

# SOCIAL EXPERIMENTATION IN AN UNJUST WORLD

Jacob Barrett and Allen Buchanan

Abstract: There has been a recent resurgence of interest in social experimentation as a means of promoting social progress, including progress in justice. In this paper, we first advance an argument in favor of social experimentation drawing on its capacity to resolve uncertainty both about how to achieve socially valuable goals and about what goals are worth pursuing. We then identify four challenges: the information problem (experiments may not yield relevant information), the selection bias problem (potentially informative experiments may not be undertaken), the uptake problem (the information generated by experiments may not be put to good use), and the risk problem (experiments may carry serious risks). Finally, we argue that existing injustices can exacerbate all four problems, rendering social experimentation a less reliable path to progress, and, in cases of severe injustice, perhaps even a regressive force. The upshot is not that we should give up on social experimentation, but that we should temper our expectations and focus on constructing conditions under which experimentation is more progressive. Specifically, in order to render social experimentation a more reliable engine for social progress of any sort, we must remedy or mitigate existing injustices.

## I. Introduction

Social reform is a rather uncertain endeavor. The reforms we implement often fail to achieve their desired effects or have unanticipated deleterious effects. And even when attempts at reform do work out as planned, we may later revise our standards of evaluation, and so conclude that we did not produce genuine improvements after all.

Reflecting on this uncertainty has led many—from historical figures such as John Stuart Mill, Karl Popper, and especially John Dewey, to their legion of contemporary intellectual descendants—

to endorse an experimentalist orientation toward social improvement.<sup>1</sup> The basic case for experimentalism is elegant and straightforward. Since the institutional or policy alternatives that we currently deem optimal frequently prove otherwise, trying out a variety of promising alternatives often uncovers better options than our current best guess. And even when experiments don't work out so well, they can provide valuable information about how better to pursue other reforms in the future. So social experimentation sometimes leads us to stumble upon better options, and more generally teaches us how to improve our society going forward—either through revealing better ways of implementing our current conception of what we ought to strive for, or through helping us to improve

---

<sup>1</sup> John Dewey, *The Public and its Problems* (New York: Henry Holt and Company, 1927), John Stuart Mill, *On Liberty*, in J. M. Robson (ed.) *The Collected Works of John Stuart Mill*, vol. 18 (Toronto: University of Toronto Press, 1977), and Karl Popper, *The Open Society and its Enemies* (Princeton: Princeton University Press, 2013). In recent years, experimentalism has experienced a resurgence in political philosophy. See, for example, Elizabeth Anderson, “The Epistemology of Democracy,” *Episteme* 3 (2006): 8-22, Jacob Barrett, “Social Reform in a Complex World,” *Journal of Ethics and Social Philosophy* 17 (2020): 103-132, Jack Knight and James Johnson, *The Priority of Democracy* (Princeton: Princeton University Press, 2011), Ryan Muldoon, *Social Contract Theory for a Diverse World* (New York: Routledge, 2016), Julian F. Müller, *Political Pluralism, Disagreement, and Justice* (New York: Routledge, 2019), Gregory Robson, “The Rationality of Political Experimentation,” *Politics, Philosophy & Economics* 20 (2021): 67-98, and Charles F. Sabel and Jonathan Zeitlin, “Experimentalist Governance,” in D. Levi Faur (ed.) *The Oxford Handbook of Governance* (New York: Oxford University Press, 2012). Though not our focus here, related approaches have also become popular in economics. For example, see Abhijit Banerjee and Esther Duflo, *Poor Economics* (New York: Public Affairs, 2012), who (with Michael Kremer) won the 2019 Nobel Prize in Economics “for their experimental approach to alleviating global poverty.”

this conception. Other things being equal, a society that engages in a wide range of social experimentation should therefore be a more progressive society: one that is better at getting better, more conducive to improvement across time.<sup>2</sup> This is the thesis of progressive experimentalism.

Our goal in this paper is to cast a critical eye on progressive experimentalism—not in order to refute the thesis that social experimentation promotes social progress, but to take it seriously enough to ask some hard questions. More specifically, our focus will be on the way that existing injustices can make social experimentation a less reliable path to progress, and, in some cases, may even render social experimentation a regressive force. This is ironic, given that many progressive experimentalists are especially enthusiastic about the ability of social experimentation to promote improvements in justice. Our analysis therefore suggests that the relationship between social experimentation and social justice is more complicated than is typically appreciated: in the very circumstances where progress in justice is most needed, experimentation is least likely to be a reliable path to improvements of any sort, including improvements in justice.

In Part II of this paper, we clarify the ambiguous term “social experimentation.” In Part III we articulate the basic argument that social experimentation promotes progress, including progress in justice. In Part IV, we identify four challenges to this argument: (i) the information problem (experimentation may not produce information that is useful for making progress), (ii) the selection bias problem (potentially informative experiments may not be undertaken), (iii) the uptake problem (even if an experiment yields valuable information, that information may not be put to good use), and (iv) the risk problem (the risks of an experiment may outweigh the benefits, or be too unjustly distributed). Although we will suggest that these challenges serve as a helpful lens for evaluating the

---

<sup>2</sup> We borrow this notion of a “progressive” society as one with a tendency to improve from Barrett, “Social Reform in a Complex World.”

prospects of social experimentation in general, our primary aim will be to show how existing injustices, in particular, can exacerbate all four challenges. This task will occupy us in Part V, where we defend our headline conclusion that injustice undermines the progressiveness of social experimentation.

In Part VI, we take stock, drawing out some implications of our analysis. One implication is that, if injustices are too severe, social experimentation may fail to make things better or may even make things worse. This is a crucial limitation that has been overlooked by advocates of social experimentation. But even in reasonably just societies, where experimentation is on balance progressive, injustice reduces the effectiveness of social experimentation as a means of achieving progress not only in justice, but also with respect to other social goods. Although this needn't lead one to abandon a commitment to experimentation, it should encourage more realistic expectations as well as careful reflection on how we might mitigate the problems in question. It also suggests an important practical upshot. To render social experimentation a more reliable engine for social improvement of any sort, it is imperative that we remedy or mitigate existing injustices. So, no matter one's normative priorities, there is a strong instrumental case that remedying injustices in the short term will help one to achieve those priorities in the long run.

