Just Governance and the Right to the Algorithmic City

Comments on Seth Lazar's 'Algorithmic Governance and Political Philosophy'

Renée Jørgensen Sept 2023

Perhaps one of the most important insights in these rich lectures is that the various choices made in the course of designing and managing 'algorithmic intermediaries' are *governance*, and so properly subject to the normative constraints on appropriate governance. That is, in addition to evaluating the aims for which algorithmic intermediary power is employed, we should hold its exercise accountable to both procedural and substantive constraints on just governance. Lazar tells us "If political philosophy has a central question, it should be: what gives some the right to govern others?" His efforts in these lectures are accordingly organized around the all-things considered justification of power, organized by of three questions: what are the right ends, what is the right way, and *who* has the right, to exercise intermediary power? He finds little in contemporary or historical work in analytic political philosophy that can be helpfully extended because, as he writes, "Political philosophy has mostly understood authority as the right to exercise coercive power to make and impose laws. Although states and other powerful entities have always also used pre-emptive power to shape behavior, this has been largely overlooked in the justification of authority."[1] So, Lazar embarks on the mammoth task of excavating the underlying values of just governance in order to build a political philosophy for normative computing from the ground up.

I agree that theories of just governance that focus exclusively on the exercise of coercive state power are ill-suited to evaluating the power dynamics that characterize our algorithmically-mediated relationships ('the Algorithmic City'). But I suspect Lazar's choice of starting point—focusing on the agents *exercising* governing power—substantially explains his difficulty in finding relevant and extendable political philosophy. Starting with authority, as he does, is a very familiar path for theorizing State use of coercive power. But governing power is power *over someone*, and that means there is a second natural place to start a robust inquiry into just governance: by asking how the rights of *those over whom* power is held constrain its legitimate exercise. This approach is in many ways a better fit for Lazar's project, and there is a rich tradition of political philosophers using it to theorize problems strikingly similar to the ones he identifies in the algorithmic city. Several prominent feminist political philosophers have contributed to work in this vein (including Marilyn Frye, Susan Okin, Patricia Hill Collins, Iris Marion Young, and Ann Cudd[2]), urging that issues previously relegated to the 'interpersonal' or 'private' are *political* problems precisely because they structure agents' options (constraining them from the inside out), exercise power *between* agents, and---much like the intermediaries that are Lazar's focus—largely constitute our shared social structure.[3] Each of these scholars is emphatic that looking only at the state's exercise of coercive legal power leaves out a great number of power relations relevant to questions of justice.

---

[1] Lazar THIS VOLUME, 44.

[2] See e.g. Frye, *The Politics of Reality* (1983, Crossing Press: New York); Okin, *Justice, Gender, and the Family* (1989, Basic Books); Collins, *Black Feminist Thought* (1990, Routledge); Young, *Justice and the Politics of Difference* (1990, Princeton University Press); Cudd, *Analyzing Oppression* (2006,Oxford University Press).

[3] In a 1987 paper, ('Justice and Gender', *Philosophy & Public Affairs* 16), Susan Okin criticized Rawls' *Theory of Justice* and Walzer's *Spheres of Justice* specifically for failing to recognize that social institutions like family structure, interpersonal normative expectations concerning the gendered division of labor, and market structures are all non-state forces that constrain women's

Starting with the *subjects* of governance led Young to give a very different characterization of the central task for political philosophy: to diagnose "how institutions and social relationships differentially conspire to restrict the opportunities of some people to develop and exercise their capacities"[4] or to "participat[e] in determining […]the conditions of their actions".[5] She centrally focuses on the "institutional relations and processes of the society"[6], because it is these, more than the extrinsically coercive forces of law, that do the most to shape our lives, protecting us from or exposing us to exploitation, marginalization, and systemic violence; either empowering or rendering us relatively powerless to shape our relations with others. Other theorists (saliently Charles Mills[7]) use this approach to centrally analyze the *mediating* power of law—in defining who counts as a person, and hence who can claim the protections of law—as an urgent site for theorizing justice.

