

NONIDEAL JUSTICE, FAIRNESS, AND AFFIRMATIVE ACTION

Matthew Adams

AFFIRMATIVE-ACTION policies aim to increase the representation of a target group. If such policies try to realize this aim by giving preference to members of a target group, then public controversy is often aroused on the basis of a perceived unfairness.

This controversy is current again. In 2014, Students for Fair Admissions filed a lawsuit against Harvard University, alleging that the admissions preference given to African and Latinx Americans results in unfair discrimination against Asian Americans.¹ A US District Court ruled, in 2019, that Harvard does not discriminate in this way. But Students for Fair Admissions plan to appeal and it is expected that the case will eventually be heard by the Supreme Court.² Significantly, the changing ideological profile of the Supreme Court has led to speculation that “affirmative action could be dead not only at public schools but also at private ones whose practices have largely escaped legal scrutiny until now.”³

In addition to being politically pressing, affirmative action raises a parallel set of theoretical issues. The appeal of affirmative-action policies is that they can be an effective means of, at least partially, overcoming legacies of injustice. But they can also be contested because they require *prima facie* unfair treatment, at least in some instances. Affirmative action is, therefore, a good test case for the adequacy of a theoretical conception of justice. An adequate conception must specify the conditions (if any) under which affirmative action is just; a successful philosophical defense must explain how the unfairness objection can be overcome. The topic of affirmative action thus invites careful reflection on the nature of justice in unjust conditions. In particular, at least given the core commitments of nonconsequentialist liberalism that I presuppose, a compelling explanation as to why it is permissible for affirmative-action policies to treat

1 See Moses, “After *Fisher*.”

2 Following Anderson, “Federal Judge Rules Harvard Does Not Discriminate against Asian Americans in Admission.”

3 Gerstein and Haberkorn, “It’s Not Just Abortion.”

certain individuals in a *prima facie* unfair way is required. The mere fact that such treatment would be an effective means of realizing a more just society in the future is not a sufficient explanation.⁴

In this paper I argue for two related claims, one substantive and the other methodological. First, I forge a new justification for affirmative action, via a basic-liberties argument. And second, in doing so, I illustrate the value of a new conceptual innovation that I term “nonideal principles of justice.” More precisely, I defend affirmative action on the ground that it increases certain comparatively disadvantaged people’s ability to exercise their basic liberties. I argue that, given the particular empirical conditions that obtain in the contemporary US, this basic-liberties justification supports attaching special weight to being African or Latinx American in admissions procedures. Furthermore, it can provide a compelling response to the unfairness objection.

My approach has a certain affinity with so-called integrationist justifications, insofar as it construes the appropriate function of affirmative action as overcoming legacies of injustice rather than promoting diversity.⁵ But my argument does not presuppose that achieving racially integrated educational institutions is necessary for democratic legitimacy, or that racial integration is an imperative of social justice.⁶ Racial integration plays a function in my argument to the contingent extent that it is an effective and fair means of promoting certain people’s ability to exercise their basic liberties.

As I noted above, my argument has a second, methodological upshot. In order to tackle racial injustice in particular, a number of philosophers argue that the Rawlsian paradigm of justice should be abandoned. Elizabeth Anderson claims that Rawls’s theory of justice is inadequate because it focuses on ideal principles of justice, which specify how a perfectly just society should be arranged. She argues instead that philosophers should take an empirically informed “bottom-up”

4 In contrast, consequentialists might acknowledge that affirmative-action policies are unfair to certain individuals while arguing that, within certain empirical parameters, the benefits of affirmative-action policies outweigh this cost of unfairness. See Beauchamp, “In Defense of Affirmative Action.”

5 The Supreme Court has ruled that colleges can consider race to enhance student diversity. *Regents of the Univ. of Cal. v. Bakke*, 338 US 265 (1978) at 300. However, the diversity justification does not support anything like the scope of actual affirmative-action policies. Such policies standardly focus on promoting a racially diverse student body. But if the justification for affirmative-action policies is that they promote a diverse student body, then there is no reason why racial diversity should be the only type of diversity that is given significant weight. A commitment to diversity also supports attaching significant weight to other sources of diversity within the student body, such as creationism, Scientology, and climate-change denial. Following Anderson, *The Imperative of Integration*, 142.

6 See *Grutter v. Bollinger*, 539 US 306 (2003) at 336; and Anderson, *The Imperative of Integration*.

approach that addresses the pressing problems that they actually face, such as racial injustice in the US.⁷

I agree with Anderson that Rawlsian ideal principles of justice are not (at least in and of themselves) the best way of tackling pressing problems, such as whether affirmative action should be used to ameliorate racial injustice. But I argue that her bottom-up approach is also inadequate: it does not provide a sufficiently determinate conception of justice to overcome the unfairness objection. I show how a sufficiently determinate conception can be developed by using a Rawlsian contractualist framework to forge what I term a “nonideal principle of justice.”

My paper has the following structure. I begin by presenting the unfairness objection and clarifying its scope (section 1). I then examine an unsatisfactory response to the objection (section 2), and I survey the fertile—but limited—implications of Rawls’s theory of justice for affirmative action (section 3). After that, I forge a nonideal principle of justice (section 4) that supports affirmative-action policies like those in the contemporary US (section 5) and blocks the unfairness objection (section 6). I close by showing how my account can be used to refine some features of contemporary affirmative-action policies, and I reflect more generally on the value of nonideal principles of justice for tackling exigent topics (section 7).

A few preliminary clarifications are in order: I focus on affirmative action in a contemporary US educational context. In the US, the term “affirmative action” has been used to label a disparate set of policies. Such policies range from measures that simply outlaw group-based discrimination, to soft (i.e., non-explicit) quotas and hard (i.e., explicit) quotas for members of target groups.⁸ I put affirmative-action policies that simply outlaw group-based discrimination to one side because they can be given a straightforward defense: they block group-based discrimination and safeguard equality.

I use the term “ideal theory/justice” to refer to a conception of how a perfectly just society should be structured and “nonideal theory/justice” to refer to a conception of what justice requires in conditions that fail to realize ideal justice.⁹

My argument is exclusively forward looking; I am neutral about whether a backward-looking justification of affirmative action can also be provided.¹⁰

7 Anderson, *The Imperative of Integration*, 3–7. Relatedly, Charles Mills argues that a nonideal contract to end racial domination should supplant Rawlsian ideal theory (see *The Racial Contract*).

8 Following Nagel, “Equal Treatment and Compensatory Discrimination,” 349–51.

9 My explication of the distinction between ideal and nonideal justice follows Simmons, “Ideal and Nonideal Theory,” 7.

10 See Thomson, “Preferential Hiring.”

1. THE UNFAIRNESS OBJECTION

Many critics of affirmative action argue that such policies are *unfair*, at least in some instances of their application.¹¹ The following hypothetical example illustrates this charge of unfairness: a white man called Simon applies to Texas Law School and is rejected; Simon, however, would have been admitted but for a soft-quota affirmative-action policy that “added points” to favor the admission of a target African American group. A proponent of the unfairness objection can grant that African Americans are underrepresented in Texas Law School but urge that people like Simon are not directly responsible for perpetrating the historic conditions that led to such underrepresentation. It is not, therefore, fair for the admissions procedure to employ such a quota—even if this would be an effective means of realizing a more just society. This imposes the unfair burden of non-admission on Simon.

Ronald Dworkin argues that the unfairness objection presupposes a commitment to meritocracy: an affirmative-action policy is only unfair to Simon if he deserves to be admitted because he is an intellectually superior candidate. Such a presupposition is false, according to Dworkin, because no one deserves to be admitted to an academic institution because they possess some particular combination of talents.¹²

Note, however, that my intuitive presentation of the unfairness objection does not presuppose any particular independent standard of merit. Someone defending the unfairness objection can remain neutral about what standard or procedure (e.g., academic merit, or a type of lottery) should be used to determine admission. They are merely committed to the claim that it is wrong for something like membership in a particular race to have significant weight.

Simon’s predicament can be generalized into the following formulation of the conditions under which affirmative action can *prima facie* plausibly be contested as unfair, in any particular admissions procedure:

Membership in a target group—that is not directly relevant to an ability to complete/excel in the program of study—is given preferential weight. This results in the non-admission of a subset of people who are not members of the target group—who would have been admitted but for the affirmative-action policy.

Caveat: The preferential weight does not merely block/partially block the

11 See Cohen, “Why Race Preference Is Wrong and Bad,” 33–37; Lynch, *Invisible Victims*; Pojman, “The Case Against Affirmative Action,” 98–105; and Scalia, “The Disease as Cure,” 153–57.

12 See Dworkin, *A Matter of Principle*, 299–300.

discrimination that members of the target group standardly experience in the admissions procedure.