## II. What is Social Experimentation?

Before proceeding, it is important to clarify terms. We can distinguish between the experimental stance and the deliberate production of experiments. To adopt the experimental stance is to regard our current opinions about the desirability of different policies or institutions as hypotheses that are subject to critical appraisal and possible abandonment or modification in the light of empirical evidence—and actively to seek such evidence. One can adopt the experimental stance without engaging in the deliberate production of experiments, that is, without designing and constructing conditions in which such hypotheses will be confirmed or disconfirmed by evidence—evidence that

would not likely be available had those conditions not been constructed.

In this paper, our focus is on the deliberate production of experiments. This is what we mean by “social experimentation.” The thesis of progressive experimentalism we will investigate is not merely that it is wise to take the experimental stance and view existing and past arrangements as “natural experiments,” but that the deliberate production of experiments fuels social improvement. However, it is worth distinguishing two quite different ways of deliberately producing experiments. First, a society might produce experiments in a direct or “top-down” way, as when officials at the highest-level political authority design and impose different experimental policies in different places in order to test and compare their effects. Second, a society might produce experiments in an indirect or “bottom-up” way by deliberately encouraging or incentivizing experimentation, for example, by delegating authority to different political or administrative bodies in a way that predictably leads them to try out a wide range of policies. Although there are important differences between these approaches, our primary concern throughout this paper is the deliberate production of experiments in general, rather than any particular regime of social experimentation—be it highly direct or highly indirect.

We also emphasize that our focus here is not on personal experiments that individuals undertake in their own lives but is rather on social experiments—experiments with different policies or institutions, which we understand broadly to encompass not only randomized controlled trials but also what are sometimes referred to as “quasi-experiments” because they lack either randomization or a control group.<sup>3</sup> Our concern, in other words, is not with individualistic experiments in living but rather with experiments in politics or, if one prefers, experiments in living together. So, while the next two sections will draw on analogies between personal and social experimentation, this is only for

---

<sup>3</sup> See, for example, Shadish, William R., Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Boston: Houghton Mifflin, 2002)

purposes of illustration. The question we will investigate is whether deliberately producing social experiments is a reliable means of producing social progress. Here, we understand “social progress” broadly to encompass the achievement or enhancement of what are widely accepted as social goods, including justice as well as improvements in well-being, freedom, efficiency, and the like.

### III. The Case for Progressive Experimentalism

The root assumption that motivates progressive experimentalism is the idea that achieving progress isn't simply a matter of implementing what we already know to be better states of affairs. Rather, we have much to learn, not only about how to achieve improvement as we currently understand it, but also about what should count as an improvement. For the experimentalist, then, making progress is not simply a matter of putting into practice what we know to be desirable. It is also a matter of finding out what is desirable by engaging in practice—of learning through trial and error.

When we are uncertain, both about what we should seek to achieve and about how to achieve it, we face a tradeoff between exploitation and exploration: between exploiting our current understanding of how to achieve greater value and exploring the space of possibilities to find better ways of doing so. This tradeoff is familiar from daily life. Suppose you are at a restaurant, deciding what to order. Last time, you had the Pad Thai, and you liked it very much. So you are tempted to order it again. But you are unfamiliar with Thai food, and there may be other dishes you would like even more, if only you were to try them. If you always order Pad Thai, you will never discover these options. So you are tempted to try something else instead.

If you settle on the Pad Thai, you are *exploiting* your current information about how likely you are to enjoy the various items on the menu. If you try something else, you are *exploring* in the hopes of discovering something better (for example, that you prefer coconut curry) or at least of obtaining general information that will help you to make better choices in the future (for example, that curry is

not for you). Although exploiting has greater expected value if we focus only on its immediate effects, in many contexts, people who always exploit get stuck choosing options that, were they to have explored more, they would have forgone in favor of something else. In the long run, they achieve less than they would had they sometimes explored. But people who persist in seeking new information through exploration, while never putting that information to use, also achieve less than they would if they limited their exploration in order sometimes to exploit what they already know. So exploitation and exploration must be balanced against each other. This is the exploration/exploitation tradeoff.

There is nothing special about the above example. The same tradeoff arises in any context where we have incomplete information about the value of options, and where trying out various options can teach us more about their value in a way that informs our future choices. A classic illustration is the “multi-armed bandit problem.”<sup>4</sup> You face some number of slot-machines which you may use some number of times. You want to get as much money as possible, but you are uncertain about each machine’s cash payoff. What should you do? One approach is always to use the machine that you currently believe has the highest expected cash payoff. This amounts to always *exploiting*. But there is a good chance that there are other machines with an even higher payoff, so you will typically do better sometimes to *explore*: to try out a machine you currently expect less from, but which you could discover has a higher cash payoff upon trying it. Indeed, for a wide range of parameters, agents that play as simple a strategy as exploiting the machine they assign the highest expected cash payoff to most of the time but occasionally choosing some other machine at random tend to get much richer, over the long run, than those who always exploit. More sophisticated strategies—for example, positively correlating the frequency with which one uses a machine with its expected cash payoff, but

---

<sup>4</sup> For a helpful introduction, see Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning*, 2nd ed., (Cambridge: The MIT Press, 2018), ch. 2.

never abandoning any machine altogether—often do even better.

Our basic argument for the long-term benefits of social experimentation, then, is first, that balancing exploration against exploitation produces greater long-run value than always exploiting in contexts of uncertainty where we can learn through trial and error about the value of the options we will later face; and, second, that the pursuit of social improvement is one such context. We now consider an example that illustrates the second point.

