

# Fairness inside and out: A situated approach to algorithmic allocation in complex sociotechnical systems

Bertilla Fabris, AI Policy Lab, Department of Computing Science, Umeå University, Umeå, Sweden

Mayesha Tasnim, Civic AI Lab, Socially Intelligent Artificial Systems, Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

## Abstract

This article outlines a framework for modeling and simulating complex sociotechnical systems in which an allocation mechanism acts as the interface (and sometimes a barrier) between the public and institutions. Two examples contextualise algorithmic allocation: the Amsterdam school choice and organ allocation for transplant patients. We highlight how algorithmic fairness does not guarantee systemic fairness, and propose a situated approach with institutional modeling and social simulations. With the increasing adoption of AI in decision making, a situated simulation approach provides an alternative to opaque institutional governance.

## The mechanism, institutions, and people

Sociotechnical systems are formed of technical and social components; in present-day societies, most technology is used within a social environment. Mindful of algorithmic discrimination cases, researchers have sought ways to measure the *fairness of technical systems*. The impact of a technology on its surrounding environment, however, is not always clear, especially when complex social dynamics are at play. Designing a “fair” technical component is not a guarantee of bringing fairness to the system as a whole. Selbst et al. argue that this is due to a focus on solutions (i.e. “make the outcome fair”) instead of a focus on the process (i.e. “make the system fair-er”) [1]. We now need ways to understand the impact of a technical system on fairness, within the technology, and without it.

Resource allocation problems are routinely faced by public institutions, be it in the form of determining who gets access to benefits, or which infrastructure to renovate among a set of proposals. The decision making process forms in institutions, bound by laws of fairness and non-discrimination. However, institutional decisions change the dynamics of the system before they take action, while they are communicated, and after they take hold, altering the behaviour of the social system.

The field of institutional modeling investigates the interplay of individual and institutional agents within complex sociotechnical systems. This simulation-based approach can estimate the effect of policy changes by modeling the social and

institutional aspects of sociotechnical systems [2]. Social and institutional simulations might thus guide the process of creating fairness throughout the system and not only within the mechanism. We use the framing of simulations to explore allocation mechanisms and their impact on sociotechnical systems.

## Algorithmic allocation

This document proposes an approach to modeling sociotechnical systems under specific conditions: resource allocation with a registry (here called mechanism) facing the public in one direction and based on rules determined by the institutions. The communication between public and institution occurs through the mechanism, which is designed to optimize for a certain outcome. Value preferences inform the outcome, but are not always explicit, nor uniform across parties. The institution designs the rules of the allocation mechanism based on assumptions of values in the population. Without communication between public and institutions, the rules are decided based on inferred preference models, i.e., the institution designs the mechanism to please the assumed value preference of the people and reach a “fair” outcome (Fig. 1).

In this setting, which we contextualise in two examples below, the mechanism design affects the behaviour of agents independently of its inherent degree of algorithmic fairness. Specifically, we encountered two cases that pose what seem to be “already-solved” allocation problems. Upon further scrutiny, we found underlying, more complex dynamics.

