

Modeling Assumptions Clash with the Real World: Transparency, Equity, and Community Challenges for Student Assignment Algorithms

*Sam Robertson
Tonya Nguyen
Niloufar Salehi*



Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2021-93

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2021/EECS-2021-93.html>

May 14, 2021

Copyright © 2021, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Acknowledgement

Thank you to my co-advisors Professors Niloufar Salehi and Moritz Hardt for their support. Thank you to Tonya Nguyen for her research collaboration and friendship. We are indebted to our research participants who generously shared their experiences and insights. We are also grateful to Berkeley AFOG, the participants at MD4SG 2020, and the anonymous reviewers at CHI 2021 for their helpful feedback on this work. Finally, thank you to my family for their love and support.

Modeling Assumptions Clash with the Real World: Transparency, Equity, and Community Challenges for Student Assignment Algorithms

by Samantha Robertson

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences,
University of California at Berkeley, in partial satisfaction of the requirements for the
degree of **Master of Science, Plan II.**

Approval for the Report and Comprehensive Examination:

Committee:



Professor Moritz Hardt
Research Advisor

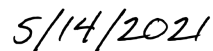


(Date)

* * * * *



Professor Niloufar Salehi
Second Reader



(Date)

Modeling Assumptions Clash with the Real World: Transparency, Equity, and Community Challenges for Student Assignment Algorithms

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

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

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

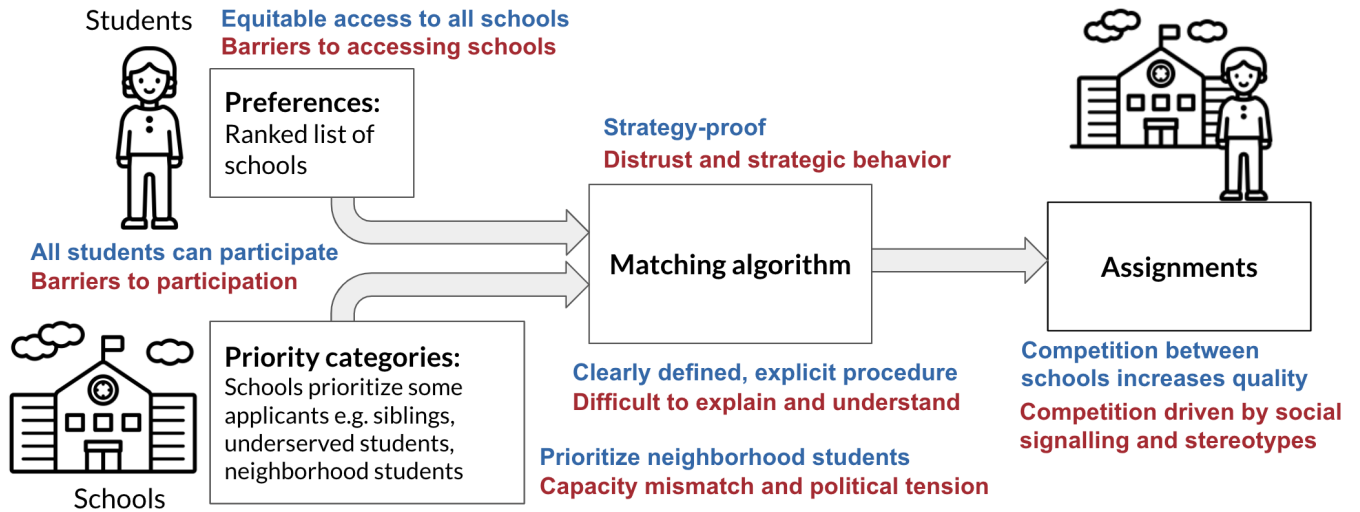


Figure 1: Student assignment algorithms were designed meet school district values based on modeling assumptions (blue/top) that clash with the constraints of the real world (red/bottom). Students are expected to have predefined preferences over all schools, which they report truthfully. The procedure is intended to be easy to explain and optimally satisfies student preferences. In practice however, these assumptions clash with the real world characterized by unequal access to information, resource constraints (e.g. commuting), and distrust.

ABSTRACT

Across the United States, a growing number of school districts are turning to matching algorithms to assign students to public schools. The designers of these algorithms aimed to promote values such as transparency, equity, and community in the process. However, school districts have encountered practical challenges in their deployment. In fact, San Francisco Unified School District voted to stop using and completely redesign their student assignment algorithm because it was frustrating for families and it was not promoting educational equity in practice. We analyze this system using a Value Sensitive Design approach and find that one reason values are not met in practice is that the system relies on modeling assumptions about families' priorities, constraints, and goals that clash with the real world. These assumptions overlook the complex barriers to ideal participation that many families face, particularly because of socioeconomic inequalities. We argue that direct, ongoing engagement with stakeholders is central to aligning algorithmic values with real world conditions. In doing so we must broaden how we evaluate algorithms while recognizing the limitations of

purely algorithmic solutions in addressing complex socio-political problems.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

KEYWORDS

student assignment, mechanism design, value sensitive design

1 INTRODUCTION

Algorithmic systems are increasingly involved in high-stakes decision-making such as child welfare [12, 73], credit scoring [51], medicine [34, 63], and law enforcement [7]. Documented instances of discriminatory algorithmic decision-making [4, 7, 19, 63] and biased system performance [9, 13, 62, 80] have prompted a growing interest in designing systems that reflect the values and needs of the communities in which they are embedded [31, 82]. However, even when systems are designed to support shared values, they do not always promote those values in practice [81]. One reason

why an algorithmic system may not support values as expected is that *these expectations rely on modeling assumptions about the world that clash with how the world actually works*. In this paper, we examine one such breakdown, the San Francisco Unified School District’s student assignment algorithm, to study where and how those clashes occur and to offer paths forward.

San Francisco has a long history of heavily segregated neighborhoods which has resulted in segregated schools when students attend their neighborhood school [43]. In 2011, in an effort to promote educational equity and racially integrated classrooms, San Francisco Unified School District joined many cities across the country who were turning to assignment algorithms to determine where students go to school [43]. Rather than enrolling in their neighborhood school, students submit their ranked preference list of schools to the district, and the algorithm uses those preferences along with school priorities and capacity constraints to match students to schools. These algorithms have been met with great excitement for their potential to provide more equitable access to public education and give families more flexibility compared to a neighborhood-based assignment system [50]. By 2018, however, diversity in schools had instead decreased and parents were frustrated by an opaque and unpredictable process [43]. In fact, many schools were now more segregated than the neighborhoods they were in [70]. The algorithm had failed to support the values its designers had intended and the San Francisco Board of Education voted for a complete overhaul and redesign of the system [43].

Following a Value Sensitive Design approach, we ask two central questions: 1) What values were designers and policy-makers hoping this algorithm would support? 2) Why were those values not met in practice? To answer these questions we first analyzed the school district’s publicly available policy documents on student assignment and conducted a review of the relevant economics literature where matching algorithms for student assignment have been developed. To answer the second question, we conducted an empirical investigation into how the algorithm is used in practice. We conducted 13 semi-structured interviews with parents in San Francisco who have used the assignment system and performed content analysis of 12 Reddit threads where parents discussed the algorithm. We complement our qualitative findings with quantitative analysis of application and enrollment data from 4,594 incoming kindergartners in 2017. This triangulation of methods enables us to paint a richer picture of the whole ecosystem in which the algorithm is embedded.

We found that the algorithm failed to support its intended values in practice because it’s theoretical promise depended on modeling assumptions that oversimplify and idealize how families will behave and what they seek to achieve. These assumptions overlook the complex barriers to ideal participation that many families face, particularly because of socioeconomic inequalities. Additionally, the system designers vastly underestimated the cost of information acquisition and overestimated the explainability and predictability of the algorithm. In contrast to expectations that the algorithm would ensure an transparent, equitable student assignment process, we find widespread strategic behavior, a lack of trust, and high levels of stress and frustration among families.

Student assignment algorithms promise a clear, mathematically elegant solution to what is in reality a messy, socio-political problem. Our findings show that this clash can not only prevent the algorithm from supporting stakeholders’ values, but can even cause it to work against them. Human-centered approaches may help algorithm designers build systems that are better aligned with stakeholders’ values in practice. However, algorithmic systems will never be perfect nor sufficient to address complex social and political challenges. For this reason, we must also design systems that are adaptable to complex, evolving community needs and seek alternatives where appropriate.

2 RELATED WORK

In this work we build on two major areas of related work: work in economics on designing and evaluating matching algorithms for student assignment; and literature in HCI on Value Sensitive Design. We end with a review of literature that examines the role of modeling assumptions in algorithmic systems. In this paper we use the term “algorithmic system” or “student assignment system” to broadly refer to the matching algorithm as well as the district’s processes and families’ practices that make up a part of the application and enrollment process.

