bodies of knowledge with engineering practices [72] and to draw from diverse community resources and from multilingual literacy practices to attain forms of economic, cultural, and social capital in science [73]. In addition to their existing ideas around fairness, we propose that adolescents will be able to leverage their funds of knowledge in reasoning about algorithmic fairness. A key objective of this study is to better characterize the funds of knowledge that adolescents of marginalized ethnic, linguistic, and socioeconomic backgrounds already have with respect to algorithmic fairness. As this knowledge is often undervalued within educational systems, it is crucial that we identify and leverage this knowledge to further socially responsible computing education.

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, as opposed to simply receiving information through direct instruction. In computing, sensemaking practices allow youth and

adolescents, regardless of background, to play an active role in learning various concepts, such as programming [51, 69], AI [18, 19], data literacy [54], and online safety [75]. As our goal is to provide opportunities for adolescents to draw from their funds of knowledge, we ground our methods in sensemaking theory to make space for different paths of achieving understanding. Sensemaking activities may offer an approach to leverage adolescents' funds of knowledge in algorithmic bias learning experiences in formal classroom settings.

3 SLOW REVEALING ALGORITHMIC (UN)FAIRNESS

Consistent with the funds of knowledge approach, we wanted to allow our students to use their existing skills and knowledge, which are often devalued in academic environments [72]. Congruent with this stance, sensemaking theory emphasizes the processing of information from various sources to achieve understanding. Since there is limited work on the funds of knowledge of adolescents of color with respect to algorithmic fairness, we developed sensemaking discussions as *design probes*. Design probes are an approach from human-computer interaction (HCI) research for learning about how people interact with and conceptualize technology in the world; what makes design probes unique is that they are specifically contextual. Design probes have three key characteristics [44]. First, probes are based on user participation, typically comprised of a collection of assignments through which users can document their experiences, thoughts, and ideas. Second, probes focus on users' individual contexts to characterize human phenomena and users. Lastly, probes are exploratory in nature, meaning that they delve into new opportunities, rather than solve problems that are already known. This final characteristic distinguishes *design probes* from the more commonly used design-based method in computing education research, *design-based research*, where researchers iteratively refine an intervention to solve a known learning problem [43].

3.1 Design of Sensemaking Discussions

We designed our sensemaking discussions to fulfil the design probe characteristics as follows. First, our discussions supported *user participation* by allowing students to express their reasoning through worksheets and design activities. Second, they focused on *users' contexts* because these discussions took place in the authentic learning context of a classroom. Lastly, our discussions are *exploratory* because there is not much work on the funds of knowledge of students with these ages, ethnic, linguistic, and socioeconomic backgrounds in regard to algorithmic bias. However, our design probe differed from traditional design probes in that we only minimally supported reflection from an important user group, namely the instructors (first, third, and fourth authors). Future work could investigate the experiences of instructors using these discussions in different contexts.

In addition to serving as an instrument for exploration, we developed these sensemaking discussions as a possible blueprint for how instructors may scaffold students in drawing upon their funds of knowledge in their classroom. In our design, we drew from sensemaking practices in math and data science education, in particular

slow reveal graphs [35]. Slow reveal graphs are instructional routines that uses scaffolded visuals and discourse to make sense of data. These routines start with a graph with minimal information, prompting students to develop hypotheses about the graph. At each incremental reveal of more information on the graph, students are scaffolded in making sense of the information and refining their interpretations. In our discussions, we adapt this routine for different layers of algorithmic bias, encouraging students to form their own interpretations at each revealed layer.

Table 1 shows the three sensemaking discussions that students engaged in. 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, similar to slow reveal graphs — 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 (Table 1).

To facilitate students' sensemaking and provide artifacts for us to analyze, each sensemaking discussion involved (1) a *warm-up question* before introducing the scenario, (2) a *worksheet* with reflection questions for each layer revealed, and (3) a semi-structured big paper *design activity* where students brainstorm ideas on big paper to support the unconstrained generation of ideas [12, 26].

Throughout the discussions, we made intentional choices to prioritize the safety and agency of students, accounting for our students' backgrounds and the power imbalance between the researchers/instructors and students. Therefore, we did not directly ask students about their harmful experiences with technology, since that might cause distress or force participants to disclose experiences they wished to keep private. Instead, we selected scenarios that might resonate with them based on prior literature. If they brought up their own experiences, we wanted them to do so on their terms. We also made deliberate terminology choices to minimize reliance on prior computing knowledge. As with "fairness" (see footnote in Section 1), we used the term "rules" to describe the algorithm and contextualized the data used in each scenario (e.g. voice recordings in the Smart Speaker scenario), rather than simply using the term "data".

3.1.1 Scenario Design. We created three scenarios of algorithmic decision-making that surfaced potential fairness issues to seed sensemaking discussions. These scenarios were selected precisely because they do not have straightforward conceptions of fairness and thus, would encourage debate among students.

- (1) The Search Engine ('Search') scenario was based on biases in representation from search results [48, 62].
- (2) The Smart Speaker ('Speaker') scenario was based on the failure of many voice recognition systems to recognize other languages or accents [36].
- (3) The School ('School') scenario was adapted from the scenario used in [23] to understand youth's perceptions of social resource inequality to reflect algorithmic redlining [57].

We presented scenarios in this order to highlight an increasing scope of harm. 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 2 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, with the Search Engine following 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 in the Speaker scenario, they differed based on English or non-English country of origin.

3.1.2 Warm-up Questions. At the beginning of each sensemaking discussion, students discussed a warm-up question together. These questions asked students to share their own experiences and were intended to help them get comfortable reflecting and voicing their perspectives. Table 2 shows the question used in the School scenario.

3.1.3 Worksheets. After warm-ups, students completed worksheets designed to scaffold the sensemaking process. For each layer (e.g. algorithm, data) that was revealed, we prompted students with accompanying questions that were focused on either: (1) understanding or (2) evaluating the decisions in each scenario to encourage divergent or convergent thinking, respectively (Table 2). Following each question, students wrote down their responses individually on the worksheet before discussing them as a group.

3.1.4 Design Activities. Once the worksheet was complete, students then brainstormed ideas about the fairness of the algorithm and the data as a group. Consistent with the big paper method [26], they wrote their ideas either directly on a big piece of paper or on sticky notes, which were then placed on the big paper (Figure 1).

For the design activities, we prompted students 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 these discussions, pilot participants sometimes fixated on the level of agency they had to effect change in each scenario, getting preoccupied with whom they would answer to instead of the task at hand. Throughout the activity, we prompted students to consider the fairness of the different layers of decision-making they designed (see Table 2 for specific prompts).

4 METHODS

Building on prior literature and existing classroom practice, our study sought to scaffold adolescent sensemaking to better understand their reasoning around algorithmic bias, with participants of different ages, ethnic, linguistic and socioeconomic backgrounds from previous work.

