Is Affirmative Action Racist? Reflections Toward a Theory of Institutional Racism

The concept of institutional racism is a staple of contemporary discourse on race. Anti-racist activists often use this concept to explain the resilience of racial inequality in the post-civil rights era. Coined by Kwame Ture and Charles Hamilton, the concept of institutional racism emerged in reaction to the dominant framework for understanding racism in terms of individual racial prejudice (Ture and Hamilton 1992). Instead, the concept of institutional racism focuses on the role that institutions—public (e.g. the police, the court system, prisons, the education system) and private (e.g. corporations, banks, real estate agencies)—play in the reproduction of racial inequality. With its emphasis on institutional dynamics that operate even in the absence of individual racial prejudice, the concept of institutional racism aims to explain the durability of racial inequality in the post-civil rights era. Moreover, proponents of this conceptual framework often emphasize that institutions can be racist partly in virtue of their negative impact on subordinate racialized groups. I refer to this as the impact-based model of institutional racism.

In light of the renewed interest in the concept, it is important that philosophical accounts of institutional racism incorporate the aforementioned insights from anti-racist theory and practice. At the same time, philosophical accounts of institutional racism—especially those that embrace the impact-based model—must be careful not to over-inflate the concept to the point where it loses its analytical value. Tommie Shelby's recent work on institutional racism—specifically his distinction between intrinsic and extrinsic institutional racism—is a welcome contribution to this conceptual
Shelby’s two-pronged theory of institutional racism makes room for the insight that institutions can be racist in virtue of their impact (extrinsic case), and not only in virtue of their constitutive features (intrinsic case). However, Shelby’s account requires revision with respect to the conditions for attributing racism to institutions—particularly in the extrinsic case. The problem with Shelby’s account lies in his understanding of the connection between racist ideology and extrinsic institutional racism. A key aim of this paper is to show that Shelby’s view of this connection leads to an over-inclusive account of institutional racism. My argument for this claim takes the form of a *reductio ad absurdum*. I show that Shelby’s account leads to the counterintuitive claim that institutions that adopt affirmative action policies are extrinsically racist. This result is problematic from the point of view of conceptual analysis because it reinforces the worry that the impact-based model of institutional racism over-inflates the meaning of “racism”. However, the idea that institutions can be racist partly in virtue of their negative racial impact is key to contemporary anti-racist theory and practice. Rather than giving up on the impact-based model of institutional racism, I propose an alternative account that does not lead to the counterintuitive claim that affirmative action institutions are racist. In order to do so, I clarify how racist ideology operates in extrinsically racist institutions. In addition, I clarify how the concept of institutional racism contributes to the anti-racist project of explaining the reproduction of racial inequality.

The paper proceeds as follows. In section 1, I briefly describe Shelby’s account of racism as ideology. In section 2, I lay out his account of institutional racism, focusing on his distinction between intrinsically and extrinsically racist institutions. In section 3, I use the case of affirmative-action
institutions to perform a reductio against his view of the connection between racist ideology and extrinsically racist institutions. In section 4, I give my own view of that connection in order to salvage the impact-based model of institutional racism. In a nutshell, what makes institutions extrinsically racist is not the work they do for racist ideology, but the work that racist ideology does for them.

1. Shelby’s account of racism as ideology

According to Shelby, racism is first and foremost an ideology. Shelby defines ideology as “a widely held set of loosely associated beliefs and implicit judgments that misrepresent significant social realities and that function, through this distortion, to bring about or perpetuate unjust social relations” (Shelby 2014, 66). Although Shelby identifies ideologies primarily with sets of beliefs, he also claims that ideologies tend to become deeply entrenched in discourse and cultural products (slogans, jokes, film, music, art, television programming, advertisements, etc.), as well as in society’s common sense. By this, I take him to mean that ideologies alter the shared frames through which individuals make sense of social reality and their social identities. In other words, ideologies include not just sets of beliefs, but also the conceptual repertoires that color people’s experience of the world. Ideologies that become part of a society’s background assumptions not only distort the way social actors perceive the social world; they also affect how they organize their lives and coordinate their actions.

Drawing on Raymond Geuss’ account of ideology, Shelby identifies three properties that all ideologies share (Guess 1981; Shelby 2003, 164). First, ideologies are epistemically flawed. Most ideologies are not simply false; in fact, because ideologies distort the way we perceive reality, the observed “facts” often appear to confirm the content of the ideology. Secondly, ideologies serve a hegemonic function insofar as they contribute to the reproduction of unjust social arrangements, which are often rooted in oppressive social relations that benefit a hegemonic group. The most common way in which ideologies perform this function is through legitimation—that is, by creating the appearance that the unjust social arrangements are justified. Finally, ideologies have a genetic dimension insofar as there is...
something troubling about their genesis or history. In particular, ideologies are subject to genetic criticism on the basis of how and why agents come to adopt them. The genetic critique applies when agents tend to embrace an ideology out of a desire to maintain a positive self-conception or a dominant social position, rather than out of purely epistemic reasons. Importantly, the agent that adopts the ideology need not be aware that she is doing so out of non-epistemic reasons because these often operate behind our backs.

Building on his account of ideology, Shelby defines racism as “a set of misleading beliefs and implicit attitudes about “races” or race relations whose wide currency serves a hegemonic social function” (Shelby 2014, 66). As an ideology, racism has epistemic, functional and genetic properties that call for criticism. For example, classical racist ideology consists of a set of beliefs, central to which is the biological concept of race, which sorts human populations hierarchically on the basis of phenotypical traits, hereditary lines, and continental origin. Classical racist ideology is epistemically flawed because of the pseudo-scientific basis that underlies this sorting. Classical racism can also be criticized on functionalist grounds insofar as it serves the function of justifying projects of racial domination, such as the transatlantic slave trade and European colonialism. The illusory belief in the natural inferiority of Blacks and Native Americans functioned to legitimize the oppression and near extermination of these populations. We may also subject classical racist ideology to genetic critique in virtue of how and why it became widely accepted. For example, the slave-holding aristocracy of the American South adopted classical racist ideology not out of purely epistemic reasons, but because it benefited their social position in the local slave economy (Shelby 2003, 182).