2.1 Matching Algorithms for Student Assignment

Economists have developed matching algorithms to find optimal assignments between two sides of a market based on each side’s preferences [33, 77]. These algorithms have since been applied to numerous real world markets, such as university admissions and organ donation [69]. Abdulkadiroğlu and Sönmez proposed two variants of matching algorithms¹ for assigning students to public schools [2]. In these systems, each student submits a ranked list of schools that they would like to attend. Schools may have priority categories for students, such as siblings or neighborhood priorities. Students’ preferences are used in conjunction with school priorities to assign each student to an available school seat. These algorithms have promising theoretical properties that should ensure a fair and efficient allocation of seats. For example, they are strategy-proof, meaning students cannot misrepresent their preferences to guarantee an improved outcome. They also produce assignments that efficiently satisfy students’ preferences. Student assignment systems based on matching algorithms have been championed for their potential to advance equitable access to high quality education, create more diverse classrooms, and provide more flexibility to families compared to a traditional neighborhood system [50].

As these systems have been implemented in the real world they have faced new types of challenges, such as confusion for families and decreasing classroom diversity. Pathak suggests that early theoretical literature overlooked or oversimplified challenges of practical importance [66]. Economists have employed empirical methods to further understand strategic behavior [25, 26, 39, 41, 45, 49, 65, 68], information needs [17, 40, 47], and diversity constraints

¹Deferred Acceptance (DA) [33] and Top-Trading Cycles (TTC) [77] are both used for student assignment. Student-optimal DA finds the stable matching that most efficiently satisfies student preferences, while TTC finds a matching that is Pareto-efficient in the satisfaction of student preferences but is not guaranteed to be stable.

[35, 36, 42, 46, 53, 60, 64]. While these approaches improve some technical shortcomings of the systems, they do not study the values supported by the design of system itself as well as the human factors that shape how it is used in practice.

In this paper we take a human-centered approach and study parents and policy-makers to gain a deeper understanding of their values, attitudes, understandings, and uses of the student assignment system in practice. Kasman and Valant warn that student assignment algorithms are subject to strong political forces and are easily misunderstood [50]. They argue that the ultimate success of matching algorithms for student assignment will depend on how people interact with them [50]. Prior work in HCI has studied human values with respect to matching algorithms in experimental settings [55, 56]. Central concerns for participants included the algorithms’ inability to account for social context, the difficulty of quantifying their preferences, and the lack of opportunities for compromise [55]. We build on this work and study stakeholders’ values with respect to a high-stakes matching algorithm that has been in use for almost a decade to assign students to public schools. Further, we focus on why the values that these algorithms theoretically support, like transparency and equity, have not been promoted in practice.

2.2 Value Sensitive Design

Value Sensitive Design (VSD) is a theoretically grounded methodology to identify and account for stakeholders’ values in the design of new technologies [32]. In Value Sensitive Design, “values” are broadly defined as “what a person or group of people consider important in life,” although values with ethical import are considered especially important [31]. VSD is a tripartite methodology, involving conceptual, empirical and technical investigations in an iterative and integrative procedure [32]. In the conceptual stage, designers identify stakeholders’ relevant values. Empirical investigations examine stakeholders’ interactions with the technology and how they apprehend values in practice [23, 30]. Technical investigations explore how the properties and mechanisms of a particular technology support or hinder values. VSD takes a proactive stance: values should ideally be considered early on and throughout the design process [23]. However, VSD can also be applied retrospectively to evaluate deployed systems with respect to human values [32]. We apply VSD methodology to understand what values San Francisco Unified School District’s assignment algorithm was designed to support, and why it has not supported those values in practice, leading to its redesign.

Zhu et al. adapt the VSD framework to the design and analysis of algorithmic systems through “Value-Sensitive Algorithm Design” (VSAD) [82]. VSAD emphasizes the need to evaluate algorithms based on whether they are acceptable to stakeholders’ values, whether they effectively address the problem they were designed for, and whether they have had positive broader impacts [82]. This is in contrast to traditional evaluation procedures for algorithmic systems, which depend heavily on narrow, quantitative success metrics [82]. Subsequent work has applied the VSAD framework to reveal stakeholder values in the context of a machine learning algorithm used to predict the quality of editor contributions on Wikipedia [79]. The authors emphasize the need to integrate values

not only into the design of the algorithm itself, but also into the user interface and work practices that form a part of the algorithmic ecosystem [79]. This is consistent with the interactional principle in VSD, which dictates that “values are not embedded within a technology; rather, they are implicated through engagement” [23].

As VSD has been developed and more widely adopted, researchers have encountered some challenges [10, 23, 54]. One challenge is resolving value conflicts, both between stakeholders with different beliefs [29] and between competing values [78]. However, even when stakeholders agree on important values, it can be difficult to predict whether a technology that supports a value in theory will actually uphold that value when the system is deployed in the real world. Zhu et al. apply VSAD to design and evaluate an algorithm to recruit new editors to Wikipedia communities [82]. They found that their algorithm was acceptable and helpful to the community, but also discovered unanticipated shortcomings. For instance, only more experienced newcomers increased their contributions in response to the recruitment outreach [82]. Ames offers another example of values breakdown, contrasting the intended values of the One Laptop Per Child project, such as productivity, with the consumptive values that were enacted in practice [5].

Researchers have identified various causes of breakdowns between intended values and values in practice. Ames’s work highlights the importance of understanding local needs in the context where a technology is to be deployed. Manders-Huits argues that problems can arise when designers misinterpret stakeholders’ values, or because stakeholders’ values changed over time [58]. Similarly to this work, Volda et al. find that tension arises from a misalignment between how a computational system operationalizes a value and how the people who use the system understand that value [81]. We build on these findings by examining a clash between algorithmic logics and real-world goals and practices. We connect these challenges to emerging work studying the role of modeling assumptions and abstraction in algorithmic breakdown.

2.3 Modeling Assumptions in Algorithmic Systems

All algorithmic systems rely on an implicit model of the world in order to compute on it. Any model is a simplified abstraction of reality but the simplifying assumptions often go unstated [11]. For example, Selbst et al. describe the *algorithmic frame* in supervised machine learning, in which each observation in labelled training data represents an abstraction of some real-world entity, often a human being [76]. The authors warn that algorithmic systems can break down if they rely on abstractions that do not capture important aspects of the interactions between technical and social systems. Researchers have documented challenges both when assumptions are too broad, and when they are overly narrow. For instance, Chancellor et al. identified significant inconsistency in how researchers conceptualize and model humans when using machine learning to predict mental health [15]. In contrast, Saxena et al. found an overly narrow focus on risk prediction in the U.S. child welfare system that oversimplifies the complexity of the domain’s needs [73].

In the student assignment context, Hitzig identified how matching algorithms rely on an abstraction of the world that makes strong,

unstated normative assumptions regarding distributive justice [48], or the appropriate distribution of benefits and burdens in a group. The matching paradigm assumes that the ideal outcome is the one where every student is assigned to their first choice school. Hitzig points out that this emphasis on efficiency may not align with school districts’ goals, but is often framed in economics as objectively optimal rather than only one of many ways to distribute resources.

This work demonstrates how unstated, erroneous modeling assumptions about the world can break an algorithmic system. Baumer argues that this breakdown can occur when an algorithm’s designers and stakeholders do not share a common understanding of the system’s goals and limitations [8]. We expand on this work by exploring how the designers of matching algorithms for student assignment relied on certain modeling assumptions about the world in order to justify their designs with respect to values like equity and transparency. We analyze the breakdown of the student assignment algorithm in San Francisco as a case study of what happens when these assumptions clash with stakeholders’ real world goals and constraints.

3 METHODS

Our goal in this research is to understand the values that San Francisco Unified School District’s (SFUSD) student assignment system was designed to support and compare and contrast these to parents’ experiences in practice. Following Value Sensitive Design methodology [31], we begin with a conceptual investigation drawing on prior literature in economics and SFUSD policy documents to identify the values the system was intended to promote. Then, we conduct a mixed-method empirical investigation to understand why the system ultimately did not support those values and needed to be redesigned.

3.1 Data Collection

We collected data from three sources to understand the district’s policy goals (how the system was *intended* to work) and parent experiences (how the system has *actually* worked).

3.1.1 District Policies. We collected two official documents from SFUSD to understand the district’s policy goals, their justification for their original design in 2011, and the reasons they voted for a redesign in 2018. We accessed the official policy describing the existing assignment system [72] and the resolution that approved the ongoing redesign [43] from the enrollment section of SFUSD’s website.²

3.1.2 Parent Perspectives. We collected parent experiences in two primary formats: through interviews with parents, and from public online discussions on social media. The interviews allowed us to ask questions and prompt parents to reflect on and dig deeper into their experiences with the assignment system. The online discussions provide potentially less filtered reflections shared without the presence of researchers and reveal how parents seek and share information online. We supplement this data with a presentation titled “Reflections on Student Assignment” by the African American

Parents Advisory Council (AAPAC) [3], which was also downloaded from the enrollment section of SFUSD’s website.

We conducted semi-structured interviews with 13 parents who have used the student assignment system to apply for elementary schools in SFUSD. We recruited parents through four parenting email and Facebook groups by contacting group administrators who shared a brief recruitment survey on our behalf. During the interview, we asked participants to describe their application and enrollment experiences, and to reflect on their understanding of the assignment algorithm. Interviews were 45 minutes and participants received a \$30 gift card. All interviews were conducted over the phone in English between February and August 2020.