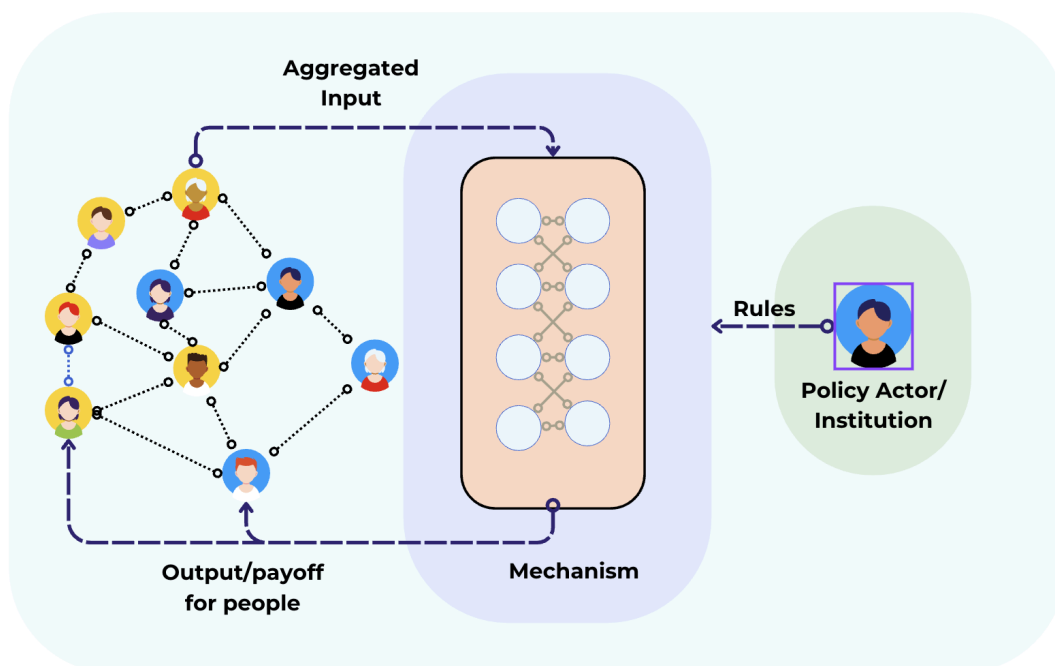


Figure 1. Motivating setup: A sociotechnical system consisting of agents, a mechanism, and policy actors. The agents provide an input to the mechanism, and are affected by its output, while a policy actor sets the rules for the mechanism.

As shown in the examples, the mismatch between the allocation rules and the preference model of the population influences the mechanism's efficacy, prompting gamification for those who have the resources to invest in understanding and playing the mechanism's rules. In the second example, the complexity of conflicting institutional values limits the mechanism's fairness and transparency.

This raises questions of prosociality, trust, governance and fairness beyond algorithmic bounds. How do people respond to a mechanism that does not align with their preference model (and perhaps their values)? How can institutions consider the longitudinal effects of algorithmic allocation within the process of policymaking?

## Amsterdam school choice

The Amsterdam school choice system is an especially useful example of a sociotechnical system where the existence of an algorithm led to unexpected changes in the behavior of the social system, leading to continuous updates to the algorithm and subsequent degradation of trust in the system. It is also a good example of when a poorly designed "fair" mechanism leads to diminishing prosocial behavior.

In Amsterdam, there is an open-choice policy when moving from primary to secondary school. Group 8 students (equivalent to 6th grade in the US) are allowed to pick any school within their education level (e.g., vocational, general secondary, pre-university, etc.) anywhere in the city, unrestricted by their zone of residence. Every year, students are asked to rank 8 to 12 schools (based on education level), and then a centralized matching algorithm automatically matches students to schools.

The system is communicated to be based on the famous Deferred Acceptance algorithm [3], a New York based school choice algorithm which won its author the Nobel prize in Economics in 2012 [4]. However, the algorithm is modified in several key places to fit the open-choice policy in Amsterdam.

Firstly, the Deferred Acceptance algorithm is a two-sided matching method, meaning that students and schools both have preferences over each other and are then matched accordingly. In Amsterdam, schools do not have a preference over students and therefore this preference is simulated with the means of a random lottery number. This choice was also **grounded in fairness**, as with a random lottery, every student has an equal chance of getting a good lottery number.

Secondly, the Deferred Acceptance algorithm does not specify a fixed number of schools to rank. In Amsterdam, however, a policy named “placement guarantee” was introduced, where students are guaranteed a position if they rank a fixed number of schools. This guarantee is ensured by increasing the capacity of schools by a marginal amount in a second round of allocation, where students who listed the required number of schools in the first round and still did not get a placement are eligible, and are then allocated using the increased capacity in the second round.

These changes unsurprisingly led to the matching algorithm performing very differently from its Nobel-prize-winning counterpart. This created problems of inefficiency and inequality.

Random lottery numbers led to an inefficient allocation of schools, increasing dissatisfaction among students and their parents [5]. Moreover, the “placement guarantee” policy led to **strategic gaming** from the parents, where they are incentivized to report schools they do not prefer just so they can ensure eligibility for the second round. This further degraded the system’s efficiency, as well as its fairness, as parents from privileged backgrounds are better able to apply strategies than parents from marginalized communities.

