Expressiveness, Cost, and Collectivism: How the Design of Preference Languages Shapes Participation in Algorithmic Decision-Making

Samantha Robertson
University of California, Berkeley
Berkeley, California, U.S.A
samantha_robertson@berkeley.edu

Tonya Nguyen
University of California, Berkeley
Berkeley, California, U.S.A
tonyanguyen@berkeley.edu

Cathy Hu
University of California, Berkeley
Berkeley, California, U.S.A
cathy_hu@berkeley.edu

Catherine Albiston
University of California, Berkeley
Berkeley, California, U.S.A
calbiston@berkeley.edu

Afshin Nikzad
University of Southern California
Los Angeles, California, U.S.A
afshinni@usc.edu

Niloufar Salehi
University of California, Berkeley
Berkeley, California, U.S.A
nsalehi@berkeley.edu

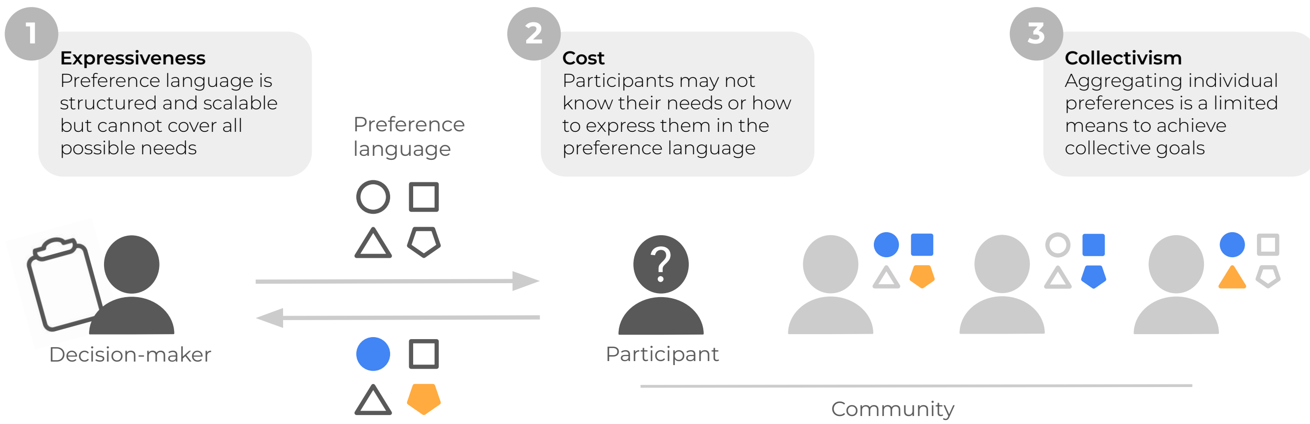

Figure 1: Preference-based systems consist of a preference language, in which participants express their needs and goals to a decision-maker, and an aggregation algorithm, which aggregates individuals' preferences into a collective decision. We identify three ways that preference languages shape opportunities for meaningful participation in algorithmic decision-making: 1) expressiveness, the range of needs that participants can communicate; 2) cost, the effort it takes for participants to express their needs and goals in the preference language; and 3) collectivism, the extent to which aggregating individuals' preferences can achieve collective goals.

ABSTRACT

Emerging methods for participatory algorithm design have proposed collecting and aggregating individual stakeholders' preferences to create algorithmic systems that account for those stakeholders' values. Drawing on two years of research across two public school districts in the United States, we study how families and school districts use students' preferences for schools to meet their goals in the context of algorithmic student assignment systems. We find that the design of the preference language, i.e. the structure in which participants must express their needs and goals to the decision-maker, shapes the opportunities for meaningful participation. We define three properties of preference languages – expressiveness, cost, and collectivism – and discuss how these factors shape who is able to participate, and the extent to which they are able to effectively communicate their needs to the decision-maker. Reflecting on these findings, we offer implications and paths forward for researchers and practitioners who are considering applying a preference-based model for participation in algorithmic decision making.

CCS CONCEPTS

• **Human-centered computing → HCI theory, concepts and models**; **Empirical studies in HCI**.

KEYWORDS

Participatory design, preference elicitation, preference language, algorithmic systems

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3580996

ACM Reference Format:
Samantha Robertson, Tonya Nguyen, Cathy Hu, Catherine Albiston, Afshin Nikzad, and Niloufar Salehi. 2023. Expressiveness, Cost, and Collectivism: How the Design of Preference Languages Shapes Participation in Algorithmic Decision-Making. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3544548.3580996

1 INTRODUCTION

Algorithmic systems increasingly impact peoples' lives by mediating their access to resources and by making high-stakes decisions in domains like education, employment, healthcare, and child welfare [18, 43, 69, 90, 96]. Documented issues of discrimination [5, 9, 28, 89], biased system performance [15, 19, 86, 117], and dissatisfaction among key stakeholders [95, 114] have increased pressure to improve these systems by accounting for the values and needs of those who use or are affected by them. Building on social choice theory [70], emerging methods for participatory algorithm design have proposed collecting and aggregating individual stakeholders' preferences to create algorithmic systems that represent the values and goals of those stakeholders [22, 37, 39, 40, 64, 66, 72, 78, 79, 87, 125, 126]. In this paper, we study how the *design of the preference language*, i.e. the language in which participants are asked to express their preferences, shapes the opportunities for meaningful participation.

One area in which participant preferences are incorporated toward algorithmic decisions is assigning students to public schools [101]. Algorithmic student assignment systems, which are widely used across the U.S., require students to submit a ranked list of schools that they would like to attend. The school district then uses a matching algorithm to assign students to schools in a way that optimally satisfies the students' preferences [3]. School districts reason that compared to neighborhood-based assignments, these preference-based assignment systems provide more flexibility to families, create more diverse classrooms, and promote educational equity [109]. However, many school districts have found that the algorithms do not meet these expectations in practice [98]. For example, San Francisco Unified School District found that families had difficulty navigating their system, and that segregation increased since the system was introduced, with students from historically marginalized backgrounds concentrated in under-served schools [53]. These challenges became so severe that the school district voted to stop using and redesign the system in 2018.

We studied student assignment processes in two neighboring school districts in the U.S. to understand how families and school districts use preferences to meet their goals. This case study offers insight into the challenges and limitations of the preference-based approach to incorporating stakeholder participation in algorithmic decision-making. Our data include 27 semi-structured interviews with parents and with community members who helped parents through the enrollment process (e.g. district, school, and non-profit staff), in addition to several informal conversations with relevant stakeholders over the course of two years.

We find that the design of the preference language defines the opportunities for participation (Fig. 1). In student assignment systems, participants submit a ranked list over schools, possibly of limited length. Other common preference languages can include pairwise comparisons between real or hypothetical options, and providing weights over features of the decision outcome [79]. We define three properties of preference languages – **expressiveness**, **cost**, and **collectivism** – and discuss how the design of the preference language with respect to each of these factors can shape and limit meaningful participation. Preference languages are often designed to be structured and scalable so that large numbers of participants can be involved at low cost. This means that a preference language cannot cover all possible needs, and some participants may be able to express their needs more than others (**expressiveness**). Second, it takes time and effort for participants to translate their complex and often vague needs into the structured preference language (**cost**). This can be especially costly for participants who do not already know what they need, or how to express their needs in the preference language. These costs create disparities between people based on their access to time and resources. In the school assignment case, parents with fewer resources face greater barriers to convey all of their needs and goals through their ranked list of schools. This is partly due to the cost of gathering information about the schools [99]. Finally, we find that aggregating individual preferences is often a limited means to achieve complex collective goals (**collectivism**). In the case of student assignment, school districts' goals such as integration and educational equity have been very difficult to achieve through the aggregation of individual preferences, which can only express families' self-interested priorities.

Reflecting on these findings, we offer implications and paths forward for researchers and practitioners who are considering applying the preference-based model for participation in algorithmic decision making. First, we discuss opportunities to improve expressiveness and reduce costs associated with a preference language. These include improving the resources available, providing support (e.g. information support) for stakeholders to use the preference language, and simplifying or re-designing the preference language to make it a more natural representation of how people already think about their needs. Second, we discuss paths forward for engaging community members in deliberation to co-define collective goals, drawing insight from procedural justice theory. Procedural justice theory considers how the process through which a decision is made impacts satisfaction, perceptions of fairness and legitimacy, and compliance with that decision [119, 122]. We also propose technical mechanisms for better aligning preference aggregation mechanisms with community-defined collective goals, even while only eliciting self-interested preferences. Finally, we discuss the limitations of preference-based systems, and the potential for their emphasis on individualism and free market values to cause harm. Ultimately, we argue that preference elicitation and aggregation mechanisms will never reach the ideals of participatory methodology [83] without careful attention to how the preference language and sociotechnical infrastructure supporting it enable equitable and meaningful participation.

2 RELATED WORK

In this paper, we use student assignment algorithms as a case study to examine the challenges for using individual preference aggregation as a form of participation in algorithmic decision-making. In this section we provide an overview of participatory algorithm design and the preference-based approach to algorithm design. Then, we provide a brief overview of the literature on matching algorithms for student assignment.[1]

2.1 Stakeholder Participation in Algorithmic Decision-Making

Growing awareness that algorithmic decision-making may harm marginalized people has driven researchers to seek methods for directly involving stakeholders in the design of algorithmic systems. Incorporating direct stakeholder participation is intended to bring diverse knowledge and perspectives into design, and to build technologies that have a more positive influence on people's lives [14, 24, 30]. Recent work has drawn on methods and practices from frameworks like participatory design [82], human-centered design [88], and value-sensitive design [35]. Common methods include design workshops [6, 18, 93] and interviews [34, 77, 104, 112, 127], while further work is needed to define what participation can and should look like in the context of algorithmic systems. Delgado et al. [30] catalogued 9 different approaches to increasing stakeholder participation in algorithmic systems, noting that they vary substantially in terms of the degree of power different stakeholders are afforded, and when in the design and deployment of a system they are afforded that power. Birhane et al. [14] developed standards for evaluating whether particular approaches are aligned with the goals and values of participatory AI, such as the degree of reciprocity and participant empowerment.

Prior work in participatory and human-centered design methods foreshadows some of the challenges that arise when attempting to engage stakeholders to build more beneficial and just algorithmic technologies. For instance, it can be very difficult to subvert power dynamics, both between designers and participants and among participants, to promote genuine and equal participation [30, 54, 61]. The format and outcomes of participation are also often constrained in a way that presumes that there must be a technical solution and only allows participants to tinker around the edges of it, rather than offering meaningful decision-making power about what the system should do and whether it should exist at all [30, 61]. In the worst case, seemingly participatory practices can offer a guise of legitimacy to harmful technologies, making it more difficult to challenge those systems in the long run [13, 14, 33, 111].