Scenario	Seed Text	Situation	Computer	Algorithm	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 Maximó, who just moved to the US from Mexico.	✓		✓	✓	✓
School ('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 1: Seed Text & Layers Discussed in Each Scenario.

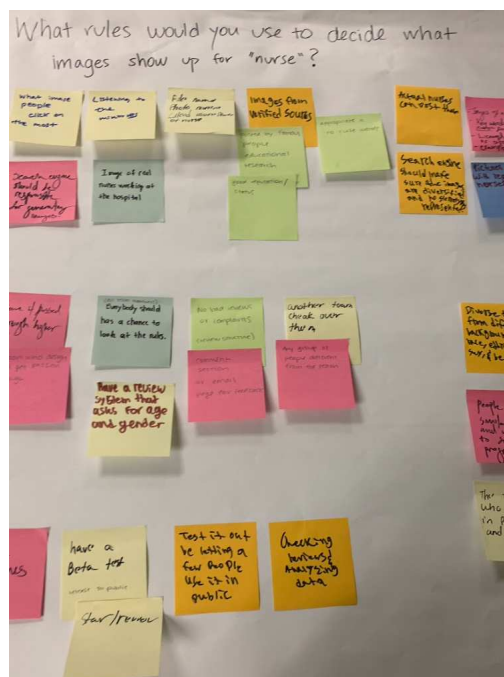


Figure 1: Example from a Big Paper Design Activity

4.1 Study Context

As the goal of this study was to contribute a formative understanding of adolescents’ sensemaking of algorithmic fairness for socially responsible computing education design, it was important that our study occurred in an authentic learning environment [33]. We conducted this study during a 6-week summer program (June-August 2022) at a northwest United States university aimed at students ages 14-18 from local under-resourced schools who were low-income

and/or the first in their family to pursue a post-secondary education (i.e. first-generation). Though students received official school credit for classes they took in the summer program, it differed from the academic school year in several ways. First, teachers at local schools nominated students to join the program. Second, the program provided students lunch allowances and a stipend to offset the opportunity cost of a summer job. Lastly, it was designed to build rapport among students through study groups and field trips.

The first author was the lead instructor of a computing-based elective class in the program, while the third and fourth authors

Stage	Phase	School: Questions & Revealed Layers
Warm-Up	—	<i>Ask:</i> 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		<i>Reveal seed text (Table 1))</i>
Computer	Understanding	<i>Ask:</i> Why do you think School A has more supplies than School B? <i>Reveal:</i> A computer decided how much supplies each school should get. <i>Ask:</i>
	Evaluation	a. What do you think of a computer making that decision?
Algorithm	Understanding	b. Why do you think a computer decided to give School A more supplies than School B? <i>Reveal:</i> 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.” <i>Ask:</i>
	Evaluation	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		<i>Reveal:</i> 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. <i>Ask:</i>
	Evaluation	a. What do you think of the data that the computer used?
Team	Evaluation	b. How fair is it that the computer used past data? Why? <i>Reveal:</i> The team who designed the rules and data the computer used was made up of all White people. <i>Ask:</i>
	Evaluation	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 Activity		<i>Ask:</i>
	Brainstorming	- 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?
	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 2: School Sensemaking Discussion Questions in full. *Italics* denote actions performed by facilitators.

were co-instructors. Since the study was conducted as part of instruction, our university institutional review board granted this study exemption. We managed informed assent by allowing students to opt out or assent to different levels of participation after describing the nature of the research.

4.2 Participant Demographics

Out of the 20 students in the computing class, 15 assented to their classwork being analyzed for research through a form on the first day of instruction. The form also included free-response questions for research participants to self-disclose their gender identity, ethnic identity, languages spoken at home, disabilities, and other aspects

of their identity they would like the instructors to know. Participants who assented were also asked to choose their pseudonym; if they did not provide a pseudonym, we use their initial (Table 3). All participants had personal smartphones and school-administered laptops at home because of pandemic remote learning. However, we did not ask for any more information about their prior education experiences with AI, data, or computing more broadly. Prior work has linked learners’ perceptions of having less prior computing education experience than their peers with a lower sense of belonging and confidence, especially for learners who identify as women or underrepresented minorities [42, 59]. As almost all our participants had either or both of those identities, asking them about their

prior computing experiences may have negatively influenced their participation in both the study and the class.

4.3 Study Timeline

In the weeks preceding the class, the lead instructor and co-instructors (first, third, and fourth authors) trained to facilitate sensemaking discussions by creating detailed protocols and piloting with other undergraduate and graduate researchers. This resulted in several iterative refinements to the protocols and scenarios until all authors felt confident in their ability to lead a discussion.

To acclimate participants to the sensemaking discussion structure, Days 1 and 2 of the class included practice activities. On Day 1, participants did big paper design activities to collaboratively decide upon class expectations. On Day 2, they completed a worksheet reflecting on their own experiences with computing that had a similar structure to our sensemaking worksheets. Facilitating these activities also provided the instructors additional practice.

On Days 3–6 of the class, the first and third author facilitated a series of sensemaking discussions with two groups of 7 participants. The fourth author facilitated a group in which only one student assented, so the third group's design activity data was discarded. Participant discussed Search and Speaker scenarios on Days 3 and 4 respectively for 60 minutes each, followed by the School Scenario on Day 5 for 60 minutes and Day 6 for 30 minutes (Table 1). Once data collection was over after Day 6, we debriefed with participants on the discussions in subsequent classes, allowing for additional instruction and clarification of any misconceptions about computing or potentially harmful stereotypes that arose during discussions.

4.4 Data Collection & Analysis

For each sensemaking discussion, we collected worksheets from 15 participants, except for the Speaker scenario where one participant lost their worksheet (giving us 14 worksheets for that discussion). This provided us 44 total worksheets to analyze. We also collected a big paper board from two groups (seven participants each) for each discussion's design activity, for a total of six design boards.

To analyze the worksheets, we took an inductive thematic analysis approach using participants' responses as our data source. 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 [28] on qualitative coding that uses codes as an organizational aid for thematic claims about the data.

The first four authors collaboratively developed our codebook (Table 4) over three rounds of inductive coding, allowing the worksheet data to guide our analysis [61] and discussing major themes that arose as we became more familiar with the data. Our codebook stabilized into three "facets" representing different aspects of our participants' sensemaking: the *salient factors* participants used while reasoning about fairness, the *evaluations* they made, and the *alternatives* they brainstormed about. Each of these facets in turn encompassed several codes representing low-level patterns present in participants' responses. Once the codebook was stable, the third and fourth authors then coded the scenario worksheets, adopting the practice of taking participants' responses "literally" to

minimize inference. After coding batches of 5–7 worksheets independently, the third and fourth authors resolved disagreements and built consensus for each batch of worksheets. The descriptions for both the *evaluations* and *alternatives* (Table 4) required the most discussion because they required coders to summarize them without information loss.