Suppose you are a public official with a clear mandate and broad authority to tackle some social problem, say, to reduce poverty or ameliorate racial discrimination. You have surveyed the available theories, models, and evidence bearing on various alternatives and now believe that some particular reform would best serve this aim. If your sole goal is to mitigate the social problem, it might seem that what you ought to do is obvious: you should implement that reform wherever possible. But this would be only to exploit, and so may lead you to miss out on better options in the long run. If instead you were to explore, or to engage in social experimentation—say, by trying out a variety of reforms in a variety of places—you might very well discover that some other reform is in fact better, or at least learn something that leads you to revise your theories or models so that you can come up with better reforms in the future.

Perhaps, for example, your current evidence suggests that a suite of means-tested benefit policies best reduces poverty, but if you give some weight to exploration, you might discover that a basic income policy, say, does much better. Or perhaps your choice between reforms depends in part on your appeal to a model on which providing individuals with unconditional benefits reduces their motivation to work, and social experimentation would allow you to test and calibrate this model, which would in turn facilitate the development of better reforms going forward. More fundamentally, experimentation might lead you to reconceive the problem you are addressing. For example, you might come to believe that the major cause of some form of racial discrimination is institutional and that



priority should be given to institutional reforms rather than to educational efforts to combat individuals' racial prejudices.

In the cases considered so far, what we hope to achieve through exploration is the better realization of antecedently specified goals. In particular, we can obtain either information about how well the specific options we have tried achieve these goals or more general information that bears on the evaluation of other options we may implement in the future. This alone would be enough to demonstrate the long-term value of social experimentation. But it is worth stressing that in social contexts, we similarly face uncertainty regarding the ends we should be striving for, and that exploration can also help us to resolve such uncertainty.<sup>5</sup>

That exploration has this further benefit follows from one additional premise, namely, that the full or partial realization of goals in practice provides us with evidence not only about how better to realize them, but also about whether they are indeed worth realizing, how much weight they merit, or what other goals we should pursue instead. This can occur in a number of ways. For example, if a social experiment produces an outcome that our current conception of morality registers as an unambiguous improvement, but those living under the experiment credibly protest their new conditions as unjust, their testimony may lead us to rethink our moral stance. Alternatively, a social experiment might produce some new outcome that appears either highly morally desirable or undesirable to us upon reflection, but that we otherwise would not have thought to consider (or whose importance we wouldn't have appreciated when considered only in the abstract). It is also common

---

<sup>5</sup>This is arguably a point at which the experimentalism of Popper departs from that of Mill and Dewey. For Popper (and his followers) the primary benefit of social experimentation is the generation of instrumental knowledge; for Mill and Dewey (and their followers), experimentation also helps us to revise our ends. Thanks to Joshua Forstenzer for helpful discussion.

to find that those who live under new arrangements are forced to confront various preconceptions—say, that members of some groups have lower moral status than others, since they are naturally inferior and so would not take advantage of greater educational or economic opportunities even if they were available—and so to revise their moral beliefs accordingly.<sup>6</sup>

We take the occurrence of this sort of moral learning (if not the particular examples) to be uncontroversial. Indeed, we find it plausible that social experimentation can promote learning not only with respect to the sort of mid-level or proximate goals we typically aim to pursue through policy or institutional design, but even with respect to fundamental moral principles: even if, at a foundational level, values float free of facts, it is widely accepted that the plausibility of values depends in part on their concrete implications in practice, and that we understand such implications better the more experience we have with them.<sup>7</sup> Regardless of whether experimentation can teach us about fundamental principles or only mid-level goals, however, the lesson is clear. Exploration not only helps us better to pursue, and better understand how to pursue, the goals we accept. It also helps us to understand what goals we should be pursuing.

There remains an important disanalogy between the examples we have discussed and real-life contexts of social reform. In the above cases, choices are made by a single individual, but in actual practice, social change is a complex process over which no individual has complete control. However,

---

<sup>6</sup> For a careful discussion of how empirical observation may inform our moral beliefs, on which we have drawn here, see Sarah McGrath, *Moral Knowledge* (Oxford University Press: 2019), ch. 4.

<sup>7</sup> Thus, even G. A. Cohen, “Facts and Principles,” *Philosophy & Public Affairs* 31(2003): 211-245, 227, who famously argues that fundamental normative principles are “fact-insensitive,” affirms that facts about how principles play out in practice crucially inform our evaluation of such principles, since “responses to actual facts reveal our principles better than our responses to hypothesized facts do.”

the exploration/exploitation tradeoff does not only arise in contexts of individual decision-making; it also arises at the system level. Indeed, two major literatures on the exploration/exploitation tradeoff concern how to structure firms so that they balance exploration against exploitation, and how complex adaptive systems more generally strike the same balance.<sup>8</sup> As these literatures reveal, contexts involving multiple decisionmakers often provide additional methods for resolving the tradeoff, since they do not require single agents to internalize the tradeoff in their own decisions. Instead, the tradeoff may be handled at the system level, if some devote their time to exploiting and others to exploring, or even if decisionmakers each engage in what they see as exploitation of existing information in a context where this results in a wide range of options being tried out because decisionmakers have different information or disagree about what courses of action their information supports.

Similarly, when it comes to social improvement, we need not think of experimentation as being conducted by a centralized experimenter. Instead, experiments may be overseen by the decisions of a large number of voters, whose representatives engage in experiments in policymaking—say, through widespread use of pilot programs, or the employment of social scientific advisors who help them to design and evaluate policy experiments in ways that better facilitate learning.<sup>9</sup> Or, as we noted above, we might achieve the benefits of social experimentation more indirectly, for example, through establishing a “polycentric democracy” in which different political jurisdictions, though each operating under the same overarching democratic framework, each serves as its own locus of decision-making about which reforms to experiment with, such that there is widespread variance in which reforms get

---

<sup>8</sup> For pioneering discussions, see, respectively, James G. March, “Exploration and Exploitation in Organizational Learning,” *Organizational Science* 2(1991): 71-87 and John H. Holland, *Adaptation in Natural and Artificial Systems* (Cambridge: MIT Press, 1992).