Even if a participatory process effectively and meaningfully engages marginalized stakeholders, challenges remain. For example, researchers or designers often have to aggregate input from many participants and translate this input into a technical design specification, leaving room for unequal representation, misinterpretation, or even disregard of stakeholders' views. Some researchers have tried to bridge this gap by allowing participants to play a more direct role in building algorithmic systems. For example, the ORES system allows Wikipedia editors to directly specify and request machine learning models that meet their editing needs [52]. In this paper we focus on another line of work, which draws on social choice theory [70] to develop what we refer to as "preference-based systems." This approach incorporates direct input from stakeholders by eliciting individuals' preferences over some available alternatives, and then aggregating those preferences using an algorithm to make a decision. Preference-based systems are appealing because they easily scale to allow a large number of people to directly contribute their views, and there exist a range of well studied aggregation procedures to translate this input into a decision. We next discuss how this class of algorithms has been used to increase participation in algorithmic decision-making, and what challenges remain.

2.2 Preference-based Algorithmic Systems

There are two main stages of any preference-based system, *preference elicitation* and *preference aggregation*. In the preference elicitation stage, a decision-maker asks participants to quantify their relative value for some available options in a structured *preference language*. A preference language consists of a set of features and a way for participants to express their priorities across those features, typically either by ranking them, making pairwise comparisons, or providing weights over features [70]. Features may be a set of real alternatives (e.g., a list of schools in a school district), or attributes of those alternatives (e.g., a school's location, start time, or language programs).

Once the decision-maker has collected preferences from stakeholders, they then need to aggregate those preferences to make a decision. Researchers in social choice theory[2] have developed various aggregation procedures that are designed to make decisions that satisfy some normatively justified axioms or formal definitions of optimal assignment [16, 70]. For example, matching algorithms are guaranteed to produce assignments that are stable — no two students can swap assignments in a way such that both the students and the schools are better off — or Pareto efficient — there is no way to improve one student's assignment without making another worse off [3]. In other contexts, researchers have advocated for voting procedures on the basis of properties like robustness to noisy preference information [12, 42, 66].

Systems based on preference elicitation have long been used in market and mechanism design [101] and participatory democracy [23]. More recently, researchers have explored how these methods could incorporate stakeholders' perspectives and values more directly into the design and function of algorithmic systems. Some of this work fall under the umbrella of computational social choice [17]. For example, Noothigattu et al. [87] proposed collecting people's preferences in hypothetical scenarios, using that data to build personalized preference models, and then aggregating predicted preferences to make decisions in new situations. The authors suggested that this approach could be used to build autonomous vehicles

[1]There is a large literature on these algorithms, a complete review of which is out of scope for this paper. We refer interested readers to Abdulkadiroğlu and Sönmez [3], which introduced the mechanism design approach to the student assignment domain, and Roth [101] for an overview of applications of mechanism design.

[2]A substantial body of work in social choice theory across economics and philosophy seeks to understand how individual preferences should be aggregated to inform decisions on behalf of a group. Kenneth J. Arrow [70] provides an overview. In this work we apply human-centered methods to understand experiences with a matching algorithm in practice and compare these experiences to the ideals of participatory design. We leave a detailed comparison of these findings to theoretical results and philosophical discussions of preferences to future work.

that align emergency decision-making with people's ethical preferences. Researchers have subsequently explored how this approach can improve efficiency and fairness in distributing donations for a non-profit organization [78], align shift scheduling with workers' and managers' preferences [79], and account for citizens' ethical preferences in automated flood management decisions [39]. Freedman et al. [40] suggested a related approach, modifying a matching algorithm for organ transplants to weight matches according to peoples' preferences about which patients should be prioritized. In experiments, their algorithm improved outcomes for typically under-demanded patients who are currently less likely to receive an organ match. Other areas where researchers have studied how individuals' preferences can be elicited and aggregated to guide algorithmic decision-making include patient triage in hospitals [64], selecting appropriate performance trade-offs for ML models [22, 125, 126], allocating public budgets and resources [23, 72], defining diversity quotas for elections [37], selecting student volunteers for conferences [94], dividing goods and labor among groups of people [75, 76], aligning recommender systems with users' values [115], and balancing conflicting perspectives in content moderation [46].

This research highlights the potential for these kinds of systems to align algorithmic decision-making with people's needs and values at scale. The process of quantifying one's preferences can help people better understand their own needs [78, 79], and this approach makes explicit to decision-makers the pluralism of priorities and values in a group of stakeholders [46, 72]. However, it is difficult to design effective and easy-to-use preference languages and aggregation procedures. For example, Lee and Baykal [75] conducted an experiment where a group of people used a matching algorithm to divide up tasks according to their preferences. Participants found it difficult to quantify their preferences using the given preference language. As a result, they relied on error-prone cognitive heuristics to simplify the task. Individual preferences also cannot account for cooperative and altruistic behavior. In the same study and in a related experiment involving goods division [76], participants wanted to deliberate and cooperate to find more acceptable compromises than the algorithm's allocation.

Matching algorithms are a kind of preference-based system that have been used in real world markets, such as organ transplant matching, assigning medical students to residency programs, and assigning students to public schools, for decades [100]. In this time, researchers have been able to observe how these systems function in the real world. This has provided insight into how to build effective systems, as well as what issues remain unsolved. Our goal in this work is to draw lessons from the student assignment context that illuminate key considerations for building effective, equitable preference-based systems. We also draw on this newer body of work developing other kinds of preference-based systems to inform the design of student assignment algorithms. In particular, we study how the *preference languages* in these different systems shape who can participate, what they can communicate about their needs, values, and priorities, and how collective goals can be achieved using the submitted preferences. We conclude this section with an introduction to student assignment algorithms.

2.3 Student Assignment Algorithms

In most public school districts in the U.S., students are assigned to schools based on where they live. Therefore, racial segregation and economic inequalities result in segregated and unequal schools. Increasingly, school districts have been introducing school choice systems that allow students to apply to schools across the district. Students submit a ranked list of schools they would like to attend and the district uses an algorithm to match students to schools based on those preferences. Many school districts implemented these systems for their potential to advance equitable access to high quality education, create more diverse classrooms, and provide flexibility to families [68].

Economists in the field of market design have developed a number of matching algorithms that model the student assignment problem as a two-sided market and seek to find an optimal matching based on each side's preferences [41, 110]. The incoming students and available seats are the two sides of the market. Students report their preferences by ranking the schools, and schools can define priority categories for students, such as priority for younger siblings of continuing students or priority for students living in the school's surrounding neighborhood. The matching algorithms used in the student assignment context[3] are *student-optimal* in the sense that they are optimized to satisfy student preferences as efficiently as possible,[4] subject to each school's capacity constraints [3]. Often, school priorities only determine the order in which students are offered over-demanded seats.

These systems are a useful case study of preference-based algorithmic systems because they have been used in school districts across the country for several decades [101]. This has given researchers the opportunity to observe challenges in practice, such as confusion for families and decreasing classroom diversity. For instance, through working with school districts to implement these algorithms, Pathak [95] noticed that early theoretical literature overlooked or oversimplifed challenges of practical importance like strategic incentives, transparency, and coordinating offers to improve the efficiency of waitlists. Economists have since applied empirical and experimental methods to understand strategic behavior [31, 32, 48, 50, 55, 67, 92, 97], information needs [8, 20, 26, 27, 29, 49, 57, 59, 80], and diversity constraints [36, 45, 51, 56, 73, 85, 91].

However, researchers have raised concerns that tinkering with the technical implementation of these algorithms is insufficient to improve the enrollment systems overall. Kasman and Valant [68] discussed the strong political forces shaping how these algorithms are used, understood, and accepted in school districts. They argued that these algorithms are easily misunderstood by stakeholders, and that adoption will depend more on how people interact with these systems than their underlying theoretical properties. Hitzig [60] pointed out that matching algorithms' emphasis on efficiency makes strong implicit assumptions about the optimal distribution of assignments, namely that the ideal outcome is the one where every student is assigned to their first choice school. This is often

[3]Deferred Acceptance [41] and Top-Trading Cycles [110] are commonly used for student assignment [3].

[4]For more details about properties of matching mechanisms and matchings, such as strategy-proofness, and trade-offs between stability and efficiency, see Abdulkadiroğlu and Sönmez [3]. For the purposes of this paper, it is most important to keep in mind that the primary goal of these algorithms is to satisfy student preferences.

framed in economics as objectively optimal rather than only one of many ways to distribute resources, and one which may not align with school districts' collective goals. In our prior work exploring why San Francisco Unified School District decided to re-design their assignment algorithm, we argued that the modeling assumptions underpinning the system's design overlooked the complex barriers to ideal participation that many families face, particularly because of socioeconomic inequalities [98]. In follow up work in Oakland Unified School District, we highlighted the importance of information sharing among trusted community networks for supporting marginalized families' participation in school choice [99].

We build on this prior literature by bringing together the interview data we have collected in San Francisco and Oakland, California to discuss the role of the preference language in shaping opportunities for participation. We compare and contrast this case study to other kinds of preference-based systems to identify broader implications for the design of preference languages and propose key paths forward for the emerging field of participatory algorithm design. To end this section, we briefly provide context for each of the two school districts studied in this work. In the next section we describe our research method.

San Francisco Unified School District. San Francisco Unified School District (SFUSD) introduced a student assignment system based on a matching algorithm[5] in 2011 in the hopes of promoting equitable access to educational opportunity and diverse classrooms [109]. However, by 2018 the Board had voted to redesign the system in response to widespread dissatisfaction among families and clear evidence that the system was not serving the district's goals [53]. Under this system, families could apply to any school in the district, and could list as many schools as they wanted to in their application. The algorithm gave priority to students living near schools that performed poorly on statewide standardized tests [109]. Although this system, in theory, offers equitable access to all of the district's schools, the algorithm had been unable to promote diverse classrooms and equitable access to education in practice, largely due to racial and socioeconomic disparities in participation rates and segregation in the preferences of those families who do participate [107]. The newly proposed assignment system in San Francisco will allow each student to choose from within geographical zones, rather than every school in the district [108]. The student's zone will be determined by where they live, and each zone will be designed to reflect the diversity of the city in terms of factors including socioeconomic status, race and ethnicity, and English proficiency [103]. Our work in San Francisco focused on understanding why this system has not met expectations [98].

Oakland Unified School District. Oakland Unified School District (OUSD) also has an open enrollment system, where families can apply to any public school in the district and the district uses a matching algorithm to assign students to schools. Unlike San Francisco, families can only rank up to 6 schools on their application, and the algorithm gives top priority to students living in each school's surrounding attendance zone. Almost 90% of OUSD students are students of color, and over 70% are eligible for free and reduced price lunch,[6] but schools have remained segregated and unequal, with several of the highest resource schools serving almost 50% white students. In response to this problem, OUSD is piloting a new priority category to make more space for students who live in poorer parts of the city to attend these high-resource schools. The district is also concerned about families' access to information about schools and the application process, particularly for parents who are new to the district and/or have limited English proficiency. Our work in Oakland focused on learning about these participation challenges [99].

3 METHODS

This work draws on qualitative data collected as part of separate, but closely related, research projects that we conducted in collaboration with two neighboring school districts, San Francisco Unified School District (SFUSD) and Oakland Unified School District (OUSD) [98, 99].