12 parents completed a demographic survey. Parents reported their income as low income (1), middle income (5), and upper-middle to high income (4) and identified their race or ethnicity as white (4), Asian (3), Chinese (2), white and Hispanic (1), white and Middle Eastern (1), and Vietnamese (1). The 12 respondents reside in six different zip codes in the city. In all 12 households one or more parents had a Bachelor’s degree and in nine households the highest level of education was a graduate degree. To preserve participant privacy, we identify participants in this paper by unique identifiers P1 through P13.

We supplement the interview data with twelve Reddit threads posted on the *r/sanfrancisco* subreddit³ between 2016 and 2020. These threads were selected by conducting a comprehensive search of *r/sanfrancisco* using the search term “school lottery,” as it is commonly known to parents.⁴ Each post was reviewed to ensure that it was a discussion of the current SFUSD assignment algorithm. From the twelve threads made up of 678 posts and comments, we manually coded content where the author demonstrated first-hand experience with the assignment algorithm, resulting in a final dataset of 128 posts from 83 contributors. Excluded posts were those that were off topic or presented the author’s political view rather than their personal experiences with the system. We paraphrase this content to protect the users’ privacy.

3.1.3 Application and Enrollment Data. We complement our qualitative data about parent experiences with publicly available, de-identified kindergarten application data from 2017 to understand higher-level trends in how parents use the system.⁵ For each of the 4,594 applicants, the data includes their ranked list of schools, the school they were assigned to, and the school they enrolled in. It also includes the student’s zipcode, race, and whether the student resides in a census tract with the lowest performing schools (CTIP1 area), which makes them eligible for priority at their preferred schools. Applicants are 28% Asian or Pacific Islander, 24% white, 23% Hispanic and 3.2% Black. 21% declined to state their race. Approximately 15% of applicants were eligible for CTIP1 priority, 45% of whom are Hispanic. 11% of CTIP1-eligible students are Black, which is 53% of all Black applicants.

3.1.4 Limitations. We recruited interview participants through convenience sampling online and complemented the interviews

³<https://reddit.com/r/sanfrancisco>

⁴Search conducted using the PushShift Reddit repository at <https://redditscraper.io/>

⁵The data was collected as part of a public records request by local journalist Pickoff-White for a story about how parents try to game the system [67]. The data is available at <https://github.com/pickoffwhite/San-Francisco-Kindergarten-Lottery>.

²<https://www.sfusd.edu/schools/enroll/ad-hoc-committee>. Accessed April, 2020.

with existing online data, which biases our data towards those who have the time and motivation to participate in research studies, online discussions, and district focus groups. Our dataset lacks sufficient representation of low-income families and Black and Hispanic families. It is important that future work addresses this limitation, particularly considering that integration is a key goal for the school district, and that these families are underrepresented in existing discourses. In future work we will focus on understanding the experiences of historically underserved families with student assignment algorithms, specifically families of color, low-income families, and families with low English proficiency.

3.2 Data Analysis

In order to understand the district’s values for student assignment and the reasons why the assignment algorithm has not supported these values, we conduct inductive, qualitative content analysis [59] and quantitative data analysis.

3.2.1 Qualitative Analysis. Our qualitative dataset was made up of district policy documents and community input, interview transcripts, and Reddit content. We performed an open-ended inductive analysis, drawing on elements of grounded theory method [16]. We began with two separate analyses: one to understand the district’s values and policies; and a second to understand parent experiences and perspectives. The authors met regularly throughout the analysis to discuss codes and emerging themes. In both analyses we began by conducting open coding on a line-by-line basis using separate code books [16]. We then conducted axial coding to identify relationships between codes and higher level themes. In the axial coding stage for the SFUSD policy documents, we identified three high level codes relevant to our research questions: Values: What are the district’s values and goals for student assignment?; Mechanism: How was the district’s current system expected to support their values?; and Challenges: Why did the district ultimately decide to redesign the system?. Next, we analyzed parent perspectives from the community input documents, interview transcripts, and Reddit content. We conducted two rounds of open coding. First, we focused only on these three data sources. We identified codes that included "priorities," "algorithmic theories," and "challenges." Then, we linked the open codes from the first round to the challenges identified in the policy documents. We found that challenges parents described in our parent perspectives dataset were relatively consistent with those described in the policy documents and we reached theoretical saturation after approximately ten interviews.

3.2.2 Quantitative Analysis. We linked the application dataset to publicly available school-level standardized test results in order to understand how families use the system to access educational opportunities. We accessed third grade results in the California Smarter Balanced Summative Assessments in 2017-2018, provided by the California Department of Education.⁶ We conducted exploratory data visualization to investigate trends in preferences.

⁶Data available at [urlhttps://caaspp-elpac.cde.ca.gov/caaspp/ResearchFileList](https://caaspp-elpac.cde.ca.gov/caaspp/ResearchFileList). We link the school achievement data to the applications by state-level (CDS) code. The preference data contains only school numbers, a district-level coding scheme. SFUSD has published a document linking these district school numbers to the school name and state-level (CDS) codes [http://web.sfusd.edu/Services/research_public/rpadc_lib/SFUSD%20CDS%20Codes%20SchYr2012-13_\(08-20-12\).pdf](http://web.sfusd.edu/Services/research_public/rpadc_lib/SFUSD%20CDS%20Codes%20SchYr2012-13_(08-20-12).pdf).

We measure variation in preferences by race and CTIP1 priority status in order to gain insight into if and how participation varies across groups differently impacted by structural oppression and historical exclusion from high quality education. We present quantitative findings using visualizations to include all students. When comparing summary statistics we use the bootstrap⁷ method to estimate statistical significance [27]. For this analysis we used third grade standardized test results as a rough estimate of resources and opportunities at each elementary school. We recognize that there are many ways in which schools provide value to children that are not reflected in standardized test results.

4 STUDENT ASSIGNMENT IN SAN FRANCISCO: INTENDED VALUES

In this section, we present our findings on the values that San Francisco Unified School District (SFUSD) intended their student assignment system to support. In the next section we analyze why this system did not realize those values in practice.

SFUSD has been utilizing different choice-based systems to address educational inequality in the district for almost forty years [70]. Although the mechanism for assigning students to schools has changed significantly over time, SFUSD has been consistent in their values and goals for student assignment. Their current policy designates three primary goals:

- (1) "Reverse the trend of racial isolation and the concentration of underserved students in the same school;
- (2) Provide equitable access to the range of opportunities offered to students; and
- (3) Provide transparency at every stage of the process." [72]

In addition, they emphasize the importance of efficiently utilizing limited district resources, ensuring predictability and ease of use for families, and creating robust enrollments at all schools.

In SFUSD’s current assignment system [70], students or their parents apply for schools by submitting their *preferences*: a ranked list of schools they would like to attend (Figure 2). To increase flexibility and access to opportunities, students can rank any school in the district and there is no limit on the number of schools they can rank. The district also defines priority categories. Elementary schools give top priority to siblings of continuing students and then to underserved students. Underserved students are defined as those living in neighborhoods with the schools that have the lowest performance on standardized tests, known as *CTIP1* areas. The matching algorithm⁸ then takes student preferences and school priorities and produces the best possible assignments for the students subject to the schools’ priorities and capacity constraints. Importantly, the resulting assignments from this algorithm are guaranteed to efficiently satisfy *student preferences* not school priorities. School priorities are only used to determine which students are assigned to over-demanded seats. The matching algorithm is also strategy-proof, meaning that it can be theoretically proven

⁷We use percentile intervals to estimate confidence intervals and the bootstrapped t-test to estimate p-values for differences in means using 10,000 re-samples, following [27]. Groups (race and CTIP1) are re-sampled independently.

⁸SFUSD uses a variant of the Top Trading Cycles algorithm [77]. See [2] for a technical analysis of Top Trading Cycles in the student assignment context or [69] for a more broadly accessible introduction to market design.

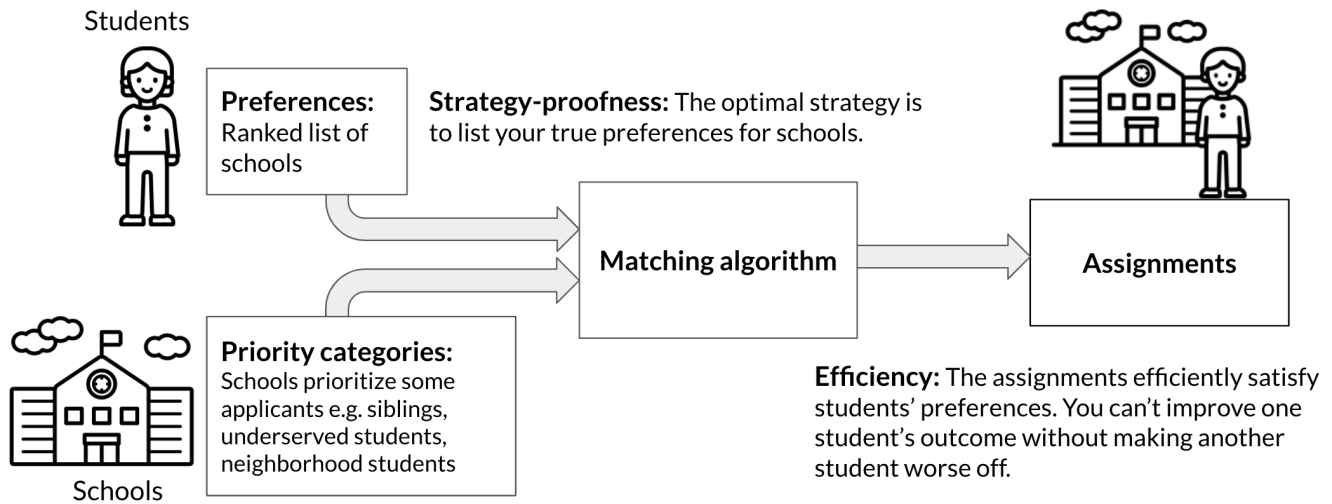


Figure 2: The matching algorithm takes students’ preferences over schools and schools’ pre-defined priority categories as inputs and outputs the most efficient assignment of students to schools.

that families do not benefit from manipulating their preferences to game the system.