<sup>9</sup> Compare Donald T. Campbell, “Reforms as Experiments,” *American Psychologist*, 24(1969): 409–429.

tried out.<sup>10</sup> Another option is to establish an “experimentalist governance” framework in which local agencies have autonomy to implement their own policy experiments in pursuit of common goals, on the condition that they participate in a collective process of subjecting the results of their experiments to peer review and revising them (and their common goals) accordingly.<sup>11</sup>

Our aim here is not to compare the merits of these different frameworks for social experimentation, but merely to make plausible the claim that social reform is a context, like many others, in which exploration can and should be balanced against exploitation. The key point is that without exploration, we will often fail to discover better options, and will obtain less general information bearing on the evaluation of options we have yet to try. This is the basic argument for progressive experimentalism: social experimentation is an engine for long-term social improvement for the simple reason that exploration generally conduces to long-run value.

#### IV. Four Challenges to Progressive Experimentalism

The argument we have just sketched is *prima facie* persuasive, but it rests on several assumptions that warrant further scrutiny. In this Part, we outline four such assumptions, explain how each raises a distinct challenge to progressive experimentalism, and illustrate how these four challenges provide a helpful framing for further research into social experimentation. In the next, we narrow our focus to how injustice exacerbates all four challenges.

The first assumption is that we will indeed learn valuable information from social experimentation—information that we can put to use or “exploit” in pursuit of better options going forward. But this is not obvious. If I never go back to a Thai restaurant, then exploring the menu and

---

<sup>10</sup> Müller, *Political Pluralism, Disagreement, and Justice*.

<sup>11</sup> Sabel and Zeitlin, “Experimentalist Governance.”

learning more about what dishes I like will not help. Or, even if I learn something about the relative tastiness of various dishes prepared in one Thai restaurant, this might not generalize to others. And, similarly, one might worry that experiments lack external validity: that we cannot learn much from how well a social experiment works in one context about how well it will work in other contexts, or once circumstances change, if its effects depend on the complex interplay of factors that differ from place to place and time to time. Or perhaps biases at the individual or institutional level will contaminate the interpretation of whatever valuable information experiments would otherwise yield. This is the *information problem*, and it calls into question the value of social experimentation as an informative form of exploration on which the basic argument for progressive experimentalism relies.

Second, the basic argument presupposes that, in practice, social experiments will be undertaken in a way that genuinely helps us to explore regions of social space that are relevant to social improvement. But this may not be how things play out. If our dining club's decisions about what to order are determined by majority rule, and if the majority has a bias against food that sounds too "ethnic," then we may end up only exploring the kids' menu rather than the richer array of options available. And a similar problem may arise in the social case. A society that institutes a regime of social experimentation may fail to undertake certain valuable experiments, not because they are less important than others, but for morally arbitrary or pernicious reasons. So even if social experiments can be informative, the sort of experimentation we actually undertake may differ starkly from experimentation that would genuinely inform the pursuit of social progress. This is the problem of *selection bias*—bias regarding which experiments are undertaken.

Third, our argument for social experimentation only succeeds if the information generated through such exploration is actually exploited—that is, put to good use in determining which further reforms to undertake. But we cannot assume this will occur. Even if we successfully explore a wide range of options at a restaurant, this exploration may lead us to order food that fails to better satisfy,

or even worse satisfies, our preferences in the future, if our decisions about what to eat are determined randomly or by a socially dominant individual whose tastes negatively correlate with the preferences of the rest of the group. And, in the same way, social experimentation may fail to lead to improvement, or may even lead to regression, depending on whether the information gathered from it is employed in the pursuit of good or bad ends. This is the *uptake* problem, and it is especially troubling in the following sense: it suggests that even under conditions where social experimentation is undertaken in a genuinely informative way, it may fail to be beneficial, or may even make things worse.

Finally, our argument presupposes that the risks of experimentation are not so great that we would be better off avoiding it. But this cannot be assumed either. If I have a life-threatening allergy, and I don't know which dishes might trigger it, then I may be better off sticking with what I know is safe than trying out other options. And likewise, social experimentation might be too risky—either because experiments threaten to have terribly negative consequences in the aggregate, or because such risks are too unjustly distributed. This is the *risk problem*: if the risks are great enough, the benefits of social experimentation may not be worth the costs even when experimentation yields valuable information that is put to good use in promoting social improvement.

Of course, it is one thing to raise these four problems in the abstract, and another to investigate the extent to which they genuinely undermine the progressiveness of experimentation in the real world. This is an important task on the agenda for future research on social experimentation. In fact, we believe that the four challenges provide a helpful framing not only for evaluating the basic case for progressive experimentalism, but also for comparing specific frameworks for social experimentation.

For example, one plausible hypothesis is that more top-down frameworks do better on the information problem, since they may be able to more carefully design experiments in ways that produce valuable and generalizable information, say, by running a range of subtly different experiments in different jurisdictions while leaving others as “controls.” But at the same time, another

promising hypothesis is that more “bottom-up” frameworks generally have an easier time mitigating the selection bias problem: in a polycentric system where diverse jurisdictions each have authority to engage in their own reforms, we will likely have a wider range of experiments, at least if such jurisdictions differ considerably in their decision-making procedures or goals.<sup>12</sup>

Similarly, when it comes to the uptake problem, competing hypotheses present themselves. On the one hand, more top-down frameworks may have an easier time rapidly implementing policies that experiments have shown to be valuable.<sup>13</sup> But on the other, this very capacity may be a curse if the central government is prone to capture or to pursuing objectionable goals (such as increasingly effective surveillance of its citizens). And when it comes to the risk problem, the case again seems mixed. Two domain-general risk-reduction heuristics are containability and reversibility.<sup>14</sup> If the effects of an experiment will be limited to one locale, then, other things being equal it is less risky because whatever damage it produces will be relatively limited. If an experiment is reversible—if it can be halted fairly promptly and in such a way that its effects largely cease—then, other things being equal, it is less risky. A plausible hypothesis here is that bottom-up frameworks produce more containable experiments,<sup>15</sup> but that top-down frameworks are better at reversing damaging experiments. If so, it

---

<sup>12</sup> Müller, *Political Pluralism, Disagreement, and Justice*, ch. 10.