3.1 Data Collection

We conducted 27 semi-structured interviews with parents (13 SFUSD; 10 OUSD) and staff in schools (2 OUSD), district offices (1 OUSD), and community-based organizations (1 OUSD) to understand how families use their preferences over schools to communicate their needs and constraints to school districts, and how the districts aggregate those preferences to satisfy parents' needs and meet district goals such as school diversity. In San Francisco, we recruited parents through four parenting email and Facebook groups by contacting group administrators who shared a brief recruitment survey on our behalf. In Oakland, we recruited parents on Twitter and through community-based organizations. We grew the initial sample of participants using snowball sampling. The interviews in San Francisco mainly focused on the kindergarten application process, while the parents from Oakland had enrolled children in elementary, middle, and high schools.[7]

During the interview, we asked participants to describe their application and enrollment experiences. In addition to interviews with parents, we interviewed 4 people who work in schools, school district offices, and non-profit organizations in Oakland. These were people who had actively worked to support families through the enrollment process in their role. During these interviews, we asked participants to describe their work related to student enrollment. Three of these participants were also parents, in which case we were asked a mix of questions about their experiences enrolling their own students in schools and their experiences supporting other families through the enrollment process. The interviews with parents provided insight into individual families' experiences with and perceptions of the system. The interviews with staff members who work with families to support enrollment gave a broader view of the types of challenges that families face and how schools, districts, and communities have responded to address these challenges.

Interviews were between 30 and 50 minutes and participants received a $30 gift card. Interviews were conducted over the phone

[5]SFUSD uses a variant of Top-Trading Cycles.

[6]Source: OUSD Fast Facts 2021-2022

[7]The enrollment process is slightly simpler for families at later stages of education, as there are fewer available schools and some districts have feeder patterns between stages; otherwise, there were no major differences in reported experiences by educational stage.

in English and Spanish (with an interpreter) between February 2020 and March 2021. At the end of the interview, we asked participants a set of demographic questions. Participant demographics are shown in Table 1.

3.1.1 Limitations. Due to convenience and snowball sampling methods, the interviews are not representative of the full breadth of experiences that families have when enrolling in public schools in U.S. school districts with school choice. The interviews in San Francisco, in particular, lacked representation of marginalized families. This was a limitation that we prioritized in follow up work in Oakland, where we specifically recruited families of color, low-income families, and families with limited English proficiency, who have been and continue to be underserved by the public education system in the United States.

3.2 Data Analysis

We conducted inductive, qualitative analysis on the interview transcripts [81]. First, we conducted open coding on a line-by-line basis [25] to understand how parents use the system and what challenges they face in meeting their goals. We then conducted axial coding to identify relationships between codes and higher level themes. We first analyzed the data from San Francisco, before conducting interviews in Oakland. At this stage, the first author conducted open coding and identified a common set of parents' priorities (e.g., "test scores," "resources," and "travel logistics") and challenges (e.g., "stress," "information needs") when participating in school choice. Next, we analyzed the data from Oakland, building on the codes we used when analyzing the San Francisco data. The first two authors worked together to code two transcripts, then discussed findings and resolved misaligned interpretations through discussion. Each author then analyzed half of the remaining transcripts. The initial codebook from San Francisco was used as a reference, but codes were adjusted, added, and removed as necessary to best fit the data. We then conducted axial coding, and grouped 25 of our 39 codes into two higher level groups, "finding information" and "considering priorities." Other codes included "building relationships," "worrying about availability/scarcity," and "voicing concern." Finally, we conducted a final round of coding on the full dataset to compare and synthesize findings across the two districts.

3.3 Researcher Positionality

We recognize that our personal and professional backgrounds and contexts shape our approach to this research, our interactions with participants, and our interpretations of the findings [11]. The first and second authors conducted recruitment and interviews. Neither author has first-hand experience of a public school choice system, nor is either author a member of the communities in San Francisco and Oakland with whom we conducted this research. Throughout the course of this work, we shifted from direct engagement with parents to working with trusted community leaders, finding that this was a better way to show respect towards the relationships and work already happening in those communities [99]. Our interpretation of our findings is certainly shaped by both our positions as outsiders, as well as our disciplinary backgrounds, which between the authors include human-computer interaction, sociology, law, and economics.

4 DESIGNING PREFERENCE LANGUAGES

In this section, we discuss how the design of a preference-based system shapes (and can limit) participation in algorithmic decision-making. We argue that a core consideration should be the design of the *preference language*, which consists of a set of features and a way for participants to express their priorities across those features. First, we discuss the implications of the preference language for individual participants: **expressiveness** and **cost**. Then we explore the implications of the preference language for achieving collective goals (**collectivism**). For each of these three factors, we illustrate relevant challenges using examples from the student assignment context, then discuss how those challenges appear and are addressed in other kinds of preference-based systems.

4.1 Expressiveness: No preference language can cover all possible needs

Preference-based systems offer participants a fixed set of alternatives over which they can express their preferences. A core part of designing a preference language is selecting what alternatives or factors are available to participants. Because preference languages must be structured and scalable to large groups of people, no preference language will perfectly capture every possible dimension of everyone's needs and values. Certain preference languages will be more expressive for some participants than others. It is therefore important to attend to not only how expressive a given preference language is, but *for whom* it is more or less expressive.

4.1.1 Student Assignment. School choice systems increase access to educational opportunities by offering each student a wider range of schools to choose from, rather than being directly assigned to their neighborhood school. However, even if students can apply to more schools this does not mean they have access to those opportunities in practice. Most families want a school that is close to home and that will provide their child with a high-quality education [21, 74]. However, in the United States, a long history of racist housing and education policies have concentrated schools with more resources and higher academic performance in high-income, predominantly white neighborhoods [74, 102]. As a result, families with more resources are able to more heavily prioritize academic factors, whereas other families face difficult trade-offs between economic and social factors, such as transportation logistics or their child's safety and sense of belonging [56, 73].

Low-income families are more likely to face this trade-off between school resources and proximity to home, since higher resource schools are mostly located in wealthier areas of the city [2, 20, 58, 74]. Proximity was especially important to parents of younger children, parents who don't have access to a car, and parents who were concerned about the safety of their neighborhood. O12 *"used to walk with him* [to school] *and then I started to learn to drive. Because when I walk or when I wait for the bus, the bus would take so long."* Discrimination and segregation in classrooms also mean that a school that offers high quality opportunities to white students may not provide a safe and supportive environment for students of color [4, 21, 105]. For instance, O14 saw that her son was *"becoming a different person, I couldn't even recognize him,"* due to his treatment by his first grade teacher. Resolving this situation

Table 1: Parent demographics. Interviews with participants marked with † were conducted in Spanish and English with an interpreter. Incomes marked with * are estimates based on median income in home zip code, all other fields are self-described.

	Role	Race/Ethnicity	Income	Education
S1	Parent	White and Hispanic	Low	Bachelor's
S2	Parent	-	-	-
S3	Parent	Chinese	Middle	Graduate
S4	Prospective parent	Chinese	Middle	Bachelor's
S5	Parent	Vietnamese	Middle	Graduate
S6	Parent	Asian	Middle	Bachelor's
S7	Parent	White	High	Graduate
S8	Parent	White and Middle Eastern	Middle	Graduate
S9	Parent	White	Upper middle	Graduate
S10	Parent	Asian-American	Upper middle	Graduate
S11	Parent	Asian	High	Graduate
S12	Parent	White	Above median	Graduate
S13	Parent	White	Upper middle	Graduate
O1	Parent	African-American	Low*	Graduate
O2	Parent	Filipino	Low*	-
O3	Parent, Educator	White	Middle*	-
O4	Parent, Non-profit staff	-	-	-
O5	Parent	African-American / Black	Low*	-
O6	Parent	-	-	-
O7	Parent, School-parent liaison	Latino	Above median	Bachelor's
O8	Parent, School district staff	White	Above median	Graduate
O9†	Parent	Honduran	Low*	
O10†	Parent	Latino	Below median	Middle school
O11†	Parent	-	Unemployed	None
O12†	Parent	Guatemala	Unemployed	2nd grade
O13	Parent	-	-	-
O14	Parent	African-American	Below median	Some college

was a slow and stressful process with serious, lasting impacts on her child, *"[the teacher] was fired but I need to build my son's self-esteem back."* When O7's daughter started crying on the way to school every day, he faced a language barrier when raising concerns with the school principal. This experience in part motivates his current work as a bilingual parent liaison: *"I don't want any other parents going through the nightmare that I went through."*

Empowering parents to decide what is best for their child is one of the central arguments in support of school choice, but that empowerment is often elusive in reality [106]. A system that offers parents the ability to rank any of the schools in their district may appear to maximize expressiveness and empowerment.[8] However, this approach assumes that all families receive the same utility from being assigned their first choice school, second choice school, and so on. This makes invisible the trade-offs that families face, and the fact that for some families there may be no school that currently meets their needs. This becomes particularly problematic for evaluating and improving these systems. For example, statistics like the percentage of students assigned to their first choice school can overlook persisting segregation and inequalities in access to resources. Since ranked lists do not explicitly convey what factors families value in schools, it is also difficult to reduce these inequalities. For instance, districts trying to increase the enrollment of underserved students at higher resourced schools may assume that those families simply didn't know about those schools and jump to informational interventions, when really families need better transportation options to make those schools viable.

4.1.2 Other preference-based systems. The expressiveness of a preference language is constrained by its features, and to a lesser extent the format in which participants express their priorities over those features. Designers of preference-based systems have carefully considered how the selection of features over which participants can express their preferences shapes the information that those preferences convey. For instance, researchers have been exploring how to ask users explicitly what kind of online content brings them value to guide recommendation systems, rather than relying on implicit signals like likes and engagement time, which may not be correlated with value and well-being [115]. Park et al. [94] highlight the importance of considering the ways in which relations of power and positionality influence disparities in expressivity between stakeholders. For example, a participant in their study

[8] Note that even a complete ranked list of every school in a school district does not fully specify an individual's preferred decision outcome. The full space of decision outcomes is intractable, roughly n^s, where n is the number of students needing an assignment, and s is the number of schools available. Even this huge space of outcomes does not express any needs that are not met by the existing schools. As we emphasize throughout this work, even preference languages that appear "natural" have been designed and impose a particular set of constraints and costs.

suggested that students who need to secure visas for international conference travel could be prioritized for student volunteer positions, but pointed out that if the organizers had not dealt with this challenge personally it may not occur to them to include this factor in the selection process. Participants in Lee and Baykal [75] observed that the preference language assumed equal total utility for different participants, which did not account for the fact that some participants were indifferent about which task they would have to complete, and were happy to cede their preferred task to someone who had very strong preferences.

In response to these challenges, researchers have acknowledged that participation must begin at the stage of designing the preference language in order to ensure that it is expressive and morally acceptable [37, 39, 94]. For example, Lee et al. [78] and Lee et al. [79] began by conducting interviews with stakeholders to collaboratively define key factors over which they would later be asked their preferences. Freedman et al. [40] used an open-ended survey to elicit factors that people felt were and were not morally acceptable for prioritizing organ transplant recipients. However, developing participatory processes to define the preference language brings with it many of the challenges that preference-based systems aimed to address in the first place. For example, Park et al. [94] and Freedman et al. [40] found that for almost every factor they considered, some participants thought it was very important to consider in decision-making, while others felt strongly that it should not be considered. It is not clear how to resolve these disagreements in designing a preference language. As demonstrated by the student assignment case study, choices at this stage have significant consequences for participants' ultimate ability to express their needs to the decision-maker.