We then used a subset of our codebook to analyze the design boards. The first author coded the ideas written on each board, with the second author verifying their codes and resolving uncertainties through consensus. During analysis, we found that the "salient factors" and the "brainstorming alternatives" facets were the only ones that applied to design boards. We also added a "subject" to the "brainstorming alternatives" facet for design board analysis, since the focus of the board activities was brainstorming and we found that more granularity was needed.

Finally, the first and second authors conducted a post-hoc analysis of codes to synthesize higher-level themes. We started with the three levels of human factors from Brown et al.'s work on undergraduate ethics education in AI, where they classified the different considerations of human factors in AI solution design, namely (1) no human factors used, (2) human factors used, but limited to those given in the example, and (3) human factors used beyond those explicitly included in the example [7]. We drew from Brown et al. because similar to us, they studied an ethics activity with multiple conceptions of fairness within an authentic classroom environment. While we drew from Brown et al.'s framework for human factors in our analysis, we did not draw from their technical factors as their framework was designed for advanced undergraduates in computing, which would not be suitable for our adolescent participants with limited computing experience. As the participants considered many factors that were not explicitly included in the scenario, the first and second authors further analyzed the factors that supported participants' sensemaking, iterating over different ways of representing them to best illustrate the relationships between them. After 3 rounds of iteration lasting approximately 2 hours each, we decided that the metaphor of a camera would best organize the relationship between the factors, which we will further elaborate on in the Results section.

4.5 Author Positionality

Positionality statements allow for transparency in how the identities of the authors relate to the research topic and the identities of the participants [31, 39, 53]. Each of the following statements was written by the respective author to describe experiences and perspectives that impacted their engagement with this project.

The first author saw her own experiences reflected in many of the participants. Like most of the participants, she is an immigrant, grew up with economic difficulties, and was the first in her family to pursue a post-secondary education in the US. She also shared a racial identity with half the participants. In her decade in the field, this was the first computing classroom she had been in that comprised entirely of students of color. She felt immense responsibility as not only the lead instructor, but also the only instructor of color. These facets of her identity and her lived experiences with systemic marginalization in computing and society led to her research interests in critical computing literacies for youth and adolescents. She

Pseudonym	Age	Gender	Ethnicity	Languages Spoken at Home	Disability	Other Identities
A	15	Male	Hispanic/Latino	English & Spanish	—	—
Batman	16	Woman/Female	Black, Somali	English & Somali	Fatigue & Nausea	Muslim
Becky	16	Female	Mexican	Spanish	—	—
Bob	15	Female	Vietnamese	Vietnamese & English	—	—
Dairy	17	Male	Black/Somalia	English/Somalia	—	—
J	16	Female	Half Korean, Half Argentinian	English & Korean	—	—
Jake	17	Male	Somali	English/Somali	—	—
Kevin	15	Female	Vietnamese	Vietnamese	—	—
Miranda	15	Female	Somali	English	—	—
R	17	Girl (she/her)	Asian	Vietnamese	—	—
Sophia	16	Girl (she/her)	Chinese American	English & Chinese (Toisanese/Taishanese)	—	—
Stewart	15	Male	Laos	Laos	I don't know	—
T	16	He/him	Asian	Vietnamese	—	—
Z	16	Male	Afghan	Pashto, Dari, English	—	—
Zoro	16	Male	Vietnamese, Asian	English & Vietnamese	—	—

Table 3: Participant Demographics. “—” indicates that the participant declined to disclose disability status and/or other identities.

Facet	Explanation	Example Coding
Salient Factors ^{wd}	What factors are participants using to sensemake?	<i>“There are probably more well-funded schools in the school a area, meaning that school B has more kids making supplies more spread thin.”</i> - Batman Salient factors: “funding”, “geographic location”, “number of students”
Evaluation (Subject) ^w	What are participants evaluating?	<i>“I think they didn't think about other people when making the product. And focused more on satisfying people similar to them.”</i> - Jake Subject: “Team”
Evaluation (Description) ^w	What is the evaluation?	Description: “Inconsiderate of users unlike them”
Brainstorming Alternatives (Subject) ^d	What is the subject of the alternative?	“what images people click on the most” Subject: “rules”
Brainstorming Alternatives (Description) ^{wd}	What alternatives are they proposing	Description “use most clicked images”

Table 4: Codebook. Facets used for worksheets and design boards are denoted with “w” and “d”, respectively.

was the driver of this research project in understanding adolescents' perspectives.

The second author was brought onto the project for their expertise in inductive qualitative methods. With several years of experience studying the overlap of HCI, design methods, and computing education, they supported the first, third, and fourth authors in initial analyses and worked closely with the first author to synthesize broader themes. They are motivated by their own lived experiences growing up in a low-income, predominantly White, rural area that lacked access to computing education, and later experiences navigating systemic marginalization within the computing field based on these and other identity facets. They approach this work from a design justice perspective, with the conviction that enabling systematically marginalized folks to better understand the technologies around them is a critical first step toward liberation.

The third author is an undergraduate studying cognitive science and computer science with an interest in how algorithms impact different groups of people. She has had previous teaching experience

in classrooms composed of mostly students of color and believes in the importance of creating positive experiences for all students in computing. Her experiences with gender discrimination in computer science motivated her to partake in this research project to create more inclusive spaces and technologies within computing.

Like the participants, the fourth author was educated in a public urban school district. That said, most of the classrooms she was educated in were majority White, due to the districts' tracking practices. This, along with her experiences of gender discrimination in STEM spaces, motivates her to work to improve access to computing education. The fourth author also shared some experiences with participants who are children of immigrants, since one of her parents immigrated to the US as an adult. She comes to the work with the goal of helping adolescents develop positive relationships with computing.

The fifth author comes to this work with an interest in critical literacies about computing, a lifetime of working with youth and adolescents in education contexts, and a curiosity about how they

invoke moral and ethical ideas to reason about non-human decision making in society. Her interests in this space largely stem from her own marginalization in computing and society, and lived experiences with the unfair and oppressive ways that algorithms and data have shaped her rights and opportunities. She also approached the work as a facilitator and advisor, supporting the other authors in exploring adolescents' perspectives.

5 RESULTS

Through our analysis, we identified many different salient factors in our participants' sensemaking of algorithmic fairness. To relay our results, we organize the following section using the metaphor of a camera, specifically lenses and filters (Figure 2). Upholding Hammer & Berland's stance that the purpose of qualitative analysis is to generate claims about themes present within situated data [28], the role of this metaphor is not to propose a framework or generalizable model for sensemaking, but instead to make explicit the connections between the factors and to structure the presentation of the following results. Future work may explore the suitability of this metaphor for adolescents' sensemaking of algorithmic fairness in different contexts, though that was not the goal of this investigation.