There are advantages to starting inquiry by asking what opportunities and options the *governed must have* to avoid oppression. It sets us up to meet *any* proposed limitation or restructuring of their options by asking what justifies it; we would not even be tempted to think that justification is required only when limitations come from the state, or through the application of extrinsic coercive power. We would also see that there are many cases where governance occurs—that is, the constitutive norms of a social institution are established and enforced—but power *over* is diffuse, not exercised by any specific agent. When the unintentionally coordinated systems or actions of many people constrain and enable social groups differently, the fact that the occupants of some social positions are resultantly subject to exploitation, rendered comparatively unfree, or made powerless to participate in decision-making *is* a problem for justice, even when it does not take the form of *an agent* governing without right. These after all constitute threats both to relational equality and collective self-governance.

None of this undermines the importance of specifically theorizing how justice constrains the exercise of intermediary power in algorithmic governance—it is crucial, for many of the reasons that Lazar eloquently identified, to attend to it. But I do think that our efforts can and should be informed by the rich insights from the related work that has already been done. Precisely because the problem is broader than the exercise of coercive power by a central authority, when looking for relevant work we should cast our net wider than those who give the authority question pride of place.

Lazar's discussion is organized by two crucial but slightly cross-cutting distinctions between modalities of power—extrinsic vs. intermediary, and coercive vs. preemptive. *Extrinsic* power shapes behavior by imposing external constraints, while *intermediary* power "constitutes (in whole or part) the social relations that it mediates" (28). He presents extrinsic power as paradigmatically shaping behavior by threat of coercion, while intermediary power enables *preemptive* governance, controlling behavior by structuring the available options, designing away non-compliance and nudging behavior in preferred directions. While acknowledging that extrinsic coercive power and intermediary coercive power can blur into each other, Lazar's taxonomy is silent concerning the possibility of extrinsic preemptive power—

options in ways that leave them socially subordinated to men, and so are relevant to justice. She generalized this criticism two years later in her book *Justice, Gender, and the Family*.

[4] Iris Marion Young, 'Equality of Whom? Social Groups and Judgments of Justice' *Journal of Political Philosophy* (2001):16.
[5] Young, *Justice and the Politics of Difference* (1990): 38
[6] Young (2001): 2
[7] Mills, *The Racial Contract* (1997, Cornell University Press)

power which shapes behavior by constructing external constraints which structure the options available to a person, rather than threatening them with penalties. I suggest that it *is* possible, exemplified in the design of physical infrastructure and paradigmatically in hostile architecture, like park benches with armrests positioned to make it impossible to lay down to sleep.

The taxonomy offers helpful language to distinguish challenges that are unique to *intermediary* power from those that attend to all forms of preemptive and productive governance. Lazar leverages it to highlight two features of the Algorithmic City that he sees as making the challenges of just governance particularly acute, and extant models of just governance inadequate to theorize it. First, because algorithmic governance can be dynamically preemptive, it can constrain without coercing, and moreover can (at least in principle) render non-compliance impossible. Second, because (as he quotes Lessig), "Code is never found; it is only ever made, and only ever made by us"[8] it is not neutral; "every option has to be designed" and specifically enabled. Both of these features have major effects on the lives of the governed, and easily slip from view if we focus narrowly on state coercion. Lazar casts these features of algorithmic governance largely as *contrasting* with legal governance, writing for instance that while "Law has to be public to be effective; pre-emptive governance can be wholly invisible"(47). I find this emphasis more distracting than illuminating. Careful attention to power relations in the algorithmic city is valuable not because they are novel or entirely unlike analogue social structures, but precisely because they are deeply similar.

To see this, it is crucially important to recognize a subdivision between what I'll call 'formally social options', on the one hand, and 'generic options', on the other. A formally social option is constitutively formed and defined by social practices concerning it. Such an option in principle *cannot* be exercised unless made socially possible.[9] On digital platforms, these are often specific interactions with other users (attacking their avatar, stealing their digital property, calling them directly in real time, etc.). In the analogue world, these include having a status like 'married', or 'legal guardian', as well as doing things like maintaining credit, scoring a touchdown, or winning the World Cup. Formally social options can be made impossible, for all or some, by simply withholding the prerequisite scaffolding. By contrast, a generic option is an all-purpose option that can be exercised in a variety of ways. These can be 'social' in that they involve others (as buyers, or audiences, or subjects), but are not *constituted* by social recognition. Whether in the analogue or digital realm, it is difficult to enable or disable just some *instances* of a generic option (e.g. prohibiting death threats while permitting speech generally; or prohibiting the sale of specific goods).