4.2 Cost: Participants have to translate their values and needs into the preference language

The assumption underpinning preference-based systems is that each individual has a latent utility function over some space of alternative decision outcomes, and the preference elicitation mechanism is a way of extracting partial information to estimate that function. In reality, people's preferences are constructed in and shaped by their particular social context, and shift over time. The costs of forming preferences depend on each individual's relevant knowledge, the time they have spent reflecting on their preferences, and how directly they can map their conception of their preferences to the constraints of the given preference language. These costs can exclude participants with less time and fewer resources to dedicate to the process, and can lead participants to rely heavily on social learning and heuristics that exacerbate bias and stereotypes in choice patterns.

4.2.1 Student Assignment. In the student assignment setting, researching the available schools, forming a ranked preference list, and submitting it to the school district can be difficult and time-consuming for families. First, families need to be aware that they have a choice to apply to different schools, and then understand what they are looking for in a school. Parents with fewer years of formal education and/or who had more recently arrived in the district often did not know that they could apply to several schools or what they should look for [99]. Even for parents who know what they are looking for, researching schools to transform those known preferences for *types* of schools into preferences for *actual* schools can be very time-consuming and frustrating [98, 99].

For example, when O5 first enrolled her child in kindergarten, *"it was very hard. A single parent, never done this before, never been in a public school, so I don't know how everything is run. People just assume that you know what you're doing and I had not a clue."* To learn more about the process, she enrolled in *"a week-long class of learning how to do the application process and how to pick schools, and what makes you want to pick a school."* While this helped her figure out what factors to consider (e.g., academics, sports, other extracurricular activities), and find schools that best met her criteria, this process required a significant time investment.

This challenge is exacerbated when information is not equally available and useful to all families. For example, online information was disorganized and inconsistent across schools [98, 99]. Schools with more resources tended to have more detailed and up-to-date websites, and the information available appeared tailored to the interests of higher-resourced families [99]. School tours were often cited as the most useful way to learn about a school. However, these tours can be extremely time consuming and logistically challenging, making them inaccessible to many parents, for instance, those who cannot take time off of work during school hours, or those who cannot secure childcare during the tour [98]. Beyond logistical challenges, tours may also offer less helpful information to families from marginalized backgrounds if tour guides cannot speak to the experiences of children from those backgrounds [99].

As a result of this lack of information, parents rely heavily on others in their network, including family and friends, as well as in online social networks, to inform their preferences. However, in relying on social learning to cope with a lack of authoritative and relevant information, people may fall back on stereotypes and generalizations to make judgments about schools [98]. For instance, one parent in San Francisco said that,

> *Facebook has been the most helpful because I feel like it's an insider's look into the actual school quality and parent experiences there. Parents there are very biased, but it's better to have information than no information.* (S3)

Until recently, parents in San Francisco needed to submit their preference list in person at the district's enrollment office by a set deadline. On-time participation significantly improves the chances of being assigned a preferable option, as the number of open seats dwindles in subsequent rounds. However, this requires keeping track of deadlines and finding the time to visit the district office, which creates additional information, time, and language barriers [107, 113].

4.2.2 Other preference-based systems. A key challenge in designing preference elicitation mechanisms is balancing the trade-off between how much information is extracted, and the costs imposed on participants (e.g., time and cognitive load). For example, Johnston et al. [64] propose a dynamic preference elicitation mechanism with the goal of minimizing the number of pairwise comparisons that each participant is asked to make. Usability issues not only make the task more difficult for participants, but also undermine the informativeness of the preferences they provide. For instance, Lee and Baykal [75] describe how participants turned to error-prone heuristics when they found it difficult to express their preferences using the provided interface. Studies that have crowd-sourced preference information also raise concerns about the quality of those judgments when participants lack information and a sense of personal investment [10].

These concerns may be mitigated when the participants involved already have substantial expertise and experience making similar kinds of judgments. For example, Lee et al. [78] elicited preferences regarding food donation matching from stakeholders who had personal experience with this process: staff at the non-profit that matches donors and recipients, volunteers who transport donations, and staff at donor and recipient organizations. Another approach is to provide participants with additional information to help them form their judgments. Researchers have shown experimentally that information provision can change how people report their preferences in the school choice context [8, 26, 29, 57, 84]. However, this information also must contend with limits on participants' time and cognitive resources, and there is likely always going to be additional context that could influence people's judgments that is not available. For example, Freedman et al. [40] asked people to make pairwise comparisons between potential organ transplant recipients on the basis of factors like whether they had skin cancer, but the researchers admit that there is far more information that someone may need to judge this situation, such as the prognosis for the disease.

The structure of the preference language can impose higher or lower informational and cognitive load on participants. For instance, Lee et al. [79] argue that pairwise comparisons are easier than ranking tasks for people who have not already formalized their preferences. In fact, the process of making pairwise comparisons actually helped participants understand their own perspectives better by forcing them to make concrete judgments and trade-offs. Ultimately, designing low cost preference languages requires considering participants' access to time and resources, their knowledge about the decision domain, and how much they have reflected on their needs and priorities.

4.3 Collectivism: Aggregating individual preferences is a limited means to achieve collective goals

The challenges with preference elicitation are further compounded when preferences are aggregated. In addition to satisfying individuals' preferences, decision-makers often have collective-level goals for how they distribute resources. For example, many school districts want to promote diverse classrooms and increase equitable access to educational opportunities [109]. However, optimizing for satisfying participants' *individual* preferences does not necessarily align with collective goals.

4.3.1 Student Assignment. In the school assignment context, preference-based systems circumvent the issue of defining and working towards collective goals by implicitly defining the socially optimal outcome as the one where each individual receives their most preferred option. In the case of San Francisco, this means that regardless of how well the system works, segregated choice patterns can lead to segregated schools. Even if SFUSD were able to assign every student to their first choice school, schools would remain heavily racially and economically segregated, with students from low-income and historically marginalized backgrounds concentrated in under-served schools [53]. In other words, the ultimate decisions remain heavily shaped by individuals' submitted preferences.

Although individuals' preferences shape both individual and collective outcomes, families have no formal avenues to participate in student assignment beyond independently submitting their individual, self-interested preferences. The introduction of school choice policies however, have been closely tied to desegregation efforts, an explicitly collective goal. Most parents in our sample were aware that the system was designed to promote integration and educational equity, and many supported those goals in the abstract. Several parents said that student diversity was an important factor that they looked for in a school (S5, S6, O1, O3). However, student assignment mechanisms model parents as consumers in a market for schools, rather than citizens participating in and negotiating a political process in which their own choices impact other members of their community. Parents are expected to secure a high quality education for their own child, while at the same time accepting that there are not sufficient seats at high-resourced schools for every child to have such an opportunity. As O4 put it, referring to under-resourced schools in Oakland, *"If you don't want to put your baby there, why would it be okay for me to have to put my baby there?"* Over time, these systems entrench the idea that individual choice is the only or best way to provide families access to schools, leaving school districts with fewer politically viable opportunities to promote community-level goals.

4.3.2 Other preference-based systems. In some systems, like assigning students to schools, dividing goods or tasks in a group of people [75, 76], or scheduling shifts [79], elicit preferences from people who will be directly, immediately impacted by the decision. In this case, it is possible to ask each person which outcome they would prefer for themselves. In other cases, preference elicitation is used to gather a wide variety of perspectives about hypothetical decision scenarios. This data is then used to train algorithms to make similar decisions that represent those perspectives in real scenarios, for example, who should receive a transplant [40] or public housing [72]. In the latter examples, people still contribute individual preferences, but those preferences are in regards to how *others* should be treated, rather than what they want for *themselves*.

However, asking people directly about their preferences for collective outcomes is still insufficient for ensuring that decisions meet collective goals or principles. For instance, people hold conflicting ethical views [10, 40, 72], in which case the choice of aggregation procedures can have critical influence over whose views are represented in the final decision [46]. Gordon et al. [46] point out that majoritarianism has silently ruled in machine learning data collection, and argue that making the pluralism of annotators' opinions more explicit can better enable decision-makers to prioritize particularly relevant or informed perspectives in a given decision.

In contexts where a decision-maker elicits people's preferences for their own outcome, we may not expect participants to be willing to trade-off any personal gain to advance collective goals. However, empirical evidence refutes this assumption. For example, in experimental settings, participants wanted to discuss and negotiate the final division of labor and goods after using a splitting algorithm, making room for cooperation and altruism [75, 76]. Giving participants insight into the complexity and constraints of the decision problem can also help them consider what trade-offs they are willing to accept, and why those trade-offs are needed [76, 79]. Further research is needed to understand how negotiation could scale beyond small groups, and how systems could engage participants in considering broader collective outcomes and trade-offs without significantly increasing the time and cognitive costs of engagement.

5 ADDRESSING THE CHALLENGES WITH PREFERENCES

In the previous section we described how the design of a preference language shapes what participants can convey through their preferences, how costly it is for them to do so, and the extent to which decision-makers can use those preferences to reach collective goals. Preference-based systems are gaining popularity because preferences can be a cheap and scalable way to quantify the relative value that each person has for some available alternatives, and there are a wide range of well-studied aggregation procedures that can translate individual preferences into a collective decision [16]. When the stakeholders involved are well-informed about the decision context and have relatively equal power and access to resources, this process has been shown to improve both quantitative measures of efficiency and fairness, and stakeholders' perceptions of the decisions [78]. However, problems arise in settings characterized by political tensions, power imbalances among participants, and high-stakes decision outcomes. The student assignment case study illustrates the challenges for designing preference-based systems that support equitable participation and work towards collective goals in addition to meeting individuals' needs. This case study also highlights the ways in which privileging individual choice can cause harm, particularly to those who are marginalized in the process and face greater barriers to participating.

The challenges in the student assignment domain are longstanding and complex, highlighting the need for new approaches to designing such systems. To this end, we discuss the implications of our findings for researchers and practitioners who are considering or implementing preference-based algorithmic systems. Figure 2 summarizes our recommendations for addressing each of the three considerations we discussed in the previous section. First, we discuss paths forward in the design of preference elicitation mechanisms that improve expressiveness and reduce costs (Section 5.1). Second, we draw on theories of procedural justice to discuss how to engage participants in co-defining collective goals, and propose technical mechanisms to account for those collective goals in aggregation procedures (Section 5.2). Finally, even with interventions to improve expressiveness, reduce costs, and account for collective goals, preference-based systems can cause harm, and are not appropriate in every setting. We conclude by discussing these risks (Section 5.3).

5.1 Preference elicitation: improve the options, simplify the preference language, and provide support

In Section 4, we discussed two ways that the design of the preference language shapes participation: it determines which of their needs participants can communicate, and how costly it is for them to do so. As a result, student assignment systems offer some families access to schools they prefer over their neighborhood school, but for others this choice is more elusive than meaningful. In this section we discuss three alternatives to mitigate these issues: improving the underlying set of options, simplifying the preference language, and providing support to reduce costs for participants.

5.1.1 Improving the options. In some cases, the preference language can only express people's preferences for a set of *existing* options. For example, the set of schools in a school district. This makes it less expressive for people who benefit less from that existing set. One path forward is to spend resources improving the underlying options to ensure every participant has a high quality option realistically available to them. A similar argument could be applied in settings like medical triage, where funds could be spent on obtaining more equipment rather than developing more complex methods for rationing the limited existing stock. Of course, this option is usually appealing to all stakeholders, but is often infeasible in the short-term. Regardless, it is worth reiterating that there can be a trade-off between spending time and money on new technology for distributing resources and improving that pool of resources.