The authorities responsible for the school choice system have repeatedly come under public scrutiny for continuous changes to the system without successful results [6]. This has led to a degradation of **trust** in the institution.

Underneath this allocation problem is a complex sociotechnical issue, one that requires:

1. Making clear what values the system is actually trying to serve, and whether these align with what parents themselves care about.
2. Understanding how the rules of the mechanism change social behavior, so that reported preferences are seen not just as choices, but also as responses to risk, incentives, and unequal access to information.
3. Identifying problems that are easy to miss if we only look at fairness within the mechanism itself, such as distorted preferences, unequal strategic burden, and declining trust in the system.
4. Looking beyond the matching rule alone, and instead supporting better communication, better alignment between stakeholders, and more situated policy design.

This requires a multidisciplinary sociotechnical approach, one in which a social

simulation model can help make the problem and its tradeoffs easier to communicate and discuss. It also requires a systematic way of eliciting stakeholder values and using those discussions to work toward a shared understanding of what the system should aim to achieve.

## Organ allocation systems

Fairness plays an important role in organ allocation systems. For patients on the waiting list to receive a transplant, waiting time is a strong determinant of short-term survival, and long-term quality of life. Multiple factors influence the structure of the waiting list and the distribution of donated organs, particularly those from deceased donors. Initially, biological compatibility between donor and recipient seems to provide a baseline allocation rule: the most compatible donor-recipient match (considering blood type and HLA presence) should be prioritised to ensure the best graft survival outcomes. However, the length of the waiting list and persistent scarcity of organs complicates the allocation problem.

Cold ischemia time restrictions (meaning how long an organ is deprived of blood flow before its quality deteriorates beyond utility for transplantation) brings into consideration logistics, donor and recipient location and a corollary of geographical factors such as regional allocation rules, coordination practices and transplant capacity. Allocation policies have been studied via simulations and algorithmic optimization for a long time, with awareness of the complex interplay of medical, economic, political and legal factors.

The other component of the system is the prioritization of patients within the waiting list structure. The state of the patients is dynamic, with probabilities of becoming too ill to receive a transplant. The waiting list design attempts to accommodate for this by modeling the disease progression and setting thresholds for transplant eligibility.

In the U.S., Organ Procurement Organizations (OPO) and Transplant Centers's performance is evaluated (competitively) based on how many organs were retrieved and how many successful transplants were performed [7], adding economic incentives to the mix. This influences the local center's decision to accept available organs or reject them in the hope of a better quality alternative.

Even seemingly aligned bioethical values can create conflicting interests, especially if multiple institutional actors are pushing for their preferred outcomes. Beside the responsibility to ensure "fairness" in allocation, the meaning of "do no harm" manifests differently for Transplant Centers (i.e. reject lower quality organ offers) and OPO (i.e. provide as many quality organs as possible), exacerbating inefficiencies. Even System Dynamics can fail to account for institutional preferences in the micro scale [8]. Concerning the value preferences of the patients themselves, little is

known. Without a channel of communication to the algorithmic rule setters, and with conflicting interests among institutions, we risk optimizing for less relevant values.

Most importantly, the algorithm of allocation can be gamed by **multilisting**. Multilisting is the practice of assigning one patient to multiple waiting lists across Donor Service Areas and Transplant Centers, which initially seemed to alleviate the length of the waiting list at the national scale in the U.S. However, this came at the cost of *individual fairness* (not everyone can be listed in multiple registries as not everyone can afford to quickly travel across states to receive a life-saving transplant) and *collective fairness*: while the national wait time decreased, regional variance in wait time can increase [9]. Multiple organ offers also seem to produce a utilitarian gain in efficiency (by reducing organ wastage and time to transplant), but may do so at the cost of societal trust in the system [10].

Within this critical healthcare system, the allocation mechanisms endlessly attempt to **reconcile fairness and efficiency**. The value prioritization shifts from *face-value fairness* (e.g. centralised FIFO allocation) to *utilitarian fairness* (prioritizing patients with “the most to gain” from a transplant) or *need-based fairness* (prioritizing the most sick patients based on disease models with thresholds for exclusion), or introducing market logic to the system (monetary rewards to centers, multilisting, multioffer). Formalizing any value-rule into algorithmic allocation mechanisms creates, as in the example above, new ways for the system to be strategised upon, shifting the locus of inequality and obscuring its methods.