The caveat is necessary because people's biases against members of the target group might be so strong that the only way to cancel out (or reduce the effect of) such biases is to give preference to members of the target group. Essentially, an affirmative-action policy that cancels out, or reduces, an unfair advantage that Simon enjoys *qua* white male in the selection process is not unfair to Simon. (Moreover, *not* to cancel out that advantage is unfair to everybody who does not belong to his group.)¹³

Early social-scientific research focused on the burdens that affirmative action imposes on white men.¹⁴ More recent empirical research concludes that—at least in the context of admission to elite educational institutions—the burdens fall heaviest on Asian Americans. Thomas Espenshade and Alexandria Radford calculate that, *ceteris paribus*, an Asian American needs an SAT score 140 points higher than a white American and 450 points higher than an African American to have the same chance of admission.¹⁵

Yet, this empirical evidence is contested. In response to the lawsuit filed by

13 Some argue that the caveat needs to be expanded as follows: the preferential weight does not merely block the advantage that members of the non-target group standardly benefit from because of discrimination against members of the target group in the past, rather than in the present admissions procedure. See Boxill, "The Morality of Preferential Hiring"; and Thomson, "Preferential Hiring." This extension of the caveat faces the non-identity problem. See Morris, "Existential Limits to the Rectification of Past Wrongs." Furthermore, Kasper Lippert-Rasmussen argues that it is hard to cash out the relevant counterfactuals for this extension of the caveat to provide additional support for race-based affirmative action in the contemporary US. He notes that if there had been no past racial injustice, the US would now be a much richer society. This is because, for instance, slavery resulted in a suboptimal use of the large pool of talents among African Americans. Accordingly, if there was no past racial injustice there would have been a greater number of university places because the US would have been richer. Consequently, there would have been more African Americans in universities but *also* more white Americans in universities. "Hence, if beneficiaries are those individuals who are better off given the relevant past injustice than without it ... then there *might* be no beneficiaries of past injustice, even if some contemporary people have been harmed less than others" ("Affirmative Action, Historical Injustice, and the Concept of Beneficiaries," 82). I will show that the unfairness objection can be defeated, under the range of conditions that I specify, without expanding the caveat in this way.

14 See Lynch, *Invisible Victims*.

15 Espenshade and Radford, *No Longer Separate, Not Yet Equal*, 92. A simulation that determines the effect of race-based preferences at private institutions predicts that in 1997 Asian Americans would have comprised nearly 40 percent of all accepted students compared to less than 25 percent under current policies (Espenshade and Radford, *No Longer Separate, Not Yet Equal*, 344–46).

Students for Fair Admissions, Harvard University vigorously denied that their admissions policy made it harder for Asian Americans to be admitted. One way in which the evidence has been disputed is by arguing that, despite best efforts, empirical studies standardly fail to control sufficiently for variables such as legacy and athletic status.¹⁶ Essentially, the primary reason that it is harder for Asian Americans to be admitted is not the preference given to African American and Latinx American students but the preference given to predominantly white legacy applicants and recruited athletes.¹⁷ More generally, of course, it is difficult for even rigorous empirical studies to measure the variable of bias.

It is difficult to maintain, however, that no actual affirmative-action policies can *prima facie* plausibly be contested as unfair. After all, many admissions procedures are primarily dependent on standardized discrete data—the scrutiny of which leaves relatively little room for bias; for example, admission to law school is primarily dependent on an applicant’s LSAT score and undergraduate GPA. It is also important to highlight that a number of affirmative-action policies take a particularly strong form. In the University of Texas Law School’s affirmative-action policy under challenge in the case of *Hopwood v. State of Texas* the presumptive admit score for African Americans (a combination of undergraduate GPA and LSAT score) was lower than the presumptive deny score for white Americans.¹⁸ It seems very unlikely that this preference given to African Americans is merely blocking the unfair discrimination that they standardly face in the University of Texas Law School’s admissions procedure. Furthermore, perhaps some empirical studies of admissions procedures fail to control sufficiently for variables such as legacy status. Even still, legacy status has a comparatively minor impact in some educational admissions procedures, such as those in law schools.¹⁹

Accordingly, my argument presupposes the relatively uncontroversial claim that some actual affirmative-action policies can *prima facie* plausibly be contested as unfair under the conditions that I have specified.

16 See Espenshade, Chung, and Walling, “Admission Preferences for Minority Students, Athletes, and Legacies at Elite Universities.” In this study the authors tried to control for variables such as legacy preference.

17 See the review of Harvard’s admissions policy by the US Department of Education’s Office for Civil Rights: United States Commission on Civil Rights, “Civil Rights Issues Facing Asian Americans in the 1990s,” 104.

18 *Hopwood v. Texas*, 78 F.3d 932 (5th Cir. 1996).

19 See Schmidt, “A History of Legacy Preferences and Privilege,” 57–59.

2. AN UNSATISFACTORY RESPONSE TO THE UNFAIRNESS OBJECTION

Anderson writes that the unfairness objection

neglects the fact that as long as discrimination or its effects persist, there will be innocent victims suffering unjust burdens. The only question is whether these burdens should be borne exclusively by disadvantaged racial groups or more widely shared. There is no injustice in sharing the costs of widespread injustice.²⁰

Anderson's analysis highlights that there is nothing problematic with the government legislating to share the burdensome effects of injustice in an appropriate way, and that just policies can impose certain costs on private individuals.

But these general claims about what is permissible are not sufficient to establish that the particular burdens imposed by affirmative-action policies are fair. Indeed, a proponent of the unfairness objection can grant these general claims and simultaneously argue that the particular burdens imposed by affirmative-action policies are not fair to private individuals such as Simon; essentially, by granting that something should be done to distribute the burdens of injustice more evenly while denying that affirmative action is a permissible means of achieving such an end.²¹

Kwame Anthony Appiah tries to block this move by arguing that the burdens that are in fact imposed by affirmative-action policies are analogous to the burdens imposed by other clearly permissible policies. He writes: "If justice requires restitution to Japanese Americans for the wrongs they suffered in internment in World War II, I cannot complain, when my taxes are raised to pay this restitution, that I did not do the interring."²²

The problem with Appiah's argument is that the burdens imposed by affirmative-action policies do not seem analogous to such clearly permissible policies in the relevant sense. In order to see why, it is instructive to consider why critics argue that affirmative action is particularly objectionable: it imposes heavy burdens on a small subset of innocent individuals (such as Simon) in admissions procedures.²³ This is disanalogous to—and comparatively more controversial than—the US paying out reparations to the victims of state injustice, and the cost of these reparations being evenly distributed among all innocent taxpayers. Indeed, the case of affirmative action seems more analogous to the following

20 Anderson, *The Imperative of Integration*, 139–40.

21 See Pojman, "The Case against Affirmative Action," 108.

22 Appiah, "'Group Right' and Racial Affirmative Action," 273.

23 See Cohen, "Why Race Preference Is Wrong and Bad," 33–34.

modified version of Appiah's example: imagine that in order to compensate the interned Japanese Americans, a heavy tax was exclusively levied on a group of randomly selected non-Japanese people who comprised 5 percent of the population. This would distribute, at least in one sense, the costs of injustice more evenly. But it is an arrangement that the 5 percent group could plausibly contest as unfair—not because it imposes some costs on them, but because it imposes particularly heavy costs exclusively on them.²⁴

Affirmative action and just taxation are disanalogous in a further sense. In a just scheme of progressive taxation relative privilege and relative burden are correlated, in the sense that the rich pay more and the poor pay less. But this correlation does not hold with respect to affirmative action, at least in a contemporary US context. Indeed, many argue that affirmative action is particularly unfair because this correlation is inverted: the least privileged members of the non-target group, such as poor white men from Appalachia, are more likely to lose out on admission than comparative privileged members of the non-target group.²⁵

Anderson's and Appiah's combined attempt to overcome the unfairness objection, therefore, fails because it faces the problem of "under-theorization." It does not provide sufficient theoretical resources to determine whether affirmative action is a just policy that imposes a fair set of burdens: the claim that the demands of justice can impose certain costs on private individuals is not sufficient to establish that a set of actual costs is fair, and affirmative action is not relevantly analogous to other clearly permissible policies.

3. RAWLSIAN JUSTICE AND AFFIRMATIVE ACTION

The problem of under-theorization can be overcome by developing Rawls's non-ideal theory in a novel way. In this section, I pave the way for this endeavor by surveying the limitations of Rawls's theory of justice as it stands.

Rawls's large corpus of work contains little explicit discussion of group-based

24 Relatedly, James Sterba tries to diffuse the unfairness objection by arguing that, from the perspective of fairness, the preference given to legacy students is at least as bad as affirmative action ("Defending Affirmative Action, Defending Preference," 266–67). Sterba's argument can be used to present an *ad hominem* objection against some conservatives: it is inconsistent to object to race-based affirmative action but to approve of legacy preferences—for predominantly upper-middle-class, white Americans. But this is not sufficient to show that the unfairness objection to affirmative action does not have real moral force. Even if affirmative-action policies are no worse (or even better) than legacy preferences this does not establish that they are defensible. After all, legacy preferences can plausibly be contested as a deeply unfair feature of a society structured by economic class.