5.1.2 Simplifying the preference language. A second alternative is to change the preference language that the system uses to collect participants' preferences to make it simpler or a closer match for how people already conceptualize their needs and values. For example, in the student assignment domain, one option is to reduce the choices available in meaningful ways. Currently, systems rely on each individual participant solving a complex task to reflect their preferences: they need to search among a very large set of potential schools and form a rank-ordered list of a handful of schools as their preference list. This contributes to inequality as the costs are higher for participants with less background knowledge about the school system, and more burdensome for those with less time and resources to spend on the process (as discussed in Section 4.2).

One approach to alleviate the adverse effect of informational asymmetry is offering *choice menus*, i.e., a limited subset of schools offered to each participant from which they can select their preferred schools. The basic idea is that searching in a larger pool puts those with higher search costs at a higher disadvantage compared to those with lower search costs.[9] Such choice menus could be constructed considering an applicant's characteristics (such as background and priorities). They can also act as a lever to balance the population characteristics in a school, by offering that school to a more balanced population of students.

[9] As an extreme case, consider menus of length one, where the asymmetry across search costs has no effect on those with higher costs.

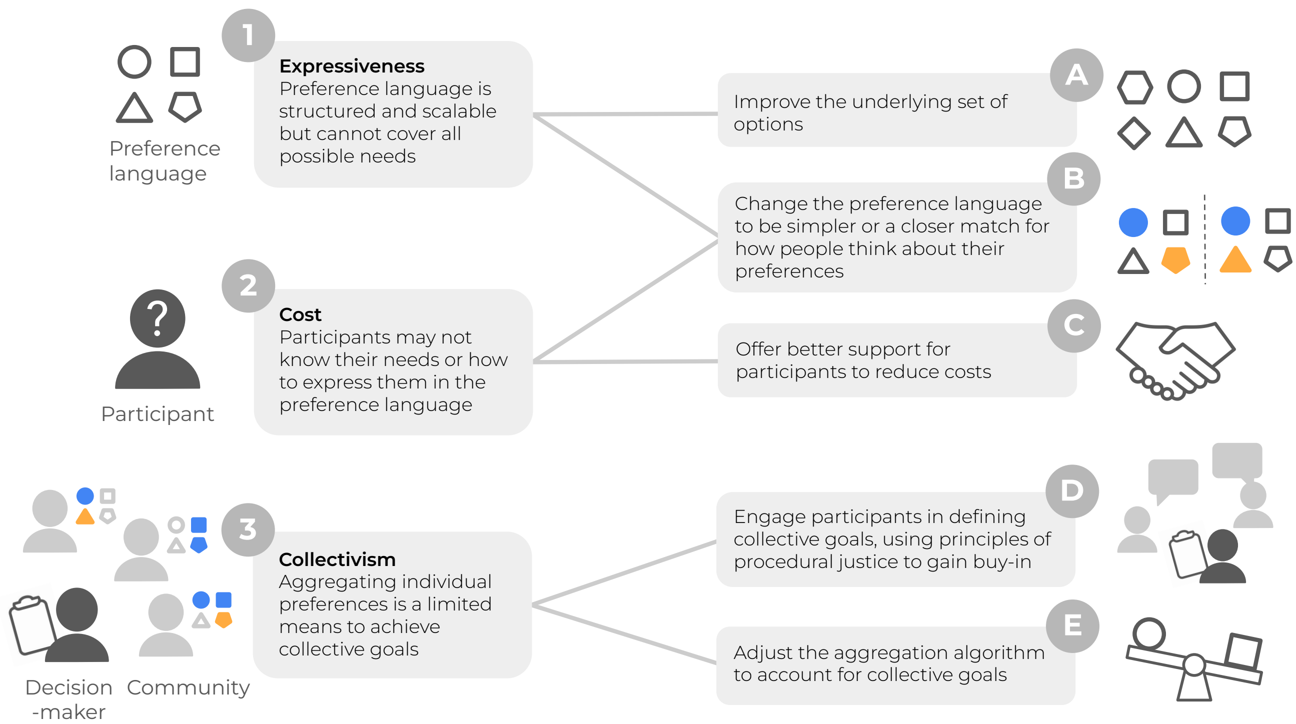

Figure 2: In Section 4 we identified three considerations for the design of preference languages that shape participation: 1) Expressiveness (4.1); 2) Cost (4.2); and 3) Collectivism (4.3). If the preference language makes it difficult and costly for marginalized stakeholders to communicate their needs, then the system offers unequal opportunities to participate. Systems that only account for individual preferences have limited means to reach collective goals. In Section 5 we discuss paths forward for designing preference elicitation mechanisms that are more expressive and less costly (A-C; 5.1) and aligning preference aggregation algorithms with co-defined collective goals (D-E; 5.2).

This is in line with the approach that SFUSD has taken in their ongoing redesign of their student assignment system [108]. It is important that decision-makers consider how changes to a preference language reduce or appear to reduce people's agency and choice, and the political ramifications of that decision. For example, SFUSD has engaged families in defining geographical zone and worked to communicate with them about the reasons for the changes.

5.1.3 Providing support. Even with a simple and carefully designed preference language, there will be participants who need support to fully participate. For example, most people are not experts in education, and many first-time parents do not know what to look for in a school. These participants will need informational support, even with fewer options to choose from. Support interventions must account for potential participants' knowledge of the context, technology access, language proficiency, and time constraints. In our prior work [99], we advocated for an assets-based design approach, which asks how we can amplify existing resources and strategies in a community [71]. We emphasized the important of personalized, one-on-one support through trusting relationships [99]. Other options for reducing the costs of using the preference language could include making it easier for people to apply online using devices they already own, e.g., through a mobile-friendly website.

In the next section we expand on the need to engage participants more deeply than through preference elicitation alone in order to define collective goals, and discuss technical mechanisms for working towards those collective goals in practice.

5.2 Preference aggregation: co-defining and accounting for collective goals

Preference-based systems have mostly assumed an individual model of preferences, where each person has some self-interested preferences for the available options, which they report to the decision-maker. However, in many distributive decisions, there are collective-level outcomes that are important but not captured in people's individual preferences. For example, many school districts have prioritized racial and socioeconomic diversity in schools, but assignment algorithms offer limited opportunities to trade-off between individual preferences for schools and collective goals.

As discussed in Section 4.3.2, adapting the preference language to explicitly ask people for their preferences over collective outcomes is likely not sufficient to address this problem. For one, such an approach is likely to make the preference language more complex and increase the costs of participation, in opposition to our recommendations in Section 5.1. Further, people's preferences are likely to be conflicting, leaving the decision-maker to choose explicitly or implicitly (through the choice of an aggregation procedure) which perspectives to prioritize [10, 46]. Instead, decision-makers may wish to enforce certain ethical principles or collective values, even if those do not align with the views of some stakeholders [87]. Therefore, there are two challenges to improving the alignment of preference-based systems with collective goals: defining which collective goals are important, and then ensuring that the aggregation procedure respects those goals.

5.2.1 Co-defining collective goals. Achieving collective goals with preference-based systems is only a problem when those goals do not align exactly with individuals' preferences. Therefore, working towards those goals will require that some stakeholders receive less personally preferable outcomes. In this way, advancing an individual preference-based system to pursue collective goals creates a deep tension between individual outcomes and collective values. It is thus especially important to consider who has a voice in shaping the collective goals, and how the decision-maker can build buy-in to those goals to establish legitimacy and trust. The principles of procedural justice offer paths forward.

Procedural justice theory teaches that the process through which a decision is made is as important as the outcome to satisfaction, perceptions of legitimacy, and compliance with that decision [119, 122]. The term "procedural justice" refers to individuals' perceptions that the procedures used to make decisions are fair. Importantly, procedural justice is a measure of subjective perceptions of fairness, not a measure of the objective fairness or equity of a decision [7]. Empirical determinants of procedural justice include the degree of voice and control individuals have in the process, whether the decision maker was respectful and unbiased, whether individuals identify with decision-makers and their values, and whether they understand the decision-making process [62, 119, 120]. Procedures that incorporate these process elements result in more satisfaction with the decision, greater perceptions of legitimacy of the decision maker, and more compliance with the decision, even when the outcome is unfavorable [118, 121]. At the same time, because procedural justice represents only the perception of fairness, yet mobilizes satisfaction and compliance with negative outcomes, it risks legitimating unfair decisions [44, 65, 116].

Rather than take individual preference satisfaction as the primary goal by hard wiring it into the assignment system, an alternative approach could encourage community deliberation and discussion to define collective goals. This process could acknowledge explicitly that in some instances, individual preferences must give way to collective values, and enable participants to determine appropriate trade-offs. With appropriate voice, participation, inclusion, and information, procedural justice principles suggest that policy makers would have a cushion of support for creating a process in which not everyone gets their first-choice school, but important collective values receive appropriate weight [123]. Procedural justice research suggests that satisfaction with outcomes, perceptions of the legitimacy of the decision makers, and acceptance of the decisions will follow.

5.2.2 Adjusting algorithms to account for collective goals. Once decision-makers have determined their collective goals, ideally in collaboration with the community, they need to ensure that the mechanism will account for those goals in determining outcomes. Most existing preference languages do not account for collective goals. Systems thus can feature undesirable equilibrium outcomes. For example, segregated schools can remain segregated, in part due to their reputation and the absence of means through which the participants can coordinate their decisions. In the context of assigning candidates to pre-military academies, "preference-specification" languages have been proposed that allow academies to express their preferences over the diversity of candidates they admit by specifying lower or upper quota constraints on subpopulations of candidates. Building on this work, we next discuss two approaches to address this issue, using student assignment algorithms as an example: one based on making minimal changes to Deferred Acceptance, and the other based on Mathematical Programming approaches that remove focus from Deferred Acceptance and stability as solution concepts.

School-specific priority scores are often used in the course of Deferred Acceptance to ration seats among students when there is excess demand at a school. One way to shift away from segregated equilibrium outcomes is the dynamic adaptation of these priority scores. When the policy maker has a different target population distribution at a school than at status quo, then higher priority scores can be given to the most absent subpopulations. These scores should be adjusted dynamically as the population distribution gets closer to the target over time. This approach can balance the population distribution induced by the algorithm at a school while preserving stability overall.

There are other technical approaches that could more directly implement collective goals at the expense of dismissing stability. One example is methods based on Mathematical Programming which, e.g., could associate a binary variable to each student-school pair and find an assignment that respects feasibility constraints (such as schools' capacities) while taking other collective goals into account, such as a statistical distance of the population distribution at a school from a target distribution.

While we have offered paths forward for designing preference-based systems that address the challenges discussed in Section 4, we do not believe that preference-based systems are always appropriate. We conclude this section by discussing the risks and limitations of preference-based systems as a whole.

5.3 Limitations and risks of preference-based systems

In deciding whether a preference-based system is appropriate, one should determine the extent to which the system offers meaningful choice (and to whom), and weigh the costs involved in providing adequate support to participants. It is also important to consider how introducing choice can make alternative or complementary avenues for promoting positive change more difficult.