In the metaphor that emerged from the data, we focus on the lenses and filters of a camera. Lenses allow photographers to change the scale and resolution of a shot, while filters allow photographers to change the kinds of light in a shot. Photographers can attach different filters to a lens to capture the same view but with different lights, making the final image appear different. In this metaphor, our participants are photographers, capturing algorithmic (un)fairness in different scales and lights to make sense of them. Each participant has their own camera and their own set of lenses and filters, which can grow over time. We identified two different *lenses* participants used to make sense of (un)fairness at different scales and resolutions: (1) a human lens, which ranged from individual to societal factors, and (2) a technical lens, which ranged from individual technology creators to broader technical factors. In addition to adjusting the scale and resolution with their chosen lens, we found that participants used different characteristics, such as gender and class, as *filters* to change what was most salient to their sensemaking for each scenario.

We describe the lenses followed by the filters because just like photographers who decide which lens to use before attaching filters, participants first have to decide what they are sensemaking about (lenses) before deciding which characteristics are most salient (filters). We present data in two ways: (1) worksheet response quotes attributable to a single participant and (2) ideas from each scenario's design boards only attributable to groups. Individuals' quotes will be *italicized*, while group ideas will use monospace font to distinguish them from each other. We also report participant counts out of 15 total participants (14 for Speaker scenario), though we caution against inferring anything past observed frequency from these provided numbers.

5.1 Human Lens: Different Scales of Human Groups

Under the human lens, we found that factors were relevant in increasing group sizes: individual, community, and society.

5.1.1 Individual. This level encompasses factors attributed to individuals, usually the participants themselves. Participants used stereotypes they espoused themselves (Search: 4/15), lived experiences (Search: 3/14), and principles (Search: 12/15; Speaker: 10/14; School: 14/15). When discerning why most of the images for nurses were women in the Search scenario, T stated the stereotype of women as caregivers:

"Women are better than men at taking care of people."

Others drew from their own lived experiences, such as J in the Speaker scenario:

"My mom gets upset so she just switches to typing than using speech."

Participants also used their principles, or their views of right and wrong. When evaluating an all-men team for the Search scenario, R asserted:

"If it comes from men, the information will not be complete and diverse."

While participants often utilized their own backgrounds to ground their understanding and evaluation of unfairness when filling out worksheets, that same grounding did not always translate to the brainstorming during the design activities. When deciding on individuals to test their algorithms with in the Speaker scenario, participants wrote "average consumers or everyday consumers" and "people off streets". The vagueness of these responses given during brainstorming stood in stark contrast with the situated reasoning participants displayed when they were understanding or evaluating. It was not clear who these individuals were, if participants viewed themselves as the "average consumers" or if they were referring to other individuals.

5.1.2 Community. The "community" level of the human lens includes factors that participants discussed about a collective group of people. This level only emerged in the School scenario, where community was made especially salient, but all 15 participants engaged in community-level reasoning throughout the whole School discussion.

Participants wrestled with the connections between a community's geographic location and wealth, both of which were explicitly referenced in the School scenario. For example, Dairy hypothesized:

"School A area is a rich place where I assume doctor, engineer or business people but area two is where low income people lives."

Not only did they engage with a community's characteristics as explicitly stated, some participants also conceptualized *who* was in a community, a factor that they introduced to the discussion. When brainstorming alternatives to the algorithm in the School scenario, some participants wanted to include various community members, with ideas like "teachers, staff, neighborhood should be involved". However, similar to the individual level, a tendency towards ambiguity emerged in the brainstorming phase that was not present in the understanding or evaluation phases. For instance,

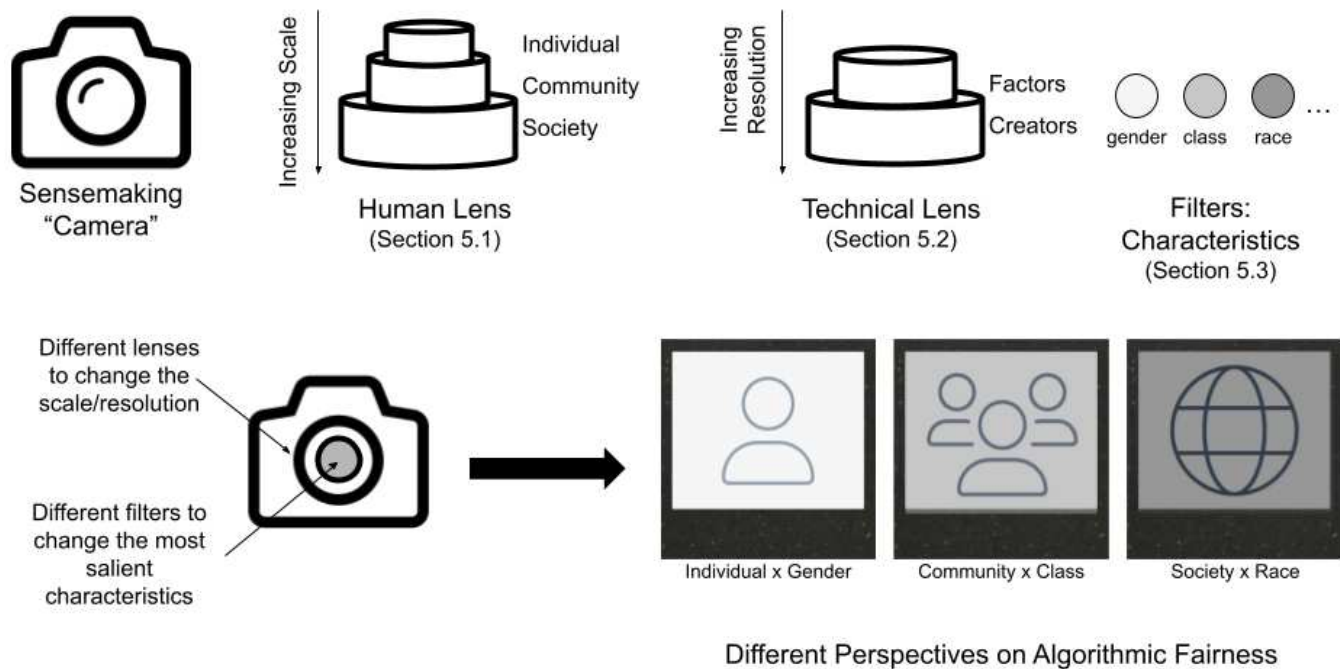


Figure 2: Camera Metaphor as an Organizational Aid for the Factors in Adolescents of Color's Sensemaking of Algorithmic Fairness

one idea was to “have people fill out suggestion for rules anonymously and have the community look over it”, but did not specify *who* the suggesters or community members might be.

5.1.3 Society. This level covers factors participants ascribed to structural issues. Participants reasoned about societal issues through history (Search: 8/15), societal stereotypes (Search: 9/15), power distribution (School: 7/15), and systemic marginalization (Search: 4/15; Speaker: 2/14; School: 8/15).