This division cross-cuts means of governance, and I suggest that it, rather than whether power is algorithmic or legal, determines the degree of preemptive control available. The illusion of asymmetry arises, I suggest, from focusing on generic options when discussing the limits of law, but on formally social options when illustrating the preemptive power of algorithmic intermediaries. Though he recognizes constitutive laws as a form of intermediary power, Lazar contends—in discussing the example of marriage—that they are still fundamentally dependent on enforcement: they "can determine whether a

[8] Lazar 51:n146

[9] Lazar calls these 'socially recognized options', (see pp. 29, 40) but I hesitate to adopt this terminology as it implies the antecedent existence of an option which society then does or doesn't recognize, rather than options that *are constitutively formed* by specific social practices concerning them.

given kind of social relation is legally recognised but cannot (without enforcement) prevent people from participating in the legally-unrecognised counterpart" (40). But formally social options rarely have substitutable unrecognized counterparts, and consequently can be made impossible in analogue spaces just as effectively as in digital ones.

As Young, Mills, and others emphasize, it can hard to see the preemptive legal constraints that don't directly affect you; they are often invisible by design (much like the algorithmic preemptive controls that worry Lazar). So let me offer a few examples. A great deal of legal oppression in the United States has been accomplished through omission and preemption. Black Americans were excluded from voting, testifying in court, or standing for election to public office well before the exclusion was made legally explicit in the 1860s.[10] These forms of civic participation were previously simply not enabled for Black Americans: they were not positively defined as eligible to perform these roles. Similarly, well into the 1880s coverture doctrines precluded married women from retaining property in their own name or having personal bank accounts, not by *prohibition* but simply by defining married women as an extension of their husband's legal person. And though credit cards have been commercially common since the 1950s, this option was not enabled for women—they were not eligible to have credit in their own name—until 1974. Lest we begin to think that oppression by omission is a thing of the past, in present-day Italy (as of January 2023) the government only recognizes the legal parenthood of the biological parent in same-sex couples, and Padua has retroactively removed non-gestational mothers from their children's birth certificates.

We can disobey legal prohibitions; *criminal* law might well have to be public to be effective. But there's a great deal of legal governance that doesn't work this way, and there are many actions that we simply can't do without being recognized as able to do them. Running for public office, building credit, exercising the rights of a legal parent, and being married are just some among this set. If not enabled by the relevant social or legal structures, formally social options are simply unavailable. The efficacy of this kind of legal constraint does not depend on internalization or compliance, and can be implemented with frightening speed (as anyone following the news in Italy knows all too well). Omission is a powerful thing. Whether legal or algorithmic, an agent with the power to withdraw the prerequisite scaffolding can determine whether such options are literally possible or impossible for those affected; this power requires neither publicity nor coercive enforcement for efficacy.

Generic options are the inverse. We can target undesired instances with bans and content filters, or try to design them away by withholding tools that would make it easy to express the banned content (or to manufacture the banned goods)—but all of these measures are evadable, by tactics that work just as well in digital as in analogue spaces. The recent "blank paper protests" in Russia (2022), Hong Kong (2020), and China (2022-3) nicely illustrate the deep similarities between the strategies for evading content filters and moderation in digital space, and those used to circumvent legal restrictions on speech. In the wake of aggressively enforced State suppression of criticism, people gathered in public spaces

[10] Mississippi and South Carolina passed the first 'Black Codes' (more widely known as 'Jim Crow' laws), formally prohibiting Black Americans from many forms of civic participation in 1865. These laws were enforced with the penalty of unpaid labor for white plantation owners. So for the two years until the passage of the 14th Amendment in 1867, Black Americans were *coercively* barred from the exercise of these options. For the rest, they were excluded preemptively, typically by not being granted the status of legal persons.

holding *blank* sheets of white paper or poster board, in silence—and were successfully understood as expressing condemnation for both the State's salient recent policies[11], and its willingness to interpret *any* semantic content as justifying repressive punishment. The whitepaper protests in China (or "A4Revolution") were accompanied by an online component: people posted their support for the protest—while evading aggressive algorithmic content filters—by posting walls of text simply repeating the characters for "good", "yes", "correct".[12] Users on Facebook and Instagram signaled support by posting blank white images. Both digital and analogue forms of speech were violating anti-criticism regulations from the state, and both successfully evaded the censors' imperfect control. Of course, an algorithmic content filter can learn, and scrub records of past protest that escaped censorship—but so can the enforcement arm of a powerful authoritarian government.