Preference-based systems can only improve people's access to resources if there are high-quality options *realistically* available to every participant. As discussed in Section 4.1, offering participants a choice among the same set of alternatives does not necessarily offer every participant the same value. Some would argue that offering some degree of choice is better than offering none, even if that choice can only partially reduce disparities in access to resources. For instance, many parents in our sample were able to use the choice system to access a school they preferred over their neighborhood school, while admitting that others in their community were still not afforded that opportunity. While this limitation may not be a reason to forgo a preference-based system altogether, awareness of the realistic capacity for such a system to promote social change and equalize access to resources is important so that decision-makers and advocates can ensure complementary strategies are in place.

In any participatory process there will be unequal costs associated with participating, which if not properly addressed will exclude those already at the margins. As discussed above and in our prior work [99], personalized, one-on-one support is crucial for ensuring the full and equal participation of marginalized stakeholders in preference-based systems (5.1.3). This call to prioritize personalized support over one-size-fits-all solutions is at odds with some of the motivations for preference-based systems, e.g., that they easily scale to large groups of people. However, in order to reduce costs for participants, the decision-maker must be prepared to take on some of those costs on their behalf. To genuinely promote longer-term social change and progress towards distributive justice, we must recognize the necessary frictions and ongoing maintenance required to create inclusive, democratic, participatory processes [111]. Preference-based systems will not function equitably at scale if resources are not provided for this maintenance and support work.

A higher level political risk is that individual preference-based systems, e.g. school choice systems, can entrench the belief that individuals have a *right* to choose, and that the only optimal allocation of resources is that which efficiently maximizes people's individual preferences [60]. These systems resonate with individualist values, belief in markets, and popular neoliberal policy solutions. Collective goals, which may have been the stated motivation for the system in the first place, become tangential to the process. Further, when used to distribute public resources, this ideology shifts responsibility onto individual stakeholders to secure their access to resources to which they would otherwise be entitled, and normalizes and even exacerbates inequality. Once such a system is in place, it becomes increasingly difficult to implement policies that reduce, or seem to reduce, the degree to which the decision-maker respects individuals' preferences.

Alternatives to choice exist. For example, one alternative to ensure equitable opportunities in public education is a redistributive approach, which would create high quality programs in low-resource neighborhoods and protect local students' access to these programs [4]. Parent participation in school governance and strong teacher's unions also have a history of promoting change, for instance, fighting school closures in Black communities [38, 47], and pushing for improvements at schools like smaller class sizes, experienced teachers, and more equitable funding structures [1, 63, 106]. However, offering choice risks reducing the effectiveness or political viability of these alternatives by promoting individualism and free market values, and shifting the responsibility for ensuring the quality of children's education onto their guardians.

Finally, a question remains about whether preference-based systems should ever be considered truly "participatory" systems [83]. Our findings highlight that preference elicitation *alone* does not align with the values of participatory design. For instance, in the student assignment setting, families have little to no say over the conditions or structure of their participation, or how the system should work overall. Recent work somewhat addresses this problem by integrating other forms of engagement with participants, such as conducting interviews or surveys to define what the preference language should look like [40, 77]. As discussed above, another opportunity for integrating deeper participation is in defining collective goals that the system should support (Section 5.2.1). In summary, the specific configuration and conditions of participation are critical factors in determining whether a particular preference-based system is truly increasing people's voice and power in algorithmic decision-making [124].

6 CONCLUSION

A convenient way to incorporate stakeholders' input into algorithmic decision-making is by collecting and aggregating individual preferences. Implementing a preference-based system requires designing a preference language, in which participants will convey their needs and goals to the decision-maker. We used student assignment algorithms as a case study to illuminate three properties of preferences languages that shape opportunities for meaningful participation: expressiveness, cost, and collectivism. When these factors are not appropriately accounted for, preference-based systems can exacerbate inequality and fail to promote collective goals. Based on our findings, we offered implications and paths forward to increase the expressiveness of preference languages, reduce costs for participants, and work towards co-defined collective goals. With these paths forward comes the warning that preference-based systems are not appropriate in every setting, and that preference elicitation alone will never be sufficient to engage stakeholders in meaningful sharing of power and agency in algorithmic decision-making.

ACKNOWLEDGMENTS

We would like to thank our research participants. We are grateful to staff members at SFUSD, OUSD, and Oakland Enrolls for their insight and support throughout this research. We would also like to thank Moonhawk Kim for help with coordinating research collaboration and outreach, and Lucinda Matias Mendoza for interpretation support. This work was supported by the National Science Foundation award CCF-2131519 and a William T. Grant Foundation scholarship.

REFERENCES

[1] IntegrateNYC n.d.. *Real Integration*. IntegrateNYC. https://integratenyc.org/mission
[2] Atila Abdulkadiroglu, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. 2019. Do Parents Value School Effectiveness? *NBER Working Paper 23912* (2019). https://www.nber.org/papers/w23912
[3] Atila Abdulkadiroğlu and Tayfun Sönmez. 2003. School Choice: A Mechanism Design Approach. *American Economic Review* 93, 3 (June 2003), 729–747. https://doi.org/10.1257/000282803322157061
[4] African American Parent Advisory Council. 2017. AAPAC Reflections on SFUSD's Student Assignment Policy. https://archive.sfusd.edu/en/assets/sfusd-

staff/enroll/files/AAPAC_Student_Assignment_Presentation_3.8.17.pdf?_ga=2.81786516.225077105.1586807928-1027271765.1579115407
[5] Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, and Aaron Rieke. 2019. Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 199 (Nov. 2019), 30 pages. https://doi.org/10.1145/3359301
[6] Oscar Alvarado and Annika Waern. 2018. *Towards Algorithmic Experience: Initial Efforts for Social Media Contexts.* Association for Computing Machinery, New York, NY, USA, 1–12.
[7] Maureen L. Ambrose and Anke Arnaud. 2005. *Are procedural justice and distributive justice conceptually distinct?* Lawrence Erlbaum Associates Publishers, 59–84.
[8] Tahir Andrabi, Jishnu Das, and Asim Ijaz Khwaja. 2017. Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets. *American Economic Review* 107, 6 (June 2017), 1535–1563. https://doi.org/10.1257/aer.20140774
[9] Ben Casselman Anna Maria Barry-Jester and Dana Goldstein. 2015. The New Science of Sentencing. (2015). https://www.themarshallproject.org/2015/08/04/the-new-science-of-sentencing
[10] Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-Francois Bonnefon, and Iyad Rahwan. 2018. The Moral Machine experiment. *Nature* 563 (2018), 59–64. https://doi.org/10.1038/s41586-018-0637-6
[11] Shaowen Bardzell and Jeffrey Bardzell. 2011. Towards a Feminist HCI Methodology: Social Science, Feminism, and HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 675–684. https://doi.org/10.1145/1978942.1979041
[12] Gerdus Benade, Swaprava Nath, Ariel D. Procaccia, and Nisarg Shah. 2017. Preference Elicitation for Participatory Budgeting. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (San Francisco, California, USA) *(AAAI'17)*. AAAI Press, 376–382.
[13] Ruha Benjamin. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code.* Polity Press.
[14] Abeba Birhane, William Isaac, Vinodkumar Prabhakaran, Mark Diaz, Madeleine Clare Elish, Iason Gabriel, and Shakir Mohamed. 2022. Power to the People? Opportunities and Challenges for Participatory AI. In *Equity and Access in Algorithms, Mechanisms, and Optimization* (Arlington, VA, USA) *(EAAMO '22)*. Association for Computing Machinery, New York, NY, USA, Article 6, 8 pages. https://doi.org/10.1145/3551624.3555290
[15] Tolga Bolukbasi, Kai-Wei Chang, James Y Zou, Venkatesh Saligrama, and Adam T Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Eds.). Curran Associates, Inc., 4349–4357. http://papers.nips.cc/paper/6228-man-is-to-computer-programmer-as-woman-is-to-homemaker-debiasing-word-embeddings.pdf
[16] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. 2016. *Handbook of Computational Social Choice* (1st ed.). Cambridge University Press, USA.
[17] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. 2016. *Handbook of computational social choice.* Cambridge University Press.
[18] Anna Brown, Alexandra Chouldechova, Emily Putnam-Hornstein, Andrew Tobin, and Rhema Vaithianathan. 2019. Toward Algorithmic Accountability in Public Services: A Qualitative Study of Affected Community Perspectives on Algorithmic Decision-Making in Child Welfare Services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 41, 12 pages. https://doi.org/10.1145/3290605.3300271
[19] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Proceedings of Machine Learning Research, Vol. 81)*, Sorelle A. Friedler and Christo Wilson (Eds.). PMLR, New York, NY, USA, 77–91. http://proceedings.mlr.press/v81/buolamwini18a.html
[20] Simon Burgess, Ellen Greaves, Anna Vignoles, and Deborah Wilson. 2009. Parental choice of primary school in England: what 'type' of school do parents choose? *The Centre for Market and Public Organisation* (2009).
[21] Simon Burgess, Ellen Greaves, Anna Vignoles, and Deborah Wilson. 2015. What Parents Want: School Preferences and School Choice. *The Economic Journal* 125, 587 (2015), 1262–1289. https://doi.org/10.1111/ecoj.12153
[22] Robin Burke, Amy Voida, Nicholas Mattei, and Nasim Sonboli. 2020. Algorithmic Fairness, Institutional Logics, and Social Choice. *AI for Social Good Workshop* (2020).
[23] Yves Cabannes. 2004. Participatory budgeting: a significant contribution to participatory democracy. *Environment and Urbanization* 16, 1 (2004), 27–46. https://doi.org/10.1177/095624780401600104 arXiv:https://doi.org/10.1177/095624780401600104
[24] John M. Carroll and Mary Beth Rosson. 2007. Participatory design in community informatics. *Design Studies* 28, 3 (2007), 243 – 261. http://www.sciencedirect.com/science/article/pii/S0142694X07000191
[25] Kathy Charmaz. 2006. *Constructing grounded theory: A practical guide through qualitative research.* SAGE.
[26] Yan Chen. 2017. Information Acquisition and Provision in School Choice : An Experimental Study.
[27] Yan Chen and Yinghua He. 2020. *Information Acquisition and Provision in School Choice : An Experimental Study.* Working Paper. http://yanchen.people.si.umich.edu/papers/Chen_He_2020_09_Distribute.pdf
[28] Alexandra Chouldechova. 2017. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big Data* 5, 2 (2017), 153–163. https://doi.org/10.1089/big.2016.0047 PMID: 28632438.
[29] Sean P Corcoran, Jennifer L Jennings, Sarah R Cohodes, and Carolyn Sattin-Bajaj. 2018. *Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions.* Working Paper 24471. National Bureau of Economic Research. https://doi.org/10.3386/w24471 Series: Working Paper Series.
[30] Fernando Delgado, Stephen Yang, Michael Madaio, and Qian Yang. 2021. Stakeholder Participation in AI: Beyond "Add Diverse Stakeholders and Stir". arXiv:2111.01122 [cs.AI]
[31] Tingting Ding and Andrew Schotter. 2017. Matching and chatting: An experimental study of the impact of network communication on school-matching mechanisms. *Games and Economic Behavior* 103 (2017), 94 – 115. https://doi.org/10.1016/j.geb.2016.02.004 John Nash Memorial.
[32] Tingting Ding and Andrew Schotter. 2019. Learning and Mechanism Design: An Experimental Test of School Matching Mechanisms with Intergenerational Advice. *The Economic Journal* 129, 623 (05 2019), 2779–2804.
[33] Joseph Donia and Jay Shaw. 2021. Co-Design and Ethical Artificial Intelligence for Health: Myths and Misconceptions. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) *(AIES '21)*. Association for Computing Machinery, New York, NY, USA, 77. https://doi.org/10.1145/3461702.3462537
[34] Allison Woodruff et al. 2018. A Qualitative Exploration of Perceptions of Algorithmic Fairness. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press.
[35] Haiyi Zhu et al. 2018. Value-Sensitive Algorithm Design. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 1–23.
[36] Yannai A. Gonczarowski et al. 2019. Matching for the Israeli "Mechinot" Gap-Year Programs: Handling Rich Diversity Requirements. In *EC '19*. ACM Press.
[37] Florian Evequoz, Johan Rochel, Vijay Keswani, and L. Elisa Celis. 2022. Diverse Representation via Computational Participatory Elections - Lessons from a Case Study. In *Equity and Access in Algorithms, Mechanisms, and Optimization* (Arlington, VA, USA) *(EAAMO '22)*. Association for Computing Machinery, New York, NY, USA, Article 12, 11 pages. https://doi.org/10.1145/3551624.3555297
[38] Eve L. Ewing. 2018. *Ghosts in the Schoolyard: Racism and School Closings on Chicago's South Side.* University of Chicago Press.
[39] Gregory Ewing and Ibrahim Demir. 2021. An ethical decision-making framework with serious gaming: a smart water case study on flooding. *Journal of Hydroinformatics* 23, 3 (May 2021), 466–482. https://doi.org/10.2166/hydro.2021.097
[40] Rachel Freedman, Jana Schaich Borg, Walter Sinnott-Armstrong, John P. Dickerson, and Vincent Conitzer. 2018. Adapting a Kidney Exchange Algorithm to Align with Human Values. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (New Orleans, LA, USA) *(AIES '18)*. Association for Computing Machinery, New York, NY, USA, 115. https://doi.org/10.1145/3278721.3278727
[41] D. Gale and L. S. Shapley. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly* 69, 2 (1 1962), 9–15.
[42] Nikhil Garg, Lodewijk Gelauff, Sukolsak Sakshuwong, and Ashish Goel. 2019. Who is in Your Top Three? Optimizing Learing in Elections with Many Candidates. In *AAAI Conference on Human Computation and Crowdsourcing (HCOMP '19)*.
[43] Marzyeh Ghassemi, Tristan Naumann, Peter Schulam, Andrew L. Beam, Irene Ya-Ping Chen, and Rajesh Ranganath. 2019. A Review of Challenges and Opportunities in Machine Learning for Health. *arXiv preprint* (2019).
[44] James L. Gibson. 1989. Understandings of Justice: Institutional Legitimacy, Procedural Justice, and Political Tolerance. *Law & Society Review* 23, 3 (1989), 469–496. http://www.jstor.org/stable/3053830
[45] Steven Glazerman and Dallas Dotter. 2017. Market Signals: Evidence on the Determinants and Consequences of School Choice From a Citywide Lottery. *Educational Evaluation and Policy Analysis* 39, 4 (2017), 593–619. https://doi.org/10.3102/0162373717702964
[46] Mitchell L. Gordon, Michelle S. Lam, Joon Sung Park, Kayur Patel, Jeff Hancock, Tatsunori Hashimoto, and Michael S. Bernstein. 2022. Jury Learning: Integrating Dissenting Voices into Machine Learning Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article