We consolidated the school district’s stated goals for student assignment into four high-level values: (1) transparency, predictability and simplicity; (2) equity and diversity; (3) quality schools; and (4) community and continuity (Table 1). In this section, we described the system that was expected to support these values. In the next section, we explore why these expectations were not met in practice.

5 ALGORITHMIC BREAKDOWN: VALUES IN PRACTICE

In December 2018, San Francisco Board of Education determined that the the algorithm was not working as intended [43]. While the number one stated goal of the algorithm was to “reverse the trend of racial isolation and the concentration of underserved students in the same school,” the Board found that segregation had *increased* since the algorithm was introduced and there was widespread dissatisfaction amongst parents [71, 72]. The assignment algorithm had failed to respect the values that it was designed to support and the Board voted to stop using it and to design a new system. In this section we present our findings that help explain why.

For each of the district’s four high-level values for student assignment (Table 1), we first review the theoretical properties and promises of the algorithm related to that value: why would economists and district policy-makers expect that the system would respect that value? Next, we analyze what implicit modeling assumptions those expectations depend on. Finally, we explain how families’ needs, constraints, and values in the real world clashed with system designers’ assumptions about them, which prevented the algorithm from meeting its theoretical promises and enacting the district’s values in practice.⁹

⁹In this work we identify the school district’s values and draw on families’ experiences to explain why they haven’t been supported. The district’s values may not completely

5.1 Transparency, Predictability, and Simplicity

5.1.1 Theoretical promises. Matching algorithms are clearly and explicitly defined procedures. This differentiates them from assignment systems based on imprecise admissions criteria, which have historically been more difficult to justify and have led to legal disputes [2]. If a student wants to understand why they did not receive an assignment they were hoping for, the algorithm’s decision can be explained. Matching algorithms are also provably strategy-proof. That is, students cannot guarantee a more preferable assignment by strategically misrepresenting their preferences. Strategic behavior requires time and effort, so preventing strategic advantages is critical not only for simplicity and efficiency, but also for ensuring that all families can participate equally.

5.1.2 Modeling assumptions: families will accept their assignment as fair and legitimate as long as the algorithm’s logic is explained to them. This assumes that the school district provides an accessible, comprehensible explanation and that families would seek out, understand, and trust this explanation. Families have known preferences for schools and recognize that they should report those preferences truthfully.

5.1.3 Real world challenges. In practice, families find the assignment system difficult to navigate and struggle to find relevant, clear, and consistent information. Some parents engage in strategic behavior to try to improve their child’s assignment, contrary to theoretical incentives. Rather than seeking and accepting an explanation, families who are dissatisfied with their assignment seek to change it. Families’ trust in the system is eroded by the lack of clear

align with families’ values. We assume that satisfying families is one of the district’s priorities, and we find substantial overlap between the four district values and what parents in our sample find important. We leave a detailed analysis of families’ values to future work.

Table 1: We consolidated the San Francisco Unified School District’s goals for student assignment into four overarching values. Assignment algorithms have theoretical properties aligned with these values. However, the San Francisco assignment algorithm’s theoretical promises have not been realized because they rely on modeling assumptions that clash with real world challenges.

<i>Value</i>	<i>Promises and Properties</i>	<i>Modeling Assumptions</i>	<i>Real World Challenges</i>
Transparency, predictability, and simplicity	Algorithm has a clearly defined procedure. Assignments are explainable.	The district provides accessible, clear information. Families want and understand explanations. Families do not try to game the system.	Finding and understanding information is difficult. Some parents try to game the system. There is a lack of trust: assignments are perceived as unpredictable and unfair.
Equity and diversity	Any student can apply to any school. Underserved students are given priority access.	All families participate equally and the all-choice system offers identical opportunities to all families.	Time, language and economic constraints create participation barriers for lower resourced families.
Quality schools	Competition for applicants will drive up the overall quality of schools in the district.	Families base their preferences on accurate estimates of school quality. Schools can respond to competitive pressures.	Competition is driven by social signalling and negative stereotypes. Underserved schools lack resources to attract applicants.
Community and continuity	Priority for siblings and students in the school’s attendance area.	Schools have sufficient capacity to handle demand from siblings and neighborhood children.	A lack of guaranteed access to local schools frustrates families living in neighborhoods with very popular schools.

information and the belief that some parents are able to game the system.

Parents face a significant information cost to understand the various opportunities available across the city. There are 72 elementary school programs in SFUSD [43]. Parents indicated that researching schools is a burdensome time-commitment. In-person school visits are a popular source of information when forming preferences, but these visits are time-intensive and logistically difficult.

[...]It’s like a full time job doing all the school tours.
(P7)

Online information is another widely used source, but school information is not centralized, nor is it consistent across schools. A number of parents mentioned the difficulty of navigating online district resources:

[...]Finding and gathering the information about the schools from the district is a mess. (P11)

None of the parents we interviewed felt that they had a clear understanding of how the algorithm works. The algorithm is colloquially known to parents as “the lottery.” Although the algorithm has only a small lottery aspect to break ties between students with the same priority, many believe it is mostly or entirely random.

I’m not really that confident in their actual lottery system. It could be bingo in the background for all I know.
(P4)

This leaves families feeling a lack of agency and control over their child’s education.

I mean, the word itself, lottery, most of it is random. I don’t feel like we can do anything at all. (P5)

Confused and frustrated by district resources, parents frequently seek advice from other parents online and in-person. Reddit users sought and shared complex strategies, sometimes relying on substantial independent research. This is consistent with prior work showing that advice sharing in social networks can encourage strategic behavior [25, 26]. Advice from other families is often conflicting and unclear, further exacerbating confusion about the system.

[W]e also got different advice from different parents. They’re very, very different from each other. Some people say, “Put in as many schools as possible,” and some people say, “No, just put two schools that you really wanted, and then you have a higher chance of getting those.” (P5)

The 2017 application data indicates that strategic behavior may be more widespread amongst more privileged families. On average, families who were eligible for the CTIP1 priority for underserved students ranked 5.5 schools in their application (95% confidence interval (CI): 5.0–6.2 schools), while families in other areas of the city ranked an average of 11.6 (95% CI: 11.2–12.1 schools; difference in means: $p = 0.00$) (Figure 3). 96% of families eligible for CTIP1 priority were assigned their first choice, so this difference may reflect these families’ confidence that they will get one of their top choices. On the other hand, it may reflect disparities in access to the time and resources needed to research schools and strategies. White students submitted especially long preference lists (mean = 16.5;

95% CI: 15.6–17.6),¹⁰ a further indication that strategic behavior is more popular with families with more structural advantages.

Receiving an unfavorable assignment was a major concern for families in our sample. The district offers multiple rounds of the assignment algorithm, which many parents participate in if they are dissatisfied with their child’s assignment. However, this process can be long, uncertain, and frustrating. Some parents received the same assignment every round with no further explanation or assistance.

[... T]he first announcement that we got [...], I actually wasn’t that upset. I said, “You know what, there’s more rounds. [...] We could stick it out.” But I was really upset at the second one because there was literally no change. And that really had me questioning, “I’m just trying to play by the rules. Should I not trust this any more than it’s going to work out?” (P9)

Parents on Reddit recommended unofficial avenues for recourse, many of which require substantial time and resources. These include going in person to the enrollment office repeatedly to request reassignment, remaining on waiting lists up to ten days into the school year, and opting out of the public school system altogether.

Overall, a complicated algorithm together with a shortage of transparent and accessible information has fostered distrust and frustration amongst parents in the district. Distrust is fuelled by perceptions that the system is random and unscientific, and that it allows parents with more time and resources to gain an unfair advantage.

It’s definitely convoluted. It’s definitely multilayered, it’s complex. And that favors people who have the time and the wherewithal to figure it out. [...] The complexity invites accusations of [corruption] and does not inspire trust (P9)

5.2 Diversity and Equity

5.2.1 Theoretical promises. The assignment system is an all-choice system with unrestricted preference lists, so any student can apply to any school in the district. Compared to a neighborhood system, or even more restricted choice systems, this design has the potential to enable more equitable access to educational opportunity. In an effort to promote equitable access to education and diverse schools, SFUSD has added the CTIP1 priority category, which gives priority admission at over-demanded schools to students from neighborhoods with under-performing schools.