Seeing this reveals that the features Lazar presents as the accidental virtues of law—publicity and resistibility—aren't of *law*, per se; they're accidental virtues of extrinsic governance which depends for efficacy on internalization and compliance. Framing the difference between coercive and preemptive power as even roughly mapping onto the distinction between legal and algorithmic power obscures the underlying truth that both forms of governance exercise huge amounts of *both* forms of power, corresponding to the regulation of formally social options and generic options, respectively. The lesson is not that preemptive interference in the digital realm is unproblematic, but rather that it is also a source of injustice in its analogue forms. The ability to control which options are enabled—to determine which people, relationships, or actions are recognized—is tremendous power over and between, even (and especially) when it can be exercised without threat of coercive force.

This is not to say that all theorizing work has already been done, or that there are no new challenges posed by algorithmic intermediaries. Some of the features Lazar emphasizes merit special attention concerning the regulation of generic options. The long arm of the state grows substantially longer when it can link traces of individuals across large datasets. And insofar as they are automated, the edicts of algorithmic censors are less dependent on the buy-in of street-level bureaucrats, so there are fewer opportunities for discretionary non-enforcement than in analogue equivalents. Some of the ways of resisting increasing authoritarianism are less available within algorithmic governance. Most dangerously, we may be more complacent about censorship when enforced by denial of service rather than by imprisonment.[13] But over-emphasizing the novelty of the problems evident in algorithmic contexts can lead us to miss the opportunity it offers for better understanding the parallel problems in ordinary contexts.

A similar lesson awaits concerning justificatory neutrality. Lazar argues that it is harder to justify the decisions of algorithmic governance on terms that all can accept, because it must be more opinionated than legal governance. So even if it is possible to give *reasons* for design decisions that appeal equally to a wide range of reasonable views of the good, the decisions involved in algorithmic governance are so

[11] e.g. the invasion of Ukraine, passage of strict anti-criticism measures, and covid lockdowns, respectively.
[12] Chang Che and Amy Chang Chien, 'Memes, Puns, and Blank Sheets of Paper: China's Creative Acts of Protest' *New York Times*, Nov 28 2022. (https://www.nytimes.com/2022/11/28/world/asia/china-protests-blank-sheets.html)
[13] This might assume an overly optimistic picture of our likelihood to object to violent authoritarian censorship. As exercised, persecution for violation of state prohibitions against criticism aren't advertised as such; violators simply disappear, and often a cover story is offered to create just enough uncertainty that bystanders conclude they lack adequate evidence of flagrant injustice to risk taking a stand against power.

numerous and so specific that they cannot avoid privileging some views of the good and disadvantaging others. This involves two claims: first, that because in analogue contexts "our ability to perform a wide range of the most fundamental actions is not conditioned on the law making those actions possible" (53), the state can simply establish "a set of starting-gate conditions and get out of the way"(51), taking no view on specific substantive questions. And second, that in contrast "in the Algorithmic City, the default is that you are *not able to X* unless it is explicitly enabled" (53, emphasis in original), so neutrality is not an option because "the only way not to endorse some option is not to enable it, which is expressively close to prohibiting it."(52)

It is easier to evaluate these claims with concrete examples in mind; let's take Twitter and start with the second claim. If we define 'option' broadly enough to count *<post text strings that pass a content filter>* as an option, then it is true that every option must be explicitly designed in or out. But equipping users with the ability to exercise such a generic option cannot fairly be read as "endorsing" *every specific purpose* to which they put it (certainly not *endorsing* every meaningful sentence the users assert *in* the text strings.) Once we've fixed a consistent level of specificity for 'option', either it is false that every possible option must be explicitly enabled or disabled, or it is false that the enable/disable choice determines *what one can do with the option* to such a degree that it can be read as *endorsing* the particular exercise. So the second claim looks far too strong.