<sup>13</sup> Compare Daniel A. Bell’s, *The China Model* (Princeton: Princeton University Press, 2015), 134 defense of China’s political system on these grounds: “In a federal system, for example, what works well in one subunit cannot necessarily be generalized to the rest of the country if the central government does not have the constitutional power to do so. But the power to spread desirable local innovations is an important advantage of Chinese-style political meritocracy.”

<sup>14</sup> Allen Buchanan, *Beyond Humanity?* (New York: Oxford University Press, 2011), ch. 6.

<sup>15</sup> Müller, *Political Pluralism, Disagreement, and Justice*, 183.

is far from obvious which sort of framework produces less risks on net.

Furthermore, we believe that the four challenges may help us to frame and investigate other objections to progressive experimentalism than those we can focus on here. For example, one possible objection is that a society might simply fund more social science research—say, theory and model construction and laboratory experiments—as a less costly substitute for social experimentation. Although we are unpersuaded by this objection because we believe that social experiments are needed to test theories and models and to determine whether the results of laboratory experiments extend to the real world, this ultimately turns on the information and risk problems: in principle, the information value of social experiments might be sufficiently low that the information they provide over and above social scientific research isn't worth the risk.

#### V. How Injustice Exacerbates all Four Problems

The last few paragraphs represent some first thoughts intended only to illustrate the value of having our four challenges in mind when theorizing about different frameworks for experimentation or progressive experimentalism in general. For present purposes, however, we will not attempt to address such issues in anything like a comprehensive matter. Rather, in the remainder of this paper we will use these challenges as a lens through which we can investigate the ways that background injustices may undermine the case for progressive experimentalism. Although injustices are not the only sources of the problems, we think they are serious and underexplored. Indeed, as we will now explain, injustices may exacerbate all four of the problems we have mentioned.

##### *The Information Problem*

The primary benefit of social experimentation is epistemic: it can generate information that bears on how well particular experimental reforms work out, or that is useful for assessing the plausibility of



relevant social theories, models, or standards of evaluation (including principles of justice). However, since societies are highly complex, it is typically far from obvious what lessons to draw from an experiment.<sup>16</sup> In general, the effects of an institution or policy depend not on a simple one-to-one causal relation, but on its interaction with other policies or institutions, as well as with other background conditions. So, although an experiment may appear to show how a desired outcome may be produced, that outcome may instead be the result of other factors or may not be achievable through similar means in other contexts.<sup>17</sup>

This difficulty is exacerbated by the normal human proclivity to bias. Observations are notoriously theory-dependent, and we are all subject to confirmation bias and motivated reasoning. So, two people with different prior beliefs or interests might disagree about how well an experiment has played out, or even if they agree it has failed or succeeded, they might disagree about why this is so, and so each take it to confirm their prior expectations or self-serving hopes. If they are in the grip of rival ideologies, such disagreements may be even more pronounced—and harder to resolve. The point is that information generated from social experimentation is hard to gather and interpret,

---

<sup>16</sup> For a powerful recent articulation of this worry, see Gerald Gaus, *The Open Society and its Complexities* (New York: Oxford University Press, forthcoming), Part III.

<sup>17</sup> Nancy Cartwright and Jeremy Hardie, *Evidence-Based Policy* (New York: Oxford University Press, 2012) provide an overview of the difficulty of determining whether an experiment that works in one context will work in another. For a recent survey and analysis of evidence from international development, see Eva Vivalt, “How Much Can We Generalize from Impact Evaluations?” *Journal of the European Economic Association* 18 (2020): 3045-3089.

rendering misinterpretation likely given cognitive biases and the distortions that ideologies foster.<sup>18</sup>

This is where injustice enters the picture. In the Marxist tradition of ideology critique as developed by the Frankfurt School and by contemporary critical theorists, ideologies function to uphold unjust social orders.<sup>19</sup> Although the details of this picture are controversial, the basic idea that less just societies tend to contain more distortive ideologies is fairly commonplace—we need not defend it, for example, by reference to the functionalist claim that ideologies flourish *because* they function to support injustice, but might merely note that this correlation appears to hold empirically, perhaps because ideologies paradigmatically represent the status quo as fundamentally just, and it takes a greater distortion to obscure larger departures from justice. As a result, we should generally expect the problem of ideological bias in the interpretation of the results of experiments to be graver in less just societies. Importantly, however, ideologies do not only include beliefs that disguise or deny existing injustices. They also feature “belief immune systems”: patterns of thinking, and in particular, epistemically flawed cognitive dissonance resolution mechanisms, that cause their adherents to ignore or discount information that contradicts any of their connected beliefs.<sup>20</sup>

If ideologies function to support existing injustices, then one would expect the ideological belief immune system to be especially potent in screening out or discounting information that, if acted on, would challenge existing injustices. If that is the case, then in the very circumstances in which experiments are most needed to find ways of reducing injustice, they are likely to be of limited value

---

<sup>18</sup> Elizabeth Anderson, “The Quest for Free Labor: Pragmatism and Experiments in Emancipation,” *The Amherst Lecture in Philosophy* 9 (2014): 1-44 explores other biases that may undermine the information value of experiments.

<sup>19</sup> Raymond Geuss, *The Idea of a Critical Theory* (Cambridge: Cambridge University Press, 1981).

<sup>20</sup> Allen Buchanan, *Our Moral Fate* (Cambridge: MIT Press, 2020), ch. 8.

for doing so because the interpretation of their results will be biased toward the perpetuation of injustice. More generally, so long as ideologies tend to be more prevalent and distortive in less just societies, the risk of experiments being misinterpreted will be greater, other things being equal, in less just societies. Such ideological biases can prevent experiments from yielding information that is useful for making improvements of other sorts, not just those regarding justice.

A related problem is that the information value of experiments may be undercut by “testimonial epistemic injustice”: the testimony of some people may be discounted or disregarded entirely, simply because they are identified as members of groups that are regarded as inferior, and this includes their testimony about the effects of social experiments.<sup>21</sup> In some cases, members of such groups may be best positioned to make an accurate appraisal of the information yielded by an experiment, yet their voices may be ignored. Or, because they are generally excluded from the predominant practices of knowledge production in their society, they may not even have the opportunity to evaluate the results of an experiment. In either case, valuable information about the results of an experiment, especially as it impacts disadvantaged groups, may not be made available. Much like ideological distortion, epistemic injustice can therefore exacerbate the information problem.