5.2.2 Modeling assumptions: all families participate equally in the system and the all-choice system offers identical opportunities to all families. CTIP1 students prefer to attend over-demanded schools if they can access them. Applicant pools reflect the racial and socioeconomic diversity of the city.

5.2.3 Real world challenges. Although an all-choice system offers greater flexibility than a neighborhood system, our results show that families with fewer resources face significant barriers to ideal participation in SFUSD’s choice system. Although families can rank any school on their application, some families are not able to choose

the schools that offer the greatest opportunities. Preferences are segregated by race and income, preventing the algorithm from creating diverse assignments.

Our results indicate that the all-choice system does not offer identical opportunities to all families. Every family can apply to any school, but that does not mean that every family can actually access every school. For example, transportation logistics can be a significant challenge. When choosing a kindergarten for their child, P1 met with an education placement counselor at SFUSD to understand the special education services offered across the district. P1 recalled their response to one of the counselor’s suggestions:

So, you are telling me this school is [...] three blocks uphill and we’re supposed to do that with a kindergartner and no car? [...] There’s no way that on my worst day that I would be able to drag my kindergartner with special needs uphill in the rain. (P1)

The CTIP1 priority is potentially a useful advantage for underserved students. In 2017, 96% of students who were eligible for this priority were assigned their first choice school, compared to 58% of students without this priority. However, CTIP1 priority is only useful for advancing educational equity if these students can actually use it to enroll in well-resourced schools. In 2017, students with CTIP1 priority enrolled in schools with lower academic outcomes than other students (Figure 4). On average, underserved students enrolled in a school where 45.0% of third graders met or exceeded expectations in the English Language Arts/Literacy exams¹¹ (95% CI: 43.2% – 46.7%), compared to 57.2% (95% CI: 56.5% – 57.9%) of students at the average school that other students enrolled in (difference in means: $p = 0.00$). This difference points to persisting inequities in access to higher resource schools that priority assignment is insufficient to address. CTIP1 priority cannot, for example, help students access schools that are physically inaccessible for them. Social factors may also influence choice patterns. For instance, the African American Parent Advisory Council (AAPAC) has raised concerns that Black students in San Francisco continue to face racism and biases in racially diverse classrooms [3].

These findings are consistent with prior work showing that while proximity and academics are important to most families, more privileged parents tend to put more emphasis on a school’s academic performance [1, 14, 46], while parents from low-income or racialized backgrounds may be more likely to prioritize proximity [53] or representation of students from a similar background [46]. As a result of differences in students’ preferences, applicant pools at schools across the city are segregated by race and income. This prevents the algorithm from creating diverse assignments [43, 53].

5.3 Quality Schools

5.3.1 Theoretical promises. System designers have suggested that choice systems indirectly improve school quality. For instance, Pathak argues that matching mechanisms create competition between schools, which pushes under-demanded schools to improve in order to attract applicants and sustain their enrollment [66]. In addition, Pathak points out that an algorithmic system based on student preferences creates a useful source of demand data for the

¹⁰Differences in means between white students and Black, Asian or Pacific Islander, and Hispanic students is highly statistically significant, even with conservative adjustments for multiple hypothesis testing.

¹¹Qualitatively similar results to those presented in this section hold for Mathematics results.

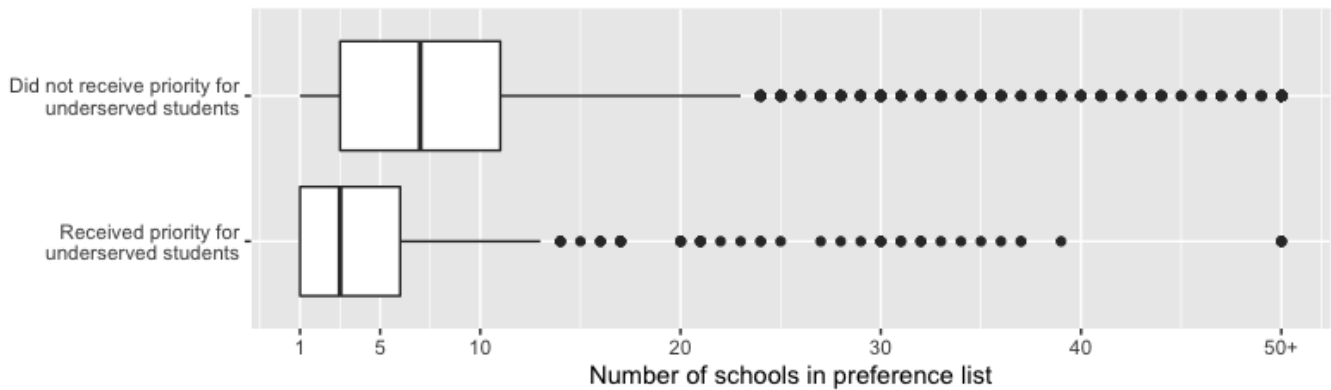


Figure 3: Families who were eligible for priority for underserved students ranked fewer schools on average (mean: 5.5; 95% CI: 5.0-6.2) than other families in the city (mean: 11.6; 95% CI: 11.2-12.1; difference in means: $p=0.00$). This may suggest that strategic behavior is more widespread amongst higher resource families.

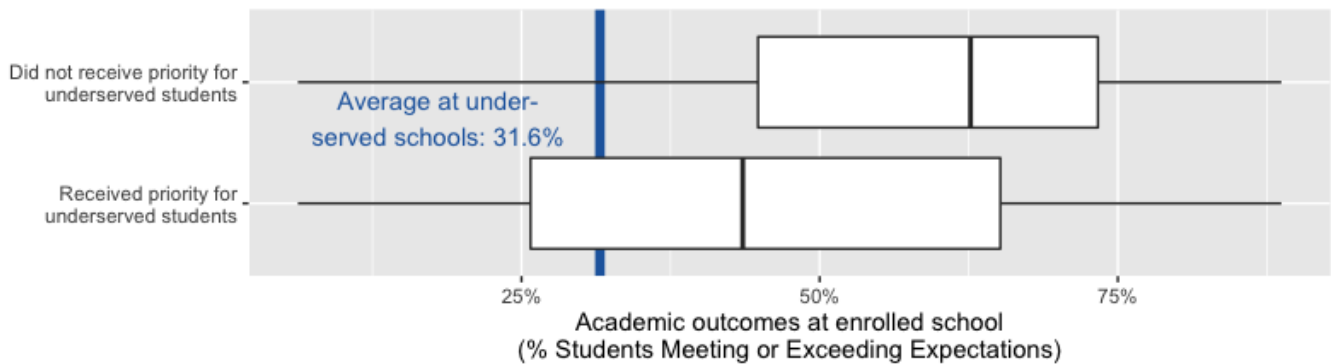


Figure 4: The priority for underserved students helps those students access educational opportunity, but there remain inequities that priority enrollment cannot address. Students with priority enrolled in higher performing schools (mean: 45.0% of students met or exceeded expectations on standardized tests; 95% CI: 43.2% – 46.7%), than their average neighborhood school (mean: 31.6%). However, they still enrolled in lower performing schools on average than students who were not eligible for priority (mean: 57.2%; 95% CI: 56.5%–57.9%) (difference in means: $p = 0.00$). Academic outcomes are measured as the percentage of third grade students at the enrolled school who met or exceeded expectations in the 2017-18 statewide assessments.

district to target interventions or closures at underenrolled schools [66].

5.3.2 Modeling assumptions: a competitive market will drive up the overall quality of offerings. This assumes that demand is driven by accurate estimates of school quality.

5.3.3 Real world challenges. Unfortunately, competition in SFUSD has not resulted in an improvement in educational opportunities and outcomes across the district [43]. Our findings reveal that parents base their preferences on noisy signals of school quality. Still, some students depend on under-demanded schools and are harmed by under-enrollment and school closures.

Our results suggest that parents' preferences are strongly shaped by social learning and stereotypes. Many parents reported using other parents' opinions and experiences of schools to inform their preferences. Some feel that a few schools are disproportionately

regarded as the “best” schools in the city. Parents on Reddit attested that many good schools are unfairly dismissed by more advantaged parents, sometimes on the basis of thinly veiled racist and classist stereotypes. Standardized test scores or aggregate scores like those reported by *greatschools.org* are another popular source of information. Though seemingly more objective, these measure are heavily correlated with resources and demographics at schools [6], further exacerbating preference segregation. In the presence of these types of competitive pressures, well-resourced schools are heavily over-demanded while under-resourced schools struggle to maintain robust enrollments [43]. SFUSD believes the algorithm has created “unhealthy competition” between schools, resulting in schools ranging in size from 100 to nearly 700 students [71].

While Pathak argues that choice patterns are useful in determining which schools to close and which to support and expand [66],

this overlooks the correlation between demand patterns and existing patterns of inequality. Under-enrollment and school closures can seriously harm the communities at those schools, which often serve predominantly poor students of color [28, 37]. SFUSD has acknowledged the need to more equitably distribute resources, but it can be politically difficult to direct resources to schools with low demand and enrollment [72].

