5-2021

# Moral Duties Amid Legal Bias

Benjamin Klafter
bklafter@bates.edu

# Moral Duties Amid Legal Bias

An Honors Thesis

Presented to

The Faculty of the Department of Philosophy

Bates College

In partial fulfillment of the requirements for the

Degree of Bachelor of Arts

By

Benjamin K. Klafter

Lewiston, Maine

May 5th, 2021

**Acknowledgements**

I would like to acknowledge and extend my sincerest gratitude to all those who have helped me along the way in the process of writing this thesis. First and foremost, I would like to acknowledge Professor David Cummiskey, without whose guidance I would never have been able to complete this thesis. I would also like to acknowledge Professors Susan Stark, Paul Schofield, Lauren Ashwell, Mike Dacey, and Mark Okrent, whose instruction in philosophy gave me the foundation of knowledge which I needed in order to undertake this project. In addition, I would like to acknowledge those who gave me personal support and encouragement while I was writing this thesis: namely, I would like to thank my parents, my sister, and my friends and classmates at Bates College.

**Table of Contents**


Acknowledgements 2

Table of Contents 3

Abstract 4

Introduction 5

Contemporary Legal Theory 8

Background Claims 11

Chapter 1: The Moral Duty to Effect Substantive Justice 19

Chapter 1.1: The Fugitive Slave Act 19

Chapter 1.2: In Defense of Role Infidelity 22

Chapter 1.3: The Dworkinian New Naturalist Objection 26

Chapter 1.4: Ideal and Non-Ideal Theory 31

Chapter 1.5: The "Insurgency" Objection 33

Chapter 2: The Value of Role Fidelity Examined 38

Chapter 2.1: The "Threat to Democracy" Objection 38

Chapter 2.2: The Rawlsian Objection 42

Chapter 2.3: The Value and Nature of Authority 47

Chapter 2.4: Evaluating Unrestrained Role Infidelity 58

Conclusion 71

Bibliography 75

**Abstract**

The United States Constitution and the American legal system which has stemmed from it specify the roles of judges, prosecutors, jurors, and other actors within the criminal justice system. These roles, although they include a certain amount of discretion, prohibit their occupants from taking justice "into their own hands" by unilaterally deciding whether the law ought to apply in a given case. These constraints on the roles within the criminal justice system are justified by the broader democratic, egalitarian political system within which they exist. However, given the increasing body of evidence which suggests that, far from being egalitarian, the criminal justice system is heavily influenced by various types of bias, including racial bias, the constraints on individuals' roles within the criminal justice system must be critically reexamined. This thesis, by way of performing such a reexamination, defends the position that to the extent that racial bias is likely to affect the outcome of a given criminal case, the individual actors within the system who deal with that case (specifically, the judges and prosecutors), have the moral right and, in some cases, the moral duty to promote the most substantively just outcome, even if doing so would require stepping outside the boundaries of their defined roles. Despite the moral importance of remaining faithful to the boundaries of one's role, I argue that in some cases, stepping outside the boundaries of one's role may actually better serve the purposes of that role than would remaining faithful to the role's boundaries.

**Introduction**

Under conditions of social injustice, individual moral duties and professional duties may come into conflict with one another. In this thesis I will examine this topic by focusing on critical race theory and the moral dilemma for individual actors - judges and prosecutors in particular - in the United States legal system that results from racialized legal decision-making. I will ground my thesis in the following empirical claim, largely drawn from critical race theory (CRT): the law is created, enforced, and adjudicated within a racially unequal society, and as such tends to reinforce the racial inequities of the society.

This topic is philosophically important because there is generally a disconnect between the majesty of legal reasoning as it is enshrined in judicial opinions and described in academic literature, on the one hand, and the historical and current trends in actual criminal procedure, on the other hand. Although legal reasoning relies on terms and concepts which are, in themselves, non-discriminatory, and such reasoning is governed by principles which explicitly prohibit discrimination based on certain factors, the legal system frequently produces outcomes which cannot be plausibly explained as the results of non-discriminatory reasoning. Conversely, the variety of legal principles and the open texture of legal language allows actors in the legal system to justify discriminatory decisions by appealing to non-discriminatory concepts.

I will defend the conclusion that there is a disconnect between legal reasoning and criminal procedure so stark that it cannot be attributed to mere coincidence or to the incompetence of a few specific actors, and rather calls into question whether criminal procedure, as it is practiced, is based on unbiased legal reasoning at all. My thesis will posit that this disconnect stems from the underdetermination of outcomes by law, which legal realism describes, combined with the significant role which racial bias plays in determining outcomes.

In examining the moral import of this claim about the influence of racial bias in criminal procedure, I will be focusing on judges and prosecutors and the moral duties relevant to their professional decisions, rather than addressing the question of whether a given citizen is morally permitted to break the law. Intuition and legal philosophy suggest that actors in the legal system, judges in particular, have a special allegiance to legal doctrines, greater than that of other citizens, so by analyzing the moral duties of judges and prosecutors in their roles as such, I plan to engage with a more controversial, substantive question than that of whether ordinary citizens can break unjust laws.

The conclusion about this question which I will defend is that in many cases, a judge or prosecutor is morally permitted or even obligated to act in a way which best accords with justice, even if such an action would mean contradicting relevant legal doctrines. In order to defend this conclusion, I will first provide background about contemporary legal theories and their relationship to my position. Next, I will lay out the set of ethical and empirical background claims which will serve as the foundation of my argument. After that, in the chapter "The Moral Duty to Effect Substantive Justice," I will make a series of arguments defending the position that individual judges and prosecutors are morally compelled to promote substantively just outcomes in criminal cases and that the discretion built into the judicial and prosecutorial roles is insufficient to allow those actors to properly promote substantive justice. The first such argument will use Robert Cover's writing about the Fugitive Slave Acts to illustrate that roles within the criminal justice system are not self-actualizing and that, therefore, each actor must decide to follow or not to follow their role in each case. After establishing that there is a choice between maintaining what Cover calls "role fidelity" and violating that role fidelity, I will use Lysander Spooner's writing to argue that one ought to choose to violate role fidelity in certain

circumstances, so as not to allow oneself to be complicit in the perpetration of moral wrongdoing. Following this, I will elaborate the objection that my argument assumes a positivist legal philosophy and that under a Dworkinian philosophy in which moral consideration is internal to the judicial role, my argument fails. I will respond to this objection by showing that Dworkin's legal philosophy does not grant enough discretion to allow judges and prosecutors to promote substantive justice in the contexts in which I recommend it. Next, I will address the objection that my response to the previously-described objection criticizes ideal theory as such, even though my position is rooted in ideal theory. To conclude the chapter, I will develop the objection that I am encouraging people to act as "insurgents" infiltrating the criminal justice system and then respond to that objection using the doctrine of double effect.

In the next chapter, titled "The Value of Role Fidelity Examined," I will put forward several arguments in favor of maintaining role fidelity and respond to those arguments. I will begin the chapter by creating an argument that role fidelity on the part of judges and prosecutors is necessary in order to preserve American democracy. I will respond to this argument by employing Cover's arguments about the tendency of judges in Fugitive Slave Act cases to inflate the negative consequences of role infidelity. Next, I will explain John Rawls' argument that rules, rather than providing reasons for or against specific actions, define the practices within which actions are performed, leaving the actor no choice but to follow the rules if they want to perform that action. I will respond to Rawls' argument, which clearly challenges the justifiability of role infidelity, with several arguments of my own. Following this, I will explore Scott Shapiro's critical evaluations of different conceptions of authority in order to illustrate the ways in which my position would jeopardize the valuable functions of authority, as well as the ways in which my position could coexist with or even improve the functions of authority. After that, I

will elaborate the views of Kimberley Brownlee, who advocates for role infidelity even more strenuously than I do, both in order to draw on her arguments for support and in order to mark off the limits of my position.

**Contemporary Legal Theory**

I begin with a brief overview of modern conceptions of law and their associated theories of legal reasoning, which will ground my claims about the role of legal reasoning in criminal procedure. H.L.A. Hart held a positivist conception of law, under which the law is the set of rules which are laid down by the legislature in accordance with pre-existing, secondary rules about how the legislature should operate. Under Hart's stance, the role of a judge is to make decisions which accord with the duly-enacted law, regardless of the moral content of the law. However, due to the open texture of legal language, there are bound to be cases in which the law does not clearly indicate one outcome. In these cases, Hart asserts, a judge must act as "deputy legislator" and opt for the outcome which best advances the social and moral aims of the law.

Ronald Dworkin posits a new naturalist conception of law, which parallels Hart's positivism but includes a few key differences. Dworkin agrees with Hart that much of law is made up of rules, whose validity comes from secondary rules which govern legislative and judicial processes. Dworkin, however, does not believe that in hard cases, judges become legislators, creating their own law to fill in gaps in the existing law; rather, under his view, there are principles, whose validity does not rely on secondary rules, embedded in the law, and judges must uphold these principles in all cases, even when the rules created by the legislature seek to violate those principles. If, for example, the legislature enacts a rule which limits the right to vote based on racial criteria, Hart would hold that a judge should strike that rule down, because the

rule violates the 15th Amendment, which is a secondary rule which constrains the powers of the legislature. Dworkin would also hold that a judge should strike such a rule down, not only because the rule would violate the text of the 15th Amendment, but also because the rule would contradict the binding principle of racial equality enshrined in the law, a principle which draws some of its force from enumerated anti-racist statutes, but which is not reducible to the set of those statutes. While Dworkin asserts that principles, such as the one which prohibits racially discriminatory legislation, are "irreducibly normative," he does not paint them as mere reflections of each judge's moral sensibilities, because a judge must balance competing principles in a way that is both justifiable from an external perspective and consistent with the relevant body of precedent.