#### *The Selection Bias Problem*

Even if experiments can be undertaken in an informative way, unjust societies that have arrangements in place to encourage experiments and extract valuable information from them may not undertake the full range of experiments needed to yield relevant information about how to remedy injustice or other social problems. This may occur for a number of reasons, including lack of effective advocates. Those most likely to benefit from an experiment and to see the value of undertaking it may, because they are

---

<sup>21</sup> Miranda Fricker, *Epistemic Injustice* (New York: Oxford University Press, 2007).

the victims of injustice, lack the political, material, or educational resources needed to conduct the experiment or to mobilize others to conduct it.

Alternatively, an unjust social order may, through the power of the ideology it contains, encourage the victims of injustice either to fail to recognize the injustices they are suffering or to discount their seriousness. This, again, is often seen as the function of ideologies: to uphold an unjust status quo by distorting people's beliefs and perceptions so that they fail to appreciate the injustice they are subject to. If one fails to recognize an injustice or underestimates its seriousness, then one may not even get to the point of thinking about experiments that might yield knowledge of how to overcome it. Valuable experiments will not be undertaken because no one sees the need for them. In each case, the problem is that the very injustices that experimentation is supposed to help us overcome can limit the value of experimentation by preventing informative experiments from being conducted.

Putting this point more generally, if the distribution of resources and power in a society is characterized by unjust inequalities, then there will also be injustice in the production of experiments—regardless of whether the experiments themselves pertain to injustice.<sup>22</sup> Those who are systematically disadvantaged by social practices and institutions will also be at a disadvantage when it

---

<sup>22</sup> As we will go on to argue, inequalities in resource and power also generate what we will call the uptake problem. Our emphasis on power inequalities echoes a common objection to John Dewey, namely, that he does not adequately take power into account when defending his experimental account of democracy. Whether Dewey himself was so blind to power is a subject of ongoing scholarly debate. For one explicitly Deweyan approach that aims to accommodate such concerns, see, Knight and Johnson, *The Priority of Democracy*, who argue that departures from their principle of “equal opportunity of political influence” interfere with the proper functioning of the democratic process—which they interpret in Deweyan, experimentalist terms.

comes to proposing and gaining support for experiments. The process of social experimentation may be captured by elites who use it to limit experimentation to issues where greater information would serve their particular goals. An analogous phenomenon is familiar in the case of medical experimentation: resources are disproportionately funneled into research on “diseases of affluence” rather than more widespread and serious medical conditions that are common among the worse off.

Ironically, other things being equal, the selection bias problem is more serious, the greater the need for improvements in justice. This is because the very flaws—and in particular the maldistribution of resources and political power and influence—that are most in need of reform are likely to prevent valuable experiments from being conducted.<sup>23</sup> However, it is worth pausing here to note an important point. Although the information and the selection bias problem reduce the benefits of social experimentation, the case for progressive experimentalism can succeed even under social conditions where these two problems are only very imperfectly mitigated. So long as they are engaging in *enough* experiments, and these experiments are informative *enough*, societies that experiment may still be more progressive than those that don't. This assumes, however, that relevant information will be put to good use, and that, even if it is, offsetting costs are not too significant. The uptake problem and the

---

<sup>23</sup> It might be thought that there will be more experimentation in seriously unjust societies, because people will be more receptive to the idea that innovation is needed; there may be low-hanging fruit, or cases of patently unjust practices or policies that are widely recognized as such. This will not occur, however, if ideology prevents people from perceiving how unjust their society is or that change is feasible. And even if there is dissatisfaction with the status quo, it will not result in experimentation that seriously challenges the status quo if it is only the victims of injustice who are dissatisfied and the unjust distribution of power renders their dissatisfaction impotent. In any event, once the lowest-hanging fruit have been harvested, the problem of selection bias may be quite formidable.

risk problem respectively call these assumptions into question.

### *The Uptake Problem*

Recall that the basic case for progressive experimentalism was disarmingly simple: we must balance the exploitation of existing information about how to improve our society against the exploration of alternatives in order to generate information about how to improve our society going forward. So far, we have considered two worries about the extent to which social experimentation is likely to proceed in an adequately informative way, as the “exploration” side of the tradeoff requires. But there is a further worry about whether such information will be appropriately exploited. Unless the information generated by social experimentation is put to good use—instead of sitting idle or being put to nefarious purposes—social experimentation will not conduce to social improvement.

To appreciate this problem, suppose we live in a society that encourages a wide range of social experiments by giving different jurisdictions the authority to try out their own reforms in pursuit of various social goals. Suppose also that, through experimentation with different interventions in different jurisdictions, it becomes clear that some genuinely help to improve our society, while others make things worse while promoting the interests of undeserving social elites. This knowledge will only lead to social improvement if it is exploited in pursuit of such improvements. Among other things, this means that there should be some positive correlation between the discovery of better alternatives and their uptake: if experimental reform  $x$  but not  $y$  improves society, then  $x$  should have a greater tendency than  $y$  to spread to other jurisdictions that would also benefit from it. But if the elites in question have inordinate influence on this process, this may result not in reforms that make society better, but instead in interventions that benefit the elites at the expense of social value. In this case, we may even find a negative correlation between the discovery of better alternatives through experimentation and their uptake. Social experimentation may be rendered an engine for sustaining

the status quo or even regression rather than for progress.

Another way to put this point is in terms of cultural selection. Cultural selection occurs when three conditions are satisfied: there are variants in cultural phenomena, these variants can be transmitted over time, and there is some mechanism or “filter” that prevents some variants from being transmitted and allows others to be transmitted. With some simplification, we can say that cultural selection occurs when some variants outcompete others because they are more amenable to transmission according to the existing mechanisms of transmission.<sup>24</sup> Although social experimentation might seem to speed up the rate of improvement by introducing more variants on which cultural selection can operate, the worry is that unjust power inequalities may introduce a severe bias into our mechanisms of cultural transmission from the standpoint of what we ought to value. Policies may be sustained and spread simply because they are favored by those who have disproportionate control over public policy or educational practices or the deployment of information technologies.