25 See Hurst, Fitz Gibbon, and Nurse, *Social Inequality*.

injustices, such as racism, that affirmative-action policies are designed to ameliorate. That said, Tommie Shelby has shown that Rawls's ideal theory of justice has substantive implications for such injustices. Rawls's liberty principle condemns race-based slavery and apartheid: under either institutional arrangement, *not* everyone would have access to a fully adequate scheme of basic liberties.²⁶ Similarly, the fair equality of opportunity (FEO) principle condemns any educational procedure that discriminates against a racial group.²⁷ Essentially, from the perspective of Rawlsian ideal theory, these group-based injustices are objectionable because they are deviations from ideal justice.

In order to determine whether affirmative-action policies are an acceptable way of ameliorating group-based injustice, we must turn to Rawls's nonideal theory of justice. Drawing on *The Law of Peoples*, A. John Simmons argues that Rawls's nonideal theory has the following content and structure:

The specific "policies and courses of action" it mandates must be (i) "morally permissible," (ii) "politically possible," (iii) "likely to be effective" in moving society toward the ideal of perfect justice.²⁸

Rawlsian nonideal theory is transitional. From the perspective of Rawlsian nonideal theory the goals of affirmative-action policies are good, at least insofar as they are an effective means of transitioning toward ideal justice. The crucial question is whether such policies satisfy the "moral permissibility" condition. This condition entails that not all paths that would be an effective means of transitioning toward ideal justice are necessarily permissible.

But the limitation of this permissibility condition is that neither Rawls nor Simmons offer any real guidance for determining which transitional paths are permissible. Perhaps they intend for permissibility to be judged intuitively. This clearly can be done in some cases: for instance, it seems obvious that despotic rule by a dictator should be judged impermissible, even if (surprisingly) this would be an effective means of bringing about ideal justice in the very long term.

It is not, however, always so easy to determine the permissibility of certain possible transitional paths. As I noted above, affirmative action is a difficult test case for nonideal theorists. It can *prima facie* plausibly be contested as unfair. But—in contrast to the example of dictator rule—it is not intuitively clear that

26 See Rawls, *Political Liberalism*, 292. Rawls's conception of the liberty principle evolved between *A Theory of Justice* and *Political Liberalism*. Most important, given my present purposes, "a fully adequate scheme" replaced "the most extensive total liberty."

27 Following Shelby, "Race and Ethnicity, Race and Social Justice. See also Rawls, *A Theory of Justice*, 266.

28 Simmons, "Ideal and Non-ideal Theory," 18. Following Rawls, *The Law of Peoples*, 89.

this unfairness objection is sufficient to rule out affirmative action. Consequently, in order to determine whether affirmative action is permissible a more normatively determinate conception of permissibility is required. On Simmons's reconstruction, at least, Rawlsian nonideal theory cannot determine whether affirmative action is in fact just, for (like Anderson's and Appiah's approach) it faces the problem of under-theorization.

4. A NONIDEAL PRINCIPLE OF JUSTICE

Given the limitations of Rawls's theory of justice, philosophers who wish to offer a broadly Rawlsian treatment of affirmative action need to be creative. My way of developing Rawls's theory has two stages. First, I use a contractualist framework to derive a nonideal principle of justice that applies in all empirical conditions. Second, in section 5, I argue that, given the particular empirical conditions that obtain in the contemporary US, this nonideal principle of justice supports affirmative action.²⁹

Nonideal principles of justice are "idealized" in the sense that they abstract away from certain feasibility constraints and specify what justice *simpliciter* requires.³⁰ But they are "nonideal" in the sense that they specify how an unjust society should transition toward becoming a perfectly just society, rather than how a perfectly just society should, itself, be structured. The innovation of nonideal principles of justice is, I suggest, the key to giving substantive normative content to the under-theorized "moral permissibility" condition in Rawls's nonideal theory. I will not derive a complete set of nonideal principles; rather I will derive a single principle that justifies affirmative action under a broad range of conditions.

Although my approach is Rawlsian, it (arguably) abandons one Rawlsian orthodoxy. Rawls argues that justice exclusively regulates the basic structure: the main institutions of society. The precise scope of the basic structure is disputed. But universities are (standardly) not construed as part of it.³¹ Consequently, some Rawlsians would argue that justice is silent about whether affirmative action should be used in universities.³² In contrast, I assume—at least in nonideal

29 For different ways of developing a nonideal theory within a contractualist framework, see Arvan, "First Steps toward a Nonideal Theory of Justice"; and Mills, *The Racial Contract*.

30 I am neutral about whether justice *simpliciter* depends on some feasibility constraints. See Wiens, "Motivational Limitations on the Demands of Justice."

31 Following Hodgson, "Why the Basic Structure?" 303–32.

32 Even for such Rawlsians my argument has some value: it shows that if universities choose to implement affirmative-action policies that are supported by my nonideal principle of justice, then such policies cannot plausibly be contested as unfair.

conditions—that justice has direct implications for university policies, like affirmative action.³³ This assumption is motivated by the claim that if institutions—even those outside the basic structure—are capable of ameliorating injustice, then they ought to do so.³⁴

4.1. *The Basic Liberties and Two Distinctions*

The requisite nonideal principle of justice hinges on the concept of “basic liberties” that are protected by Rawls’s ideal liberty principle. Rawls argues that a liberty should be classified as basic if and only if it is essential for the adequate development and full exercise of the two moral powers: the capacity for a sense of justice and the capacity for a conception of the good.³⁵ Such liberties fit into five categories: freedom of thought and liberty of conscience, freedom of association, equal political liberty, rights and liberties protecting the integrity and freedom of the person, and rights and liberties covered by the rule of law.³⁶

Rawls notes that the various basic liberties are bound to conflict with one another; consequently, particular liberties can be restricted so that a complete and coherent scheme of liberties is generated.³⁷ I acknowledge the need for such holistic specification. But, given my present purposes, the rights and liberties protecting the integrity and freedom of the person are particularly important. Such liberties are valuable in themselves and also, as Samuel Freeman notes, because they are instrumental to the exercise of the other basic liberties.³⁸ To explain Freeman’s point, suppose that someone’s rights protecting the integrity and freedom of their person are infringed. Then, plausibly, they will also lack the ability to engage effectively in politics and hence fail to have equal political liberty.

It is necessary to make two related distinctions concerning the basic liberties that are salient in nonideal conditions but not in ideal conditions. Both distinctions can be illustrated using the same example. Between 2005 and 2012, the New York Police Department increasingly implemented a “stop-and-frisk” practice. As the name of this practice suggests, it allowed police officers to stop,

33 This assumption is shared by a number of philosophers who extend Rawlsian justice to the regulations of corporate entities outside the basic structure. See Donaldson, *Corporations and Morality*; Donaldson and Dunfee, *Ties That Bind*; and Taylor, “Rawlsian Affirmative Action.”

34 For related discussion, see Berkey, “Rawlsian Institutionalism and Business Ethics.”

35 Rawls, *Political Liberalism*, 293, 332–33.

36 Rawls, *Political Liberalism*, 291.

37 Rawls, *Political Liberalism*, 295.

38 Freeman, *Rawls*, 56.

question, and frisk pedestrians under a standard of reasonable suspicion.³⁹ This practice—particularly given that the standard of “reasonable suspicion” was so vague that it could be interpreted to apply to almost any case—clearly violated the basic liberty of freedom and integrity of the stopped person. This is because, as Rawls notes, this basic liberty includes freedom from psychological oppression.⁴⁰ And, quite understandably given its nature, the practice induced a great deal of fear, which prevented innocent citizens from exercising this basic liberty in a public space.

The first distinction to note is between the state’s official recognition of basic liberties in documents such as a written constitution and the degree to which people are in fact able to exercise their basic liberties because of their reasonable reaction to practices such as stop-and-frisk.⁴¹ A consequence of this distinction is that a state can be in nonideal conditions even if its official recognition of basic liberties conforms to the ideal liberty principle. This is because people may not actually be able to exercise their basic liberties to a fully adequate degree.

Being able to exercise one’s basic liberties to a fully adequate degree is a threshold sufficientarian concept. But—as a second, related distinction illustrates—if it is not reached then this can also give rise to certain egalitarian concerns: in nonideal conditions an inability to exercise one’s basic liberties is a burden that could fall disproportionately on certain types of people. For example, 90 percent of the people who were stopped and frisked in New York City were Black or Latinx and had committed no crime.⁴² Due to the stop-and-frisk practice, African American and Latinx American New Yorkers were *ceteris paribus* less able to exercise their basic liberties than other citizens because of either the direct effects of this practice or the fear that it induced. Such an inequality seems problematic in itself. And, from a Rawlsian perspective, it is also problematic in a deeper sense. For a central Rawlsian commitment is that political liberty must be (at least approximately) equal.⁴³ But as I noted above, the rights and liberties protecting the integrity and freedom of the person are instrumental to the realization of equal political liberty. Consequently, if, for example, African Ameri-

39 NY Criminal Procedure Law §140.50.

40 See Rawls, *A Theory of Justice*, 53.

41 The caveat of “a reasonable reaction” rules out cases in which actual people feel psychological oppression; however, this oppression should be judged as either psychologically eccentric or stemming from an unjustifiable set of beliefs. Hosein makes a similar move in “Racial Profiling and a Reasonable Sense of Interior Political Status,” e6.