While Hart and Dworkin's competing stances operate on the premise that the law is that which ideally shapes judicial decision-making, the legal realist conception of law rejects this premise and asserts that the law is that which in fact determines how judges decide cases. Legal realism, pioneered by Karl Llewelyn and Justice Oliver Wendell Holmes, Jr., holds that in difficult or novel cases, outcomes are underdetermined by the law, so the judges in such cases end up making decisions based on personal ideology, psychological bias, or mere arbitrariness.

Critical legal studies takes the phenomenon described by legal realism in which non-legal factors determine outcomes and ascribes it to the legal system more broadly; whereas legal realism focuses on the ways in which judges are largely-unconsciously influenced by non-legal factors, CLS asserts that the legislature, the police, prosecutors, jurors, and the judiciary all tend to use their discretion to reinforce existing socioeconomic inequities. Critical race theory holds a similar position, except that it examines the ways in which the legal system tends to reinforce racism (as well as making more broadly historical and anthropological claims about the role of

racism in American society). Feminist jurisprudence uses the same framework as CLS and CRT to examine the ways in which the legal system tends to reinforce misogyny. My thesis will be motivated by the claims of CRT in particular.

Of the above-described conceptions of law, Dworkin's new naturalism will likely provide the strongest objection to my argument. The discretion which Dworkin's view gives judges to interpret and enforce unwritten principles breaks down the positivist division between the validity of law and the morality of law, which thereby threatens the distinction I make between legal reasoning and legal practice; if principles, which attempt to ensure that legal practice is substantively just, are an essential element of legal reasoning, legal reasoning and legal practice cannot be disconnected in the morally relevant way that I assert they are. While I will delve deeper into this objection in my thesis, I will likely respond to the objection in a way which highlights the difference between the idealist conceptions of law presented by positivism and new naturalism and the realist conceptions of law (including CRT). The principles which ensure that legal reasoning translates into just practice are binding and valid elements of the law, under Dworkin's view, but they are not self-actuating; they require judges to implement them, which leaves this Dworkinian objection vulnerable to the above-described legal realist argument that theories of legal reasoning are not in fact outcome-determinative in every case. This normative/empirical distinction appears to hold "all the way down:" from Dworkin's standpoint, a judge who violated the principle of racial non-discrimination would be contradicting a *bona fide* element of the law, but this normative claim does not change the empirical claims about racial discrimination which ground my thesis, so the normative claim does not preempt the moral claims about judges and prosecutors which I will make in my thesis. In my thesis, I will elaborate further Dworkinian objections which could be leveled against the foregoing response.

**Background Claims**

I will now describe the ethical and empirical claims against which the argument of this thesis will be grounded. By stipulating a set of ethical claims, including both metaethical and substantive moral assertions, I will preempt objections which would hold that under certain metaethical or substantive moral views, my argument is incorrect or simply irrelevant. In addition, by stipulating a set of empirical claims, I will preempt potential objections which would hold that my argument is simply based on false empirical claims.

Beginning with the ethical claims, I will stipulate that the justice system should not be racist and, more broadly, that one's race should not affect their life prospects. These substantive moral claims, intuitive as they may be, are necessary to my thesis because if preserving racial bias were taken to be morally neutral or good, my arguments would be *prima facie* incorrect. Beyond this, I will leave open the range of normative theories and substantive moral claims, because I do not wish to beg the question or paint my arguments as self-evident.

The core empirical claim of my argument is that racial bias[1] plays a significant role in determining outcomes within the criminal justice system. I will not venture to make statistical claims about the likelihood of racial bias influencing a given criminal case or the percent of judges and prosecutors who perform their jobs in a racially biased manner, but I will provide some support for the above general claim about racial bias.

Firstly, there is a body of evidence which shows that outcomes in the criminal justice system are racially disparate: currently, white people make up 42% of death row inmates and black people make up 41% of death row inmates, despite the fact that 73% of the United States

[1] When I use the term "bias," here and throughout the thesis, I am using it in an expansive sense, referring to implicit and explicit psychological bias, as well as prejudice, bigotry, discrimination, and the like. Although each of these terms has a distinct meaning, I have chosen to group them all together under the term "bias," because this thesis is about the moral import of racially-impacted decision-making, not about the ontology or psychological basis of such decision-making.

population is white, while only 13% of the population is black[2]. Not only does the race of the defendant appear to determine outcomes, the race of the victim in capital crimes does as well: according to FBI statistics, approximately 51% of murder victims are white[3], yet 82% of death row inmates were sentenced to death for murders involving white victims[4]. Looking at the United States penal system more broadly, black people make up 38.6% of the prison population, while white people make up 57.6% of that population[5]. Although this statistic shows a lower degree of disproportionality than the above statistic about the race of death row inmates, the total prison population is far greater than the death row population, so the statistic about the racial disparity in the prison population as a whole is more robust evidence of disparate outcomes in the criminal justice system. However, as Jennifer Eberhardt points out in the book *Biased: Uncovering the Hidden Prejudice that Shapes What We See, Think, and Do*, disparate outcomes in the criminal justice system do not necessarily point to biased decision-making or unfair procedures; rather, the criminal justice system may be unbiased and disparate outcomes may simply be evidence that black people commit more crimes, including capital crimes, than white people do[6]. Further empirical evidence, however, makes it more plausible to conclude that the criminal justice system produces outcomes in a racially biased manner than it is to conclude that white and black people have disparate rates of criminal behavior: as of 2015, 95% of elected prosecutors were white and 14 states did not have any non-white elected prosecutors[7]. Similar

[2] "Racial Demographics," https://deathpenaltyinfo.org/death-row/overview/demographics.
[3] "2018: Crime in the United States," *Federal Bureau of Investigations,* https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/tables/expanded-homicide-data-table-6.xls.
[4] "Race and the Death Penalty," *American Civil Liberties Union,* https://www.prisonpolicy.org/scans/aclu_dp_factsheet4.pdf.
[5] "Inmate Race," *Bureau of Prisons*, October 24, 2020, https://www.bop.gov/about/statistics/statistics_inmate_race.jsp.
[6] In *Biased*, Eberhardt focuses on racial bias within police activity, rather than focusing on racially disparate outcomes within sentencing and incarceration, but her point, that the presence or absence of racial bias in the criminal justice system is underdetermined by disparate outcomes, applies equally to either focus.
[7] "Study: Most States Have No Black Prosecutors," *The Atlantic*, July 7, 2015, https://www.theatlantic.com/politics/archive/2015/07/american-prosecutors-are-incredible-whitedoes-it-matter/397847/.

racial disparities hold among the judiciary, where, as of the summer of 2019, 23 state supreme courts were entirely white[8], and where only 27% of U.S. District Court and Court of Appeals judges are non-white[9]. In addition, as of late 2019, federal judges were four times as likely to be former prosecutors than they were to be former criminal defense lawyers[10]. Beyond simply representing demographic disparities, prosecutors and judges tend to apply their discretion in a racially disparate manner: studies show that prosecutors charge black people for crimes which carry mandatory minimum sentences or enhanced sentences for repeat offenders at a higher rate than they charge white people under those statutes, even adjusting for the nature of the offense[11].

Although I have presented evidence that the criminal justice system produces racially disparate outcomes and that this disparity is the result of bias within the system, I have not yet presented evidence which bears on whether the bias in the criminal justice system is implicit or explicit. Jennifer Eberhardt, in the aforementioned book *Biased*, asserts that the bias in the criminal justice system is largely implicit; Eberhardt asserts that humans' tendency to unconsciously rely on preconceived stereotypes, combined with this country's history of racial injustices, results in a population of people who, despite mostly holding no explicitly racist beliefs, tend to treat people of different races differently. Under Eberhardt's view, implicit racial bias influences the criminal justice system because actors in the criminal justice system come

---

[8] "In State with 27% Black Population, All Alabama Appellate Judges Are White," *Equal Justice Initiative*, March 11, 2020, https://eji.org/news/all-alabama-appellate-judges-are-white-in-state-with-27-percent-black-population/.

[9] "Examining the Demographic Compositions of U.S. Circuit and District Courts," *Center for American Progress*, February 13, 2020, https://www.americanprogress.org/issues/courts/reports/2020/02/13/480112/examining-demographic-compositions-u-s-circuit-district-courts/.

[10] "There Are Way Too Many Prosecutors in the Federal Judiciary," *Slate*, October 14, 2019, https://slate.com/news-and-politics/2019/10/too-many-prosecutors-federal-judiciary.html. - One could argue that former prosecutors are not actually overrepresented on the federal bench, because there are more prosecutors than there are public defenders and private criminal defense lawyers, but it is nevertheless true that federal judges are far more likely to come from a career of advocating for the government, rather than for defendants, in criminal cases.

[11] "Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System," *Sentencing Project*, April 19, 2018, https://www.sentencingproject.org/publications/un-report-on-racial-disparities/.

from a society in which this bias is unnoticed but ever present. Eberhardt acknowledges that implicit racial bias is not monolithic across all members of our society, because, for example, police officers are disproportionately exposed to people committing crimes or being perceived to commit crimes, and these experiences are filtered through the officers' existing implicit bias, molding and reinforcing that bias.

The same logic can be applied to prosecutors and judges, because prosecutors choose people to charge from among those detained by law enforcement officers, whose above-described implicit bias shapes the group of people they detain, which reinforces prosecutors' implicit associations between people of color and criminality. These implicit associations held by prosecutors result in disparate patterns of charging, plea bargain negotiation, proposed sentences, etc., which disparities in turn mold the implicit biases of judges, whose courtrooms are disproportionately full of black defendants.