When understanding the reasons for the disproportionate representation of women nurses in the Search scenario, some participants attributed it to history, such as A:

“They probably made woman work as nurses historically as they wanted men to go to war.”

Other participants blamed societal stereotypes of nurses. In this case, participants did not espouse the stereotypes themselves and instead, identified the problematic ideas as “stereotypes”, such as Kevin:

“Our society associates nurses as women and not men.”

Participants were particularly attuned to the uneven distribution of power in the School scenario. In evaluating the algorithm used to distribute supplies, Zoro lamented the lack of agency the students had:

“Some families [...] can barely provide for themselves. It's just unfair for the kids because they have no control.”

Participants identified systemic marginalization across all three scenarios. In the Search scenario, Kevin criticized the algorithm:

“When searching up nurses you might see majority White women and less of women or people of color. I think the rules create a dominant narrative.”

While participants were adept at integrating societal issues in their understanding and evaluation of unfairness, they had more difficulty incorporating this knowledge in brainstorming alternatives. Most ideas of how to account for broader society were ill-defined, such as “ask everyone about the rules” (Search); “see if the public agrees” (Speaker); and “Come up with a final solution including every point of view” (School). These vague mentions of “everyone”, “public”, and “every point of view” used during brainstorming starkly contrast with detailed descriptions participants developed when understanding or evaluating aspects of fairness.

5.2 Technical Lens: Different Resolutions of Technology-Related Factors

Participants employed the technical lens to make sense of technology-related factors resulting in algorithmic unfairness at different resolutions. This theme was not as well-characterized or often used as the human lens, which was expected given that we did not require prior computing knowledge of our participants.

5.2.1 Technology Creators. We use the term “technology creators” to cover programmers, designers, engineers, and other creators because participants did not meaningfully distinguish between these roles. Participants accounted for the qualifications (Search: 5/15; School: 4/15) and biases of the creators (Search: 6/15; Speaker:

5/14; School: 8/15), as well as the power dynamics within teams of creators (Search: 4/15).

Participants had varied ideas for what made creators “qualified” for the job, ranging from pure technical skill to shared backgrounds with users. When evaluating a team from non-English speaking countries in the Speaker scenario, Becky emphasized the importance of local, cultural knowledge in creators:

“its still important to have someone from English speaking countries for them to put slang and add others things only people from those countries would know.”

Participants also attributed unfairness in the scenarios to biases held by the creators. When it was unveiled that a computer was responsible for supply allocation in the School scenario, R placed the blame solely on the creator:

“It only follows what the programmer does and maybe the programmer is an unfair person.”

Lastly, participants accounted for the power dynamics within a group of creators, specifically in the Search scenario. Stewart cited gendered workplace issues when evaluating an all-men team:

“It’s unfair because females can also do the work they are doing it’s just they don’t get noticed as much.”

In this response, Stewart used socialized gender norms in his sense-making process.

While not as thorough as the human lens, this was the more detailed of the two levels as the most “human” part of the technical lens, suggesting its potential as a gateway for adolescents to further exploring technology-related factors when sensemaking around algorithmic fairness.

5.2.2 Other Technical Factors. This level covers factors attributed to the computer or other aspects of technology, such as accuracy and efficiency. Participants engaged with this level sparingly, which was unsurprising given their limited computing backgrounds. However, an interesting tension emerged when participants did use it — a tension between simplicity of representation for the computer versus the complexity of humanity, which emerged in the Search scenario (2/15). In her evaluation of the all-men team, Miranda wrote:

“The rules for the computer are just the most accessible and straightforward designing. But in general, it may benefit from diversity.”

A design board idea echoed the same sentiment: “Letting everyone give their opinions and seeing which ones can be added into the rules without complicating anything”. These responses revealed a desire to incorporate various human considerations conflicting with the limitations of technology, though also an uncertainty about how best to do so.

5.3 Filters: Characteristics Considered in Sensemaking

Revisiting our camera metaphor, once participants selected a human or technical lens as the basis for their sensemaking process, they then selected a filter that made different aspects of unfairness more salient. Participants often used *characteristics* from both the prompts and their own conceptions to make sense of algorithmic

unfairness in the scenarios. Participants rarely considered characteristics in isolation, so highlighting a quote or idea for one characteristic does not mean that there were not other characteristics in them. We start with characteristics only used in one scenario and close with the two characteristics that cut across multiple scenarios: economic status/class, and race/ethnicity.

5.3.1 Gender. All 15 participants used gender to make sense of unfairness in the Search scenario, where it was explicit in the prompt. For example, Miranda considered gender when evaluating the team, but expressed skepticism of the motives for gender diversity:

“If the group was more diverse when it came to gender that’s good. I don’t think it will change their work ethics, especially if this is just a tech company and not PR type of work.”

Participants accounted for gender when making sense of the scenario, but typically examined it with other contextual factors and rarely on its own.

5.3.2 Country of Origin, Language, & Accent. When making sense of the Speaker scenario, many participants accounted for two explicitly mentioned characteristics, country of origin (11/14) and language (13/14), and a closely-related though not explicit characteristic, accent (14/14). When critiquing the voices from English-speaking countries used to train Blurty, Jake cited the context of its audience:

“There are people living in America that come from many different countries. So appealing to those people also is a must since the device is located here in America.”

All of the participants had immigrant backgrounds like Jake and nearly all of them spoke a language other than English at home. They may have experienced similar situations where voice recognition software did not recognize different languages and accents, which could have factored into their sensemaking.

5.3.3 Ability & Age. Some participants accounted for ability (2/14) and age (2/14), which were not mentioned in the Speaker scenario prompt. When evaluating the impact the algorithm and the training data would have on different people, Bob hypothesized:

“It’ll affect people with an accent or people who have lisps.”

Similarly, Sophia reflected on users on both ends of the age range:

“It might be used to the voice of an English speaker adult and that can exclude younger children, elders.”

Ability and age were neither explicitly mentioned nor had a strong link to the characteristics explicit in the scenario, suggesting participants might have leveraged their own knowledge and experiences to bring this characteristic into the discussion.

5.3.4 Academic Performance. Many participants (9/15) introduced academic performance as a characteristic in the School scenario. When making sense of the unequal distribution of supplies, Z conceptualized academic performance in terms of not only numerical test scores, but also student behavior:

“School A might have better statistics in terms of test scores, behaviors, and private funding.”

Academic performance was especially salient, which may be associated with the nature of the schools participants attended or this study’s formal classroom context.

5.3.5 Economic Status & Class. Unlike the previous characteristics, participants used economic status and class to make sense of unfairness in two scenarios, the Search (2/15) and School scenarios (15/15), only the latter of which referenced it explicitly.

In the Search scenario, when it was revealed that the Search Engine algorithm was designed by an all-men team, Sophia considered class alongside gender in assessing the team:

“When these rules were established computers were still fairly new and only the top ‘bracket’ of people had access to computers. Women were also probably not common in that workforce around that time.”