42 Center for Constitutional Rights, “Racial Disparity in NYPD Stops-and-Frisks.”

43 Rawls, *Justice as Fairness*, 148–49.

can and Latinx American New Yorkers are disproportionately unable to exercise such rights and liberties, then this will also undermine equal political liberty.

Stop-and-frisk practices provide an especially clear illustration of the two distinctions. The distinctions, however, also apply to practices that take place on a more diffuse social level. For example, middle-class African Americans often report that they are avoided like criminals even when they dress in respectable clothing.⁴⁴ As Anderson notes, “to be subject as a matter of public reputation to the default presumption of criminal suspicion is . . . to be publicly dishonored and degraded. . . . Even those with thick skins and high self-esteem suffer harm to their public standing due to racial stigmatization.”⁴⁵ Essentially, diffuse racial stigmatization constitutes a type of psychological oppression. African Americans are, consequently, *ceteris paribus* less able to exercise their basic liberties to a fully adequate degree than other citizens because of their reasonable reaction to such psychological oppression.

Some philosophers, such as Iris Marion Young, argue that Rawls’s theory of justice is insensitive to many modes of social oppression.⁴⁶ The two distinctions, concerning the basic liberties, are not explicitly articulated by Rawls. But, I suggest, they are in keeping with the spirit of his theory; furthermore, once added they help to illuminate how Rawlsian theory can be sensitive to an important type of oppression in nonideal conditions.

4.2. *Deriving and Defending a Nonideal Principle of Justice*

In order to derive the nonideal principle, I begin by clarifying how the contractualist framework and parties are modeled. The nonideal contracting parties—exactly like Rawls’s ideal contracting parties—are rational, in the sense that they want to advance their ends as effectively as possible. The parties are placed behind a veil of ignorance. This veil precludes knowledge of the particular social position that they will actually occupy when the veil is lifted; it thereby prevents the parties from tailoring principles of justice to advance the particular social position they will occupy, such as a particular race or social class.⁴⁷

The nonideal original position has an intergenerational component: the parties are ignorant of when they will be born prior to the realization of ideal justice. This stipulation is introduced to ensure that the path to ideal justice is intergenerationally fair—as opposed to merely intragenerationally fair.⁴⁸

44 See Feagin, “The Continuing Significance of Race.”

45 Anderson, *The Imperative of Integration*, 55.

46 See Young, *Justice and the Politics of Difference*.

47 Rawls, *A Theory of Justice*, 17.

48 I remain neutral about how, precisely, this intergenerational component is modeled. I favor

The nonideal contracting parties are presented with the following “nonideal” scenario, which accounts for the two previously drawn distinctions: not all actual people will be able to exercise their basic liberties to a fully adequate degree because of significant noncompliance with justice and/or structural injustice. Furthermore, there is a chance that different people will be unable to exercise their basic liberties to different degrees.⁴⁹

The contracting parties must select a nonideal principle of justice for this nonideal scenario: a principle that it is rational for them to adopt, given that they do not know which social position they will occupy or when they will be born.

Furthermore, they select the principle against the backdrop of three presuppositions. First, they assume that Rawls’s ideal principles of justice are correct. Second, they assume that there are sufficient economic resources for it to be possible to increase people’s ability to exercise their basic liberties to a significant degree. Third, they assume that all actual people will strictly comply with the nonideal principle of justice that they select.

To clarify this third presupposition, one of the causal reasons that a society can be in nonideal conditions is that actual people have failed to comply fully with the demands of justice. (Stop-and-frisk illustrates this point.) But this causal genesis is compatible with the claim that the contracting parties should assume that actual people will strictly comply with the nonideal principle of justice that they select. This assumption is not realistic. But it is adopted because it allows the contracting parties to select a principle that specifies what justice *simpliciter* requires, without that selection being tainted by actual people’s expected noncompliance.⁵⁰ Although I make this assumption, I grant that the derived

the “narrowing choice” model: the parties know that they all belong to the same generation but they do not know to which generation they belong. They are mutually disinterested and their selection of principles is constrained by certain formal features, such as universality. For a good defense of this model, and a survey of competing models, see Attas, “A Trans-generational Difference Principle.”

49 In the context of ideal theory, Rawls describes a four-stage sequence in which the veil is gradually lifted in order to determine principles of justice, then a constitution, then laws, and then the application of laws to particular cases. In this sequence, each stage is guided and constrained by the results of the previous stages (*A Theory of Justice*, 171–74). My nonideal scenario is comparable to the later stages of this sequence, insofar as the veil is partially lifted because more information is introduced. But it is different from the four-stage sequence because more information is introduced in order to determine a *sui generis* nonideal principle of justice rather than to guide the application of ideal principles of justice to things such as the constitution and law.

50 It might be objected that this idealizing assumption of strict compliance is inappropriate in the context of Rawlsian nonideal theory. After all, Rawls sometimes defines nonideal theory as partial compliance theory. See Rawls, *A Theory of Justice*, 215. But Rawls defines nonideal

principle is not directly action guiding. For it is necessary to consider how actual people will react to policies that are supported by the principle in order to assess the efficacy and feasibility of such policies.⁵¹ Accordingly, in section 5 I consider the possibility that affirmative-action policies may have negative stigmatizing effects even if they are just.

I am now in a position to present the (irreducibly baroque) nonideal principle of justice that, I contend, the contracting parties would settle on. After presenting the principle, I will outline the reasoning that leads to its selection. In the principle, “comparative disadvantage/advantage” is with respect to an ability to exercise one’s basic liberties to a fully adequate degree because of one’s reasonable reaction to other people’s noncompliance with justice and/or structural injustice.

1. Measures should be instituted to increase the standing of comparatively disadvantaged people. The required measures are specified by the following clauses.
2. If comparatively disadvantaged people are disadvantaged to different degrees, priority should be given to the most disadvantaged.
3. Comparatively disadvantaged people’s standing should be increased in such a way that it imposes as few demands on comparatively advantaged people as possible.
4. If it is not possible to increase the standing of disadvantaged people without imposing costs on comparatively advantaged people, such costs should be distributed according to the following two principles: (i) costs should be imposed evenly on comparatively advantaged people at the same level of advantage; (ii) the relative significance of these costs should be determined by the priority ordering of Rawls’s ideal principles of justice (e.g., the basic liberties have priority over equality of opportunity).⁵²
5. In determining what measures should be implemented in a particular set of nonideal conditions it *is not* sufficient to consider what measures

theory in two different ways: as a conception of what we ought to do in conditions that fail to realize ideal justice, and as partial compliance theory. These two definitions potentially cut against one another as there is no reason to think that the best theoretical account of what we ought to do in conditions that fail to realize ideal justice could not assume strict compliance, at least at some levels of theorizing. I expand on this point in Adams, “The Value of Ideal Theory.”

51 For related discussion, see Carroll, “In Defense of Strict Compliance as a Modelling Assumption.”

52 Assuming, of course, that such costs are not so great that bearing these costs would make people who were antecedently comparatively advantaged more disadvantaged than people who were antecedently comparatively disadvantaged.

would be the most effective means of increasing the standing of comparatively disadvantaged people at that time. Rather, measures should be introduced that are the most effective means of increasing comparatively disadvantaged people's standing prior to the realization of ideal justice—with priority given to the most disadvantaged, regardless of which generation they are born into. This clause also applies *mutatis mutandis* to the distribution of costs: such costs should be distributed according to clause 4, without discriminating between comparatively advantaged people because they are born into different generations.

Clause 1 is selected because—at the very minimum—the parties want, *ceteris paribus*, to increase the standing of comparatively disadvantaged people in case they end up occupying this unfortunate position. They are, however, also concerned with the costs and opportunity costs of achieving this end. Consequently, they select a set of clauses that specify the parameters under which this end should be achieved.

Clause 2 is chosen because, given the choice problem posed, it is rational for the nonideal contracting parties—like the ideal contracting parties—to be guided by maximin: to instigate measures to guarantee that their social standing is as good as possible in case they end up occupying the position of the most disadvantaged.⁵³

Clause 3 is adopted because, although the parties are prepared to impose costs on comparatively advantaged people for the sake of improving the social standing of the disadvantaged, they are not indifferent to the nature of these costs. After all, they could end up occupying the social position of comparatively advantaged people. Therefore, *ceteris paribus*, they prefer for the costs that fall on comparatively advantaged people to be as small as possible.

The first part of clause 4, (i), is adopted because the parties prefer to impose costs on the comparatively advantaged so long as this increases the standing of the disadvantaged and does not bring the overall new standing of the comparatively advantaged down to a level below the new standing of the comparatively disadvantaged. As I noted above, given the choice problem posed, it is rational for the parties to improve the standing of the most disadvantaged—given that they could end up occupying this most disadvantaged position—rather than to produce the best aggregated outcome. They decide that costs should be imposed evenly on comparatively advantaged people at the same level of advan-

53 There is a vast literature discussing both why and whether it is rational for the parties to favor the interests of the comparatively disadvantaged. For a good overview see Gaus and Thrasher, "Rational Choice in the Original Position." I will not, here, attempt to defend Rawls's position further.

tage because they are as likely to become any one of these particular advantaged people as any other; consequently, they want the costs on each comparatively advantaged person to be as small as possible.