What about the first claim, the idea that legalistic governance is more neutral than algorithmic? Lazar writes that "if the law (say) is silent on X, then by default the law permits X" (51)—but if this is true, it is equally true of algorithmic governance. That is, it is true only if we read 'permit' as merely meaning *not explicitly prohibit.* But as my earlier examples illustrate, one can be rendered thoroughly *unable to exercise* a formally social option through legal omission; explicit prohibition isn't necessary. In letting these cases slip from view, Lazar risks repeating the error for which Young, Okin, Mills, and others criticized Rawls: conflating being "*silent* on specific substantive questions" with "taking no view".

I take Lazar's argument to be not that it is impossible to provide neutral justification for some procedures or principles to guide option-design in algorithmic governance, but that "the appeal to justificatory neutrality seems like a sham if your design decisions precisely enable and constitute differential impacts on different conceptions of the good."(53) I agree. But if we consider legal governance at the same close-range at which he inspects algorithmic governance—attending to the regulatory and administrative choices, and not only the guiding principles and criminal codes—we will find that in fact no society can be governed exclusively with extrinsic coercive power. Legal government cannot entirely prescind from making choices about which relations and options to enable: it must decide who to recognize as persons, what powers to vest them with. Moreover as Kukla points out, "There is never any such thing as the unmarked, general way that a space is available to be used or repurposed; different bodies will always fit into and experience a space differently, and be afforded different possibilities for agency within it."[14] The literal spaces of a city—from park benches to underpasses and sidewalks—must all take some determinate form, and no form will serve everyone's interests equally well. The decisions we make concerning a city's physical infrastructure and architecture shape its denizens' prospects, the power relations between them, and do a lot to influence which sorts of social norms are likely to emerge. No

[14] Kukla, *City Living*, (Oxford University Press: Oxford, 2021): 261.

matter how neutrally justifiable, they will not equally enable every reasonable conception of the good. We can never *merely* set starting-gate conditions and get out of the way.

So the appeal to justificatory neutrality is no less a sham in the analogue than the algorithmic city. That's actually good news: it means that theories of just governance in physical spaces and concerning the exercise of legal power can inform and be informed by theories about justice in the design and governance of the algorithmic city. This promises to be fruitful exchange. Analysis of the physical city has recently yielded the claim that residents have "a right to the city"—which Kukla argues should be interpreted as meaning that the governed—those who live *in* a city—must be provided "not just access to suitable spaces that meet one's needs and fit one's activities, but also the ability to exercise agency not only within but *upon* these spaces."[15] There are echoes here of a concern with relational equality, something like autonomy, and self-determination—it is a promising template for theorizing the sort of participation in self-governance that Lazar has argued must be afforded to the governed in algorithmic cities.

The problems Lazar calls us to attend to are important problems, and probably do raise new challenges. They *certainly* afford new opportunities to attend to social and relational structures as *designed* sites of governance, subject to our collective (imperfect, but effective) control. If we set aside concern with whether the issues raised by algorithmic governance are different from those in legal governance, and attend carefully to the types of options being regulated, Lazar's exploration of preemptive power and the failures of justificatory neutrality in the algorithmic city underscores many important lessons for theorizing just governance generally. First, a model that assumes that the only choices in need of justification are choices to remove options by threat of violent enforcement is a deeply impoverished model of justice. Second, we should be as careful about preemptive power in legal governance as he urges us to be for algorithmic governance; we cannot rely on the accidental virtues of publicity or resistibility to safeguard us in either domain. Third, there is *no* context where we can safely approach the question of just governance as rule-from-afar, thinking that if only we select the substantively right set of aims, procedures, or norms, perfected top-down implementation will yield perfected justice.

Even identifying an appropriate ruler is not sufficient; we must ask not only who has a right to make governance decisions, but also *what rights constrain* the exercise of governing authority. Echoing Kukla, we might say we have a right to the algorithmic city: in our digital lives, as in so many other domains, just governance requires that we have appropriate influence on selecting and augmenting the rules which govern us, and urges a form of substantive pluralism rather than justificatory neutrality. There is a lot to unpack here; theorizing justice for the algorithmic city is a mammoth task, and I thank Lazar for giving us such a good start.

[15] Kukla, *City Living*, p. 260.