Participants also factored in economic status and class when critiquing the decisions in the School scenario, such as J:

“It’s great to have different backgrounds but we need people from different classes.”

Overall, economic status and class seemed to play an important role in sensemaking, with participants considering it without being prompted and often overlaying it on top of other characteristics.

5.3.6 Race/Ethnicity. Race/ethnicity was the only characteristic that participants used across all three sensemaking discussions (Search: 4/15; Speaker: 1/14; School: 13/15), although it was only explicitly mentioned in the prompt of the School scenario.

When evaluating a team comprised of women and non-binary people in the Search scenario, some participants concluded that gender diversity alone would not be a panacea, such as Batman:

“This alleviates some problems because (depending on their race and income) they probably haven’t experienced privilege when it comes to gender.”

Similarly in the Speaker scenario, when evaluating a team consisting of people from English-speaking countries, Z weighed racial representation in addition to country representation:

“It won’t represent different races and countries & they won’t be able to make Blurty more user friendly.”

Race/ethnicity was particularly relevant to participants’ sensemaking, with many bringing it into discussions without explicit prompting and often accounting for it in conjunction with other characteristics.

6 DISCUSSION

6.1 RQ1: When engaging in scaffolded sensemaking, what funds of knowledge might adolescents of color use to make sense of algorithmic fairness?

We found that all participants used factors not explicitly mentioned to make sense of algorithmic (un)fairness in the scenarios. These factors fell under three themes: (1) a human lens, which aligns with ecological systems theory that frames an individual with respect to their communities and larger society [6], (2) a technical lens, and (3) characteristics as filters. Participants used the two different lenses to modify the subject of their sensemaking, adjusting the scale of groups with the human lens and the resolution or level of detail with the technical lens. They then used different characteristics as filters to adapt what appeared most relevant to their sensemaking.

The human lens seemed to be used more often in sensemaking compared to the technical lens, which may have been due to participants’ lack of formalized computing instruction prior to the course. However, they drew upon their own funds of knowledge from lived experience to consider many different human-centric aspects of fairness. Understanding our participants’ different yet often devalued existing knowledge and skills is crucial to meaningfully including them in socially responsible computing education. These funds of knowledge could serve as anchor points for introducing more technical concepts in learning experiences. Our findings reinforce prior work that showed that a funds of knowledge approach allows learners with little formal training to bring in their skills and lived experiences when making sense of problems [63].

These results stand in contrast with the findings in Coenraad et al.’s study [11], where, without instruction, youth were only aware of visible technological biases, and only after they were instructed on common technological biases could they expand their conceptualizations to include less visible biases. This distinction is likely due to the different age ranges of our participants (8-13 years in [11] vs 15-17 years in ours), reinforcing prior work on fairness conceptions [5, 14]. This suggests a potential for various learning trajectories, or routes from existing knowledge to a more sophisticated and detailed understanding [10], for algorithmic fairness learning experiences. It may also be attributed to the different ethnic, linguistic, and socioeconomic backgrounds of our participants, or the structure of our activities, where we made space for scaffolded yet open-ended sensemaking about algorithmic fairness.

These results also stand in contrast with Brown et al.’s study [7], where most advanced undergraduate students only used the factors explicitly stated in the prompt in an AI ethics assignment. This difference may be attributed to the differing technical backgrounds of our participants and different task framings. The tasks in our study largely involved problem framing and brainstorming around algorithmic fairness, while the task in [7] was to program a solution to a specified problem. That programming framing may have encouraged learners to focus on low-level coding details before solidifying the high-level structure of their designs. Oleson et al. identified a similar pattern when investigating design learning difficulties in post-secondary HCI courses, noting that the behavior may arise from how prior computing courses train students to problem-solve ([49], “RUSH” difficulty). This presents an interesting conundrum for efforts to educate learners in the social impacts and ethics of computing, especially if there is a programming component — How might learning experiences around algorithmic fairness be designed to continuously leverage learners’ funds of knowledge, especially when there are existing computing educational practices that can prevent them from doing so?

Lastly, upholding the position of Hammer and Berland [28] that the role of qualitative coding is to understand themes within situated data, the camera metaphor used to structure our themes in this paper is currently only applicable to our data. Its role was to highlight the relationship between the factors and to organize the presentation of results, not necessarily to contribute a framework or generalizable model for sensemaking of algorithmic fairness. While we cannot contribute broad implications for algorithmic fairness sensemaking based on only these findings, this study provided rich

characterizations and initial insights into adolescents' sensemaking processes. Future studies building upon these results would be necessary to confirm, refute, or generalize this metaphor to other contexts.

6.2 RQ2: When engaging in scaffolded sensemaking, how might the identities and backgrounds of adolescents of color shape their sensemaking of algorithmic fairness?

Participants were able to bridge their existing knowledge and experiences of fairness into a computing context. When sensemaking, they introduced several characteristics not in the prompt, such as ability and age, into the Smart Speaker sensemaking discussion in particular. This may be because it was especially relatable to the participants, who drew from their lived experiences with voice recognition technologies to make sense of the scenario. Participants also introduced several kinds of stakeholders, such as teachers and parents, into the School scenario, likely because it was a context they were very familiar with. Further, participants reasoned about the unfairness in multiple scenarios using race/ethnicity, economic status, and class. Prior work has shown that adolescents are acutely aware of their own unfair circumstances [2] so they may have drawn from their own identities and experiences as members of marginalized ethnic (see Table 3) and socioeconomic backgrounds (i.e. first-generation, low-income) in computing and society.

Throughout our results, we found that participants seemed to leverage situated reasoning differently across the phases of the sensemaking process. Although participants used situated knowledge while understanding and evaluating the scenarios of algorithmic fairness, this practice did not entirely persist when brainstorming alternatives, as shown in the tendency towards ambiguity, (i.e. designing for “average users”, “everyone”, “public”, etc.) in the human lens (Section 5.1). This may be reflective of implicit assumptions about “unmarked” users, that they are assumed to be the most privileged in society (e.g. White, cis-gendered, male, abled, English-speaking, etc.) [13]. Unfortunately, even individuals from marginalized groups often make the same normative assumptions [15, 67]. This may also be attributed to the power dynamics in group design activities [29, 34]. Participants may not have felt comfortable using their situated knowledge while working in a group, leading them to “fall back” on conceptions of normative users. Participants may also simply not have the design vocabulary to go beyond the “average” user. We attempted to circumvent this tendency by designing the worksheets to scaffold contextualization of stakeholders within scenarios, but the more open-ended nature of the design activities did not provide similar scaffolds to participants. This indicates a need for scaffolding that supports learners in contextualizing their users in similar learning experiences.