So far, I have described the ideological properties of racism by referring to classical or biological racist ideology. However, like other ideologies, racism is not static. Its content is constantly changing in response to shifting cultural, political, and economic contexts, as well as social criticism (Shelby 2014, 66; Bobo, Kluegel, and Smith 1997; Taylor 2013). Classical racism, with its now debunked
doctrine of white biological superiority, has been replaced (for the most part) by cultural racism. Unlike its predecessor, cultural racism relies on non-biological themes, such as black cultural pathologies, the disintegration of the black family and blacks’ lack of individual motivation, to explain away the persistence of racial inequality. Despite this change in content, cultural racism serves the same hegemonic function as classical racism; namely, to legitimize a system of racial subordination.

We have established that a defining feature of racist ideologies is that they fulfill a hegemonic (or legitimizing) function in the broader system of racial domination. For example, racist concepts and stereotypes such as the biological and hierarchical concept of race and Blacks’ presumed lack of individual motivation serve to legitimize the reproduction of racial subordination. However, even the ways in which we think and talk about racism can serve this hegemonic function. Grant Silva has argued that defining racism in terms of intentional hatred is ideological insofar as it allows whites to distance themselves from the problem of racism (Silva 2019, 92-93). Defining racism as a problem of race-based antipathy and intentional harm invites us to think of anti-black racism as a thing of the past (the recent resurgence of white supremacist groups notwithstanding).

More importantly, the racism-as-intentional-hatred model allows whites to reap the material benefits of systemic racial inequality, while extricating themselves from the problem of racism. As such, this way of thinking about racism makes it easy for whites to lend their support (active and passive) to practices and institutions that reproduce racial inequality without questioning the legitimacy of their racially disparate impact. The impact-based model of institutional racism that this paper aims to vindicate is a reaction to the ideological effects of the racism-as-intentional-hatred model. Rather than emphasizing intentionality or antipathy, the impact-based model focuses on how institutions
Contribute to the reproduction of durable racial inequality. In the following section, I discuss a recent account of institutional racism that incorporates the impact-based model.4

2. Shelby’s account of institutional racism

Shelby distinguishes between two types of institutional racism: intrinsic and extrinsic. Although Shelby’s distinction is extremely helpful, his discussion of it is brief and opens the door to potential misunderstandings. In this section, I will fine-tune the distinction and illustrate it with some examples of my own. Institutions are intrinsically racist when their constitutive features are infected by a racist ideology.5 Shelby identifies three constitutive levels at which racism can infect an institution. First, the goals of an institution may be racist if it aims to harm the members of a racialized group and such aims are justified or rationalized by a racist ideology. Such racist goals need not be explicit or public; in fact, institutions that appear to have race-neutral goals may be covertly designed to subordinate a racialized group. An institution so designed will be intrinsically racist even if it fails to attain its ends. In other words, whether an institution is intrinsically racist is not necessarily a matter of its effects.

An institution can also be intrinsically racist if the content of its rules and/or criteria for assigning roles to its members contains racial bias or is racially discriminatory. Shelby conceives of institutions as social practices—that is, as “formal system[s] of roles and rules that enable and regulate sustained cooperative action for some specified purpose” (Shelby 2016, 25). To function appropriately, an institution requires its members to enact the role assigned to them and the rules associated with

---

4 As we will see, Shelby’s theory of institutional racism incorporates both the intentional-hatred and the impact-based models of racism. I see this as a virtue of his account. As I read Silva, what is ideological about the intentional-hatred model of racism is that it defines racism purely in terms of race-based antipathy and the intention to harm. A thorough analysis of racism must accommodate both models of racism. In fact, given the recent rise of white nationalist and neo-Nazi hate groups, it is key to also theorize racism from the intentional-hatred perspective. That said, this paper will focus on the impact-based model.

5 Note how Shelby’s account of intrinsic institutional racism resembles Garcia’s “infection model” of institutional racism. The difference is that, for Garcia, the essence of racism, which infects institutions is a set of race-based individual vices, ranging from hatred and ill-will to disregard.
the role. Institutional rules and role criteria need not be explicitly racist. All that is needed is for racially biased or discriminatory rules and role criteria to be operative in the institution, even if only at an implicit and informal level. The Ferguson Police Department is an example of an institution with racist rules. The DOJ’s Ferguson Report found that that one of the implicit common-sense rules followed by police officers was to treat Black citizens as “criminals” and as “lacking personal responsibility”. These racially biased rules that undergird the practice of policing in the FPD make the institution intrinsically racist.\(^6\)

Finally, racist ideology can make an institution intrinsically racist if it infects the application of its procedures. Even if the goals of an institution are non-racist and its rules and role criteria are free of racial bias, an institution may be intrinsically racist if its members systematically fail to apply those rules and role criteria equally and consistently because of personal prejudice. The pervasive racially discriminatory application of institutional procedures may be conscious or unconscious, but is always the result of individual actions informed by a racist ideology. For instance, even if the FPD purged its implicit and explicit rules from racial bias and discrimination it would still remain intrinsically racist so long as it does not address the police officers’ implicit biases, which lead them to use excessive force against Blacks.\(^7\) Insofar as police officers would systematically fail to apply the correct (non-racist) rules for the use of force when dealing with Black citizens—and insofar as this systematic misapplication of institutional rules would stem from the workings of racist ideology via the psychological mechanism of implicit bias—the FPD would continue to be intrinsically racist.

On the other hand, extrinsically racist institutions are racist not because of their constitutive features, but “in virtue of their policies’ effects” (Shelby 2016, 24). An institution that is not intrinsically racist may be extrinsically racist to the extent that it “perpetuate[s] the negative effects of

---

\(^6\) For a reading the US Justice Department Ferguson Report that further develops the point that racist ideology is operative behind the discriminatory practices of the FPD, see Gooding-Williams 2021.