The mere fact that the priority ordering structures the ideal principles of justice does not straightforwardly entail that the nonideal contracting parties would also choose for it to structure the nonideal principle, as clause 4 (ii) states. But the salient point concerns the relative ordering of value that the priority ordering reflects: the priority ordering reflects the fact that, for example, the ideal contracting parties attach greater value to the basic liberties than other considerations of justice. This point about value also applies in nonideal conditions in the sense that, for instance, the nonideal contracting parties—like the ideal contracting parties—would attach more value to the basic liberties than other considerations of justice. Consequently, they would prioritize measures to increase their ability to exercise their basic liberties—in case they end up in a position in which their exercise of their basic liberties is compromised—over other possible considerations. Therefore, clause 4 (ii) is endorsed by the parties because the priority ordering of the ideal principles determines the relative importance, or urgency, of different types of injustice in nonideal conditions.⁵⁴

Finally, clause 5 is selected because the parties do not know when they will be born; consequently, they reject measures that privilege the interests of a particular generation prior to the realization of ideal justice.

5. HOW THE NONIDEAL PRINCIPLE OF JUSTICE SUPPORTS AFFIRMATIVE ACTION

The nonideal principle of justice supports affirmative action if and only if the following conditions are satisfied:

- a. People who are unable to exercise their basic liberties to a fully adequate degree because of their reasonable reaction to other people's noncompliance with justice and/or structural injustice are in that position (at least partly) because they possess the characteristic(s) that affirmative-action policies are designed to target (from clause 1).

54 Following Rawls, *A Theory of Justice*, 216; and Korsgaard, *Creating the Kingdom of Ends*, 148. As noted above, I assume that there are sufficient resources for it to be possible to increase comparatively disadvantaged people's ability to exercise their basic liberties to a significant degree. Robert Taylor helpfully elaborates the threshold of resources that is necessary for the priority ordering of liberty to apply: "a society must have achieved a level of wealth sufficient for it to allow its citizens to engage in meaningful formation of life plans" ("Rawls's Defense of the Priority of Liberty, 263).

- b. Affirmative-action policies are a generally effective means of increasing comparatively disadvantaged people's ability to exercise their basic liberties in a way that gives priority to the most disadvantaged and *does not* discriminate between equally advantaged/disadvantaged people who are born into different generations (from clauses 1, 2, and 5).
- c. There is not another candidate policy that would be at least as effective a means of increasing comparatively disadvantaged people's ability to exercise their basic liberties, but would impose less significant costs on comparatively advantaged people (regardless of which generation they are born into) as specified by the priority ordering of Rawls's ideal principles of justice (from clauses 3, 4, and 5).

Given the particular empirical conditions that obtain in the contemporary US, I argue that a strong case can be made that the nonideal principle of justice supports affirmative action. For the sake of simplicity, and because there is the most relevant social scientific data on the topic, I will focus on affirmative action for African Americans. (I suggest that a similar conclusion applies for Latinx Americans.)

It is uncontroversial that condition a is satisfied. There is overwhelming evidence that racial discrimination harms African Americans.⁵⁵ One of the ways in which it does so, as my previous example of stop-and-frisk illustrates, is to interfere with African Americans' ability to exercise their basic liberties to a fully adequate degree.

Condition b rests on the following presupposition: using affirmative-action policies to ensure that a sufficient threshold of African Americans is placed in certain educational institutions can be a causally effective means of reducing racism of various sorts. And, consequently, given that racism is a causal mechanism that undermines African Americans' ability to exercise their basic liberties—affirmative action can be a causally effective means of increasing African Americans' ability to exercise their basic liberties.

This presupposition can be defended in two main ways. First, contact theory postulates that intergroup contact can reduce prejudice and discrimination if the following four conditions are satisfied: the members of the different groups have equal status (at least in certain relevant respects), they work toward common goals, they engage in cooperation, and the contact is supported and regulated by institutional authority.⁵⁶ These conditions are satisfied in educational

55 For a detailed but nontechnical overview of the relevant statistical data, see Anderson, *The Imperative of Integration*, chs. 1–3; and Sterba, *Affirmative Action for the Future*, "Introduction" and ch. 1.

56 Allport, *The Nature of Prejudice*; Dhont, Van Hiel, and Hewstone, "Changing the Ideological Roots of Prejudice"; and Pettigrew, "Intergroup Contact Theory."

institutions, in which students cooperate on equal terms (at least in certain relevant respects) to pursue educational goals that are regulated by institutional norms. Affirmative-action policies ensure that there is a greater representation of African Americans in educational institutions and, thereby, facilitate greater intergroup contact. This helps break down prejudice and discrimination within educational institutions. Clearly, education is a particularly formative time in many people's lives; consequently, the reduction of racial prejudice within educational institutions can have an impact not just within such institutions but also on graduates' subsequent professional and personal lives.⁵⁷

The presupposition underpinning b can be defended in a second way. Affirmative action can reduce discrimination in a far broader sense and, thereby, benefit African Americans who neither attend the particular institutions that practice affirmative action nor directly encounter the graduates of such institutions. Especially on a large intergenerational scale, it can do so by breaking down negative race-based stereotypes and thereby reduce so-called statistical discrimination against all African Americans. To explain, in the US being African American is correlated with variables such as having relatively low educational attainment and social class.⁵⁸ People's knowledge about these variables, with respect to particular individuals, is standardly imperfect and increasing this knowledge is costly. Race, however, is a visible feature that can be instantly assessed with relative reliability at almost no cost. Consequently, rational economic actors who do not have any racial prejudices may use race as a proxy for this knowledge. This discrimination is statistical because individuals are judged in terms of the group averages of all African Americans rather than in terms of their individual merits and level of achievements.⁵⁹

Affirmative action can reduce statistical discrimination by helping under-

57 Elizabeth Anderson also uses contact theory to argue that achieving racial integration—in all walks of American life—is an imperative of justice (*The Imperative of Integration*, 123–27). In contrast to Anderson, I defend the more restricted claim that contact theory supports affirmative action in an educational context. There are two major advantages to my approach. First, as Anderson herself acknowledges, contact theory most obviously supports racial integration in formal settings such as educational institutions and the workplace. For in such settings—in contrast to residential neighborhoods—the integration is backed by institutional authority. See Anderson, *The Imperative of Integration*, 123. Second, in contrast to Anderson, I can acknowledge that it is permissible for African Americans to resist integration for a variety of reasons, for example, in a residential context, due to the short-term threat of greater interracial conflict, or out of solidarity with other African Americans. Following Shelby, "Integration, Inequality, and Imperatives of Justice," 267–82.

58 See Jones, Schmitt, and Wilson, "50 Years after the Kerner Commission," 1–8.

59 Following Phelps, "The Statistical Theory of Racism and Sexism"; and Arrow, "Models of Job Discrimination."

mine the rational basis of such discrimination: the general correlation between being African American and having relatively low educational, occupational, and social status. It does this by increasing the number of African Americans studying and, consequently, also ultimately working in institutions of power and prestige. As Dworkin argues, in doing so it thereby decreases the degree of racial identification and by extension statistical discrimination in the US, by reducing the existing correlation between being African American and having an assumed social standing.⁶⁰

It might be objected that when all the effects of affirmative-action policies are taken into consideration they will not satisfy condition b. In particular, many argue that affirmative action has a stigmatizing effect because granting preferential treatment to African Americans implies acknowledging the inferiority of their average strength as applicants.⁶¹ This effect could be broad because it is usually impossible to identify the subset of African Americans who would not have been admitted without affirmative action. Consequently, affirmative action may have the ironic effect of making people view all African American students as inferior.⁶² This stigma could prevent affirmative-action policies from satisfying b for two different reasons. First, this stigmatizing effect could undermine my argument that contact theory supports affirmative action. This is because contact theory requires that different groups have *equal* status and affirmative action makes people view African Americans as inferior rather than as equal. Second, the psychological oppression caused by stigmatization could undermine African Americans' ability to exercise their basic liberties.

Much of the evidence in support of this alleged stigmatizing effect is anecdotal.⁶³ The most comprehensive statistical study on the effects of affirmative action by William Bowen and Derek Bok surveyed over eighty-thousand students at twenty-eight top-tier institutions. It concludes that the effects of stigmatization were comparatively low and that most alumni thought that affirmative action helped reduce stereotypes and mutual animosity.⁶⁴ Similarly, Deirdre Bowen's study finds that African Americans experience greater stigma in educational institutions located in states that have banned affirmative action.⁶⁵ Thus, it seems

60 Dworkin, *A Matter of Principle*, 294. See also Goffman, *Stigma*.

61 Following Sabbagh, *Equality and Transparency*, 109.

62 See Eastman, *Ending Affirmative Action*; and Scalia, "The Disease as Cure," 219.

63 See, in particular, Clarence Thomas's opinion in *Adarand Constructors Inc. v. Peña*, 515 US 200 (1995).

64 Bown and Bok, *The Shape of the River*.

65 Bowen, "Brilliant Disguise."

plausible to conclude that although there may be some stigmatizing effect it is not sufficient to prevent many affirmative-action policies from satisfying b.