Injustice may therefore undermine our ability to exploit information about how well particular policies work in ways that are conducive to social improvement. A related problem concerns the exploitation of more general information gleaned from social experimentation about the likely effects of other policies. The problem is that such information is typically just as useful to those who wish to improve their own position at the expense of others as to those who wish to use it to improve society. The rich and powerful may capture the policy process and use this information to line their own pockets or to further entrench their power. Because such information is itself a valuable resource, social elites may also have greater access to such information, and so may be able to successfully disguise such policies as attempts to genuinely promote socially valuable ends.

---

<sup>24</sup> For a helpful overview aimed at non-specialists, see Peter J. Richerson and Robert Boyd, *Not by Genes Alone* (Chicago: University of Chicago Press, 2005).

The uptake problem is not, of course, uniquely a problem for the progressive experimentalist. It is a more general problem about how information is put to use in society, regardless of that information's source. However, we think it important to emphasize this problem here for two reasons. First, as we have noted, in cases where the uptake problem is sufficiently severe, it not only reduces the value of experimentation, but can actually render experimentation detrimental. The problem is not just that information generated through experimentation may fail to be put to good use, but that it may be put to bad use. Second, it is all too tempting to think that in the face of a problem that we are unsure how to solve, we simply need to try out various policies to find out which works best. The uptake problem reminds us that a crucial step is missing in this picture: someone must decide which policy to implement, and they will not necessarily pick the best option. There are many familiar worries about how, under conditions of inequality, those with greater power can disproportionately influence the policy process. None of these worries go away simply once we adopt an experimentalist approach.

#### *The Risk Problem*

When a society is able to solve or at least adequately mitigate the information problem, the selection bias problem, and the uptake problem, social experimentation has important benefits: it generates valuable information about how to improve society that is exploited to this end. But this does not yet show that social experimentation is on balance progressive, because it may generate costs that outweigh these benefits. Here, we do not merely have in mind the tradeoff at the heart of the case for progressive experimentalism: that exploration always involve some short-term opportunity cost of foregone exploitation. Rather, we are concerned with the specific worry that experimentation may bring risks of severely negative or unjust consequences that cumulatively outweigh the potential benefits of such experiments, even in the long run.

The risk problem is challenging in part because solving it requires more than resolving to



refrain from excessively risky experiments. The risks of social experimentation—both the magnitude of possible harms and the probability that any such harm will occur—are typically difficult to ascertain, even by the most knowledgeable and well-motivated of experts, simply due to the complexity of the phenomena involved and the limits of our knowledge. Even genuine experts with the best of intentions may err in estimating risks. An unacceptably risky experiment might therefore be undertaken because information about the risk of harmful consequences was not available or was not heeded; hence the problem of unintended bad consequences.<sup>25</sup> But there are other sources of faulty risk assessment. For example, the risks of an experiment might be underestimated or overlooked entirely by those who are most likely to propose the experiment and expend political capital mobilizing support for it, because confirmation bias or motivated reasoning leads them to focus too much on favorable rather than unfavorable outcomes.

This problem is exacerbated by injustice because, in many cases, the benefits and risks of an experiment will be unevenly distributed.<sup>26</sup> If the people who have the most influence over whether an experiment is undertaken have the most to gain if it works out, or the least to lose if things go wrong, the result may be an underestimation of risk. In societies characterized by unjust power inequalities and epistemic injustice, the predictable result will be communities that lack political or epistemic influence disproportionately bearing the cost of social experiments, even when benefits are spread more evenly. So existing injustice may result in the undertaking of overly risky experiments, or in

---

<sup>25</sup> Indeed, in contexts of social experimentation we may be subject to “Knightian” uncertainty in the sense that we cannot estimate the probabilities of different consequences accruing. See Frank H. Knight, *Risk, Uncertainty and Profit* (Boston: Houghton Mifflin, 1921).

<sup>26</sup> Knight and Johnson, *The Priority of Democracy*, 48-49 raise the more general issue of justice in the distribution of the costs and benefits of experiments.

experimentation that itself perpetuates injustice in the form of unjustly distributed risks.

Another aspect of the risk problem is that social experiments might carry with them moral risks in the sense that they may violate independent moral requirements. In the case of medical experimentation, there is widespread agreement that informed consent is necessary to justify an experiment—but such consent may be hard to come by in the case of social experiments. A convincing defense of social experimentation would either have to accept the informed consent requirement and acknowledge corresponding limitations on social experimentation; would have to show that such limitations are not as severe as they might seem, perhaps because certain regimes of social experimentation allow individuals to opt-out of experiments by moving away from them;<sup>27</sup> or would have to show why informed consent is required for medical but not social experiments.<sup>28</sup> One promising line of argument for this last conclusion is that democratic processes can take the place of consent, when it comes to policies. Indeed, democracy can be seen as a mechanism for achieving policy decisions under conditions where the requirement of unanimous consent would be unworkable.

Regardless of how one handles this question, however, such moral risks will generally be greater, as well as less justly distributed, in less just societies. This will be true given any of the three defenses of social experimentation we have mentioned: in the presence of unjust power inequalities, epistemic injustice, and ideological distortion, there is a greater risk that those with less political or epistemic resources will be experimented on without their informed consent; such individuals will be less likely to have the meaningful ability to opt-out of experiments (for example, since they will be less

---

<sup>27</sup> Müller, *Political Pluralism, Disagreement, and Justice*, 176-177

<sup>28</sup> On the question of whether informed consent is needed to justify social experiments, see Douglas Mackay and Averi Chakrabarti, “Government Policy Experiments and Informed Consent,” *Public Health Ethics* 12(2019): 188-201.

likely to have the resources or ability to move to a different jurisdiction); and the democratic process will be less likely to adequately take their interest or opinions into account in a way that would legitimize experimentation. This is yet another way that, in the presence of injustice, social experimentation risks perpetuating rather than mitigating injustice.