115, 19 pages. https://doi.org/10.1145/3491102.3502004

[47] Mark Winston Griffith and Max Freedman. 2019. Episode 5: The Disappearing District. https://www.schoolcolorspodcast.com/episodes/episode-5-the-disappearing-district

[48] Pablo Guillen and Rustamdjan Hakimov. 2017. Not quite the best response: truth-telling, strategy-proof matching, and the manipulation of others. *Experimental Economics* 20, 3 (2017), 670–686.

[49] Pablo Guillen and Rustamdjan Hakimov. 2018. The effectiveness of top-down advice in strategy-proof mechanisms: A field experiment. *European Economic Review* 101, C (2018), 505–511.

[50] Pablo Guillen and Alexander Hing. 2014. Lying through their teeth: Third party advice and truth telling in a strategy proof mechanism. *European Economic Review* 70 (2014), 178 – 185. https://doi.org/10.1016/j.euroecorev.2014.05.002

[51] Isa E. Hafalir, M. Bumin Yenmez, and Muhammed A. Yildirim. 2013. Effective affirmative action in school choice. *Theoretical Economics* 8, 2 (2013), 325–363. https://doi.org/10.3982/TE1135

[52] Aaron Halfaker and R. Stuart Geiger. 2020. ORES: Lowering Barriers with Participatory Machine Learning in Wikipedia. *ArXiv preprint* (2020). https://arxiv.org/pdf/1909.05189.pdf

[53] Matt Haney, Steven Cook, and Rachel Norton. 2018. Developing a Community Based Student Assignment System for SFUSD.

[54] Christina N. Harrington, Lauren Wilcox, Kay Connelly, Wendy Rogers, and Jon Sanford. 2018. Designing Health and Fitness Apps with Older Adults: Examining the Value of Experience-Based Co-Design. In *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare* (New York, NY, USA) *(PervasiveHealth '18)*. Association for Computing Machinery, New York, NY, USA, 15–24. https://doi.org/10.1145/3240925.3240929

[55] Avinatan Hassidim, Assaf Romm, and Ran I. Shorrer. 2016. "Strategic" Behavior in a Strategy-Proof Environment. In *Proceedings of the 2016 ACM Conference on Economics and Computation* (Maastricht, The Netherlands) *(EC '16)*. Association for Computing Machinery, New York, NY, USA, 763–764. https://doi.org/10.1145/2940716.2940751

[56] Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. 2009. Heterogeneous Preferences and the Efficacy of Public School Choice. *NBER Working Paper* (2009).

[57] Justine S Hastings and Jeffrey M Weinstein. 2007. *Information, School Choice, and Academic Achievement: Evidence from Two Experiments.* Working Paper 13623. National Bureau of Economic Research. https://doi.org/10.3386/w13623 Series: Working Paper Series.

[58] Justine S Hastings and Jeffrey M Weinstein. 2007. *Information, School Choice, and Academic Achievement: Evidence from Two Experiments.* Working Paper 13623. National Bureau of Economic Research. https://doi.org/10.3386/w13623

[59] Yoan Hermstrüwer. 2019. *Transparency and Fairness in School Choice Mechanisms.* Technical Report. Max Planck Institute for Research on Collective Goods.

[60] Zoë Hitzig. 2019. The normative gap: mechanism design and ideal theories of justice. *Economics and Philosophy* (2019), 1–28. https://doi.org/10.1017/S0266267119000270

[61] Anna Lauren Hoffmann. 2021. Terms of inclusion: Data, discourse, violence. *New Media & Society* 23, 12 (2021), 3539–3556. https://doi.org/10.1177/1461444820958725 arXiv:https://doi.org/10.1177/1461444820958725

[62] Yuen J. Huo. 2003. Procedural Justice and Social Regulation Across Group Boundaries: Does Subgroup Identity Undermine Relationship-Based Governance? *Personality and Social Psychology Bulletin* 29, 3 (2003), 336–348. https://doi.org/10.1177/0146167202250222 arXiv:https://doi.org/10.1177/0146167202250222 PMID: 15273011.

[63] Sarah Jaffe. 2019. *The Chicago Teachers Strike Was a Lesson in 21st-Century Organizing.* The Nation. https://www.thenation.com/article/archive/chicago-ctu-strike-win

[64] Caroline M. Johnston, Simon Blessenohl, and Phebe Vayanos. 2020. Preference Elicitation and Aggregation to Aid with Patient Triage during the COVID-19 Pandemic. *Workshop on Participatory Approaches to Machine Learning at International Conference on Machine Learning (ICML)* (2020). https://participatoryml.github.io/papers/2020/33.pdf

[65] John T. Jost, Mahzarin R. Banaji, and Brian A. Nosek. 2004. A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo. *Political Psychology* 25, 6 (2004), 881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9221.2004.00402.x

[66] Anson Kahng, Min Kyung Lee, Ritesh Noothigattu, Ariel Procaccia, and Christos-Alexandros Psomas. 2019. Statistical Foundations of Virtual Democracy. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 97)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.). PMLR, Long Beach, California, USA, 3173–3182. http://proceedings.mlr.press/v97/kahng19a.html

[67] Adam J. Kapor, Christopher A. Neilson, and Seth D. Zimmerman. 2020. Heterogeneous Beliefs and School Choice Mechanisms. *American Economic Review* 110, 5 (May 2020), 1274–1315. https://doi.org/10.1257/aer.20170129

[68] Matt Kasman and Jon Valant. 2019. The opportunities and risks of K-12 student placement algorithms. *The Brookings Institute* (2019).

[69] Anna Kawakami, Venkatesh Sivaraman, Logan Stapleton, Hao-Fei Cheng, Adam Perer, Zhiwei Steven Wu, Haiyi Zhu, and Kenneth Holstein. 2022. "Why Do I Care What's Similar?" Probing Challenges in AI-Assisted Child Welfare Decision-Making through Worker-AI Interface Design Concepts. In *Designing Interactive Systems Conference*. 454–470.

[70] Kotaro Suzumura Kenneth J. Arrow, Amartya Sen. 2010. *Handbook of Social Choice and Welfare.* Elsevier.

[71] John Kretzmann and John P. McKnight. 1996. Assets-based community development. *National Civic Review* 85, 4 (1996), 23–29. https://doi.org/10.1002/ncr.4100850405 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/ncr.4100850405

[72] Amanda Kube, Sanmay Das, Patrick J. Fowler, and Yevgeniy Vorobeychik. 2022. Just Resource Allocation? How Algorithmic Predictions and Human Notions of Justice Interact. In *Proceedings of the 23rd ACM Conference on Economics and Computation* (Boulder, CO, USA) *(EC '22)*. Association for Computing Machinery, New York, NY, USA, 1184–1242. https://doi.org/10.1145/3490486.3538305

[73] Mariana Laverde. 2020. Unequal Assignments to Public Schools and the Limits of School Choice.

[74] Mariana Laverde et al. 2022. *Distance to Schools and Equal Access in School Choice Systems.* Technical Report.