Note that my nonideal principle specifies the threshold at which a stigmatizing effect would be intolerable. In order for condition b to be satisfied, priority must be given to the most disadvantaged without intergenerational discrimination. It would, therefore, be intolerable for an affirmative-action policy to impose a stigmatizing effect that has the net effect of making people the most disadvantaged. This would be the case even if the affirmative-action policy was a causally effective means of increasing certain less disadvantaged people's ability to exercise their basic liberties, who were members of a different future generation.

Finally, consider condition c. As I will explain in section 6, affirmative action suspends certain features of fair equality of opportunity. This is preferable to an alternative policy that restricts comparatively advantaged people's ability to exercise their basic liberties. This is because c requires that the relative significance of the costs must be determined by the priority ordering of Rawls's ideal principles of justice, and the ideal liberty principle is lexically prior to the FEO principle.⁶⁶

The nonideal principle of justice would, however, support abolishing affirmative action in favor of alternative policies that merely redistribute wealth, if these alternative policies were an equally effective means of increasing African Americans' ability to exercise their basic liberties. This is because such alternative policies would impose a less significant cost on comparatively advantaged people, given that the FEO principle has priority over the distribution of economic goods according to the difference principle.⁶⁷

It might be argued that the nonideal principle supports abolishing affirmative action because any benefit that could be produced by affirmative action could also be produced exclusively by a redistribution of economic goods: in the US there is a significant correlation between being African American and relative poverty. In 2018 African Americans were about 2.5 times as likely to be in poverty compared to white Americans, and in 2016 the median African American family had only 10.2 percent of the median white American family's wealth.⁶⁸ This relative poverty, it might be claimed, is the primary cause of African American underrepresentation in academic institutions. But a sufficient redistribution of economic goods would remove this relative poverty and, consequently, the primary cause of African American underrepresentation. Therefore, after sufficient

66 See Rawls, *A Theory of Justice*, 266.

67 Rawls, *A Theory of Justice*, 266.

68 See Jones, Schmitt, and Wilson, "50 Years after the Kerner Commission," 3–4.

economic redistribution, no possible affirmative-action policy could satisfy condition b. This is because a consequence of sufficient economic redistribution is that there would be enough African Americans in academic institutions such that no affirmative-action policy could increase African Americans' ability to exercise their basic liberties.

This argument can be challenged empirically. It is not clear that relative poverty is the primary cause of African American underrepresentation. Susan Mayer, for instance, argues that the degree to which children's educational achievement is dependent on the financial resources of their parents is far weaker than standardly supposed. Indeed, she argues that it has relatively little effect as long as the parents are not in extreme poverty and the basic material needs of their children are met.⁶⁹ If this is correct, then mere economic redistribution would (almost certainly) be insufficient for any affirmative-action policy to be *unable* to satisfy b.

More important, suppose for the sake of argument that a sufficient redistribution of economic goods would make it impossible for any affirmative-action policy to satisfy b. Even so, this would only be something that could be achieved in the relatively long term—plausibly, at the very minimum, after there is no significant correlation between being African American and relative poverty for at least one generation. Consequently—at least in the short term, before this is achieved—there is reason to retain affirmative-action policies.

In summary, the nonideal principle of justice supports affirmative action because, at least in the short term, it is an effective and fair means of increasing some African Americans' ability to exercise their basic liberties to a significant degree.

Two clarifications about the scope of this claim are required. First, the claim that the nonideal principle supports affirmative action in the short term is compatible with the claim that actions should be undertaken to make affirmative action unnecessary in the long term. Indeed, suppose that some possible set of policies that merely redistribute wealth would make affirmative-action unnecessary in the long term (for instance, by increasing the funding of predominantly African American high schools or alleviating African American poverty). Then, it would be obligatory to implement this set of policies. For they would impose less significant costs on comparatively advantaged people, as specified by condition c.

Second, note that I use the phrase to a "significant degree" rather than to the "fully requisite degree." Indeed, in my view, affirmative action is a small com-

69 Mayer, *What Money Can't Buy*. See also Satz, "Equality, Adequacy, and Education for Citizenship," 633.

ponent of what is required to enable African Americans to exercise their basic liberties to the fully requisite degree even in the short term. Other important courses of action will include ending practices like stop-and-frisk and the disproportionate mass incarceration of African Americans, which (among other things) interferes with African Americans' ability to exercise their basic liberties because of the stigmatizing effect that it induces.⁷⁰

Still, even if much more than affirmative action is required, this should not distract from the fact that affirmative action can perform the *sui generis* function of reducing racial discrimination within certain educational institutions and a type of statistical discrimination in society as a whole.

6. TWO GOOD-MAKING FEATURES OF MY ARGUMENT

6.1. *Blocks the Unfairness Objection*

The unfairness objection gains critical traction because when affirmative-action policies are viewed in isolation they appear unfair because individuals such as Simon experience burdens in virtue of their membership in a particular racial group. But this is not sufficient to ground a charge of unfairness. For under the conditions that I have specified, affirmative-action policies reflect a fair distribution of the burdens that are required to transition to a more just society. The explanation for why they are fair can be presented using the nonideal contractualist framework: it would be rational for parties who do not know what social position they will occupy to assent to a principle that condones affirmative action under the specified conditions. Therefore, the policy is impartial—hence fair—in the appropriate sense.

The topic of affirmative action invites reflection on the relationship between “fairness” and “justice.” Many use these two concepts interchangeably in everyday speech. However, some philosophical theories of justice render these concepts completely distinct: for instance, a theory in which justice is explicated as “nondomination.”⁷¹ Given such a conception of justice, demonstrating that affirmative action is just would not be sufficient to obviate the unfairness objection. Some additional account would have to be supplied in order to explain why considerations of justice trump considerations of fairness.

In contrast, a Rawlsian contractualist framework is uniquely suited to over-

70 In 2000, Human Rights Watch reported that in at least fifteen states African Americans constituted 80 percent to 90 percent of all drug offenders sent to prison, despite the fact that African Americans were no more likely to be guilty of drug crimes than white people (*Punishment and Prejudice*). Following Alexander, *The New Jim Crow*, 98–99.

71 See Pettit, *Republicanism*.

coming the unfairness objection because it presupposes such a tight connection between the concepts of “justice” and “fairness.” Rawls describes his view as “justice as fairness” because “the principles of justice are the result of a *fair* agreement or bargain.”⁷² It would be an exaggeration to say that, in a Rawlsian framework, fairness and justice are synonymous concepts. Even so, if conditions a–c are satisfied, it is difficult to see how a plausible charge of unfairness could be presented against such policies. For they are justified by a nonideal principle of justice that is determined by a *fair* agreement.

6.2. Not Narrowly Focused on Achieving Equality of Opportunity

Rawls’s ideal FEO principle states that “those who are at the same level of talent and ability, and have the same willingness to use them, should have the same prospects of success regardless of their initial place in the social system.”⁷³ In the contemporary US this principle is not satisfied because of factors such as extreme poverty and the effects of past racial discrimination. Many Rawlsian and non-Rawlsian philosophers argue that affirmative action can be justified because it helps to counteract the unequal opportunities that are rooted in such past discrimination and, thereby, results in admissions procedures that are closer to the ideal of fair equality of opportunity.⁷⁴ Thus, affirmative action suspends features of fair equality of opportunity in the sense that race—which would be arbitrary if the FEO principle were realized—is given preference.⁷⁵ This is done, however, in order to realize fairer equality of opportunity in the long term.

Robert Taylor mounts a powerful challenge to this approach. He argues that “we simply cannot know what the counterfactual result of a “clean” competition would look like unless we run one.”⁷⁶ For under genuinely fair equality of opportunity there could be disproportional group outcomes that are (at least in part) due to cultural reasons, such as Jewish overrepresentation among academics.⁷⁷ Given the epistemic opacity of counterfactuals about what the outcomes of genuinely fair equality of opportunity would look like, Taylor argues that we ought to

72 Rawls, *A Theory of Justice*, 11, emphasis added.

73 Rawls, *A Theory of Justice*, 63.

74 See Mason, *Levelling the Playing Field*, 38n29; Miller, *Principles of Social Justice*, 175–76. Interestingly, Samuel Freeman notes that in his lectures Rawls, himself, indicated that affirmative action could be justified in order to remedy the present effects of past discrimination (Rawls, 90–91).

75 But for an argument that race-based affirmative action is compatible with Rawlsian ideal justice, see Meshelski, “Procedural Justice and Affirmative Action.”

76 Taylor, “Rawlsian Affirmative Action,” 494.

77 Taylor, “Rawlsian Affirmative Action,” 498.

err on the side of caution. Because “at least for nonconsequentialist liberals, sins of commission should be of much greater concern than sins of omission—especially when the sinner is the state.”⁷⁸ Therefore, at least in standard cases, we cannot tinker with the result of an admissions procedure, using soft or hard quotas, in order to bring it in line with an outcome that we antecedently judge to be fair. For we lack the epistemic capacity to judge what quotas would reflect the outcomes of genuinely fair equality of opportunity, and we should err on the side of caution.