6.3 Limitations & Future Work

Although our data provided valuable insights, some elements of our study design limit the internal and external validity of our interpretations. Given the sensitive nature of the discussion topics, the positionality of the facilitators (first, third, & fourth authors) may have influenced how comfortable or open participants would be when articulating their sensemaking process. The first author

shared many identities with the participants, but was also the lead instructor and more than a decade older than the participants; she may have been perceived as an authority figure. The third and fourth authors did not share as many identities, but as undergraduate researchers, may have been perceived as more approachable. Additionally, participants were nominated by their teachers to be part of this summer program and were considered to be the most “meritorious” among their peers. Participants also selected this class among five elective classes and may have had prior interest in computing that could have raised their awareness of issues around algorithmic bias. Participants were also from an area with a large technology industry, possibly raising their awareness. We also used a specific approach, sensemaking, to engage participants with limited prior knowledge within a specific context. Due to the situated nature of our intervention, we cannot make any claims about its efficacy compared with other approaches (nor was it the goal of this study to do so). We also only highlighted certain aspects of algorithmic fairness for discussion; there are certainly others, such as scalability, that merit future work. While our participants only reflected their own ideas and experiences, the diversity in their ethnic, linguistic, and socioeconomic backgrounds help shed light on how adolescents of similar marginalized backgrounds might make sense of algorithmic fairness. Future work could study different populations, approaches, contexts, or other aspects of algorithmic fairness.

Despite the limitations, our results provide important insights for algorithmic fairness learning experiences, especially for this age group (ages 15-17) which has comparatively thinner literature than for children. First, when given the space to leverage their situated knowledge, adolescents were not only aware of algorithmic fairness, but also adept at sensemaking about it. Educational efforts addressing algorithmic fairness may benefit from the incorporation of adolescents' funds of knowledge, allowing them to bridge their existing knowledge into a new domain. While our study demonstrates one way to integrate funds of knowledge through scaffolded sensemaking, future work could investigate other techniques to do so. Further, learners' funds of knowledge change with age, experience, and social circumstances [6, 45]. We conducted this study with participants who were likely to have experienced issues of fairness, but it is also crucial to replicate this study with participants whose identities align with dominant groups in computing (e.g. White, male, affluent, able-bodied), because unfortunately, such participants are more likely to wield power in the creation of computing systems, either directly as designers and engineers or indirectly as voters or policymakers. Understanding how learners' funds of knowledge change based on development and experience will help inform the design of learning trajectories, which are especially crucial in the growing movement for socially responsible computing education. Second, learning experiences on algorithmic fairness should consider highlighting and deconstructing the pervasive notion of the “average user” [13]. One example of an educational effort to deconstruct this notion is the CIDER assumption elicitation technique [50], which has been shown to be effective for post-secondary students. Future work could explore how this notion could be deconstructed with youth and adolescent learners.

In line with recent trends toward more critical computing education [32, 76], our findings provide a foundation for future work

that aims to help adolescents critically consider both technology's benefits and harms. This study also contributes insights for the efforts of computing educators, curriculum designers, and other stakeholders to further socially responsible computing education. Learning opportunities that meaningfully incorporate learners' situated knowledge can better empower adolescents to leverage their unique skills and competencies as tools for liberation.

ACKNOWLEDGMENTS

We would like to thank Stefania Druga for her feedback on the scenario questions and design activities. We would also 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 Microsoft, Adobe, and Google.

REFERENCES

- [1] Safinah Ali, Blakeley H. Payne, Randi Williams, Hae Won Park, and Cynthia Breazeal. 2019. Constructionism, ethics, and creativity: Developing primary and middle school artificial intelligence education. In *International workshop on education in artificial intelligence k-12 (eduai'19)*. 1–4.
- [2] William F. Arsenio, Susanna Preziosi, Erica Silberstein, and Benjamin Hamburger. 2012. Adolescents' perceptions of institutional fairness: Relations with moral reasoning, emotions, and behavior. *New directions for youth development* 2012, 136 (2012), 95–110.
- [3] Alicia Barreiro, William F. Arsenio, and Cecilia Wainryb. 2019. Adolescents' conceptions of wealth and societal fairness amid extreme inequality: An Argentine sample. *Developmental Psychology* 55, 3 (2019), 498.
- [4] Ruha Benjamin. 2019. Race after technology: Abolitionist tools for the new jim code. *Social forces* (2019).
- [5] Thomas J. Berndt. 1982. Fairness and friendship. In *Peer relationships and social skills in childhood*. Springer, 253–278.
- [6] Urie Bronfenbrenner. 1992. *Ecological systems theory*. Jessica Kingsley Publishers.
- [7] Noelle Brown, Koriann South, and Eliane S. Wiese. 2022. The Shortest Path to Ethics in AI: An Integrated Assignment Where Human Concerns Guide Technical Decisions. In *Proceedings of the 2022 ACM Conference on International Computing Education Research V. 1*. 344–355.
- [8] Nicole Campione-Barr and Judith G. Smetana. 2010. "Who said you could wear my sweater?" Adolescent siblings' conflicts and associations with relationship quality. *Child Development* 81, 2 (2010), 464–471.
- [9] Eunhyung Chung and Elliot Turiel. 2022. Adolescents' judgments about resource inequality involving group disparities. *Journal of Experimental Child Psychology* 218 (2022), 105373.
- [10] Douglas H. Clements and Julie Sarama. 2012. Learning trajectories in mathematics education. In *Hypothetical Learning Trajectories*. Routledge, 81–90.
- [11] Merijke Coenraad. 2022. "That's what techquity is": youth perceptions of technological and algorithmic bias. *Information and Learning Sciences* ahead-of-print (2022).
- [12] 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.
- [13] Sasha Costanza-Chock. 2020. *Design justice: Community-led practices to build the worlds we need*. The MIT Press.
- [14] Eveline A Crone. 2013. Considerations of fairness in the adolescent brain. *Child Development Perspectives* 7, 2 (2013), 97–103.
- [15] Kieran Cutting and Erkki Hedenborg. 2019. Can personas speak? Biopolitics in design processes. In *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*. 153–157.
- [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] Laura Elenbaas and Melanie Killen. 2017. Children's perceptions of social resource inequality. *Journal of Applied Developmental Psychology* 48 (2017), 49–58.
- [24] Sheena Erete, Karla Thomas, Denise Nacu, Jessa Dickinson, Naomi Thompson, and Nichole Pinkard. 2021. Applying a transformative justice approach to encourage the participation of Black and Latina Girls in computing. *ACM Transactions on Computing Education (TOCE)* 21, 4 (2021), 1–24.
- [25] Jayne Everson, F. Megumi Kivuvu, and Amy J. Ko. 2022. "A Key to Reducing Inequities in Like, AI, is by Reducing Inequities Everywhere First" Emerging Critical Consciousness in a Co-Constructed Secondary CS Classroom. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 1*. 209–215.
- [26] 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.
- [27] Norma González, Luis C. Moll, and Cathy Amanti. 2006. *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Routledge.
- [28] 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.
- [29] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing community-based collaborative design: Towards more equitable participatory design engagements. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–25.
- [30] 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*.
- [31] 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.
- [32] Amy J. Ko, Alannah Oleson, Mara Kirdani-Ryan, Yim Register, Benjamin Xie, Mina Tari, Matthew Davidson, Stefania Druga, and Dastyini Loksa. 2020. It is time for more critical CS education. *Commun. ACM* 63, 11 (2020), 31–33.
- [33] Janet L Kolodner. 2004. The learning sciences: Past, present, future. *Educational technology* 44, 3 (2004), 34–40.
- [34] Max Krüger, Ana Bustamante Duarte, Anne Weibert, Konstantin Aal, Reem Talhouk, and Oussama Metatla. 2019. What is participation? Emerging challenges for participatory design in globalized conditions. *Interactions* 26, 3 (2019), 50–54.
- [35] Jenna Laib. 2022. <https://slowrevealgraphs.com/>
- [36] Halcyon M. Lawrence. 2013. *Speech intelligibility and accents in speech-mediated interfaces: Results and recommendations*. Illinois Institute of Technology.
- [37] Min Kyung Lee and Su Baykal. 2017. Algorithmic mediation in group decisions: Fairness perceptions of algorithmically mediated vs. discussion-based social division. In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing*. 1035–1048.
- [38] Victor R. Lee, Victoria Delaney, and Parth Sarin. 2022. Eliciting High School Students' Conceptions and Intuitions about Algorithmic Bias. In *Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 2*. 35–36.
- [39] 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.
- [40] Meghan K. Loeser, Shawn D. Whiteman, and Susan M. McHale. 2016. Siblings' perceptions of differential treatment, fairness, and jealousy and adolescent adjustment: A moderated indirect effects model. *Journal of child and family studies* 25, 8 (2016), 2405–2414.
- [41] 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.
- [42] 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.
- [43] Lauren E Margulieux, Brian Dorn, and Kristin A Searle. 2019. *Learning sciences for computing education*. Cambridge: Cambridge University Press.
- [44] Tuuli Mattelmäki et al. 2006. *Design probes*. Aalto University.
- [45] Peggy McIntosh. 1990. White privilege: Unpacking the invisible knapsack.