Decontextualised fairness as an attribute of technical (sub)systems has already been criticized by Selbst et al [1]. The algorithms designed for organ allocation systems present a clear example of the Framing and Formalism traps, in which fairness evaluation stops at the technical part of the sociotechnical system and its formalization generates new, unaccounted behaviours in the system. Some literature acknowledges these limitations, suggesting we might expand our abstraction boundary, or even consider other processes beside algorithmic allocation [11,12]. In the absence of a consensus of what “fair” organ allocation is, we might want to shift our focus to a process approach. Social and institutional modelling provide the best toolset for this endeavour.

## Our proposal

We propose a four-step approach for studying complex sociotechnical systems in which an allocation mechanism mediates between the public and institutions. The goal is to understand how to embed algorithmic allocation mechanisms in a wider system of stakeholder values, behavioral adaptation, and institutional feedback. As the Amsterdam school choice reminds us, the algorithmic allocation must be designed to accommodate a response to the mechanism itself. To this end, we consider **social/institutional simulations** as the most eligible, situated approach.

The four steps are as follows:

**1. Value elicitation:** The first step is to determine which values the mechanism is meant to serve, and whether those values are actually shared by the stakeholders affected by it. Mechanisms are often designed around an inferred preference model: institutions assume what matters to the population and encode those assumptions into rules and objectives. But in high-stakes systems, these assumptions may diverge from the lived priorities of the people who interact with them. A situated approach therefore begins by making these values explicit and contextable.

For example, in the Amsterdam school choice, policymakers may prioritize procedural fairness and avoiding unassigned students, while parents may care more about genuine access to preferred schools, reduced strategic burden, transparency, and trust. This gap matters for how the system performs in practice. Within organ allocation, value elicitation is needed to clarify which purpose the allocation system serves within the transplantation system as a whole, and what systemic fairness means to all stakeholders. Although this might not result in a univocal “solution”, it lights the fire to improve procedural awareness.

**2. Modeling and simulation:** The second step is modeling and simulation. The simulation model is not the end goal, but a structured approach to facilitate and open communication around assumptions, tradeoffs, and possible interventions.

In systems such as school choice and organ allocation, the mechanism does not act on fixed inputs. The inputs themselves change in response to the mechanism. Parents adapt their reported preferences strategically; OPOs and patients respond to waiting list structures, matching rules, and institutional incentives. Social simulation makes it possible to examine these interactions explicitly and to ask not only what outcomes a rule produces, but how that rule changes behavior across the wider system.

**3. Identifying bottlenecks and blind spots:** The third step is to identify systemic bottlenecks and blind spots. Mechanisms are often evaluated in terms of fairness or efficiency within the allocation procedure itself, but this can obscure deeper problems elsewhere. A mechanism may be fair on paper while placing informational or strategic burdens unevenly across groups. It may also optimize one part of the process while ignoring upstream or downstream failures.

A situated simulation framework helps uncover these blind spots by tracing feedback loops across the system: strategic adaptation, unequal ability to navigate rules or exploit advantages, changing interpretations of fairness, and degradation of trust over time. The modeling of heterogeneous agents and institutions allows the representation of multiple values and goals.

**4. Policy recommendations:** The final step is policy recommendations. These recommendations should go beyond a technical redesign of the mechanism itself, as the system is sociotechnical; interventions may also need to address communication, participation, institutional coordination, trust, and policy stability. As the approach is situated within a community of stakeholders (institutions and population), the policy recommendations are tailored to local needs.

This step can be expressed in policy briefs, appointment of committee and representatives to coordinate across stakeholder groups.

The approach is intended to be *iterative*. Fixed time intervals between iterations allow for policy to take action, and structure the process of altering it. Moreover, they facilitate participation of stakeholders that are not institutional, constructing the communication channel that was lacking in the initial setting (Fig. 1).