\(^7\) A recent study confirms bias in police officers’ use of non-lethal force against Blacks and Hispanics (Fryer 2016).
ongoing or past racist actions and thereby encourage[s] racist attitudes and stereotypes” (Shelby 2016, 24). This second type of institutional racism accommodates Ture and Hamilton’s original use of the concept as a tool for explaining the subtle dynamics of racial oppression. In particular, “extrinsic institutional racism” covers cases of institutions that are not intrinsically racist, but which nevertheless contribute to the reproduction of systematic racial disadvantage.

“Extrinsic institutional racism” covers attributions of racism that emphasize an institution’s 

\textit{racist impact}. However, impact-based accounts of institutional racism are not without criticism. For example, J.L.A. Garcia argues that talk of institutions having “racist impact” and “racist outcomes” conflates racism with the effects of past racism, which to him are separate phenomena—just like warfare is separate from the lingering effects of warfare (Garcia 1999, 16-18). Thus, he recommends that we reserve the term “institutional racism” to cases in which the vice of racism infects an institution's constitutive features—that is, to intrinsic institutional racism.

Garcia’s criticism of impact-based accounts of institutional racism suggests that we cannot define extrinsic institutional racism \textit{solely} in terms of an institution's effects. Social dynamics are highly complex and most institutions—even those with anti-racist constitutive features—may at one time or another have the unintended consequence of contributing to the reproduction of racial inequality. And yet, recognizing that institutions can be racist \textit{in virtue of reproducing racial inequality} is a fundamental feature of the concept of institutional racism as advanced by anti-racist movements. There is no way to offer a comprehensive theory of institutional racism without accommodating some kind of impact-based features. Moreover, I do not think that impact-based accounts of institutional racism are intrinsically problematic. I agree with Garcia that counting institutions as extrinsically racist \textit{solely} in

\footnote{The influence of Ture and Hamilton on Shelby’s account of institutional racism is most explicit in Shelby 2014, 61-62, 71.}
César Cabezas  
Temple University

virtue of their negative effects on subordinate racialized groups is too simplistic. What we need is to further specify the conditions for attributing “extrinsic institutional racism”.

Shelby introduces a further requirement for categorizing an institution as extrinsically racist—namely, that it have a connection to racist ideology. After all, for Shelby racism is primarily an ideology, so it stands to reason that predicking racism of institutions would require that they be connected to racist ideology. In fact, the same requirement applies for intrinsically racist institutions. The precise nature of the connection between racist ideology and racist institutions will be the subject of sections 3 and 4. Before moving on, I will give an example to illustrate the concept of extrinsic institutional racism.

Word-of-mouth hiring is a widely used practice whereby employers channel job offerings through the personal networks of current employees. The rationale for this practice is that it gives employers a better chance of hiring a good candidate because current employees tend to refer people who they are willing to vouch for and who they think would be a good fit for the organization. In most cases, this recruiting technique has a race-neutral purpose; namely, to increase the odds of hiring a good candidate, while also keeping hiring costs down. However even if this social practice is not intrinsically racist, it has the effect of erecting an institutional barrier to employment for Blacks. Due to formal and informal segregation, Blacks are often excluded from white personal networks, which due to past and present racist practices, tend to have more social capital. Blacks’ exclusion from white personal networks that channel job offerings through word of mouth bars them from one of the most important mechanisms for success in the labor market and is an important cause for the white-Black unemployment gap (Ezorsky 1991, 14-18; DiTomaso 2013, 73-77).9 Thus, companies that hire through personal connections can be extrinsically racist insofar as

---

9 The unemployment rate of Blacks has been consistently twice that of Whites for at least 6 decades (Desilver 2013).
they have a negative impact for Blacks that stems from past (formal residential and social segregation) and present (informal residential and social segregation) racist practices.\(^\text{10}\)

It is important to note that the negative racial impact of word-of-mouth hiring does not require that whites who refer their (mostly white) friends and family intend to harm black people or harbor racist prejudices (Stikkers 2014, 12). Nor is it necessary that the institutions hiring through personal connections have racially discriminatory goals. This is a key feature of extrinsic institutional racism—and of impact-based models of institutional racism more generally. Attributions of racism on the impact-based model do not focus on the intention to harm, but rather on the impact that an institution’s practices have on the perpetuation of racial inequality and other race-related social problems afflicting Blacks and other subordinate racialized groups (Stikkers 2014, 7).

3. What is the connection between racist ideology and institutional racism?

Given Shelby’s account of racism as primarily an ideology, referring to phenomena other than ideologies as racist may at first seem like a category mistake. However, in our everyday talk, we refer to individuals, actions, institutions and even whole societies as racist. In Shelby’s framework, predicking racism of such phenomena is not misguided so long as we can explain them in terms of “the main characteristics and effects” of ideological racism (Shelby 2016, 24). In particular, ascriptions of institutional racism require us to show how racist ideology can be said to operate through institutions. In the case of intrinsically racist institutions, the connection to racist ideology is evident; the ideology operates through these institutions by infecting their constitutive features. The connection to racist ideology is not so clear in the extrinsic case, however. As discussed above, clarifying the connection between racist ideology and extrinsically racist institutions is key for salvaging the concept of extrinsic institutional racism from Garcia-like objections raised against impact-based

\(^{10}\) On the role of segregation in perpetuating racial inequality, see Anderson 2010.
accounts of racism. An institution cannot be (extrinsically) racist merely because it has a negative impact on subordinate racialized groups. That would make theories of institutional racism too expansive and would bolster those who—like Garcia—worry about the conceptual inflation of “racism” in contemporary discourse. To avoid this problem, ascriptions of extrinsic institutional racism must also require that the institution have a particular connection to racist ideology. I now turn to the important task of clarifying the nature of this connection.