While I give credence to Eberhardt's evidence about the psychology of implicit bias and I believe that such bias plays a large role in producing the racially disparate outcomes of the criminal justice system, I believe that explicit bias plays a significant role as well. This is to say, I believe that some prosecutors and judges allow their decision-making to be influenced by explicitly-held racist beliefs and, relatedly, that some such actors use their roles in the justice system as tools to maintain racial hierarchies which already exist in our society. As evidence, I will present examples of prosecutorial discretion, as well as an individual criminal case, which serve as evidence of the effects of explicit bias within the criminal justice system.

In the book *Usual Cruelty: The Complicity of Lawyers in the Criminal Injustice System*, Alec Karakatsanis writes, "the post-arrest criminal system is not 'law enforcement,' but a bureaucracy designed and permitted to circumvent the 'rule of law' when necessary to pursue the

aims that political elites have assigned to it" (Karakatsanis, 36). Included in the "aims that political elites have assigned to [the law]," Karakatsanis asserts, is the subjugation and repression of non-white people. This does not mean, however, that political elites, such as legislators, governors, presidents, and attorneys general have autonomously decided to promote white supremacy and instructed their subordinates to act accordingly. Rather, these political actors exist within a kind of "feedback loop" with the public and with lower-level government officials: the public and lower-level government officials (including, for example, police and correctional officers' unions) express their racial bias in the way that they vote, which causes the political actors in power to take racially biased actions. These actions in turn appear to confirm the racially biased attitudes already held by the public and by lower-level government officials, which continues the cycle. As an example of this feedback loop in action, public anti-black sentiment surrounding the use of drugs may influence policy-makers to raise the mandatory minimum sentence attached to the possession of drugs more commonly used by black people than by white people, which confirms, in the minds of voters and lower-level officials, that black people engage in more destructive and criminal behavior than white people do, which leads to more anti-black sentiment affecting policy-makers.

The results of this feedback loop are evidenced both by the results of the criminal justice system, which include the aforementioned imprisonment of non-white people at a rate vastly disproportionate to their share of the population, and by the decisions made in the implementation of criminal procedure: as Karakatsanis emphasizes, when certain elements of the law are unclear or conflict with each other, judges, prosecutors, and other parties in the criminal justice system tend to enforce those elements of the law in the way which best advances the aforementioned political aims of the law. Karakatsanis, using the example of the Philadelphia

Row bombing of 1985, points to prosecutorial discretion as one way in which the indeterminacy of the law is used to advance racist political aims: in May of 1985, Philadelphia police dropped bombs provided by the Federal Bureau of Investigations (FBI) on the houses of black liberation activists. Public officials then instructed the fire department not to put out the resulting fire, and police shot at surviving activists as they fled the burning homes. As Karakatsanis points out, "exercising their discretion, local and federal prosecutors chose not to prosecute any of the people who carried out or ordered the bombing. But they did exercise their discretion to prosecute the only adult black liberation activist who survived" (Karakatsanis, 43). Implicit in the foregoing quote is the statement that prosecutors did not decide whom to prosecute and whom to abstain from prosecuting based on any neutral calculus of harm caused, evidentiary standards, or the like, but that prosecutors instead made their decisions in such a way as to advance the anti-black aims of the political environment. The Philadelphia Row bombing is simply one striking example in a larger pattern, as Karakatsanis claims; the FBI's Counterintelligence Program (COINTELPRO), which ran from 1956 until the early 1980s, consisted of the systemic infiltration of leftist and pro-civil rights groups and the persecution of the members thereof. In service of COINTELPRO, FBI agents performed illegal surveillance, made threats of violence, and even kidnapped and murdered people, but when the Program was finally exposed, no FBI agents were prosecuted for their crimes[12]. Given the substantial evidence that serious crimes were committed on a large scale, it is difficult to explain this prosecutorial decision in racially or politically neutral terms.

To give a more recent example, Rodney Reed, a black man, has been on death row in Texas for 23 years for the murder and rape of Stacey Stites, a white woman with whom Reed had been having an affair, even though there are numerous substantive and procedural problems with

[12] Karakatsanis, Alec, *Usual Cruelty: The Complicity of Lawyers in the Criminal Injustice System*, 2019.

his conviction: the forensics experts whose testimony was central to the case against Reed have since admitted that their conclusions were in error, and outside forensics experts have confirmed the falsity of the original forensic testimony offered. In addition, a detective conducting the investigation indicated on tape shortly after Stites was murdered that he and the other police investigating believed Stites' fiance, Jimmy Fennell, to be Stites' killer. This belief was supported by the fact that Fennell gave inconsistent reports about where he was at the time the murder was committed, that Fennell, a white police officer, had previously been convicted of sexually assaulting women in his custody, and that Stites was having an affair with Reed, the discovery of which would have given Fennell a motive to kill her. Nevertheless, Reed, not Fennell, was prosecuted for Stites' murder, and Reed's conviction has never been overturned, nor have any criminal charges been filed against Fennell, despite the aforementioned exculpatory facts, in addition to the fact that DNA matching one of Fennell's colleagues in the police force was found on a beer can at the site where Stites' body was found, the fact that the prosecution withheld the aforementioned DNA results from Reed's defense lawyer[13], the fact that the murder weapon has never been tested for DNA, and the fact that per a sworn affidavit signed by an inmate who met Fennell during a recent prison stay, Fennell admitted to Stites' murder[14]. Despite the quantity of substantive reasons why it appears that Reed is innocent, as well as the quantity of procedural reasons why Reed's conviction was invalid, the Texas Supreme Court only ruled to halt Reed's execution 5 days before the execution was scheduled, even though the aforementioned evidence of problems with Reed's conviction had existed and been presented by Reed's defense lawyers years earlier.

---

[13] "Rodney Reed's Appeals: Key Players," *The Austin Chronicle*, 2019. https://www.austinchronicle.com/news/2019-11-15/rodney-reeds-appeals-key-players/.
[14] "Ten Facts about Rodney Reed's Case You Need to Know," *Innocence Project*, 2019. https://www.innocenceproject.org/10-facts-you-need-to-know-about-rodney-reed-who-is-scheduled-for-execution-on-november-20/.

Reed's case supports my claim about racially-motivated decision-making on the part of prosecutors and judges because the trajectory of the case has not been determined by an individual instance of police misconduct, faulty testimony, or the like; rather, the trajectory of the case, from the original trial through the ongoing appeals process, has been determined by countless instances of prosecutorial discretion and judicial review. It seems near impossible to understand how all of these instances of decision-making have resulted in such a clear miscarriage of justice without making reference to the underlying fact that Reed was a black man known to be having an affair with a white woman. As I acknowledge above, I do not have sufficient empirical knowledge to make specific statistical claims about the likelihood of racial bias impacting a given case or about the percent of prosecutors and judges who allow racial bias to impact their decisions, but I believe that the evidence and arguments just presented are sufficient to support the empirical claim that racial bias frequently plays a non-trivial role in the decision-making of prosecutors and judges.

## Chapter 1: The Moral Duty to Effect Substantive Justice

In this section, I will argue that individual judges and prosecutors have a moral duty to use their offices within the criminal justice system to promote the greatest substantive justice possible in each case, even when doing so would require the individual to go against precedent, positive law, stated prosecutorial priorities, etc.

### Chapter 1.1: The Fugitive Slave Act

By the nature of their offices, individual prosecutors and judges have significant power to mold outcomes within the criminal justice system, which invites the question of whether these individuals can use their power to bring about the most just outcomes in each case, even if doing so would require role infidelity of the type described above. A prominent historical example in which judges wrestled with this question occurred in the mid-nineteenth century, when northern judges, many of whom were personally abolitionist, heard cases involving the Fugitive Slave Acts (FSAs). The FSAs were a series of laws, stemming from the "Bloody Clause" of the Constitution, which mandated that slaves could not achieve legal freeperson status by escaping to the North, and, relatedly, that the owners of escaped slaves could recapture those individuals within northern states and transport them back to captivity in the South. I will examine this example in detail, because it illustrates the contours of role fidelity and the context of the moral duties on judges.

As Robert Cover describes in the book *Justice Accused: Antislavery and the Judicial Process*, whereas before the implementation of the FSAs, personally abolitionist judges had been able to interpret the relevant statutes *in favorem libertatis* ("in favor of liberty"), using gaps or vagueness within the positive law and appeals to the libertarian principles enshrined elsewhere in

the law to decide in favor of the escaped slaves, the clarity of the FSAs prevented this method of interpretation and brought the question of role fidelity versus infidelity into relief. Judges in these cases tended to abide by the FSAs and rule in favor of the slave owners or the agents hired by the slave owners to recapture the escaped slaves, and tended to write in their opinions defending these rulings that they could not rule otherwise, because the positive law was binding and outcome-determinative. Cover writes that these opinions reflected a misunderstanding, on the part of the judges, of the position held by their critics and ideological opponents (namely, the escaped slaves' legal advocates): the judges' critics did not hold that the judges had misunderstood the FSAs and their relation to the judges' official role; rather, the judges' critics held that even assuming that the judges' official role would entail deciding in favor of the slave owners, the judges should have decided in favor of the escaped slaves. The key conceptual distinction is that the judges were treating the FSAs as mechanistically determinative of a specific outcome, which condition would preempt the ability to decide whether to obey the FSAs or not, while the judges' critics were pointing out that the FSAs were only determinative of the outcome which would fall within a judge's official role, so a judge did, in fact, have the ability to decide whether the moral weight of anti-slavery was greater or lesser than the moral weight of role fidelity. Cover rephrases this distinction by writing, "the formal restraints [of the judge's role] are under challenge and cannot be justified by mere reiteration" (Cover, 129).