My account sidesteps Taylor’s challenge.⁷⁹ For the epistemic judgments required on my account are significantly easier to make in a crucial respect. Innocent group preferences may make it hard to predict the outcome of genuinely fair admissions procedures. But, in contrast, it is not plausible to think that any innocent group preference could account for the fact that some members of certain groups are disproportionately unable to exercise their basic liberties to a fully adequate degree. Essentially, the demands imposed by the liberty principle are considerably less opaque than the demands imposed by the FEO principle.⁸⁰

More generally, my account provides a robust rationale for suspending features of fair equality of opportunity precisely because it defends affirmative action in terms of basic liberties rather than fair equality of opportunity. For Rawlsians, there are different features of justice; however, the basic liberties have greater value than all other features of justice. Consequently, my defense of race-based preference in admissions procedures is given the strongest possible Rawlsian defense: it is necessary to promote certain people’s ability to exercise their basic liberties.

7. CODA

7.1. *Policy Refinement*

Although the nonideal principle of justice supports something roughly like the affirmative-action policies in the contemporary US, it also requires some modification of these policies. As I noted at the outset, there is empirical evidence that affirmative action imposes the heaviest burdens on Asian Americans. Yet the nonideal principle of justice does not support this feature of actual affirma-

78 Taylor, “Rawlsian Affirmative Action,” 501

79 For a direct response to Taylor, see Matthew, “Rawlsian Affirmative Action”; and Meshelski, “Procedural Justice and Affirmative Action.”

80 Of course, this does not mean that my account does not require difficult epistemic judgments about the long-term (perhaps intergenerational) effects of candidate policies. But such epistemic challenges apply to many public policies; consequently, they do not support abolishing affirmative action policies *per se*.

tive-action policies. For it requires that the costs should be distributed *evenly* on comparatively advantaged people at the same level of advantage (clause 4/condition c). And there is no evidence that Asian Americans are in general more comparatively advantaged compared to white Americans, at least with respect to the ability to exercise their basic liberties. Therefore, there is no reason to think that they should have to bear greater burdens of affirmative-action policies than white Americans.

A particularly controversial feature of affirmative-action policies is that they impose benefits on the most privileged subset of African Americans: the upper middle class, many of whom are recent immigrants rather than descendants of slaves.⁸¹ Indeed, many argue that affirmative action is particularly unfair because it gives preference to upper-middle-class African Americans over poor white Americans from Appalachia.⁸² Note, however, that the nonideal principle provides a *pro tanto* justification of this comparative preference—at least if this privileged subset of African Americans is in the best position to reduce racial discrimination in certain educational institutions and, by extension, statistical discrimination in society as a whole.

In this context it is important to highlight, however, that although I have focused on the case of affirmative action for African Americans my nonideal principle leaves it as an open empirical question as to whether affirmative action should also be extended to other groups, such as poor white rural Americans. If “poor white rural American” is a social category that inhibits its members’ ability to exercise their basic to a fully adequate degree—and affirmative action for such a group would satisfy conditions a–c—then it should be extended.

7.2. *The Value of Nonideal Principles of Justice*

In recent years, political philosophers have become increasingly self-conscious about philosophical methodology, and an enormous amount has been written on the so-called ideal versus nonideal theory debate.⁸³ In the context of racial injustice in particular, a number of philosophers, including Anderson, reject the ideal-theory paradigm. Anderson rejects this paradigm for two reasons. First, following Amartya Sen, she argues that it is not *necessary* to work out an ideal theory of justice in order to offer an adequate nonideal treatment of racial injus-

81 At Harvard University, for instance, it is estimated that only about one-third of African American students are from families in which all four grandparents were born in the us. See Rimer and Arenson, “Top Colleges Take More Blacks, but Which Ones?”

82 Hurst, Fitz Gibbon, and Nurse, *Social Inequality*.

83 Following Valentini, “Ideal vs. Non-ideal Theory,” 654.

tice.⁸⁴ Second, she argues that political philosophers should take an empirically informed “bottom-up” approach that responds to the concrete problems that they face. Given that Anderson opposes ideal theory for this latter reason, she would likely disapprove of my innovation of nonideal principles of justice. Like Rawls’s ideal principles of justice, the nonideal principles make very abstract claims about what justice requires; accordingly, they should be classified as a top-down approach to nonideal theorizing.

I suggest that the methodological question of what types of theorizing (e.g., ideal theory) and concepts (e.g., nonideal principles of justice) are practically valuable should be settled by using the following test.⁸⁵ Consider which types of theorizing and concepts are required to craft the most compelling philosophical accounts of first-order problems (e.g., when, if at all, affirmative action is just). Essentially, the second-order methodological question should be determined by what adequate treatment of the first-order problems requires.

As I have argued, Anderson is unable to overcome the unfairness objection because she faces the problem of under-theorization. Assuming that I am right that the second-order methodological question should be determined by what is required to address first-order problems, it is incumbent on philosophers like Anderson to show how a nonideal “bottom-up” methodology can provide an adequate response to the unfairness objection. I am not optimistic that this could be achieved: given that the debate about affirmative action hinges on such deep questions about fairness and justice, I suggest that something very like my conceptual innovation of nonideal principles of justice will be indispensable.⁸⁶

Indiana University Bloomington
mra10@iu.edu

84 Anderson, *The Imperative of Integration*, 3. See also Sen, “What Do We Want from a Theory of Justice?” 218–22.

85 I leave open the possibility that some political theory has intrinsic, nonpractical value. See Estlund, “What Good Is It?”

86 Previous versions of this article were presented at the Philosophy, Politics, and Economics Society in New Orleans, San Francisco State University, Stanford University, University of Rochester, University of Virginia, and the Western Political Science Association in San Diego. Thanks to all of the audiences in attendance for their questions and suggestions. I am particularly grateful for incredibly helpful feedback from Marcia Baron, Colin Bird, Talbot Brewer, Eamonn Callan, Hannah Carnegie-Arbuthnott, Randal Curren, Harrison Frye, Laura Gillespie, Johannes Himmelreich, Donncha Maccoil, Kristina Meshelski, Anne Newman, Fay Niker, A John Simmons, and Rebecca Stangl. Finally, I thank the editors of the *Journal of Ethics and Social Philosophy* and a number of anonymous referees for helping me to improve the paper.

REFERENCES

- Adams, Matthew. "The Value of Ideal Theory." In *John Rawls: Debating the Major Questions*, edited by Jon Mandle and Sarah Roberts-Cady, 73–88. Oxford: Oxford University Press, 2020.
- Alexander, Michelle. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: New Press, 2012.
- Allport, Gordon W. *The Nature of Prejudice*. Cambridge, MA: Perseus Books, 1954.
- Anderson, Elizabeth. *The Imperative of Integration*. Princeton: Princeton University Press, 2013.
- Anderson, Nick. "Federal Judge Rules Harvard Does Not Discriminate against Asian Americans in Admission." *Washington Post*, October 1, 2019. https://www.washingtonpost.com/local/education/federal-judge-rules-harvard-does-not-discriminate-against-asian-americans-in-admissions/2019/10/01/dc106b54-a8a1-11e9-a3a6-ab670962db05_story.html.
- Appiah, Kwame Anthony. "'Group Rights' and Racial Affirmative Action." *Journal of Ethics* 15, no. 3 (September 2011): 265–80.
- Arrow, Kenneth. "Models of Job Discrimination." In *Racial Discrimination in Economic Life*, edited by Anthony Pascal, 83–102. Lexington, MA: Lexington Books, 1972.
- Arvan, Marcus. "First Steps toward a Nonideal Theory of Justice." *Ethics and Global Politics* 7, no. 3 (2014): 95–117.
- Attas, Daniel. "A Transgenerational Difference Principle." In *Intergenerational Justice*, edited by Alex Gosseries and Lukas H. Meyer, 189–218. Oxford: Oxford University Press, 2012.
- Beauchamp, Tom L. "In Defense of Affirmative Action." *Journal of Ethics* 2, no. 2 (1998): 143–58.
- Berkey, Brian. "Rawlsian Institutionalism and Business Ethics: Does It Matter Whether Corporations Are Part of the Basic Structure of Society?" *Business Ethics Quarterly* 33, no. 2 (April 2021): 179–209.
- Bowen, Deirdre M. "Brilliant Disguise: An Empirical Analysis of a Social Experiment Banning Affirmative Action." *Indiana Law Journal* 85, no. 4 (Fall 2010): 1197–254.
- Bown, William G., and Derek Bok. *The Shape of the River: Long-Term Consequence of Considering Race in College and University Admissions*. Princeton: Princeton University Press, 1998.
- Boxill, Bernard R. "The Morality of Preferential Hiring." *Philosophy and Public Affairs* 7, no. 3 (Spring 1978): 246–68.