- [46] Vonnice C. McLoyd. 2019. How children and adolescents think about, make sense of, and respond to economic inequality: Why does it matter? *Developmental psychology* 55, 3 (2019), 592.
- [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] Safiya Umoja Noble. 2018. Algorithms of oppression. In *Algorithms of Oppression*. New York University Press.
- [49] Alannah Oleson, Meron Solomon, and Amy J. Ko. 2020. Computing students' learning difficulties in HCI education. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [50] Alannah Oleson, Meron Solomon, Christopher Perdriau, and Amy J. Ko. 2022. Teaching Inclusive Design Skills with the CIDER Assumption Elicitation Technique. *ACM Transactions on Computer-Human Interaction* (2022).
- [51] Constanta Olteanu. 2020. Programming, mathematical reasoning and sense-making. *International Journal of Mathematical Education in Science and Technology* (2020), 1–19.
- [52] Cathy O'neil. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway books.
- [53] 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.
- [54] 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.
- [55] Adam Rutland and Melanie Killen. 2017. Fair resource allocation among children and adolescents: The role of group and developmental processes. *Child Development Perspectives* 11, 1 (2017), 56–62.
- [56] Jean J. Ryoo, Alicia Morris, and Jane Margolis. 2021. "What Happens to the Rasputin man in a Cash-free Society?": Teaching and Learning Socially Responsible Computing. *ACM Transactions on Computing Education (TOCE)* 21, 4 (2021), 1–28.
- [57] Sara Safransky. 2020. Geographies of algorithmic violence: Redlining the smart city. *International Journal of Urban and Regional Research* 44, 2 (2020), 200–218.
- [58] Nithya Sambasivan, Erin Arnesen, Ben Hutchinson, Tulsee Doshi, and Vinodkumar Prabhakaran. 2021. Re-imagining algorithmic fairness in india and beyond. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 315–328.
- [59] 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.
- [60] 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.
- [61] Steve Stemler. 2000. An overview of content analysis. *Practical assessment, research, and evaluation* 7, 1 (2000), 17.
- [62] 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>
- [63] Vanessa Svihla, Yan Chen, and Sung "Pil" Kang. 2022. A funds of knowledge approach to developing engineering students' design problem framing skills. *Journal of Engineering Education* 111, 2 (2022), 308–337.
- [64] 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.
- [65] Matti Tedre, Tapani Toivonen, Juho Kahila, Henriikka Vartiainen, Teemu Valtonen, Ilkka Jormanainen, and Arnold Pears. 2021. Teaching machine learning in K–12 classroom: Pedagogical and technological trajectories for artificial intelligence education. *IEEE Access* 9 (2021), 110558–110572.
- [66] David Touretzky, Christina Gardner-McCune, Fred Martin, and Deborah Seehorn. 2019. Envisioning AI for K-12: What should every child know about AI?. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 9795–9799.
- [67] Phil Turner and Susan Turner. 2011. Is stereotyping inevitable when designing with personas? *Design studies* 32, 1 (2011), 30–44.
- [68] Sepehr Vakili. 2018. Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education. *Harvard Educational Review* 88, 1 (2018), 26–52.
- [69] Andrea Valente and Emanuela Marchetti. 2011. Programming Turing Machines as a game for technology sense-making. In *2011 IEEE 11th International Conference on Advanced Learning Technologies*. IEEE, 428–430.
- [70] Lisa Wagner. 2019. Good character is what we look for in a friend: Character strengths are positively related to peer acceptance and friendship quality in early adolescents. *The Journal of Early Adolescence* 39, 6 (2019), 864–903.
- [71] 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.
- [72] Amy Wilson-Lopez, Joel Alejandro Mejia, Indhira María Hasbún, and G Sue Kasun. 2016. Latina/o adolescents' funds of knowledge related to engineering. *Journal of Engineering Education* 105, 2 (2016), 278–311.
- [73] Amy Wilson-Lopez, Christina Sias, Allen Smithee, and Indhira María Hasbún. 2018. Forms of science capital mobilized in adolescents' engineering projects. *Journal of Research in Science Teaching* 55, 2 (2018), 246–270.
- [74] Allison Woodruff, Sarah E. Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A qualitative exploration of perceptions of algorithmic fairness. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–14.
- [75] Sijia Xiao, Coye Cheshire, and Niloufar Salehi. 2022. Sensemaking, Support, Safety, Retribution, Transformation: A Restorative Justice Approach to Understanding Adolescents' Needs for Addressing Online Harm. In *CHI Conference on Human Factors in Computing Systems*. 1–15.
- [76] 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.