[75] 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* (Portland, Oregon, USA) *(CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1035–1048. https://doi.org/10.1145/2998181.2998230

[76] Min Kyung Lee, Anuraag Jain, Hea Jin Cha, Shashank Ojha, and Daniel Kusbit. 2019. Procedural Justice in Algorithmic Fairness: Leveraging Transparency and Outcome Control for Fair Algorithmic Mediation. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 182 (nov 2019), 26 pages. https://doi.org/10.1145/3359284

[77] Min Kyung Lee, Ji Tae Kim, and Leah Lizarondo. 2017. A Human-Centered Approach to Algorithmic Services: Considerations for Fair and Motivating Smart Community Service Management That Allocates Donations to Non-Profit Organizations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3365–3376. https://doi.org/10.1145/3025453.3025884

[78] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, and Ariel D. Procaccia. 2019. WeBuildAI: Participatory Framework for Algorithmic Governance. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 181 (Nov. 2019), 35 pages. https://doi.org/10.1145/3359283

[79] Min Kyung Lee, Ishan Nigam, Angie Zhang, Joel Afriyie, Zhizhen Qin, and Sicun Gao. 2021. Participatory Algorithmic Management: Elicitation Methods for Worker Well-Being Models. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) *(AIES '21)*. Association for Computing Machinery, New York, NY, USA, 715–726. https://doi.org/10.1145/3461702.3462628

[80] Margaux Luflade. 2017. *The value of information in centralized school choice systems.* Working Paper. https://economics.sas.upenn.edu/sites/default/files/filevault/event_papers/LUFLADE_JobMarketPaper.pdf

[81] Sharan B Merriam and Associates. 2002. *Introduction to qualitative research.* Jossey-Bass, 1–17.

[82] Michael J. Muller and Allison Druin. 2012. *Participatory Design: The Third Space in HCI* (3 ed.). L. Erlbaum Associates Inc., USA, 1125–1154.

[83] Michael J Muller and Sarah Kuhn. 1993. Participatory design. *Commun. ACM* 36, 6 (1993), 24–28.

[84] Yusuke Narita. 2018. *Match or Mismatch? Learning and Inertia in School Choice.* SSRN Scholarly Paper ID 3198417. Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.3198417

[85] ThÃnh Nguyen and Rakesh Vohra. 2019. Stable Matching with Proportionality Constraints. *Operations Research* 67, 6 (2019), 1503–1519. https://doi.org/10.1287/opre.2019.1909

[86] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism.* NYU Press.

[87] Ritesh Noothigattu, Snehalkumar Gaikwad, Edmond Awad, Sohan Dsouza, Iyad Rahwan, Pradeep Ravikumar, and Ariel Procaccia. 2018. A Voting-Based System for Ethical Decision Making. https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17052/15857

[88] Donald Norman. 2013. The Design of Everyday Things: Revised and Expanded Edition. (2013).

[89] Ziad Obermeyer and Sendhil Mullainathan. 2019. Dissecting Racial Bias in an Algorithm That Guides Health Decisions for 70 Million People. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA,

USA) *(FAT* '19)*. Association for Computing Machinery, New York, NY, USA, 89. https://doi.org/10.1145/3287560.3287593

[90] Cathy O'Neil. 2016. *Weapons of Math Destruction.* Crown Books.

[91] Hessel Oosterbeek, Sándor Sóvágó, and Bas Klaauw. 2019. Why are Schools Segregated? Evidence from the Secondary-School Match in Amsterdam. *CEPR Discussion Paper No. DP13462* (2019). https://ssrn.com/abstract=3319783

[92] Joana Pais and Agnes Pintér. 2008. School choice and information: An experimental study on matching mechanisms. *Games and Economic Behavior* 64, 1 (2008), 303–328.

[93] Hyanghee Park, Daehwan Ahn, Kartik Hosanagar, and Joonhwan Lee. 2022. Designing Fair AI in Human Resource Management: Understanding Tensions Surrounding Algorithmic Evaluation and Envisioning Stakeholder-Centered Solutions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 51, 22 pages. https://doi.org/10.1145/3491102.3517672

[94] Joon Sung Park, Karrie Karahalios, Niloufar Salehi, and Motahhare Eslami. 2022. Power Dynamics and Value Conflicts in Designing and Maintaining Socio-Technical Algorithmic Processes. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 110 (apr 2022), 21 pages. https://doi.org/10.1145/3512957

[95] Parag A. Pathak. 2017. *What Really Matters in Designing School Choice Mechanisms.* Econometric Society Monographs, Vol. 1. Cambridge University Press, 176–214. https://doi.org/10.1017/9781108227162.006

[96] Manish Raghavan and Solon Barocas. 2019. Challenges for mitigating bias in algorithmic hiring. (2019). https://www.brookings.edu/research/challenges-for-mitigating-bias-in-algorithmic-hiring/

[97] Alex Rees-Jones and Samuel Skowronek. 2018. An experimental investigation of preference misrepresentation in the residency match. *Proceedings of the National Academy of Sciences* 115, 45 (2018), 11471–11476. https://doi.org/10.1073/pnas.1803212115

[98] Samantha Robertson, Tonya Nguyen, and Niloufar Salehi. 2021. Modeling Assumptions Clash with the Real World: Transparency, Equity, and Community Challenges for Student Assignment Algorithms. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445748

[99] Samantha Robertson, Tonya Nguyen, and Niloufar Salehi. 2022. Not Another School Resource Map: Meeting Underserved Families' Information Needs Requires Trusting Relationships and Personalized Care. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 316 (nov 2022), 23 pages. https://doi.org/10.1145/3555207

[100] Alvin E. Roth. 2008. Deferred acceptance algorithms: history, theory, practice, and open questions. *International Journal of Game Theory* 36, 3-4 (Jan. 2008), 537–569.

[101] Alvin E. Roth. 2015. *Who Gets What And Why: the hidden world of matchmaking and market design.* Harper Collins.

[102] Richard Rothstein. 2013. Why Our Schools Are Segregated. *Educational Leadership* 70, 8 (2013), 50–55.

[103] San Francisco Unified School District and County Office of Education. 2020. Board Policy 5101.2: Elementary School Student Assignment. https://go.boarddocs.com/ca/sfusd/Board.nsf/files/BVYUGB7BF68F/\protect\T1\textdollarfile/BP%205101.2%2C%20Elementary%20School%20Student%20Assignment.pdf

[104] Devansh Saxena and Shion Guha. 2020. Conducting Participatory Design to Improve Algorithms in Public Services: Lessons and Challenges. In *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing* (Virtual Event, USA) *(CSCW '20 Companion)*. Association for Computing Machinery, New York, NY, USA, 383–388. https://doi.org/10.1145/3406865.3418331

[105] Mark Schneider and Jack Buckley. 2002. What Do Parents Want from Schools? Evidence from the Internet. *Educational Evaluation and Policy Analysis* 24, 2 (2002), 133–144. http://www.jstor.org/stable/3594140

[106] Janelle T. Scott. 2013. A Rosa Parks moment? School choice and the marketization of civil rights. *Critical Studies in Education* 54, 1 (2013), 5–18. https://doi.org/10.1080/17508487.2013.739570 arXiv:https://doi.org/10.1080/17508487.2013.739570

[107] SFUSD. 2019. Why We're Redesigning Student Assignment. (2019). https://www.sfusd.edu/studentassignment/why-were-redesigning-student-assignment

[108] SFUSD. 2020. Changes to Student Assignment for Elementary Schools. (2020). https://www.sfusd.edu/schools/enroll/student-assignment-policy/student-assignment-changes

[109] SFUSD Office of Education. 2010. Board Policy 5101: Student Assignment. https://go.boarddocs.com/ca/sfusd/Board.nsf/goto?open&id=B55QMC657423

[110] Lloyd Shapley and Herbert Scarf. 1974. On cores and indivisibility. *Journal of Mathematical Economics* 1, 1 (1974), 23 – 37. https://doi.org/10.1016/0304-4068(74)90033-0

[111] Mona Sloane, Emanuel Moss, Olaitan Awomolo, and Laura Forlano. 2020. Participation is not a Design Fix for Machine Learning. arXiv:2007.02423 [cs.CY]

[112] C. Estelle Smith, Bowen Yu, Anjali Srivastava, Aaron Halfaker, Loren Terveen, and Haiyi Zhu. 2020. Keeping Community in the Loop: Understanding Wikipedia Stakeholder Values for Machine Learning-Based Systems. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376783

[113] Jeremy Adam Smith. 2015. As Parents Get More Choice, S.F. Schools Resegregate. *San Francisco Public Press* (2015).

[114] Logan Stapleton, Min Hun Lee, Diana Qing, Marya Wright, Alexandra Chouldechova, Ken Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. Imagining new futures beyond predictive systems in child welfare: A qualitative study with impacted stakeholders. In *2022 ACM Conference on Fairness, Accountability, and Transparency.* 1162–1177.

[115] Jonathan Stray. 2020. Aligning AI Optimization to Community Well-Being. *International Journal of Community Wellbeing* 3, 4 (Nov 2020), 443–463. https://doi.org/10.1007/s42413-020-00086-3

[116] Jason Sunshine and Tom R. Tyler. 2003. The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing. *Law and Society Review* 37, 3 (2003), 513–548.

[117] Latanya Sweeney. 2013. Discrimination in Online Ad Delivery. *SSRN Scholarly Paper* (2013).

[118] Tom Tyler, Peter Degoey, and Heather Smith. 1996. Understanding why the justice of group procedures matters: A test of the psychological dynamics of the group-value model. *Journal of Personality and Social Psychology* 70, 5 (1996), 913–930. https://doi.org/10.1037/0022-3514.70.5.913

[119] Tom R. Tyler. 1988. What is Procedural Justice?: Criteria used by Citizens to Assess the Fairness of Legal Procedures. *Law & Society Review* 22, 1 (1988), 103–135. http://www.jstor.org/stable/3053563

[120] Tom R. Tyler. 2000. Multiculturalism and the Willingness of Citizens to Defer to Law and to Legal Authorities. *Law & Social Inquiry* 25, 4 (2000), 983–1019. http://www.jstor.org/stable/829122

[121] Tom R. Tyler. 2004. Affirmative Action in an Institutional Context: The Antecedents of Policy Preferences and Political Support. *Social Justice Research* 17, 1 (March 2004), 5–24. https://doi.org/10.1023/B:SORE.0000018090.84298.5b

[122] Tom R. Tyler. 2006. *Why People Obey the Law.* Princeton University Press.

[123] Tom R. Tyler and Robert J. Bies. 2015. Beyond formal procedures: The interpersonal context of procedural justice. *Applied Social Psychology and Organizational Settings* (2015), 77–98.

[124] John Vines, Rachel Clarke, Peter Wright, John McCarthy, and Patrick Olivier. 2013. Configuring participation: on how we involve people in design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. ACM Press, Paris, France, 429. https://doi.org/10.1145/2470654.2470716

[125] Mohammad Yaghini, Andreas Krause, and Hoda Heidari. 2021. A Human-in-the-Loop Framework to Construct Context-Aware Mathematical Notions of Outcome Fairness. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (Virtual Event, USA) *(AIES '21)*. Association for Computing Machinery, New York, NY, USA, 1023–1033. https://doi.org/10.1145/3461702.3462583

[126] Takuya Yokota and Yuri Nakao. 2022. Toward a Decision Process of the Best Machine Learning Model for Multi-Stakeholders: A Crowdsourcing Survey Method. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization* (Barcelona, Spain) *(UMAP '22 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 245–254. https://doi.org/10.1145/3511047.3538033

[127] Angie Zhang, Alexander Boltz, Chun Wei Wang, and Min Kyung Lee. 2022. Algorithmic Management Reimagined For Workers and By Workers: Centering Worker Well-Being in Gig Work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 14, 20 pages. https://doi.org/10.1145/3491102.3501866