3.1. Shelby’s view of the connection between racist ideology and extrinsically racist institutions

Shelby argues that racist ideology is operative in extrinsically racist institutions in the sense that these institutions reinforce racist ideology. As he puts it, an institution is extrinsically racist when its policies are race-neutral in their content and public rationale (i.e. the institution is not intrinsically racist), and yet such policies “perpetuate the negative effects of ongoing or past racist actions and thereby encourage racist attitudes and stereotypes” (Shelby 2016, 24). Given that racist attitudes and stereotypes are embodiments of racist ideology, the link he draws between extrinsically racist institutions and racist ideology is that the former contribute to the wide acceptance of the latter. For instance, an institution that employs the practice of word-of-mouth hiring reinforces racist ideology insofar as it contributes to the concentration of Black people among the unemployed, which in turn contributes to the wide acceptance of stigmatizing (i.e. ideological) representations of Black people as lacking personal motivation, responsibility, and work ethic.

However, this way of thinking about the relation between extrinsic institutional racism and racist ideology leads to an over-inclusive account of institutional racism. Following this logic, institutions that adopt race-based affirmative action policies would also count as extrinsically racist. I address the worry about the conceptual inflation of racism in Cabezas, unpublished manuscript, April 2022. Shelby makes the connection between racist ideology and racist attitudes and stereotypes explicit in Shelby 2016, 25. For the purposes of this paper, I will rely on Anderson’s definition of affirmative action as “any policy that aims to increase the participation of a disadvantaged social group in mainstream institutions, either through “outreach” or “preference”” (Anderson 2010, 135).
Social psychologists have found that race-based affirmative action leads many Whites to resent Blacks and to adopt stigmatizing views of them (Bobo 1998, Sidanius et al. 2000). Insofar as affirmative action contributes to the adoption of stigmatizing representations of Blacks, it reinforces racist ideology. Moreover, institutions that adopt affirmative action policies are not intrinsically racist. The goal of affirmative-action institutions is not to harm the members of a racialized group, but to level the playing field for members of historically disadvantaged racialized groups. The content of the rules and role criteria of these institutions is not racially biased or discriminatory.\(^{14}\) Finally, there is no evidence that racially biased administrators systematically misapply the rules and criteria of affirmative-action institutions. Therefore, on Shelby’s account, affirmative-action institutions fit the description of extrinsic institutional racism—they are not intrinsically racist and they bolster racist ideology. Since it would be absurd to call affirmative-action institutions racist (extrinsically or otherwise), clearly something has gone awry in Shelby’s account of the connection between racist ideology and extrinsically racist institutions. Let’s look at this argument in more detail.\(^{15}\)

1. (On Shelby’s view of the connection between racist ideology and extrinsically racist institutions) non-intrinsically racist institutions that reinforce racist ideology are extrinsically racist.

2. Affirmative-action institutions are not intrinsically racist.

3. Affirmative-action institutions reinforce racist ideology.

4. Therefore, (in Shelby’s account) affirmative-action institutions are extrinsically racist.


6. Since—as I will show—[2] and [3] are true, [1] must be false. Therefore, we ought to reject Shelby’s proposal of the relation between racist ideology and extrinsically racist institutions.

\(^{14}\) Critics of affirmative action often claim that it discriminates against Whites. I take up this challenge in section 3.3.

\(^{15}\) This argument is partially inspired by Garcia’s critique of impact-based accounts of institutional racism (Garcia 1997, 24).
César Cabezas
Temple University

In the remainder of this section, I fill in the details of this argument [1-6]. In 3.2, I offer an empirically informed argument in favor of [3]. I also deal with a potential misinterpretation of Elizabeth Anderson’s skepticism about affirmative action’s contribution to the stigmatization of Blacks, which could cast doubt on the truth of [3]. In 3.3, I argue for the truth of [2] by addressing the common objection that affirmative action is intrinsically racist insofar as it is racially discriminatory against Whites. In 3.4, I show why it does not make sense to categorize affirmative-action institutions as extrinsically racist, i.e. why [4] is false.

3.2. Does affirmative action reinforce racist ideology?

In the context of her defense of race-based affirmative action in *The Imperative of Integration*, Elizabeth Anderson casts doubt on a common argument against racial preferences. The argument in question claims that affirmative action harms its intended beneficiaries by causing them to experience internal and external stigma. This line of criticism predicts that minority students in affirmative-action institutions will suffer higher degrees of stigmatization than students in non-affirmative-action institutions. However, as Anderson rightly claims, recent studies have failed to confirm this hypothesis among law students and undergraduate science students (Onwuachi-Willig, Houth, and Campbell 2008; Bowen 2010). One may think that these studies cast doubt on my claim that affirmative action contributes to Whites’ adoption of stigmatizing representations of Blacks. If that were the case, it would be uncertain whether affirmative-action institutions actually reinforce racist ideology—that is, whether [3] is true. However, we ought to differentiate between two separate theoretical constructs—stigmatizing representations and stigma. Whereas my claim that affirmative action bolsters racist ideology rests on empirical research that deals with stigmatizing representations, the debate that Anderson cites deals with stigma. Moreover, the causal mechanism that underlies my claim is different from that which

---

16 Supreme Court Justice Clarence Thomas famously used this argument in *Adarand Constructors, Inc. v. Pena*, 515 U.S. 200. See also Eastland 1992.
is called into question by the research that Anderson cites. Therefore, Anderson's rightful criticism of the argument that affirmative action reinforces stigma among its beneficiaries has no bearing on my claim [3] that affirmative action bolsters racist ideology in the form of stigmatizing representations of Blacks.

Anderson introduces the concept of stigmatizing representations in her discussion of group stigmatization (Anderson 2010, 45-47). A representation of out-groups is stigmatizing when its content imputes dishonorable meanings to stereotypes about such out-groups, which in turn rationalizes antipathy towards that group. For instance, representations of Black people as dangerous and untrustworthy (dishonorable meanings) because they are presumed to be more likely to engage in crime (stereotype) rationalizes anti-Black antipathy. In this way, such representations are also stigmatizing. Moreover, the stigmatization of Black people as dangerous and untrustworthy criminals also serves to legitimize social policies that reinforce a system of racial subordination.17 This feature of stigmatizing representations—namely, that they legitimize an oppressive racial order—is what I earlier introduced as the hegemonic function of racist ideology.