5.4 Community and Continuity

5.4.1 Theoretical promises. SFUSD’s sibling and attendance area priority categories are designed to encourage a sense of community and cohesion for families. In addition, students attending PreK or Transitional Kindergarten in the attendance area are given priority to ensure continuity for students.

5.4.2 Modeling assumptions: schools have sufficient capacity to handle demand from siblings and neighborhood children.

5.4.3 Real world challenges. Many families are dissatisfied by a lack of access to their local schools. In many neighborhoods there is a mismatch between demand for the attendance area school and its capacity. In fact, current attendance area boundaries are drawn such that some schools do not have the capacity to serve every student in the attendance area [70]. As a result, the attendance area priority does not provide an acceptable level of predictability for those who want to enroll in their local school.

For parents living in neighborhoods with popular schools, access to their attendance area school is far from guaranteed. One Reddit user expressed frustration after they found out that they may not be able to enroll their child in their local school. Due to their family’s circumstances, they feared it would be impossible to get their child to a school further from home.

Parents in our sample value access to local schools for convenience and a sense of community. Under the existing system, two children who live close to each other may attend schools on opposite sides of the city. There are even neighborhoods in San Francisco where students are enrolled across all 72 elementary school programs [43]. Some parents felt that this dispersion undermines the educational experience for children:

[I]t is really important for our children to bond and build relationships in their community. And they really connect to their education and their educational environment very differently [when they do]. (P1)

By underestimating the mismatch between demand for neighborhood schools and capacity at those schools, the assignment system has generated significant dissatisfaction among parents who live near popular schools. These parents are increasingly pushing for a return to a traditional neighborhood system. However, this would restrict flexibility and access to educational opportunities for many families across the city who use the system to enroll their children in schools other than their neighborhood school.¹²

¹² A district analysis showed that 54% of kindergarten applicants did not list their attendance area school anywhere in their preference list for the 2013-14 school year [70]. This is especially true of underserved students: according to the 2017 application data, around 75% of students who received CTIP1 priority enrolled in an elementary school outside of the CTIP1 census tracts. Schools in CTIP1 census tracts were determined according to the definition updated for the 2014-15 school year https://archive.sfsd.edu/en/assets/sfsd-staff/enroll/files/Revising_CTIP1_for_2014_15_SY.pdf.

6 DESIGN IMPLICATIONS FOR STUDENT ASSIGNMENT

In the previous section we showed how incorrect or oversimplified modeling assumptions have played a role in the breakdown of the student assignment algorithm in San Francisco. In this section we draw on these findings to present four design implications for student assignment systems: (1) provide relevant and accessible information; (2) (re)align algorithmic objectives with community goals in mind; (3) reconsider how stakeholders express their needs and constraints; and (4) make appropriate, reliable avenues for recourse available. We emphasize that student assignment is a complex, socio-political problem and our results and recommendations are our first step to better understanding it. In the future, we will continue this work focusing explicitly on the needs of underserved students. In the next section we discuss broader implications of this work for the design of algorithmic systems.

6.1 Provide relevant and accessible information

When looking for a school for their child, parents need to find schools that meet their needs, and then understand how to apply. Our research shows that information acquisition is very difficult, which leaves families with a sense of distrust and perceptions of randomness, unpredictability, and unfairness. However, more information is not always better. Information about algorithmic systems should be congruent with stakeholder needs and interests and should be limited to the most relevant information in order to minimize cognitive load [24]. In the student assignment setting, we found the most salient information for families is information about the schools available to them that best meet their needs. Relevant, accurate information should be easy to find and navigate. San Francisco Unified School District has recognized this need and has committed to making this information available in a variety of languages [43]. Further work is needed to understand what kind of information about schools will be relevant and helpful without exacerbating negative stereotyping and preference segregation.

Transparency information about the algorithm itself may also reduce stress and increase trust in the system, but only if this information is clear and useful [18, 52, 61]. The algorithmic information most relevant to parents in our sample is their chances of receiving a particular assignment. This information is currently difficult to find, in part because these probabilities depend on others’ preferences. However, this information may reduce stress and increase predictability. One concrete goal moving forward could be to ensure that information about schools and admission probabilities are easily available.

6.2 (Re)Align algorithmic objectives with community goals in mind

SFUSD expected their assignment system to satisfy individual preferences *and* promote community-level goals like equitable access to education and diverse classrooms. However, the system has had limited success in promoting educational equity, and racial and economic segregation has worsened since it was introduced [43]. One

reason for this breakdown is that the primary objective of matching algorithms is to efficiently satisfy students’ preferences, and in San Francisco students’ preferences are already heavily segregated by race and income [43]. This indicates a breakdown between community goals and what the algorithm is optimizing for.

The focus on satisfying students’ preferences can also obscure other problems. For example, if we look only at preference satisfaction, then underserved students appear to have a strong advantage in the current system. 96% of incoming kindergartners who were eligible for priority for underserved students received their first choice school in 2017, compared to only 58% of other students. However, underserved students continue to enroll in lower resourced schools and an opportunity gap persists between underserved students and others in the district. Due to the limitations of our sample, we cannot conclusively explain the reasons for segregated and unequal preferences. Nevertheless, these two challenges suggest that technical system designers need to work closely with policy-makers and community members to ensure that their algorithm’s objectives and evaluation metrics are aligned with higher-level goals and values.

6.3 Reconsider how stakeholders express their needs and constraints

Another way to make progress towards community goals is to reconsider how families express their values, needs, and constraints. Matching algorithms model families as independent, self-interested agents with some inherent preferences over schools. Schools are assumed to be merely “objects to be ‘consumed’ by the students” [2]. However, our findings highlight that preferences are based on limited information and are strongly shaped by social context. Schools are also important communities for children and their families. Researchers have found that matching algorithms for group decision-making do not give participants the space to understand each others’ concerns and arrive at compromises that might be natural in a negotiation amongst humans [55, 57]. One avenue for future work is to develop alternative methods for eliciting students’ preferences that better reflect their needs and allow for compromise and community building. For example, families could submit their weighted priorities over factors like language programs or proximity to their home. In our interviews we found that parents already make these types of comparisons frequently when researching schools. Such an approach might help shift families’ focus from how high their assigned school was in their personal ranked list to how their assigned school meets their needs and constraints and contributes to progress towards community-level goals.

6.4 Make appropriate, reliable avenues for recourse available

Because there is limited space at popular schools, some students will receive a disappointing assignment. There are multiple rounds of the algorithm for students who wish to appeal their assignment. However, our results suggest that this process can be frustrating and unpredictable. One concrete recommendation is to improve communication with parents throughout the process about their application status and their available next steps. Our findings also suggest that privileged stakeholders will continue to seek unofficial channels to achieve their goals. Therefore, future work developing

fair processes for recourse should prioritize the needs of lower resourced stakeholders and design low cost appeals processes.

7 DISCUSSION AND FUTURE WORK

In the previous section, we suggested ways to improve student assignment algorithms to better support stakeholders’ values. In this section, we discuss the implications of our work for algorithm design more broadly and identify opportunities for future work.

This work presents an example of how incorrect assumptions can prevent a system from supporting intended values in practice. Direct engagement with stakeholders early on in the design process may help system designers identify incorrect or oversimplified modeling assumptions. For example, economists initially assumed that matching algorithms would be easy to explain to families and that the procedure would be perceived as fair. A value sensitive approach would have encouraged designers to engage with stakeholders early in the development process to gauge their perceptions and acceptance of the technology [82]. Economists may have discovered that stakeholders’ acceptance of matching algorithms for student assignment would depend heavily on social and political factors, such as pre-existing institutional trust in the school district.

Even with improved methods to align algorithm design with stakeholders’ values, unanticipated challenges will arise because algorithmic systems must rely on *some* abstractions and assumptions that will always be an imperfect approximation of the real world [11, 76]. Crawford analyzed sites of conflict between algorithms and humans, and has warned of the danger of understanding algorithmic logics as autocratic [22]. Instead, algorithmic systems should be accountable to community values beyond the formal design process and stakeholders should have ongoing opportunities to voice concerns, even after the system has been deployed [82]. Future work is needed to design algorithmic systems that are adaptable and flexible in response to this feedback.

In advocating for ongoing engagement with stakeholders, it is important to grapple with differences in power and participation among them [24, 82]. We need to design mechanisms for participation that are equitable and low-cost for lower resource families to voice their concerns [44]. In the student assignment setting, we found that convenience sampling strongly skewed our sample of parents towards higher resource parents with the time and motivation to voice their concerns. While building a system that serves all stakeholders is ideal, trade-offs are inevitable when systems impact a large number of stakeholders with diverse perspectives and needs [24, 82]. Avenues for participation should encourage deliberation of trade-offs and include safeguards to prevent powerful stakeholders from compromising important community values in order to design a system that better serves their own interests.

Designing systems while taking into account stakeholders with conflicting values and priorities will require a broader view of algorithmic performance. The research literature on matching algorithms has typically emphasized theoretical guarantees, such as whether assignments are efficient or stable. A human-centered analysis of algorithmic performance would involve evaluating the system in its real world context, along dimensions such as acceptance from stakeholders and broader impacts [82]. This is in contrast

to typical practices in algorithmic fields such as machine learning, where algorithms are developed and evaluated with respect to narrow, quantitative metrics such as efficiency. A broader view of algorithmic performance may identify challenges that are central to stakeholders' experiences with the system if not directly related to the algorithm's design, such as the difficulty of forming a preference list.