Cover sheds further light on the nature of the judicial role and fidelity thereto using two analogies: the first analogy compares the judicial role to a game of chess; if a chess player were to move their bishop straight and agree with their opponent that such a move was valid, the two people would no longer be playing chess, because one of the essential properties of chess is that the bishops can only move diagonally. A player could, however, move their bishop straight and

hope that their opponent did not notice, under which circumstance the two people would still be playing chess, but one of the players would simply be cheating. Under this analogy, judges who wrote that they could not rule in favor of the escaped slaves meant that their roles were defined by implementing the positive law, so by disregarding the FSAs, they would simply no longer have been acting as judges. Conversely, under the same analogy, the judges' critics were urging the judges to "cheat" and rule in favor of the escaped slaves even though to do so would have been to "break the rules" of the judicial role. The second analogy compares the judicial role to that of a person speaking a language; someone who breaks the existing rules of a language is still speaking that language, and while they may be speaking it poorly, they may instead be speaking it eloquently and contributing to a change in the language. Under this analogy, a judge who ruled in favor of an escaped slave in an FSA case would still be acting as a judge, although people could debate whether the judge was "speaking law poorly," by violating the established rules of their role, or "speaking law eloquently," by changing the nature of the judicial role so that it included the consideration of morality[15]. I believe that the language analogy is a more apt characterization of the judicial role than is the chess game analogy, because while the rules of chess are clearcut and a chess player only has discretion in choosing which moves to make within these rules, the rules of the judicial role are less clearly defined and the nature of judicial decision-making can change over time, depending on the practices of actual judges. The same analogy applies to prosecutors, and easily so, because prosecutorial discretion is not typically considered to be clause-bound in the way that judicial decision-making is considered to be.

Applying Cover's language analogy to the context of modern judges and prosecutors, if those individuals promote the greatest possible substantive justice in each case, they may shape the nature of their respective roles, so that future judges and prosecutors would simply be hewing

[15] Cover, Robert, "The Judicial 'Can't,'" *Justice Accused: Antislavery and the Judicial Process*, 1975, p. 123-130.

to the norm by promoting substantive justice, in much the same way that the first person who said "ain't" must have been considered a linguistic rule-breaker, but now the use of "ain't" is a widely-accepted and useful part of the English language. Of course, the fact that judges and prosecutors could change the norms of their respective roles does not prove anything without an argument that those norms should in fact be changed; that is to say, the foregoing argument would only have force if combined with an argument showing that it is desirable for judges and prosecutors to promote substantive justice. I will now offer such an argument.

**Chapter 1.2: In Defense of Role Infidelity**

Thus far, I have referred several times to judges' and prosecutors' moral duties to effect substantive justice in each case, but "duties" is admittedly not the most accurate term for what I am asserting; there is a difference between the weaker claim that a judge or prosecutor has a moral right to promote the maximum substantive justice each case, even at the cost of engaging in role infidelity, and the stronger claim that such an individual has a moral obligation, *mutatis mutandis*. My position is that each judge and prosecutor, at a minimum, has a moral right to engage in role infidelity if necessary to promote substantive justice in a given case, but that whether each judge and prosecutor has a moral obligation to do so will depend on how much substantive justice is at stake in a given case[16].

It is easy to argue that, *ceteris paribus*, a person should take actions which accord with justice, but the difficulty for my position comes from the fact that when considering a judge or prosecutor, all things are not equal; judges and prosecutors agree to execute the duties of their

[16] Whether a judge or prosecutor in a given case has a moral obligation to effect substantive justice, if that were to require role infidelity, would also depend on the likely practical consequences of engaging in role infidelity in that case. However, as I will dedicate a later section of this thesis to the moral and practical value of role fidelity, I will not address it here.

respective offices (in most cases, they take formal oaths of office) and their offices play specific roles within the criminal justice system. Arguably, each judge and prosecutor only possesses the above-described power to affect outcomes in virtue of their having accepted a moral obligation to remain faithful to their role. True as this may be, I believe that the moral importance of effecting substantive justice outweighs any moral obligation of role fidelity. One could retort that the moral obligation of role fidelity is a Kantian categorical imperative or something similar, and that it therefore preempts or nullifies any countervailing moral or practical considerations. I would respond to this point by appealing to Lysander Spooner, who advocated for judges in FSA cases to rule in favor of escaped slaves by arguing that insofar as the judicial oath of office bound judges to facilitate the moral ill of slavery, that oath was invalid and therefore could not exert moral force on judges.

In his 1860 work *The Unconstitutionality of Slavery*, Spooner uses the analogy of a person, B, who is offered a sword by A on the condition that they swear to use the sword to kill someone. Spooner asserts that the person offered the sword should accept it, even though they are agreeing to commit a murder, and then use the sword to defend the person whom they had sworn to kill. It would not only be acceptable but morally commendable for the person offered the sword to break their oath in this way, asserts Spooner, because the immoral character of the oath's contents renders the oath "incapable of raising the least moral obligation, of any kind whatever, on the part of B towards A." Spooner continues, "B then holds the sword on the same principle, and by the same right, that he would have done if it had been put into his hands without any oath or condition whatever" (Spooner, 150-151). Applying this analogy to judges, as Spooner did, and to prosecutors, insofar as fidelity to their respective roles would require them to perpetuate moral ills, their moral obligation of role fidelity is nullified and they are left in the

position of a person who had their power without any of the parameters of role fidelity. As I state above, it is easy to argue that someone without the restraints of role fidelity or other restraints should take actions which accord with justice, so under Spooner's argument, it is clear that prosecutors and judges ought to promote substantive justice in each case.

Spooner's analogy, however, relies on the argument that the Constitution inherently binds judges to commit immorality, because it binds judges to uphold the Bloody Clause and to uphold any duly-enacted pro-slavery laws (assuming that those laws do not contradict other enumerated Constitutional principles). However, my argument holds that the substantive injustice perpetrated by the criminal justice system does not lie in the inherent properties of the Constitution, but in the ways in which criminal procedure is applied, so my argument is substantially disanalogous with Spooner's example of the person accepting the sword. Nevertheless, to the extent that racial bias affects decision-making within the legal system, judges' and prosecutors' moral obligation of role fidelity is an obligation to uphold a system which perpetuates racism. The mere existence of racial bias within the criminal justice system as a whole, however, is not sufficient reason to consider a judge or prosecutor's obligation of role fidelity an obligation to facilitate racism, but if a judge or prosecutor has reason to believe that a case in which they are involved is tainted by racial bias, it is more clear how maintaining role fidelity would be tantamount to facilitating racism. If, for example, a prosecutor has a pattern of charging black defendants under mandatory minimum statutes at a rate disproportional to the rate at which they charge white defendants under those statutes, a judge who maintained impartiality and simply let the jury decide each defendant's guilt or innocence would be allowing the prosecutor's racially biased use of discretion to translate into racially biased outcomes. The judge's office would function as a "sword," applying Spooner's analogy, with which the judge is able to defend the black

defendants from the prosecutor's biased use of discretion. Given that the boundaries of the judge's role, in this example, would entail hearing cases brought before them due to racially biased prosecutorial discretion and, if the jury returned a verdict of "guilty," imposing sentences on defendants who are disproportionately black, the boundaries of the judge's role would thus be sapped of moral weight and the judge would have the "sword" of their office "by the same right that [they] would have done if it had been put into [their] hands without any oath or condition whatever" (Spooner, 150-151). Therefore, the judge would have the moral right, if not the obligation, to act so as to counteract the racial bias of the prosecutor and strive to achieve the most substantively just outcome possible (for example, by rendering a judgment notwithstanding the verdict, if the jury were to convict, on the grounds that there was insufficient evidence that the defendant's conduct fit within the parameters to which the mandatory minimum statute applies).

Although naturalists and positivists differ in their beliefs about whether justice is an essential characteristic of the law, legal theorists from both camps can agree that the natural aim of the law is to promote justice. The law, in the United States, is created and enforced through a set of processes each of whose aims may not be the promotion of justice - for example, allowing the citizens to vote for legislators serves the aim of promoting democratic representation and having executive and judicial branches serves the aim of preserving checks on legislative power - but the overarching purpose which those processes serve must be the promotion of justice, or else there would be no reason for those processes to exist. The roles of judges and prosecutors, therefore, exist in order to serve the aim of promoting justice. It is obvious that when such an actor in the legal system executes the duties of their office in a racially biased manner, that actor is perpetrating an injustice. Less obvious, although, I believe, equally true, is that if a judge or

prosecutor steps outside the bounds of their role in order to counteract racial bias, they are promoting justice. Therefore, paradoxical as it may be, by violating role fidelity in order to promote justice, a judge or prosecutor would actually be completing their role, insofar as the judge's or prosecutor's role exists in order to promote justice. At this point, one could object that even though judges' and prosecutors' roles exist in order to serve justice, their roles are specifically tailored to fit within a larger legislative and judicial system which serves justice, so individual judges and prosecutors must operate within those circumscribed roles, rather than acting as the "platonic guardians" of the entire criminal justice system and acting however they see fit. I would respond that the influence of racial bias on a given case or statute changes this picture, so that judges and prosecutors should, in fact, step outside of their roles and strive to promote justice. This is the case because the roles of judges and prosecutors only contribute to the effectuation of justice if they are applied in a context which is free of racial bias. As I argue above, if a judge were to operate within the bounds of their role when the prosecutor was acting in a racially biased manner, the judge would be allowing the prosecutor's biased actions to translate into biased outcomes, which would constitute an injustice.