The social psychological research that grounds my claim that affirmative action reinforces racist ideology operates with this concept of stigmatizing representations. In particular, group position theory advances that many Whites experience the Black gains that stem from affirmative action as a threat to their perceived entitlement to greater resources and status as members of a privileged racialized group.18 This perceived threat to the structure of racial privilege breeds feelings of anti-Black

17 As Michelle Alexander notes in The New Jim Crow: Mass Incarceration in the Age of Colorblindness, stigmatizing representations of blacks as criminals have replaced Jim-Crow-era stigmatizing representations of blacks as naturally inferior; today, the idea of black criminality serves to “justify a legal, social and economic boundary between “us” and “them” (Alexander 2010, 18). Her book offers a compelling narrative of how politicians, both Democrats and Republicans, have used stigmatizing representations of Black criminality to justify “though on crime” policies that enforce a system of legalized discrimination in housing, employment, education, public benefits, and jury service that harms Blacks disproportionately.

18 This argument is advanced by social psychologist Lawrence Bobo, who builds on Herbert Blumer’s theory of prejudice as a sense of group position to explain the wide acceptance of racist ideology in terms of the dynamic between feelings
hostility among many Whites, which, coupled with a motivation to protect their race-based advantages, leads them to embrace stigmatizing representations of Blacks that disqualify the validity of affirmative action—e.g. “Blacks are lazy”, “Blacks are less intelligent”, “Blacks prefer to live off of welfare” (Bobo 1998, Bobo 1999). According to group position theory, the causal chain whereby affirmative action contributes to the wide acceptance of racist ideology is as follows:

(1) Affirmative action redresses the effects of race-based disadvantage and increases Black access to greater resources and status.

(2) Many Whites perceive the Black gains that stem from affirmative action as a threat to their perceived entitlement to such resources and status, which breeds feelings of anti-Black resentment and hostility.

(3) Guided by their interest in protecting their relative group position (instrumental mechanism) and the anti-black affect that stems from the perception of threat (affective mechanism), many Whites readily embrace stigmatizing representations of Blacks (racist ideology) that cast doubt on the legitimacy of affirmative action. Therefore, according to group position theory, affirmative-action institutions have the effect of bolstering racist ideology.19

By contrast, the anti-affirmative action argument that Anderson rightly criticizes in *The Imperative of Integration* operates with the theoretical construct of *stigma*. This construct does not have to do with Whites’ stigmatizing representations of out-groups, but rather with minority students’ experience of stigmatization in affirmative-action institutions. What is at stake in this debate is the degree to which

---

19 Some may think that it is not affirmative action per se, but rather the Black gains from affirmative action that cause threatened Whites to embrace racist ideology. In this respect, it is important to remember that Whites who are threatened by Black gains are prone to embrace racist ideology precisely because they regard affirmative action as unjust. It is the fact that such gains *stem from* affirmative action, which they regard as an illegitimate practice that discriminates against Whites, that gives threatened Whites a readymade justification for adopting racist ideology. Thus, while affirmative action contributes to redressing racial inequities, it also inadvertently contributes to the wide acceptance of racist ideology.
affirmative action causes minority students to experience (a) feelings of doubt of their own qualifications (internal stigma) and (b) the burden of being treated or viewed by others as unqualified, based on the assumption that they are beneficiaries of affirmative action (external stigma) (Onwuachi-Willig, Houth, and Campbell 2008, 1301-1303). As Anderson rightly points out, recent studies found no evidence of affirmative action having a causal effect on the prevalence of internal or external stigma among minority students.

However, it would be a mistake to interpret these findings as a challenge to group position theory. As I already mentioned, both lines of research are interested in distinct phenomena—while group position theory focuses on Whites’ adoption of stigmatizing representations of Blacks, the research Anderson cites focuses on the experience of stigmatization among minority students in educational institutions. Moreover, the proposed causal mechanisms in both lines of research are substantially different. Thus, the studies in question have no bearing on the social psychological research that grounds my claim that affirmative-action institutions reinforce racist ideology. Their important findings should not be interpreted as casting doubt on the truth of [3].

3.3. Are affirmative-action institutions intrinsically racist?

Another potential objection to my argument against Shelby [1-6] is that affirmative-action institutions are intrinsically racist, and therefore [2] is false. This objection gets at the heart of my reductio by biting the bullet and denying the absurdity of calling affirmative-action institutions racist. And to be clear, this objection does not merely propose that racism is extrinsic to such institutions; it makes the stronger claim that racism is intrinsic to these institutions’ constitutive features. Of course, this is not an unheard-of position. After all, the claim that affirmative action amounts to reverse racism is widespread in everyday discourse on race. To my mind, the most defensible version of this objection states that affirmative-action institutions are intrinsically racist to the extent that their rules and criteria
discriminate against Whites. If we recall Shelby's typology of intrinsic institutional racism, this would be an example of racist ideology infecting the rules and criteria of an institution such that they contain racial bias or are racially discriminatory. On this view, affirmative-action institutions are racially discriminatory—and hence, intrinsically racist—against Whites who will see their life chances diminished through no fault of their own.

In order to address this objection, I must introduce a key distinction in the philosophical literature on discrimination—that between normatively-laden and normatively-neutral senses of discrimination (Arneson 2006, Hellman 2008, Blum 2013, Glasgow 2015, Shelby 2016). In a normatively-neutral sense, discrimination favors some over others on the basis of a distinction or preference. In this purely descriptive sense, there is no judgment as to whether the basis for such differential treatment is justified. Some forms of discrimination may be justified, or even desirable, such as favoring those with steady hands to become surgeons, even if this preference discriminates against those with shaky hands (Shelby 2016, 30). By contrast, in a normatively-laden sense, discrimination refers to a morally wrongful differential treatment on the basis of some distinction or preference. Affirmative action is indeed discriminatory in the normatively-neutral sense that it uses race as a criterion for differential treatment. However, the sense of discrimination that is at stake when affirmative action is accused of being racist is the normatively-laden one. After all, racism is a morally loaded term that implies some degree of moral wrongness. Thus, in order to assess whether affirmative is intrinsically racist, we must examine whether it constitutes wrongful discrimination.