Finally, we cannot expect that every algorithmic system can support community values if only the right design choices are made. Demand for a technology in the first place is often closely tied to particular politics, which may necessitate certain values and preclude others. For example, education researcher, Scott argues that modern school choice programs reflect a neoliberal ideology focused on empowering parents as consumers of educational opportunities for their child [74]. Advocates claim that school choice promotes educational equity by enabling underserved students to attend a school other than their neighborhood school. Assignment algorithms can support this approach to equity with technical features like priority categories or quota systems. However, this is not the only approach to educational equity. In fact, it offers limited benefits to those who do not have the time or resources to exercise informed choice [75]. A redistributive principle, on the other hand, would prioritize providing underserved students with educational opportunities in their own communities and protecting local students' access to those resources [3]. Assignment algorithms cannot effectively support such an approach: increasing enrollment at under-demanded schools using an algorithm would require violating some students' preferences and may be disruptive and harmful to the existing communities at those schools [3, 38]. Therefore, student assignment algorithms exist within and to uphold a political ideology that privileges individual choice sometimes at the cost of other values, such as democracy, resource equality, and desegregation [75]. This example shows why it is important not only to consider how certain design choices *might* support the values that stakeholders find salient, but also what values a technology *necessitates or precludes* based on the implicit politics of its existence. Value Sensitive Design does not provide an explicit ethical theory to designate what kinds of values *should* be supported [10, 58]. Therefore, in addition to an understanding of implicit values and politics, our analysis must include a commitment to justice [21] and accept refusal as a legitimate way of engaging with technology [20].

8 CONCLUSION

In this paper we conduct qualitative content analysis of parent experiences and district policies, and quantitative analysis of elementary school applications to understand why the student assignment system in place in San Francisco Unified School District has not supported the district's goals and values. We identify four values that the system was intended to support: (1) transparency, predictability and simplicity; (2) equity and diversity; (3) quality schools; and (4) community and continuity. We identify how the algorithm's theoretical promises to uphold these values depend on assumptions about how stakeholders behave and interact with the system, and explore the ways in which these assumptions clash with the properties and constraints of the real world. We discuss

the implications of this work for algorithm design that accounts for complex and possibly conflicting values and needs.

ACKNOWLEDGMENTS

We thank our study participants for sharing their experiences and insights. We also thank members of the U.C. Berkeley Algorithmic Fairness and Opacity Working Group (AFOG) and participants at the 4th Workshop on Mechanism Design for Social Good (MD4SG) for helpful feedback on an earlier version of this work. Finally, we thank the anonymous reviewers for their feedback and suggestions.

REFERENCES

- [1] Atila Abdulkadiroglu, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters. 2017. *Do Parents Value School Effectiveness?* Working Paper 23912. National Bureau of Economic Research. <https://doi.org/10.3386/w23912>
- [2] Atila Abdulkadiroglu 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>
- [3] African American Parent Advisory Council. 2017. *AAPAC Reflections on SFUSD's Student Assignment Policy*. San Francisco Unified School District. Retrieved April 17, 2020 from 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
- [4] 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>
- [5] Morgan G. Ames. 2016. Learning consumption: Media, literacy, and the legacy of One Laptop per Child. *The Information Society* 32, 2 (2016), 85–97. <https://doi.org/10.1080/01972243.2016.1130497>
- [6] Matt Barnum and Gabrielle LaMarr LeMee. 2019. *Looking for a home? You've seen GreatSchools ratings. Here's how they nudge families toward schools with fewer black and Hispanic students*. Chalkbeat. <https://www.chalkbeat.org/2019/12/5/21121858/looking-for-a-home-you-ve-seen-greatschools-ratings-here-s-how-they-nudge-families-toward-schools-wi>
- [7] Anna Maria Barry-Jester, Ben Casselman, and Dana Goldstein. 2015. *The New Science of Sentencing*. The Marshall Project. <https://www.themarshallproject.org/2015/08/04/the-new-science-of-sentencing>
- [8] Eric PS Baumer. 2017. Toward human-centered algorithm design. *Big Data & Society* 4, 2 (2017), 1–12. <https://doi.org/10.1177/2053951717718854>
- [9] Tolga Bolukbasi, Kai-Wei Chang, James Zou, Venkatesh Saligrama, and Adam Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Proceedings of the 30th International Conference on Neural Information Processing Systems (Barcelona, Spain) (NIPS'16)*. Curran Associates Inc., Red Hook, NY, USA, 4356–4364.
- [10] Alan Borning and Michael Muller. 2012. Next Steps for Value Sensitive Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Austin, Texas, USA) (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1125–1134. <https://doi.org/10.1145/2207676.2208560>
- [11] G. E. P. Box. 1979. Robustness in the Strategy of Scientific Model Building. In *Robustness in Statistics*, Robert L. Laundser and Graham N. Wilkinson (Eds.). Academic Press, Cambridge, MA, USA, 201 – 236. <https://doi.org/10.1016/B978-0-12-438150-6.50018-2>
- [12] 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>
- [13] 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)*, Sorelle A. Friedler and Christo Wilson (Eds.), Vol. 81. PMLR, New York, NY, USA, 77–91. <http://proceedings.mlr.press/v81/buolamwini18a.html>
- [14] 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/eoj.12153>
- [15] Stevie Chancellor, Eric P. S. Baumer, and Munmun De Choudhury. 2019. Who is the "Human" in Human-Centered Machine Learning: The Case of Predicting Mental Health from Social Media. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 147 (Nov. 2019), 32 pages. <https://doi.org/10.1145/3359249>