**Chapter 1.3: The Dworkinian New Naturalist Objection**

One could object to this argument on the grounds that I am assuming a positivist legal philosophy, which allows me to draw a distinction between role fidelity and following one's conscience which I would not be able to draw if I instead assumed a Dworkinian new naturalist legal philosophy; under positivism, the role of a judge is limited to enforcing enumerated constitutional rights and duly-enacted statutes, so if a judge were to decide a case based on the morality of the relevant statute, the judge would be violating role fidelity. Under Dworkin's view,

however, in enforcing enumerated constitutional rights and duly-enacted statutes, a judge ought to balance the "dimension[s] of fit [and of] justification." The dimension of fit is the extent to which a given outcome in a case would preserve the "holistic integrity" of the law, both in terms of maintaining consistency with precedent and in terms of allowing future judges to maintain consistency by relying on the current case as precedent. The dimension of justification is the moral status of the right or statute being considered, and, accordingly, a judge's consideration of the dimension of justification includes the consideration of the moral status of each possible outcome in a given case. Regardless of whether the dimensions of fit and of justification coincide or conflict in a given case, they are not completely distinct concepts, because the principles of political morality which inform the dimension of justification may be drawn from an analysis of the body of positive law and precedent, which analysis would be part of the dimension of fit[17]. To use an example which Dworkin himself elaborated, a judge in an FSA case who was trying to maintain the dimension of fit would have to consider the facts that the Constitution includes robust protections of individuals' freedom and that the Constitution lays out a form of federalism broadly inconsistent with the FSAs[18]. For a Dworkinian theorist, a judge who looked further afield than a narrow interpretation of the positive law would not be stepping outside of their role, but fulfilling their role thoroughly. Thus, the objector would continue, I have created a conflict between maintaining role fidelity and following one's conscience, but the real conflict is simply between positivism and Dworkinian new naturalism.

I will respond to this objection with two arguments, the first of which is narrowly focused on the differences between my view and Dworkin's and the second of which is more broad and conceptual. The first argument is that the view which I propose in this paper gives substantive

[17] Dworkin, Ronald, *Law's Empire*, 1986.
[18] Dworkin, Ronald, "The Law of the Slave-Catchers," *Times Literary Supplement,* December 5, 1975.

justice a significantly greater role than does Dworkin's view, so my view cannot be reduced to a rejection of positivism and an endorsement of Dworkinian naturalism. Under Dworkin's view, most of the time that judges must apply discretion to decide difficult cases, they are applying weak discretion, meaning that they are not acting as "deputy legislators," to use H.L.A. Hart's phrase, and applying their own, extra-legal opinions about the case. Rather, judges applying weak discretion are using the dimensions of fit and of justification to discern the principles and policies which are relevant to the case and to figure out which outcome the principles and policies suggest. Principles, which exist in order to protect minority rights, and policies, which exist in order to promote the well-being of the majority, may be extra-statutory, but, in contrast with Hart's view, they are not extra-legal; under Dworkin's view, the unwritten principles and policies which undergird the positive law are themselves binding elements of the law.

Whereas my view holds that a judge must sometimes decide cases according to their own understanding of the moral values at stake, Dworkin's view holds that in most difficult cases, a judge must decide according to preexisting principles and policies whose characteristics are not swayed by the judge's personal moral beliefs. Admittedly, the dimension of justification recommends that a judge not only consider which outcome allows the case to best cohere with the existing body of precedent, statutes, and rights, but also consider which outcome would lend the most morally justifiable trajectory to the coherent body of precedent. Accordingly, Dworkin described a judge in a difficult case as analogous to a writer tasked with writing one chapter of a "chain novel," a novel in which all of the previous chapters have been written by different people and all of the future chapters will be written by different people. When considering what to write in their chapter, a writer would have to consider what content would make the most sense to follow the previous chapters and to allow the future writers to continue the novel in a coherent

manner, but the writer would also have to consider what content would make the novel as a whole the best it could be. Following this analogy, Dworkin holds that a judge should decide a difficult case in a way that allows the "chain novel" of the law to follow a consistent "storyline," while also trying to make the "storyline" which the law follows as morally justifiable as possible. In Dworkin's view of a judge as temporary author of the legal "chain novel," a judge's role is largely bound to the interpretation of the positive law, even though this interpretation allows the judge some liberty to consider extra-statutory factors, including moral factors. By contrast, under my view, a judge must consider the moral values relevant to each case and, if necessary, entirely set aside positivist interpretation in order to achieve the most morally justifiable outcome.

To illustrate this difference between Dworkin's view and mine, I will return to Dworkin's writing about the FSA cases: in "the Law of the Slave-Catchers," which is a commentary on Cover's above-cited *Justice Accused*, Dworkin writes that judges in those cases "suffered not from too much formalism, but from too little" (Dworkin, 179). Although there are elements of certain FSA cases, such as the violation of the escaped slaves' Sixth and Seventh Amendment rights and the violation of Article III of the Constitution, which would allow a judge to rule in favor of the slave on purely formal, procedural grounds, Dworkin takes a different approach in this paper; he asserts that the principle of liberty is enshrined in the Constitution and that the Constitution lays out a form of federalism under which the FSAs could not prevent free states from setting their own laws to protect the freedom of escaped slaves. Dworkin holds that these two principles, of liberty and of limited federal authority, combined with the aforementioned procedural problems with FSA cases, were "more central to the law than were the particular and transitory policies of the slavery compromise" (Dworkin, 180). Dworkin's judicial philosophy does not give enough weight to the consideration of morality, because he must contort his

judicial philosophy in order to argue that the judges in the FSA cases should have ruled in favor of the escaped slaves (and thus ruled in favor of the most morally justifiable outcome); what Dworkin refers to as "particular and transitory policies" were duly-enacted statutes which put into effect Article IV Section 2 of the Constitution, the so-called "Bloody Clause." In addition, by the time of the well-known FSA cases of the 1850s, there was already a body of precedent in which judges in similar cases had decided those cases *in favorem libertatis*, but only to the extent that the ambiguity and gaps in the positive law allowed. To the extent that the positive law was clear, judges in previous cases had tended to respect it, even when it contradicted their personal moral beliefs, so the FSAs not only had a "pedigree" of validity within the positive law (to use Hart's term), they also had some amount of precedential support[19]. In order to rule in favor of an escaped slave, a judge would have had to place far more weight on the dimension of justification than on the dimension of fit, because a judge would have had to value the principles of liberty and limited federal authority highly enough that these principles could justify disrupting the coherence of the law and invalidating an unambiguous clause of the Constitution. In skewing so far toward the dimension of justification, a judge would have had to place great value in their personal, anti-slavery moral beliefs.

Taking Dworkin's stance on the FSA cases a step further, if the constitutional principle of liberty and the immorality of slavery played a large role in justifying the invalidation of the Bloody Clause and the FSAs, those two elements should also have rendered the entire institution of slavery legally invalid. Therefore, in order to justify the outcome which he believes was correct in the FSA cases, Dworkin has to give the consideration of morality a larger role than his judicial philosophy suggests, a role so large that it would have encouraged judges to act as

[19] Cover, Robert, "Statutory Interpretation: *In Favorem Libertatis?*", *Justice Accused: Antislavery and the Judicial Process*, 1975, p. 62-79.

deputy legislators (or even deputy Framers) and abolish slavery. Dworkin's stance on the FSA cases lends a primacy to substantive justice which is inconsistent with his general view but perfectly consistent with mine. Dworkin's writing about the FSA cases not only shows the significant differences between his view and mine, it shows that when considering a situation in which fundamental moral values are at stake, Dworkin adopts a view similar to mine.

My second response to the objection that the real conflict is between Dworkin's naturalism and positivism, not between maintaining role fidelity and following one's conscience, is that both Dworkin's naturalism and positivism are ideal theories of law, whereas my stance is founded on legal realism and critical race theory (CRT), both of which reject (and respond to) ideal theories of law. As I describe above, Dworkin's naturalist view differs from positivism by including some amount of moral consideration as internal to the judicial role, rather than classifying moral consideration as an external factor which may draw a judge outside of the boundaries of their role. Given, however, that my view is rooted in the legal realist claim that one should examine which factors actually influence judicial decision-making, rather than debating the ideal characteristics of the judicial role, the question of whether the consideration of morality is internal or external to the judicial role is trivial; if there is a certain amount of racial bias which in fact influences outcomes within the criminal justice system, the moral duties of a judge should not change based on whether the judge's role "really does" or "really does not" include the consideration of morality.

### Chapter 1.4: Ideal and Non-Ideal Theory

An objector could retort that while my view hinges on rejecting the ideal theories offered by Dworkin, Hart, etc. and responding using the realist theories offered by Holmes, Llewelyn,

and critical race theorists, my view is itself an ideal theory, because it holds that given the influence of racial bias on the criminal justice system, it would be ideal for judges and prosecutors to consider their own moral beliefs in deciding cases. Therefore, the objector could continue, both Dworkin's naturalism and CRT offer idealized conceptions of the judicial role, so the only relevant difference between the two theories is in which aspects of the idealized judicial role each one emphasizes (Dworkin's naturalism emphasizes the balancing of positive statutes, precedent, principles, and policies, while CRT emphasizes anti-racism). By way of analogy, the objector could continue, John Rawls conceives of justice as centered around fair equality of opportunity (FEO), which conception Tommie Shelby employs in developing an anti-racist theory. Charles Mills, however, labels Rawls' conception of justice and Shelby's subsequent anti-racist theory as instances of ideal theory, and critiques them on the grounds that non-ideal theory is necessary in order to develop a method by which to promote justice (especially racial justice) in American society. In the paper "Racial Realities and Corrective Justice: a Reply to Charles Mills," Shelby defends his theory, as well as Rawls' conception of justice, by arguing that the non-ideal approach to achieving justice which Mills recommends would need to rely on ideal theory. Shelby summarizes this argument as follows: "if we abandon the framework for ideal theorizing, how do we determine which principles of justice should guide our reform or revolutionary efforts, and how do we justify these principles if we must rely exclusively on nonideal theory?" (Shelby, 156).