A thorough discussion of what exactly makes discrimination wrongful is outside the scope of this paper. Since I am examining Shelby’s account of institutional racism, my analysis here will start...
from his own criteria for defining wrongful race-based discrimination. In his view, racial discrimination is wrongful when the race-based differential treatment is (a) based on the presumption that one race has an inferior moral status, (b) based on racial considerations where race is not a relevant principle of differentiation that furthers a legitimate end, and (c) caused by a racially biased application of institutional rules and criteria. In what follows, I will explain why race-based affirmative-action institutions fail to meet conditions (a-c), and therefore are neither racially discriminatory (in the normatively-laden sense) nor intrinsically racist.

Regarding (a), affirmative-action institutions treat members of all races as having inherent and equal moral worth. If they use race as a principle for differential treatment, it is not because they regard Whites (or any other racialized group) as moral inferiors or subpersons, but in order to rectify past and present racial injustices. Condition (c) covers cases in which wrongful racial discrimination does not stem from morally objectionable institutional rules and criteria, but from a racially biased application of morally legitimate rules and criteria by members of the institution. Affirmative-action institutions—like all institutions—are susceptible to racially biased individual members introducing racial discrimination in their application of institutional rules. However, the issue at stake here is whether affirmative action's rules and criteria are racially discriminatory, independently of how institutional agents apply them. Therefore, condition (c) is not relevant for assessing whether affirmative action is racially discriminatory or intrinsically racist.

Thus, the crucial question revolves around (b) whether affirmative action's use of race as a principle for differential treatment furthers a morally legitimate end. In addition to rectifying past racial injustices, race-based affirmative action may also be pursued with the aim of increasing institutional diversity, countering continuing discrimination, and contributing to racial integration by encouraging
César Cabezas  
Temple University  

constructive relationships across racial lines. Alas, a thorough normative analysis of the aforementioned goals of affirmative action is outside the scope of this paper. Still, I am confident that the reader will find at least one of these goals to be morally legitimate. Consequently, affirmative-action institutions fail to meet conditions (a-c), and therefore are neither racially discriminatory (in the normatively-laden sense) nor intrinsically racist. Pace the objection I introduced at the beginning of this section, [2] is true: affirmative-action institutions are not intrinsically racist.

3.4. Is Affirmative Action Extrinsic Racism?

I have already defended the claim that affirmative-action institutions are not intrinsically racist. A different strategy to counter my argument would be to claim that affirmative-action institutions are extrinsically racist, and therefore Shelby’s theory of institutional racism does not lead to an over-inclusive account after all. In order to show why it would be incorrect to attribute extrinsic racism to affirmative-action institutions, it is helpful to revisit the definition of extrinsic institutional racism. This concept refers to those institutions that, although not intrinsically racist, contribute to the reproduction of systematic racial disadvantage. Given that affirmative action institutions are not intrinsically racist (as shown in 3.3), it may turn out that they are extrinsically racist if they reinforce systematic racial disadvantage. In this section, I show that affirmative-action institutions could not be farther from fitting this description. In fact, affirmative-action institutions are designed to address the

23 For a thorough discussion of these four kinds of institutional goals behind affirmative action, see Anderson 2010, ch. 7.
24 For a defense of the moral legitimacy of the goals of affirmative action, see Anderson 2010, ch. 7, as well as Blum 2013.
25 The reader may worry that this part of my argument relies on the endorsement of Shelby’s particular view of what makes discrimination wrongful. Given that this (undoubtedly important) normative question is tangential to the overarching goal of this paper, I cannot discuss it further here. That said, while there is ample disagreement about what makes discrimination wrongful, there is a fair degree of consensus that affirmative action does not amount to wrongful discrimination (regardless of what account of wrongful discrimination one favors). For example, in a recent edited volume (Lippert-Rasmussen 2017), the reader can find four additional distinct arguments for the claim that affirmative action is not wrongfully discriminatory (chs. 7, 11, 16, 33).
twin processes of racial accumulation and disaccumulation that are at the root of the reproduction of systematic race-based disadvantage.

According to the theory of racial accumulation and disaccumulation, even very small race-based economic and social advantages compound and can, like an investment, have large cumulative effects over many generations (Brown and Wellman 2005). Conversely, Blacks’ exclusion from economic and social opportunities compound over time and result in inter-generational racial disaccumulation. In pre-civil rights America, labor market discrimination, as well as an array of discriminatory federal social policies (e.g. veterans’ programs, federal housing policy, welfare policy) resulted in White accumulation of economic advantage while contributing to Black disaccumulation through lack of access to secure and well-paying jobs, redlining, and the disaccumulation of wealth (Brown and Wellman 2005, 196-202). Since the dismantling of Jim-Crow discrimination did not include policies that redressed the inter-generational effects of the twin processes of White accumulation and Black disaccumulation, Blacks continue to be systematically disadvantaged vis-a-vis Whites in their competition for resources and opportunities—including in employment and education. As sociologist Dalton Conley shows, the higher levels of White wealth that result from inter-generational accumulation translate into educational and labor market advantages (Conley 1999). Therefore, the assumption of a post-civil rights America in which Blacks and Whites compete on a level playing field is empirically false. Without policies that rectify the effects of racial accumulation and disaccumulation, Whites continue to benefit from a systematically conferred competitive advantage over Blacks, even in the absence of formal discrimination.26

It is in this context that affirmative action’s morally legitimate goal comes in. Affirmative action works to correct—and only partially at that—the inequality of opportunity that stems from the twin processes of racial accumulation and disaccumulation. In other words, affirmative-action institutions

---

26 For further empirical evidence of this claim, see Loury 2003 and Massey 2008.
are the polar opposite of extrinsically racist institutions. Whereas extrinsically racist institutions perpetuate the negative effects of ongoing or past racist action, affirmative-action institutions are designed and have been shown to redress these effects.\(^{27}\) Therefore, affirmative-action institutions are not extrinsically racist and \([4]\) is false.