- [16] Kathy Charmaz. 2014. *Constructing grounded theory: A practical guide through qualitative research* (2 ed.). SAGE Publications Ltd, London, United Kingdom.
- [17] 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
- [18] Hao-Fei Cheng, Ruotong Wang, Zheng Zhang, Fiona O'Connell, Terrance Gray, F. Maxwell Harper, and Haiyi Zhu. 2019. Explaining Decision-Making Algorithms through UI: Strategies to Help Non-Expert Stakeholders. 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, 1–12. <https://doi.org/10.1145/3290605.3300789>
- [19] 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.
- [20] Marika Cifor, Patricia Garcia, TL Cowan, Jasmine Rault, Tonia Sutherland, Anita Say Chan, Jennifer Rode, Anna Lauren Hoffmann, Niloufar Salehi, and Lisa Nakamura. 2019. Feminist data manifest-no. <https://www.manifestno.com/>
- [21] Sasha Costanza-Chock. 2018. Design Justice: Towards an Intersectional Feminist Framework for Design Theory and Practice. In *Proceedings of the Design Research Society*. Design Research Society, London, United Kingdom, 14. <https://ssrn.com/abstract=3189696>
- [22] Kate Crawford. 2016. Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics. *Science, Technology, & Human Values* 41, 1 (2016), 77–92. <https://doi.org/10.1177/0162243915589635>
- [23] Janet Davis and Lisa P. Nathan. 2015. Value Sensitive Design: Applications, Adaptations, and Critiques. In *Handbook of Ethics, Values, and Technological Design*, Jeroen van den Hoven, Pieter E. Vermaas, and Ibo van de Poel (Eds.). Springer, Dordrecht, Netherlands, 11–40.
- [24] Thomas Dietz, Elinor Ostrom, and Paul C. Stern. 2003. The Struggle to Govern the Commons. *Science* 302, 5652 (2003), 1907–1912. <https://doi.org/10.1126/science.1091015>
- [25] 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.
- [26] 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.
- [27] Bradley Efron and Robert J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, Boca Raton, FL, USA.
- [28] Eve L. Ewing. 2018. *Ghosts in the Schoolyard: Racism and School Closings on Chicago's South Side*. University of Chicago Press, Chicago, IL, USA.
- [29] Mary Flanagan, Daniel C. Howe, and Helen Nissenbaum. 2005. Values at Play: Design Tradeoffs in Socially-Oriented Game Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, Oregon, USA) (CHI '05). Association for Computing Machinery, New York, NY, USA, 751–760. <https://doi.org/10.1145/1054972.1055076>
- [30] Batya Friedman, David G. Hendry, and Alan Borning. 2017. A Survey of Value Sensitive Design Methods. *Foundations and Trends in Human-Computer Interaction* 11, 2 (2017), 63–125. <https://doi.org/10.1561/11000000015>
- [31] Batya Friedman and Peter H. Kahn Jr. 2003. Human Values, Ethics, and Design. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, Jacko JA Sears A (Ed.). L. Erlbaum Associates Inc., USA, 1177–1201.
- [32] B Friedman, PH Jr Kahn, and A Borning. 2006. Value Sensitive Design and Information Systems. In *Human-Computer Interaction in Management Information Systems: Foundations*, Ben Shneiderman, Ping Zhang, and Dennis Galletta (Eds.). M. E. Sharpe, Inc., Armonk, NY, USA, 348–372.
- [33] D. Gale and L. S. Shapley. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly* 69, 2 (1 1962), 9–15.
- [34] Marzyeh Ghassemi, Tristan Naumann, Peter Schulam, Andrew L. Beam, Irene Ya-Ping Chen, and Rajesh Ranganath. 2020. A Review of Challenges and Opportunities in Machine Learning for Health.. In *AMIA Joint Summits on Translational Science*. American Medical Informatics Association, USA, 191–200.
- [35] 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>
- [36] Yannai A. Gonczarowski, Noam Nisan, Lior Kovalio, and Assaf Romm. 2019. Matching for the Israeli "Mechinot" Gap-Year Programs: Handling Rich Diversity Requirements. In *Proceedings of the 2019 ACM Conference on Economics and Computation* (Phoenix, AZ, USA) (EC '19). Association for Computing Machinery, New York, NY, USA, 321. <https://doi.org/10.1145/3328526.3329620>
- [37] Mark Winston Griffith and Max Freedman. 2019. School Colors (Episode 5: The Disappearing District). <https://www.schoolcolorspodcast.com/episodes/episode-5-the-disappearing-district>
- [38] Mark Winston Griffith and Max Freedman. 2019. School Colors (Episode 7: New Kids on the Block). <https://www.schoolcolorspodcast.com/episodes/episode-7-new-kids-on-the-block>
- [39] 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.
- [40] 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.
- [41] 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>
- [42] 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>
- [43] Matt Haney, Stevon Cook, and Rachel Norton. 2018. *Developing a Community Based Student Assignment System for SFUSD*. San Francisco Unified School District. Retrieved April 13, 2020 from <https://archive.sfsud.edu/en/assets/sfsud-staff/enroll/files/2019-20/Student%20Assignment%20Proposal%20189-25A1.pdf>
- [44] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing Community-Based Collaborative Design: Towards More Equitable Participatory Design Engagements. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 216 (Nov. 2019), 25 pages. <https://doi.org/10.1145/3359318>
- [45] 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>
- [46] Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. 2007. *Heterogeneous Preferences and the Efficacy of Public School Choice*. Working Paper 12145. National Bureau of Economic Research. <https://doi.org/10.3386/w12145>
- [47] Yoan Hermstrüwer. 2019. *Transparency and Fairness in School Choice Mechanisms*. Technical Report. Max Planck Institute for Research on Collective Goods.
- [48] Zoë Hitzig. 2020. The normative gap: Mechanism design and ideal theories of justice. *Economics and Philosophy* 36, 3 (2020), 407–434. <https://doi.org/10.1017/S0266267119000270>
- [49] 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>
- [50] Matt Kasman and Jon Valant. 2019. *The opportunities and risks of K-12 student placement algorithms*. The Brookings Institute. <https://www.brookings.edu/research/the-opportunities-and-risks-of-k-12-student-placement-algorithms/>
- [51] James Rufus Koren. 2015. *Some lenders are judging you on much more than finances*. Los Angeles Times. <https://www.latimes.com/business/la-fi-new-credit-score-20151220-story.html>
- [52] T. Kulesza, S. Stumpf, M. Burnett, S. Yang, I. Kwan, and W. Wong. 2013. Too much, too little, or just right? Ways explanations impact end users' mental models. In *2013 IEEE Symposium on Visual Languages and Human Centric Computing*. IEEE, San Jose, CA, USA, 3–10. <https://doi.org/10.1109/VLHCC.2013.6645235>
- [53] Mariana Laverde. 2020. *Unequal Assignments to Public Schools and the Limits of School Choice*. Working Paper. https://drive.google.com/file/d/19HJFGmaf2HA56k3rNBgB_at4SkG3Bpz/view?usp=sharing
- [54] Christopher A. Le Dantec, Erika Shehan Poole, and Susan P. Wyche. 2009. Values as Lived Experience: Evolving Value Sensitive Design in Support of Value Discovery. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) (CHI '09). Association for Computing Machinery, New York, NY, USA, 1141–1150. <https://doi.org/10.1145/1518701.1518875>
- [55] 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>
- [56] 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>
- [57] 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>
- [58] Noëmi Manders-Huits. 2011. What Values in Design? The Challenge of Incorporating Moral Values into Design. *Science and Engineering Ethics* 17, 2 (June 2011), 271–287. <https://doi.org/10.1007/s11948-010-9198-2>
- [59] Sharan B Merriam and Associates. 2002. *Introduction to qualitative research. In Qualitative research in practice: Examples for discussion and analysis*. Jossey-Bass, Hoboken, NJ, USA, 1–17.

- [60] Thành 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>
- [61] Helen Nissenbaum. 2011. A Contextual Approach to Privacy Online. *Daedalus* 140, 4 (2011), 32–48.
- [62] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press, New York, NY, USA.
- [63] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366, 6464 (2019), 447–453. <https://doi.org/10.1126/science.aax2342>
- [64] Hessel Oosterbeek, Sándor Sóvágó, and Bas Klaauw. 2019. *Why are Schools Segregated? Evidence from the Secondary-School Match in Amsterdam*. Discussion Paper DP13462. Centre for Economic Policy Research. <https://ssrn.com/abstract=3319783>
- [65] Joana Pais and Ágnes Pintér. 2008. School choice and information: An experimental study on matching mechanisms. *Games and Economic Behavior* 64, 1 (2008), 303–328.
- [66] Parag A. Pathak. 2017. What Really Matters in Designing School Choice Mechanisms. In *Advances in Economics and Econometrics: Eleventh World Congress*, Bo Honoré, Ariel Pakes, Monika Piazzesi, and Larry Samuelson (Eds.). Econometric Society Monographs, Vol. 1. Cambridge University Press, Cambridge, MA, USA, 176–214. <https://doi.org/10.1017/9781108227162.006>
- [67] Lisa Pickoff-White. 2018. *S.F.’s Kindergarten Lottery: Do Parents’ Tricks Work?* KQED. <https://www.kqed.org/news/11641019/s-f-s-kindergarten-lottery-do-parents-tricks-work>
- [68] 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>
- [69] Alvin E. Roth. 2015. *Who Gets What And Why: The Hidden World of Matchmaking and Market Design*. Houghton Mifflin Harcourt Publishing Company, New York, NY, USA.
- [70] San Francisco Unified School District. 2015. *Student Assignment: 4th Annual Report: 2014-15 School Year*. San Francisco Unified School District. Retrieved April 13, 2020 from <https://archive.sfusd.edu/en/assets/sfusd-staff/enroll/files/2015-16/4th-annual-report-arch-8-2015.pdf>
- [71] San Francisco Unified School District. 2019. *Why We’re Redesigning Student Assignment*. San Francisco Unified School District. Retrieved April 13, 2020 from <https://www.sfusd.edu/studentassignment/why-were-redesigning-student-assignment>
- [72] San Francisco Unified School District Office of Education. n.d. *Board Policy 5101: Student Assignment*. San Francisco Unified School District. Retrieved April 27, 2020 from <https://go.boarddocs.com/ca/sfusd/Board.nsf/goto?open&id=B55QMC657423>
- [73] Devansh Saxena, Karla Badillo-Urquiola, Pamela J. Wisniewski, and Shion Guha. 2020. A Human-Centered Review of Algorithms Used within the U.S. Child Welfare System. 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, 1–15. <https://doi.org/10.1145/3313831.3376229>
- [74] Janelle Scott. 2013. School Choice and the Empowerment Imperative. *Peabody Journal of Education* 88, 1 (2013), 60–73. <https://doi.org/10.1080/0161956X.2013.752635> arXiv:<https://doi.org/10.1080/0161956X.2013.752635>
- [75] Janelle T. Scott. 2011. Market-Driven Education Reform and the Racial Politics of Advocacy. *Peabody Journal of Education* 86, 5 (2011), 580–599. <https://doi.org/10.1080/0161956X.2011.616445>
- [76] Andrew D. Selbst, danah boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and Abstraction in Sociotechnical Systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) (FAT* ’19). Association for Computing Machinery, New York, NY, USA, 59–68. <https://doi.org/10.1145/3287560.3287598>
- [77] 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](https://doi.org/10.1016/0304-4068(74)90033-0)
- [78] Katie Shilton. 2013. Values Levers: Building Ethics into Design. *Science, Technology, & Human Values* 38, 3 (2013), 374–397. <https://doi.org/10.1177/0162243912436985>
- [79] 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>
- [80] Latanya Sweeney. 2013. Discrimination in Online Ad Delivery. *Commun. ACM* 56, 5 (May 2013), 44–54. <https://doi.org/10.1145/2447976.2447990>
- [81] Amy Volda, Lynn Dombrowski, Gillian R. Hayes, and Melissa Mazmanian. 2014. Shared Values/Conflicting Logics: Working around e-Government Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI ’14). Association for Computing Machinery, New York, NY, USA, 3583–3592. <https://doi.org/10.1145/2556288.2556971>
- [82] Haiyi Zhu, Bowen Yu, Aaron Halfaker, and Loren Terveen. 2018. Value-Sensitive Algorithm Design: Method, Case Study, and Lessons. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 194 (Nov. 2018), 23 pages. <https://doi.org/10.1145/3274463>