Although the debate between Shelby and Mills concerns the methods of working toward justice in general, which is far more broad than the topic of this thesis, Shelby and Mills' debate is relevant to this thesis, because it seems that without the background of an ideal theory of judicial and prosecutorial decision-making, the arguments which I advance in this thesis would

be trivial; without the body of ideal legal theory which assumes that the law is applied in a racially unbiased manner, I would have to create a theory of judicial and prosecutorial decision-making "from scratch," as would other legal realists and critical race theorists, but given that we have not tended to create such theories "from scratch," it is clear that ideal theory is part of the foundation of our non-ideal theories. Due to this point, I have to side with Shelby over Mills, because Shelby is right that Mills has to presuppose certain elements of ideal theory, namely a liberal conception of justice which strives to promote fair equality of opportunity, in order to meaningfully engage in non-ideal theorizing with the goal of achieving racial justice. The question remains, however, of what ideal theory I am taking as foundational to this thesis; while I laid out certain general metaethical and substantive moral presuppositions earlier in the thesis, I have not enumerated the features of a racially just legal system or of a racially just society which the non-ideal theory presented here aims to advance. While I will delve into this question later in the thesis, I can state for now that the ideal theory on which my argument rests is based on the liberal idea that one's race should not affect their life prospects. Returning to the Shelby-Mills debate, Shelby writes that although Rawlsian ideal theorizing assumes perfect compliance on the part of those charged with implementing any anti-racist program, this is simply a methodological choice and does not ignore the fact that individuals are not actually perfectly compliant with rules[20]. My view differs from Shelby's on this point, because whatever the contours of the ideal theory on which this thesis is grounded, that theory must account for some degree of non-compliance in order to plausibly apply to the real world.

### Chapter 1.5: The "Insurgency" Objection

[20] Shelby, Tommie, "Racial Realities and Corrective Justice: A Reply to Charles Mills," 2013, p. 150.

Returning to the content of the non-ideal theory which I advance in this thesis, thus far, I have described the moral situation of judges and prosecutors once they have already assumed their offices, but the expectations about the legal system which these actors have before assuming their offices may be morally relevant as well; a judge or prosecutor may assume their office under the assumption that the criminal justice system tends to produce substantively just outcomes or they may assume their office with the knowledge that racial bias often plays a significant role in determining outcomes in the criminal justice system. Based on this distinction, one could object to my argument by asserting that if a legal actor takes their office under the former assumption and then decides to pursue substantive justice at the cost of role fidelity, they are simply siding with their conscience, while if a legal actor takes their office under the latter assumption, they are acting as an "insurgent," to use Cover's term. "Insurgent" is an accurate description of such a person, the objection would hold, because they would accept a job in the criminal justice system and take the oath of office, all the while planning to undermine the system by trespassing the boundaries of their role whenever they see fit.

I would respond to this objection by drawing a distinction between intention and expectation: if a judge or prosecutor were to take the oath of their office with the intention of disregarding the boundaries of their role, the above objection would hold, because they would be seeking a position of power so that they could abuse it. If, however, a judge or prosecutor were to take the oath of their office not intending to disregard the boundaries of their role but expecting that there would be future cases in which justice would compel them to do so, the above objection would not hold, because the actor would have no intention to gain power in order to abuse it; rather, the actor would simply have an understanding of the system in which they plan to work and a willingness to side with their conscience over their role, when necessary. The

imagined objector could further object by rejecting the distinction between intention and expectation, asserting that one who reasonably expects to perform a certain action is no differently morally situated than someone who intends to perform that action. Therefore, the objection would hold, the person who swears the oath of their office with the expectation that there will be cases in which the pursuit of justice will lead them to step outside the boundaries of their role is morally no different from the person who takes the oath of their office with the intention of stepping outside the boundaries of their role. I would respond that the distinction between intention and expectation is a morally sound one, which one can observe by examining the same distinction in other situations.

The distinction between the moral value of intended and expected outcomes stems from the doctrine of double effect, which holds that it may be morally permissible to take a certain action while intending a good outcome and foreseeing an accompanying bad outcome, but it is impermissible to take the same action while intending the bad outcome and foreseeing the good outcome. To provide an oft-used example, it would be wrong for a doctor to give a suffering patient a lethal dose of morphine, intending to kill the patient and foreseeing that once dead, the patient will no longer be in pain. However, it would be acceptable (or at least less unacceptable) for the doctor to give the patient a lethal dose of morphine, intending to alleviate the patient's pain and foreseeing that the morphine will kill the patient.[21] The objector could accept the doctrine of double effect but assert that it does not apply to my example of judges and prosecutors, because while the doctrine of double effect distinguishes between intended and merely expected outcomes of actions, my example attempts to distinguish between an individual

[21] "Doctrine of Double Effect," *Stanford Encyclopedia of Philosophy*, December 24, 2018, https://plato.stanford.edu/entries/double-effect/.
The doctrine of double effect comes from Catholic philosophy and is often used in discussions of whether such actions as abortion and euthanasia accord with Catholic ethics, but it can be applied just as easily to secular philosophy, in which context I am using it.

intending and merely expecting to perform an action themself. Whereas one can avoid the moral onus of a merely expected bad outcome, the objector would continue, one must take full moral responsibility for any actions which one expects to perform themself. I would respond that, as this thesis contends, the action of violating role fidelity is not itself bad; rather, it is the course of action consisting of deliberately joining the criminal justice system in order to operate as an "insurgent" and violate role fidelity which is bad. The doctrine of double effect helps distinguish between this bad course of action and the acceptable course of action taken by a person who foresees sometimes being conscience-bound to violate role fidelity, but does not join the criminal justice system with the intention to do so. If the objector considers the above description and example of the doctrine of double effect and simply rejects it, the objection still does not necessarily succeed, because even if a judge or prosecutor truly is an "insurgent," they are not necessarily acting wrongly.

In some cases, an act of outright "insurgency" performed by an official may be acceptable or even laudable. If, for example, an individual in Nazi Germany were to become an SS officer with the specific purpose of undermining the SS whenever possible, they would have been acting as an insurgent, but their actions would have been heroic. However, as this example suggests, outright insurgency is only morally justifiable if there is an extreme degree of bias and moral corruption being perpetrated by the system which an individual plans to undermine. The idea that insurgency and mere role infidelity have different requirements for justification relates to Spooner's above-described analogy of the person accepting the sword; Spooner's argument that the person offered the sword should break their promise and use it to defend, rather than to kill, the innocent person rests on the fact that the promise was a promise to commit wrongdoing, and therefore that it was devoid of moral force. By analogy, Spooner argues, the Constitution is a

fundamentally pro-slavery document, so a judge's oath to uphold the Constitution is an oath to perpetuate the moral ill of slavery, and, as such, the oath is morally void. Applying this reasoning to the question of insurgency, the Nazi regime was fundamentally and explicitly devoted to anti-semitism, racism, and a host of other moral ills, so it is clear that any oaths or conditions of office for an SS officer would be morally meaningless, which means that insurgency to resist the aforementioned moral ills would be laudable. However, while there is a significant amount of racial bias in the United States criminal justice system, the system is not fundamentally devoted to upholding moral ills in the way that Spooner argued the Constitution was, let alone in the way that the Nazi regime was, so while role infidelity within the criminal justice system is justified, insurgency is likely not.

## Chapter 2: The Value of Role Fidelity Examined

In the foregoing section of this thesis, I argued that individual judges and prosecutors have a moral right, and, at times, a duty, to use their offices to counteract racial bias within the criminal justice system, even if doing so would require stepping outside of the bounds of their official roles. Throughout the previous section, I addressed several objections to the content of my arguments and touched on a few broad counterarguments to my position. In this section, I will expand on these counterarguments, which point to the dangers, both moral and practical, of stepping outside the bounds of one's role, and I will defend my position against those arguments. Although I will alternate between posing counterarguments to my position and defending my position, I will begin by developing a counterargument, in the form of a defense of role fidelity.

### Chapter 2.1: The "Threat to Democracy" Objection

When a judge or prosecutor in the United States makes a decision about how to handle a given case, that decision is not made "in a vacuum;" the individual is acting within an office of the criminal justice system, which is itself a part of the broader structure of the government, which, in turn, is itself an instantiation of a set of laws and legal principles stemming from the Constitution. The fact that the judicial and prosecutorial roles are so thoroughly enmeshed in the organization of our society means that each judicial and prosecutorial decision has a variety of impacts on our society: the actions of judges and prosecutors are representations of this country's democracy, both insofar as some judges and prosecutors are elected (or appointed by officials who are themselves elected) and insofar as the judicial and prosecutorial roles exist in order to ensure the proper enforcement of the law, which is enacted by elected officials. These roles within the criminal justice system are further linked to the functioning of democracy because

they constitute one of our society's main mechanisms for protecting democracy: if individuals, whether voters, representatives of private companies, or government officials attempt to perpetrate election fraud, it is the role of prosecutors to ensure that they are held accountable and the role of judges to oversee the resulting proceedings; in some cases in which there is difficulty determining the outcome of an election, the judiciary must decide which outcome the election supports. These aspects of the judicial and prosecutorial roles are important to describe here because they show how vital these roles are to the preservation and exercise of our democracy, which is the basis of all public and private functions in our society[22]. Therefore, judges and prosecutors must exercise the duties of their offices in ways that preserve public trust in those offices, because a decline in public trust in those offices logically carries with it a decline in public confidence in the functioning of our democracy. A further reason which spurs judges and prosecutors to faithfully execute their roles is that these individuals owe it to the public to respect and duly fulfill the results of democratic processes; the public not only expects that the results of their participation in democracy will be honored, they are so guaranteed by the Constitution, so if a law is enacted by elected representatives or through a referendum, the public has a right to see it enforced, which creates a duty in the actors of the criminal justice system to enforce it. While this argument, that the public has a right to see the results of their collective decision-making duly enforced, is one reason in support of the separation of powers which underlies our system of government, another, related reason is that the public should not be ruled by the unilateral decisions of a few unelected individuals, which is what could happen if individual actors in the criminal justice system began to decide matters of criminal law and public policy for themselves.