I have now defended each of the key claims in my argument \([1-6]\) against Shelby’s account of the connection between racist ideology and extrinsic institutional racism. I have shown that Shelby’s view of this connection leads to an over-inclusive account that would count affirmative action institutions as extrinsically racist. The original rationale for clarifying the connection between racist ideology and extrinsic institutional racism was to address Garcia’s worry that impact-based models of institutional racism are overly expansive and lead to conceptual inflation. Insofar as Shelby’s proposal leads to an over-inclusive account, it fails to address Garcia’s worry. However, my aim here is not to deny the viability of the impact-based model of institutional racism. Quite the opposite. I welcome Shelby’s category of extrinsic institutional racism precisely because it allows us to accommodate cases of institutions that are racist partly in virtue of their impact. I say partly because, even in Shelby’s account, predicating extrinsic racism of an institution requires more than contributing to systematic racial disadvantage (i.e. impact). A further necessary condition is that the institution also have a particular connection to racist ideology. Given that Shelby’s account of this connection leads to conceptual inflation, we need a different account of the connection between racist ideology and extrinsic institutional racism in order to salvage the impact-based model.

4. What, then, is the connection between racist ideology and extrinsically racist institutions?

In order to answer this question, we need to return to one of the key features of all ideologies, namely their hegemonic function. Ideologies function to support, stabilize, or legitimize unjust states

---

\(^{27}\) For empirical evidence that race-based affirmative action succeeds at redressing systematic race-based disadvantage, see Ezorsky 1991, 63-72.
of affairs that benefit a hegemonic group (Guess 1981, 15). In this context, the unjust state of affairs is durable racial inequality that benefits whites as a group, and which has its origins in state-sanctioned racist projects, such as slavery and Jim Crow. Today, this racial inequality is partially sustained by the racist impact of extrinsically racist institutions (e.g. institutions whose word-of-mouth hiring practices contribute to the black-white unemployment gap). These institutions, although not infected by racist ideology in their constitutive features, contribute to the reproduction of racial inequality by distributing resources and opportunities unequally along racial lines. I propose that we conceive of the connection between racist ideology and extrinsically racist institutions in the following way: racist ideologies stabilize the racist impact of extrinsically racist institutions by making their impact appear legitimate and unproblematic. By legitimizing the racist impact of extrinsically racist institutions, racist ideologies undermine popular support for efforts aimed at reforming or abolishing them. In this way, racist ideologies effectively work to support the reproduction of racial inequality.

According to sociologist Eduardo Bonilla-Silva, one of the main racist ideologies playing this legitimizing role today is colorblind ideology (Bonilla-Silva 2006). In the following pages, I explain how colorblind ideology works to legitimize the racist impact of extrinsically racist institutions. However, we must be careful to distinguish colorblindness as an ideology from colorblindness as an ideal and a policy proposal (Anderson 2010, ch. 8). As an ideal, colorblindness aspires for a society in which people do not draw racial distinctions. The ideal of colorblindness opposes the use of racial distinctions for the purposes of justifying inequality, stigmatization, and discrimination. Defenders of colorblindness as a policy proposal argue that it is the best means to fight against racial injustice. They contend that race-conscious policies should be avoided because they are inherently morally objectionable and have bad consequences. Colorblind ideology shares many of the features of

---

28 See the aforementioned chapter for Anderson’s refutation of the criticisms of race-conscious policies, including a point-by-point treatment of their alleged bad consequences; namely, that they are divisive, stigmatizing, inefficient, depress motivation and harm their intended beneficiaries.
colorblind ideals and policy proposals, but it distorts them and puts them to the service of sustaining durable racial inequality. Instead of asserting the need for a society in which racial distinctions no longer ground inequality, discrimination and stigmatization, colorblind ideology depicts our society as one in which this ideal has already been realized. Colorblind ideology presents racism as a thing of the past by pointing to the perceived reduction of overt racist attitudes and the dismantling of formal state-backed discrimination. In so doing, it masks the impact that past racial discrimination continues to have in the reproduction of racial inequality, as well as the need for institutional measures to rectify past racial harms (Sundstrom 2008, 43). The call to become “blind to color” undermines popular support for color-conscious programs that are necessary to establish equality of opportunity in a society with a history of unaddressed racial injustice. Having clarified the concept of colorblind ideology, I now turn to explaining how it performs its hegemonic function.

Bonilla-Silva argues that dominant frames are one of the central components of racist ideology. According to Bonilla-Silva, dominant frames are set paths for interpreting and filtering information, which when used, lead people to “explain racial phenomena following a predictable route” (Bonilla-Silva 2006, 26). Frames can be dominant (used by the dominant group to justify the status quo) or subversive (used by the subordinate group to challenge the status quo). However, dominant frames tend to become the master framework that becomes entrenched as society-wide common sense. The dominant frames of colorblind ideology function as conceptual repertoires that shape our understanding of the world on matters of race. Bonilla-Silva’s notion of dominant frames is important because it highlights how colorblind ideology performs its hegemonic function; namely, by spreading a racial common sense that condones “as much Black disadvantage and segregation as the legacy of historic discrimination and modern-day free market forces and informal social mechanisms can reproduce or even exacerbate” (Bobo, Kluegel, and Smith 1997, 7).
Agents under the influence of colorblind ideology tend to perceive and explain the reproduction of racial inequality in terms of individual dispositions and choices, rather than structural constraints to Black improvement. Instead of recognizing that colorblind practices (e.g., word-of-mouth hiring) play a key role in the reproduction of racial inequality when past racial harms have not been fully redressed, colorblind ideology explains the persistence of racial inequality in terms of non-whites presumed cultural inferiority (e.g., Blacks’ lack of work ethic and personal responsibility).