## Conclusions

Algorithmic allocation mechanisms are a component of critical sociotechnical systems where institutions and populations interact. We argue that their fairness should be evaluated inside the algorithms and outside of them, encompassing their effect on the system they become a part of. With the increasing usage of AI to automate allocation processes, aligning the values of stakeholders and opening communication channels between the population and institutions becomes our priority. A situated approach with social and institutional simulations is presented. While we do not claim that the approach is perfect, it provides a step forward in the integration of algorithmic allocation and fosters a procedural, situated view of fairness.

## References

- [1] Selbst, Andrew D., 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, 59–68. New York: Association for Computing Machinery. <https://doi.org/10.1145/3287560.3287598>.
- [2] Ghorbani, Amineh. 2022. "Institutional Modelling: Adding Social Backbone to Agent-Based Models." *MethodsX* 9: 101801. <https://doi.org/10.1016/j.mex.2022.101801>.

- [3] Abdulkadiroğlu, Atila, Parag A. Pathak, and Alvin E. Roth. 2005. "The New York City High School Match." *American Economic Review* 95 (2): 364–67. <https://doi.org/10.1257/000282805774670167>.
- [4] Nobel Prize Outreach. 2012. "The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2012." NobelPrize.org. <https://www.nobelprize.org/prizes/economic-sciences/2012/summary/>.
- [5] Het Parool. 2026. "Weer Minder Achtstegroepers Kunnen Naar Eerste Voorkeurschool, Volgend Jaar Wordt Het Beter." Parool.nl. <https://www.parool.nl/amsterdam/weer-minder-achtstegroepers-kunnen-naar-eerste-voorkeurschool-volgend-jaar-wordt-het-beter~b0ec8fd5f/>.
- [6] Het Parool. 2026. "Loting voor Middelbare School Weer Terug naar Oude Systeem: 12 in Plaats van 9 Voorkeurscholen." Parool.nl. <https://www.parool.nl/amsterdam/loting-voor-middelbare-school-weer-terug-naar-oude-systeem-12-in-plaats-van-9-voorkeurscholen~b7972bda/>.
- [7] Washburn, Kirt. 2012. "Maximizing Donor Potential: Evolving Organ Procurement Organization Metrics and Optimizing Organ Distribution and Allocation in the United States." *Liver Transplantation* 18 (suppl. 2): S1–S4. <https://doi.org/10.1002/lt.23507>.
- [8] Hirsch, Gary B., Khalid Saeed, Karl J. McCleary, and Kevin A. Myer. 2012. "Deceased Donor Potential for Organ Transplantation: A System Dynamics Framework." Paper presented at the 30th International Conference of the System Dynamics Society, St. Gallen, Switzerland, July 22–26. <https://proceedings.systemdynamics.org/2012/proceed/papers/P1428.pdf>.
- [9] Harvey, C., and J. R. Thompson. 2016. "Exploring Advantages in the Waiting List for Organ Donations." In 2016 Winter Simulation Conference (WSC), 2006–17. Washington, DC: IEEE. <https://doi.org/10.1109/WSC.2016.7822245>.
- [10] Erazo, Ignacio, David Goldsman, Pinar Keskinocak, and Joel Sokol. 2022. "A Simulation-Optimization Framework to Improve the Organ Transplantation Offering System." In 2022 Winter Simulation Conference (WSC), 1009–20. IEEE. <https://doi.org/10.1109/WSC57314.2022.10015431>.
- [11] Thompson, David, Larry Waisanen, Robert Wolfe, Robert M. Merion, Keith McCullough, and Ann Rodgers. 2004. "Simulating the Allocation of Organs for Transplantation." *Health Care Management Science* 7 (4): 331–38.

<https://doi.org/10.1007/s10729-004-7541-3>.

[12] Feccia, Mariano, Arianna Freda, Maurizio Naldi, Gaia Nicosia, and Andrea Pacifici. 2026. "Modelling and Simulating the Organ Donation Process Using Bootstrap and Event-Driven Process Chain Representation." *Journal of Simulation* 20 (2): 135–51. <https://doi.org/10.1080/17477778.2025.2486709>.

**Keywords (comma separated):**

Fairness, Algorithmic allocation, Social Simulations, Institutions, Policy recommendation, Situatedness.

**Related URL:**

<https://openresearch.amsterdam/en/page/120012/gaming-the-system-the-risk-of-strategy-in-school-choice>