[22] While I will later challenge the idea that the judicial and prosecutorial roles in fact serve the idealized function in our democracy which is described above, I will set aside for the purposes of this thesis such views as anarchism and other forms of anti-electoralism which reject the value of representative democracy as such.

This account of the importance of role fidelity for judges and prosecutors lends itself to the following counterargument to my position: my position places supreme importance on the fulfillment of substantive justice, to the exclusion of all else; by encouraging judges and prosecutors to step outside of their roles if necessary to promote substantive justice, my position encourages these actors to frustrate the efforts of their colleagues in the criminal justice system, to betray the public by usurping the democratic processes in which they have participated, and to undermine the public's trust in the criminal justice system and in the functioning of this country's democracy. Furthermore, my position encourages judges and prosecutors to contradict not only their oaths of office but their own reasons for accepting their respective offices: appreciation of the judicial and prosecutorial roles includes the knowledge that justice is better administered and society better served by a court system than by vigilantism, but judges and prosecutors who accord with my position would be partaking in a form of vigilantism. The foregoing is admittedly a particularly forceful claim to make against my position, so in the interest of not setting up a straw-person of a counterargument, I will address more circumscribed forms of this argument as I go along.

I will begin to respond to the above counterargument by returning to Cover's writing about the Fugitive Slave Act cases: Cover describes the ways in which judges, caught between the contradictory impulses of role fidelity and personal, antislavery morality, tended to exaggerate the importance of role fidelity and downplay their own discretion, so that the decision to rule in favor of the slave owners appeared definitively correct. For example, in *Miller v. McQuerry*, Judge John McLean of Ohio ruled in favor of Miller, a slave owner from Kentucky, justifying this decision by asserting that judges could not "consider the laws of nature, and the immutable principles of right." He continued, "[judges] look to the law, and to the law only. A

disregard of this by the judicial powers, would undermine and overturn the social compact" (*Miller v. McQuerry*, 1849, cited in Cover, 120). Reflecting on this quote from McLean's opinion, Cover observes, "one has not *chosen* to participate in enslaving a man because of a mere device for controlling official behavior. No, one has *been compelled* to participate in enslaving a man because of fidelity to the consensual principles underlying ordered society itself" (Cover, 231, emphasis added). In this observation, Cover touches on what he calls "elevation of the formal stakes" (229), the phenomenon in which judges tended to portray their moral duty to maintain role fidelity, and, conversely, the likely consequences of violating role fidelity, as greater than they actually were. Applying this point to the above counterargument to my position, it is important to keep in mind that an individual judicial decision to violate role fidelity, or even a pattern of decisions in certain types of cases, is unlikely to lead to the collapse of the justice system or of our democracy. One could respond that even if role infidelity does not "overturn the social compact," it may "undermine it" in specific ways, to borrow Judge McLean's words, and I will address this response later in this section. Nevertheless, Cover's point about the elevation of the formal stakes serves to illustrate that judges' assertions about the importance of role fidelity are often influenced by their own social and psychological incentives to exculpate themselves.

Beyond elevation of the formal stakes, Cover writes, judges also tried to mitigate the difficulty of their positions by "retreating to formalism" (229), a term which Cover deliberately chose due to its similarity to the oft-used phrase "appealing to formalism." The judicial role includes (and was, at the time of the FSA cases, commonly understood to include) the consideration of various factors, including the text of the positive law, the intentions of the lawmakers, the coherence of each potential outcome with the body of precedent and with the

Constitution, and the moral status of each potential outcome. While there is no one method of negotiating between these factors which is universally considered correct, each judge must apply some method in each case, so personally abolitionist judges were incorrect when they asserted that respect for their formal roles mechanistically compelled them to decide in favor of the slave owners. Referring again to the judges' impulses to ease their social and psychological plights, Cover comments "the more mechanical the judge's view of the process, the more [they] externalized responsibility for the result" (Cover, 234). Although this point of Cover's does not lend itself to a defense of true role infidelity, it serves as a reminder that the judicial and prosecutorial roles inherently involve some amount of discretion and that respect for role fidelity may be consistent with various outcomes in a given case.

Even though, as the previous paragraphs indicate, there are social and psychological reasons which may motivate actors in the criminal justice system to exaggerate the importance of role fidelity, there are nevertheless arguments to be made in favor of role fidelity which do not rely on these types of reasons. Accordingly, I will now take a step back and examine several arguments in favor of obedience to authority in general. I will also intersperse critiques of those arguments and reasons why those arguments are insufficient to morally prohibit role infidelity in the types of cases in which I encourage it.

**Chapter 2.2: The Rawlsian Objection**

The first argument in favor of obedience to authority which I will examine comes from John Rawls' paper titled "Two Concepts of Rules." In this paper, Rawls asserts a distinction between the justification of a practice in general and the justification of a particular action taken under a practice, and asserts that people tend to overlook this distinction. Rawls attributes people's failure to distinguish between these two types of justifications to people's reliance on

the wrong conception of rules: people tend to rely on a "summary conception" of rules, states Rawls, under which rules are "summaries" of previous cases in terms of the amounts of net utility produced by different courses of action in those cases. Under this conception, rules have a purely instrumental value, to be followed only in cases in which following them is likely to produce more net utility than disobeying them - including, in this utilitarian calculus, any disutility which would result from the social effects of people learning that a rule was broken in a given case. This conception, Rawls holds, is errant, because rules develop as restraints on practices, practices which, therefore, do not include the option to engage in the type of utilitarian consideration described above. Rawls offers the practice of making promises as an example: while a rule utilitarian would assert that one must keep one's promises because of the social value of general trust in the practice of making promises and an act utilitarian would respond that in many cases, the utility gained by breaking a promise would outweigh any utility lost due to the social effects of breaking that promise, Rawls contends that both of these agents would be wrong, for the following reason: while the practice of making promises may have developed within the human population due to its utility, now that the practice exists and, by definition, entails that one who makes a promise must keep it, an agent is preempted from engaging in utilitarian consideration about whether to keep their promise in a given case.

To apply this argument to my position, one would hold that while the roles of judge and of prosecutor exist for certain reasons, including the promotion of justice, the protection of society, and the fulfillment of the will of the people, now that those roles exist and have certain boundaries, it is simply not available to an individual judge or prosecutor to deliberate whether to respect or to trespass those boundaries, even if the individual were tempted to trespass the boundaries of their role in order to better serve the aforementioned reasons for their role's

existence. I would respond that the judicial and prosecutorial roles themselves are poorly equipped to handle certain types of cases: specifically, cases in which racial bias has already had a significant effect. Whereas, to use the above-described example, the practice of keeping promises is valuable, so an individual must not violate this valuable practice even in an instance in which breaking a promise would create greater utility than would keeping a promise, the judicial and prosecutorial roles have a flaw, or a gap, when it comes to racial bias; insofar as those roles are meant to prevent or to combat racial bias, they are not valuable, so an individual criminal justice official who steps outside of their role in order to combat racial bias does not commit an error in the way in which someone who decides not to keep a particular promise commits an error.

Rawls, however, preempts this type of response: he accounts for the fact that any practice is likely to have circumstances in which it does not apply by pointing out that the rule which defines the practice need not be monolithic; a well-designed rule includes certain exceptions to the practice, so that agents would still be obeying the rule even if they are not carrying out the practice in a given case. Applying this point to my argument, Rawls could point out that the judicial and prosecutorial roles are not unitary practices, blind to racial bias, but dynamic roles which include measures to counteract bias, racial or otherwise: judges, for example, preside over the selection of jurors, a process which specifically weeds out biased individuals; judges instruct jurors to make a decision based on the facts of the case alone and set aside any personal feelings about the moral value of the defendant or of the alleged crime; judges may have racially prejudicial remarks stricken from the record or may even declare a mistrial due to such remarks; both judges and prosecutors swear an oath to uphold the Constitution, which includes protecting citizens' rights to equal protection under law. All of these measures, Rawls would point out,

serve to counteract racial bias, but they are built into the boundaries of the judicial and prosecutorial roles.

Returning to the content of the paper, Rawls asserts that the performance of many practices logically depends on obedience to the rules of those practices, so that one's action would not make sense if it diverged from the rules. To illustrate this point, Rawls uses the example of the game of baseball: one can swing a bat, throw a ball, and run along dirt paths without accepting the rules of baseball, but one cannot make a double-play or get an RBI without accepting the rules of baseball, because these phenomena only having meaning within the rules of the game[23].