## VI. Implications and Conclusion

Recently, political philosophers have exhibited a resurgence of interest in social experimentation as an engine for progress. While we have argued that social experimentation can serve as a form of exploration that generates valuable information about how to improve society, we have also argued that it is subject to four major challenges: the information problem, the selection bias problem, the uptake problem, and the risk problem. The first two problems call into question how much valuable information social experimentation will generate in practice, the third whether this information will be exploited or put to good use in pursuit of social improvement, and the fourth whether the benefits of exploiting such information will outweigh the costs of social experimentation. We noted that the four challenges provide a helpful framing for future research into social experimentation, but then narrowed our focus to how injustice, in particular, may exacerbate all four problems: in less just societies, we should generally expect social experimentation to generate less valuable information, for the information it does generate to be less likely to be put to good use, and for experiments to impose greater and less justly distributed countervailing risks.

Of course, while we have identified a few plausible mechanisms by which injustice can exacerbate the four problems, the extent to which these mechanisms are genuinely operative in existing societies is a largely empirical question that requires further study. We will not attempt to answer this question here. Instead, we end by suggesting some general lessons to draw from our discussion about the role of social experimentation in promoting social progress.

One upshot of our argument is that, at least in principle, some societies may be so unjust that social experimentation would be detrimental in them. This can occur in either of two ways. First, the information problem and selection bias problem might render the process of social experimentation sufficiently uninformative, while the uptake problem renders this information sufficiently unlikely to be put to good use, that even though social experimentation would be on balance progressive if we ignored its risks, these risks indeed outweigh its benefits. This is a case where social experimentation has genuine benefits, but injustice diminishes these benefits, and exacerbates countervailing costs, to the point that the benefits are not worth the costs. Second, regardless of the other three problems, in the presence of severe injustice, the uptake problem may, all by itself, render social experimentation detrimental. This may occur in contexts where the political process has been so thoroughly captured by individuals pursuing goals that negatively correlate with justice and other social goods that additional information generated by social experimentation would only serve to help these individuals further entrench their power or better pursue their deleterious goals.

In either sort of case, the lesson is clear. If social experimentation is to be a path toward progress, the path must first be cleared by other means. But clearing this path will, of course, be especially difficult in precisely the sort of societies that are so unjust that social experimentation fails to promote progress in them. In these societies, it may take some sort of external shock, or a strike of good luck, to get things on track.

Thankfully, many societies are not so unjust. And in reasonably just societies, social experimentation may remain on balance progressive, despite the various problems we have raised. However, even here, we must worry about the extent to which injustice undermines the progressiveness of social experimentation, and so the reliability or rate at which it causes progress to occur. In these contexts, the general upshot of our argument is not that we should give up on experimentation altogether, but that we should moderate our expectations about how well social

experimentation will promote progress.

An additional implication of our analysis is that, when proposing social experiments or designing or reforming experimental frameworks, one should think carefully about the various ways in which social experimentation can go wrong in the presence of injustice, and whether it might be possible to mitigate the various problems we have raised. For instance, in a context where epistemic injustice is an especially potent worry, we might require that those charged with evaluating an experiment are drawn from diverse groups of society.<sup>29</sup> Alternatively, we might employ targeted techniques, such as the “snowball approach,” whereby those tasked with proposing or evaluating an experiment ask not only all the stakeholders that they can think of for input, but also ask those stakeholders who else they think should be consulted, and then ask those other people who else they think should be consulted, and so on, until, hopefully, all relevant stakeholders are consulted.<sup>30</sup>

Similarly, we might try to limit the impact of inequalities in wealth and political influence on biasing the selection of experiments. For example, this problem might be ameliorated by special subsidies, provided by philanthropic organizations, to be awarded to groups that propose experiments that would otherwise not be adequately funded.<sup>31</sup> Priority might be given to groups for whom it is

---

<sup>29</sup> Every contemporary proponent of experimentalism we cite in fn. 1 emphasizes (or at least mentions) the importance of including a diverse range of individuals in the experimental process. See, further, Scott E. Page, *The Difference* (Princeton: Princeton University Press, 2007) for a general discussion of the epistemic benefits of diversity.

<sup>30</sup> Peter H. Rossi, Mark W. Lipsey, and Howard E. Freeman, *Evaluation*, 7<sup>th</sup> ed., (London: Sage Publications, 2003), 87, 124.

<sup>31</sup> On the idea that philanthropic organizations may be especially well-positioned to fund social experiments, see Robert Reich, *Just Giving* (Princeton: Princeton University Press, 2018).

reasonable to assume that their lack of resources is due to injustice, since they may provide new information about how progress is achieved that would likely be unavailable without their input.

In the end, however, we worry that such mitigating strategies can only go so far. For instance, significantly reducing the negative impact of epistemic injustice on the information value of experimentation ultimately requires us to ameliorate the larger systemic injustices that foster epistemic injustice, and significantly reducing the negative effect of unjust power inequalities on the selection of experiments or the uptake of the information they generate requires us to reduce those inequalities. Doing so may require major changes in the structure of society.

This brings us to a final implication. If one accepts the progressive experimentalist argument that social experimentation can significantly promote social improvement not only in justice but also in other important domains, and if one accepts our argument that injustice can reduce the rate or reliability at which social experimentation has this beneficial effect, then one should recognize the long-run instrumental importance of ameliorating social injustice in the short term. In particular, our argument suggests that ameliorating present injustices can make social experimentation a more reliable or faster engine for promoting socially valuable outcomes, and so can help us to achieve whatever social values we have—whether they pertain to justice or not—in the long run.

If this is right, then our paper can be seen as providing an argument for the following convergence between the amelioration of existing injustices and the pursuit of long-run value: one important way to promote social value in the long term is to reduce social injustice in the short term. To the extent that social experimentation has the potential to promote long-term social progress, we should be especially concerned with ameliorating existing injustices in order better to achieve this potential.