So far, I have described how colorblind ideology performs its legitimizing (or hegemonic) function via various forms of epistemic flaws, such as distortion, misrepresentation, masking, and framing effects. In other words, my discussion up until this point has focused on two of the three forms of ideology-critique described by Raymond Guess: one based on ideology’s epistemic flaws and the other based on the function that ideology plays in the reproduction of hegemony and domination (Guess 1981, 12-22). The third kind of ideology-critique according to Guess is genetic and focuses on the origins of how an ideology becomes socially dominant. In particular, the genetic critique applies when agents tend to embrace an ideology out of a desire to maintain a positive self-conception or a dominant social position.

Grant Silva’s account on racism as a distorted form of self-love can help us flesh out the genetic critique of colorblind ideology (and racist ideology, in general). According to Silva, racially privileged agents readily embrace racist ideologies like colorblindness out of a desire to maintain the economic and social benefits they derive from systematic racial inequality. In other words, racially privileged agents’ overwhelming adoption of colorblind ideology is not only (or even mainly) due to a purely epistemic mistake. Rather, colorblind ideology has become the dominant frame for thinking about racial matters precisely because it offers racially privileged agents a worldview that legitimizes

---

29 See also section 1 above.
their perceived entitlement to economic privileges and social standing (Silva 2019, 99). Thus, another reason to criticize colorblind ideology is that its wide adoption among racially privileged agents has its origin in their “protective attachment to the racialized dimensions of [their] social status, esteem, wealth, privilege, and identity” (Silva 2019, 86).

Colorblind ideology also makes it more difficult to recognize and challenge the racist impact of extrinsically racist institutions. In a recent study of racial attitudes in the U.S., Lawrence Bobo shows that Whites’ tendency to oppose affirmative action policies is correlated with their tendency not to perceive much racial inequality, or to explain it in terms of individual motivation (Bobo 2011, 26-27). To the extent that colorblind ideology filters many Whites’ experience of durable racial inequality in this way, they become less likely to see it as a problematic state of affairs or to even acknowledge its existence. Since racism and its effects are perceived as no longer relevant, Blacks’ lack of social mobility appears to stem from their lack of personal drive in a meritocratic society that rewards those who pull themselves up by their bootstraps. By framing agents’ understanding of the mechanisms behind the reproduction of racial inequality, colorblind ideology grants legitimacy to the racist impact of extrinsically racist institutions.

This legitimizing effect of colorblind ideology allows extrinsically racist institutions to remain mostly unchallenged. And to the extent that these institutions are subject to social criticism, colorblind ideology undermines popular support for reformist efforts. For those seeing through the frames of colorblindness, these institutions are unproblematic insofar as their constitutive features are not racist. These institutions’ racist impact is not seen as a violation of fair equality of opportunity, but as the unintended outcome of Blacks’ sub-par individual choices and free market forces. By legitimizing extrinsically racist institutions in this way, colorblind ideology obscures the need for corrective measures that prevent these institutions from perpetuating the effects of past and present racist practices. To sum up, colorblind ideology is operative in extrinsically racist
institutions in the sense that it legitimates these institutions’ reproduction of racial inequality through a variety of ideological mechanisms such as distortion, misrepresentation, masking, and framing effects.

This type of ideological racism does not operate at the constitutive level of institutions as in the intrinsic case. Rather, it works by permeating the societal common sense on matters of race in a way that prevents agents from identifying the structural mechanisms behind the reproduction of racial inequality. Colorblind ideology operates at the level of the conceptual repertoires that guide agents’ understanding of racial matters, such as the persistence of racial inequality. A society in the grip of colorblind ideology perceives extrinsically racist institutions as legitimate and unproblematic, with the corollary that it tends to keep these institutions, along with their racist impact, intact.

Conclusion

In this paper, I argued that Shelby’s account of extrinsic institutional racism does not quell Garcia-like worries about the conceptual inflation of impact-based accounts of institutional racism. In order to rein in the concept, Shelby proposes that ascriptions of extrinsic institutional racism should require that the institution have a particular connection to racist ideology. In his view, that connection is that extrinsically racist institutions reinforce racist ideology in the form of racial prejudices and stereotypes. I argued that this account yields a wrong outcome—namely, that affirmative-action institutions are extrinsically racist. I agree with Shelby’ approach of specifying the connection between extrinsic institutional racism and racist ideology as a way to rein in the concept. I also agree with the empirical claim that the racist impact of extrinsically racist institutions reinforces racist prejudices and stereotypes. However, for the purposes of conceptual analysis, that particular connection cannot be a necessary condition for ascribing extrinsic institutional racism because that would lead to an over-inclusive account.
My proposal for specifying the connection between racist ideology and extrinsic institutional racism is the reverse of Shelby’s. I argue that racist ideology reinforces extrinsically racist institutions by masking and legitimizing their racist impact. To recall the slogan introduced at the outset, what makes institutions extrinsically racist is not the work they do for racist ideology, but the work that racist ideology does for them. One virtue of my account is that it accommodates a key insight from anti-racist movements—namely, that institutional racism and racist ideology are part of a broader system of racial oppression. It does so by describing the feedback loop whereby racist ideology and extrinsic institutional racism jointly reproduce racial inequality. Shelby already explained how extrinsically racist institutions reinforce racist ideology. My account completes his picture by showing how racist ideology reinforces those same institutions, thus leading to a homeostatic system of racial inequality.30 Another virtue of my account is that it centers a key feature of all ideologies—namely, their hegemonic function. My account highlights how racist ideology performs its function of sustaining a system of durable racial inequality. Finally, my proposal is better positioned to quell conceptual inflation worries because it does not lead to an over-inclusive account of extrinsic institutional racism. In my picture, affirmative-action institutions are not extrinsically racist because they are not bolstered by racist ideologies, such as colorblindness. Quite the opposite; colorblind ideology forcefully opposes affirmative action and works to undermine its legitimacy. This is the expected result given that affirmative-action institutions are the antithesis of extrinsically racist institutions, which colorblind ideology works to stabilize.

30 On racism as a homeostatic system, see Haslanger 2017.
References