- Carroll, Jeffrey. "In Defense of Strict Compliance as a Modeling Assumption." *Social Theory and Practice* 46, no. 3 (July 2020): 441–66.
- Center for Constitutional Rights. "Racial Disparity in NYPD Stops-and-Frisks: Preliminary Report on UF-250 Data from 2005 through June 2009." Center for Constitutional Rights, January 15, 2009.
- Cohen, Carl. "Why Race Preference Is Wrong and Bad." In *Affirmative Action and Racial Preference: A Debate*, by Carl Cohen and James P. Sterba, 3–181. Oxford: Oxford University Press, 2003.
- Dhont, Kristof, Alain Van Hiel, and Miles Hewstone. "Changing the Ideological Roots of Prejudice: Longitudinal Effects of Ethnic Intergroup Contact on Social Dominance Orientation." *Group Processes and Intergroup Relations* 17 (2014): 27–44.
- Donaldson, Thomas. *Corporations and Morality*. Englewood Cliffs, NJ: Prentice Hall, 1982.
- Donaldson, Thomas, and Thomas Dunfee. *Ties That Bind: A Social Contracts Approach to Business Ethics*. Boston: Harvard Business School Press, 1999.
- Dworkin, Ronald. *A Matter of Principle*. Cambridge, MA: Harvard University Press, 1985.
- Eastman, Terry. *Ending Affirmative Action: The Case for Colorblind Justice*. New York: Basic Books, 1997.
- Espenshade, Thomas J., Chang Y. Chung, and Joan L. Walling. "Admission Preferences for Minority Students, Athletes, and Legacies at Elite Universities." *Social Science Quarterly* 85, no. 5 (December 2004): 1422–46.
- Espenshade, Thomas J., and Alexandria Walton Radford. *No Longer Separate, Not Yet Equal*. Princeton: Princeton University Press, 2009.
- Estlund, David. "What Good Is It? Unrealistic Political Theory and the Value of Intellectual Work." *Analyse und Kritik* 33, no. 2 (2011): 395–416.
- Feagin, Joe R. "The Continuing Significance of Race: Antiblack Discrimination in Public Places." *American Psychology Review* 56, no. 1 (February 1991): 101–16.
- Freeman, Samuel. *Rawls*. Oxford: Routledge, 2007.
- Gaus, Gerald, and John Thrasher. "Rational Choice in the Original Position: The (Many) Models of Rawls and Harsanyi." In *The Original Position*, edited by Timothy Hinton, 256–91. Cambridge: Cambridge University Press, 2016.
- Gerstein, Josh, and Jennifer Haberkorn. "It's Not Just Abortion: 5 Issues Likely to Be Affected by Kennedy's Exit." Politico, June 27, 2018. <https://www.politico.com/story/2018/06/27/anthony-kennedy-retirement-supreme-court-cases-680104>.

- Goffman, Erving. *Stigma: Notes on the Management of Spoiled Identity*. New York: Touchstone, 2009.
- Hodgson, Louis-Philippe. "Why the Basic Structure?" *Canadian Journal of Philosophy* 42, nos. 3–4 (September/December 2012): 303–34.
- Hosein, Adam Omar. "Racial Profiling and a Reasonable Sense of Interior Political Status." *Journal of Political Philosophy* 26, no. 3 (September 2018): e1–e20.
- Human Rights Watch. "Punishment and Prejudice: Racial Disparities in the War on Drugs." *HRW Reports* 12, no. 2 (May 2000).
- Hurst, Charles E., Heather M. Fitz Gibbon, and Anne M. Nurse. *Social Inequality: Forms, Causes, and Consequences*. 9th ed. New York: Routledge, 2016.
- Jones, Janelle, John Schmitt, and Valerie Wilson. "50 years after the Kerner Commission." Economic Policy Institute, February 26, 2018. <https://www.epi.org/publication/50-years-after-the-kerner-commission/>.
- Korsgaard, Christine. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- Lippert-Rasmussen, Kasper. "Affirmative Action, Historical Injustice, and the Concept of Beneficiaries." *Journal of Political Philosophy* 25, no. 1 (2017): 72–90.
- Lynch, Frederick R. *Invisible Victims: White Males and the Crisis of Affirmative Action*. New York: Praeger, 1991.
- Mason, Andrew. *Levelling the Playing Field: The Idea of Equal Opportunity and Its Place in Egalitarian Thought*. Oxford: Oxford University Press, 2006.
- Matthew, D. C. "Rawlsian Affirmative Action." *Critical Philosophy of Race* 3, no. 2 (2015): 324–43.
- Mayer, Susan. *What Money Can't Buy: Family Income and Children's Life Chances*. Cambridge, MA: Harvard University Press, 1999.
- Meshelski, Kristina. "Procedural Justice and Affirmative Action." *Ethical Theory and Moral Practice* 19 (2016): 425–44.
- Miller, David. *Principles of Social Justice*. Cambridge, MA: Harvard University Press, 2001.
- Mills, Charles W. *The Racial Contract*. New York: Cornell University Press, 1999.
- Morris, Christopher W. "Existential Limits to the Rectification of Past Wrongs." *American Philosophical Quarterly* 21, no. 2 (April 1984): 175–82.
- Moses, Michele S. "After Fisher: Lawsuit against Harvard Goes On." *New Boston Post*, July 19, 2016.
- Nagel, Thomas. "Equal Treatment and Compensatory Discrimination." *Philosophy and Public Affairs* 2, no. 4 (Summer 1973): 348–63.
- Pettigrew, Thomas F. "Intergroup Contact Theory." *Annual Review of Psychology* 49 (1998): 65–85.

- Pettit, Philip. *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press, 1997.
- Phelps, Edmund S. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62, no. 4 (September 1972): 659–61.
- Pojman, Louis P. "The Case against Affirmative Action." *International Journal of Applied Philosophy* 12, no. 1 (1998): 97–115.
- Rawls, John. *Justice as Fairness: A Restatement*. Cambridge, MA: Belknap Press, 2003.
- . *The Law of Peoples*. Cambridge, MA: Harvard University Press, 2001.
- . *Political Liberalism: Expanded Edition*. New York: Columbia University Press, 2005.
- . *A Theory of Justice*. Rev. ed. Cambridge, MA: Belknap Press, 1999.
- Rimer, Sarah, and Karen W. Arenson, "Top Colleges Take More Blacks, but Which Ones?" *New York Times*, June 24, 2004. <https://www.nytimes.com/2004/06/24/us/top-colleges-take-more-blacks-but-which-ones.html>.
- Sabbagh, Daniel. *Equality and Transparency: A Strategic Perspective on Affirmative Action in American Law*. New York: Palgrave MacMillan, 2007.
- Satz, Debra. "Equality, Adequacy, and Education for Citizenship." *Ethics* 117, no. 4 (July 2007): 623–48.
- Scalia, Antonin. "The Disease as Cure: 'In Order to Get beyond Racism, We Must First Take Account of Race.'" *Washington University Law Quarterly* 147 (1979): 147–57.
- Schmidt, Peter. "A History of Legacy Preferences and Privilege." In *Affirmative Action for the Rich: Legacy Preferences in College Admissions*, edited by Richard D. Kahlenberg. New York: The Century Press, 2010.
- Sen, Amartya. "What Do We Want from a Theory of Justice?" *Journal of Philosophy* 103, no. 5 (May 2006): 215–38.
- Shelby, Tommie. "Integration, Inequality, and Imperatives of Justice: A Review Essay." *Philosophy and Public Affairs* 42, no. 3 (Summer 2014): 253–85.
- . "Race and Ethnicity, Race and Social Justice: Rawlsian Considerations." *Fordham Law Review* 72, no. 5 (2004): 1697–1714.
- Simmons, A. John. "Ideal and Nonideal Theory." *Philosophy and Public Affairs* 38, no. 1 (Winter 2010): 5–36.
- Sterba, James. *Affirmative Action for the Future*. Ithaca, NY: Cornell University Press, 2009.
- . "Defending Affirmative Action, Defending Preference." In *Affirmative Action and Racial Preference: A Debate*, by Carl Cohen and James P. Sterba, 189–267. Oxford: Oxford University Press, 2003.

- Taylor, Robert S. "Rawlsian Affirmative Action." *Ethics* 119, no. 3 (April 2009): 476–506.
- . "Rawls's Defense of the Priority of Liberty: A Kantian Reconstruction." *Philosophy and Public Affairs* 31, no. 3 (Summer 2003): 246–71.
- Thomson, Judith Jarvis. "Preferential Hiring." *Philosophy and Public Affairs* 2, no. 4 (Summer 1973): 364–84.
- United States Commission on Civil Rights. "Civil Rights Issues Facing Asian Americans in the 1990s." February 1992. <https://www.ojp.gov/pdffiles1/digitization/135906ncjrs.pdf>.
- Valentini, Laura. "Ideal vs. Non-ideal Theory: A Conceptual Map." *Philosophy Compass* 7, no. 9 (September 2012): 654–64.
- Wiens, David. "Motivational Limitations on the Demands of Justice." *European Journal of Political Theory* 15, no. 3 (2016): 333–52.
- Young, Iris Marion. *Justice and the Politics of Difference*. Princeton: Princeton University Press, 1990.