Now that I have explained Rawls' argument, I will point out why his argument does not effectively challenge my position. First, in response to the point that the judicial and prosecutorial roles do not include deliberation about whether to respect or trespass the boundaries of those roles, I would employ an analogue of the arguments used by escaped slaves' lawyers in FSA cases: even if the judicial and prosecutorial roles preempt such deliberation, what is at issue is whether to respect or violate those roles themselves. To restate Cover's description of the lawyers' arguments, "the formal restraints are under challenge and cannot be justified by mere reiteration" (Cover, 129).

Second, Rawls' point that well-designed rules include exceptions to the relevant practices is tantamount to expressing a hope that rule-makers include all of the exceptions that they should, but it does not respond to the problem that a given rule does not necessarily include all of these exceptions. Applying this to the Rawlsian objection to my position, while certain measures to counteract racial bias are part of the judicial and prosecutorial roles, these measures are clearly insufficient, because they coexist with the racially biased procedures and outcomes which form

[23] Rawls, John, "Two Concepts of Rules," Duke University Press, 1955.

the empirical background of this thesis. Therefore, Rawls' point that a well-designed rule ought to include certain exceptions is not responsive to my position, because my position focuses specifically on cases in which the "built-in" exceptions have failed to counteract racial bias. One could further object that there are procedures for changing the boundaries of the judicial and prosecutorial roles - such as encouraging one's elected representatives to try to change the laws governing the criminal justice system or submitting a brief to an appellate court or the Supreme Court - but these procedures do not include an individual judge or prosecutor unilaterally changing the boundaries of their role in the middle of a case. However, I am not opposing individual criminal justice officials availing themselves of these established procedures to change the boundaries of their own roles, nor am I suggesting that those individuals can simply decide to reform their roles as they see fit during each case; rather, I am suggesting that given the current state of the judicial and prosecutorial roles, in which those roles are not sufficient to combat racial bias, individuals are morally justified in stepping outside of their roles when necessary in order to prevent racially biased outcomes.

Third, while Rawls is correct in his assertion that some practices, such as playing baseball, are defined in terms of the rules which circumscribe them, not all practices are so; returning to Cover's two analogies of the law - the law as chess game and the law as language - taking another player's pawn only makes sense in the context of the rules of chess, while using a neologism or non-standard syntax still makes sense even though it skirts the rules of the language. To reiterate the relationship between these two analogies and my argument, one may feel that a judge or prosecutor is "speaking law poorly" by trespassing the boundaries of their role, but, contrary to Rawls' argument, such a judge or prosecutor's actions would not "lack meaning."

At this point, one could offer the rejoinder that while the actions of a judge or prosecutor who violated role fidelity would not "lack meaning" in the strong sense of being meaningless as actions within the criminal justice system, these actions might still "lack meaning" in the weaker sense of being divorced from the legitimate authority of the actors who performed them; this objection would hold that if a judge or prosecutor violates role fidelity in deciding whether to prosecute someone, how to sentence someone, or what evidence to admit, their decisions are still, as a practical matter, legally binding, but they are nevertheless not the actions of a judge or prosecutor. Rather, the objection would continue, these decisions are the actions of a vigilante whose authority the criminal justice system upholds due to a mechanical fluke, and in this sense these actions "lack meaning." I would have no response to such a rejoinder except to assert that as this thesis goes on, I will demonstrate that role infidelity is sometimes justified, even if it does constitute vigilantism.

**Chapter 2.3: The Value and Nature of Authority**

For another type of argument in favor of obedience to authority and to rules, I will turn to Scott Shapiro's paper "Authority." In this paper, Shapiro elaborates several models of authority, describing what obedience to each one would entail, and uses these descriptions to form critiques of the various models and comparisons between them. As "Authority" contains too many distinct arguments for me to summarize here, I will lay out the overarching themes of the paper and then address the specific points in the paper which relate to my position: Shapiro evaluates different models of authority by applying each one to the two main "paradoxes of authority." One of these is the paradox of autonomy, which stems from the inconsistency of an agent both acting on their own will and acting on the will of the authority, and the other is the paradox of rationality, which

stems from the inconsistency of an agent both acting on their own interpretation of the balance of reasons and acting on the will of the authority. Across several conceptions of authority, Shapiro shows that it is difficult to reconcile the peremptory character of authoritative directives, which is a mainstay of different conceptions of authority, with the autonomy and/or rationality of the agents who are subject to those directives.

Shapiro notes that among means of putting authoritative directives into practice, ideology is more effective than the threat of punishment; using the example of laws as authoritative directives, most people follow the law most of the time, which introspection and common sense tell us is not because most people constantly weigh the risks of getting punished against the rewards of breaking the law, but because individuals have internalized the prevailing social norms against breaking the law. If these norms were absent, the criminal justice system would become overwhelmed and be unable to create sufficient deterrents from breaking the law. Conversely, in order for the criminal justice system to create any deterrent to crime, a critical mass of actors within that system must ideologically support the authority which they work to enforce, because if each actor within the system were only to obey authority to the extent that they felt it necessary in order to avoid punishment, the criminal justice system would likely collapse[24]. Shapiro sums up the above two points by asserting that, "it is simply not possible to have 'threats all the way down'" (Shapiro, 24-25).

This point could be used as a counterargument to my position because it shows that the ability of the criminal justice system to complete its valuable role in our society is vulnerable to "insurgency," or individual actors using their offices to contradict the will of the authority. Such insurgency is a threat to the functioning of the criminal justice system not only insofar as it

[24] Shapiro, Scott, "Authority," 2000. - Although neither Shapiro nor I present empirical evidence to support these claims, they are based on fairly self-evident political and social observations, so I will assume that they are true.

directly frustrates the will of the authority, but also insofar as it disrupts coordination between actors within the system. Using Shapiro's point, if not enough actors within the system ideologically support the authority, there is no way to bring the "insurgents" back in line, because the insurgents would have compromised the very system responsible for enforcing the rules.

The inability of the criminal justice system to right itself once significantly compromised would likely play out in the form of what Shapiro terms "coordination problems;" some plans only work if all of the actors involved coordinate their behavior, so the value of an individual's conforming to the plan comes not from the direct results of their action, but from the aggregate result achieved when everyone conforms to the plan. Shapiro illustrates this phenomenon using the example of driving: there is no reason why driving on one side of the road is better than driving on the other, but once a society has arbitrarily chosen the side of the road on which people will drive, there is immense value in everyone following suit. It would be preposterous for one to suggest that given dire enough circumstances, it is acceptable to drive on the wrong side of the road, because even a tiny fraction of drivers going the wrong way would destabilize our system of transportation[25]. Similarly, directives from the authority are what lend coordination to the actions of different officials within the criminal justice system, so if insurgents were to compromise the practice of obeying the authority, the actions of everyone within the system would become uncoordinated; each actor who wanted to conform to the will of the authority would have to alter their actions so as to predict and counteract the divergent behavior of the insurgents, and the insurgents would have to modify their own actions likewise in order to counteract the measures taken by their role-conforming colleagues. The result of this mutual

[25] One could respond to this example by pointing out that in some cases, such as when a natural disaster is affecting an area and people must evacuate, it is acceptable to drive on the wrong side of the road. The rejoinder (echoing Rawls) would be that in those circumstances, the authority has made clear that the usual practice has been altered, so all drivers' behavior remains coordinated, even though it is coordinated around an abnormal practice.

distrust and attitude of gamespersonship would not be the bulk of actors conforming to the will of the authority and a few actors diverging from that will; rather, the result would be the type of chaos which wrong-way drivers would cause on the road.

I would defend my position against the above argument by stating that although coordination problems can seriously undermine the value of an institution, it seems unlikely that role infidelity, applied in the circumstances in which my position encourages it, would in fact lead to the utter breakdown in coordination such as the one described above. I believe that this is the case because actors in the criminal justice system already engage in a certain amount of gamerspersonship and the system has not, as a result, descended into chaos. Prosecutors, for example, regularly bring more charges (and more serious charges) than they will likely be able to prove in court, so that either the defendants plead guilty to some of the charges (or lesser charges) or the jury convicts on some of the charges. Conversely, defense attorneys are sometimes able to arrange for their clients to plead guilty to lesser charges than the original charges (even if their client did not actually commit the elements of the lesser charge). Judges and juries, as well, sometimes reject the harshness with which the prosecutors have decided to charge the defendants or the severity of the punishments mandated by law and accordingly find that there is insufficient evidence to convict the defendant for such serious charges, even when the evidence is actually sufficient. These are just a few of the measures which criminal justice officials take in order to "game" the system and produce outcomes which they prefer, and thus far these measures have not produced "wrong-way-driver-like" chaos, so it is safe to assume that the addition of the anti-racist role infidelity which I suggest would not cause such chaos.

The objector could further assert that even if role infidelity did not cause the type of chaos which Shapiro describes, it could still cause a certain amount of disorganization, which

would undermine the effectiveness of the criminal justice system. My response would be that in certain circumstances, such as those in which maintaining role fidelity would allow severely racially biased outcomes, the value of stepping outside of one's role would outweigh any negative effects which role infidelity would have on the system as a whole. I would maintain this position because, as I assert above, officials frequently leverage specific aspects of their roles in order to promote certain outcomes and it has not severely impacted the organization of the criminal justice system, so it is unlikely that the role infidelity which I suggest would cause enough disorganization to outweigh its anti-racist benefits. Furthermore, the role infidelity which I encourage could actually increase organization in a positive way: if, for example, prosecutors learned that when they disproportionately charge people of color under a certain statute, judges were likely to counteract their efforts through role infidelity, prosecutors might work to make their patterns of charging defendants more egalitarian, in order to avoid losing cases and getting in unnecessary conflict with judges. Similarly, if legislators saw that prosecutors were refusing to charge suspects under statutes whose impact or creation seems racially biased, the legislators might respond by removing or reforming the racially biased statutes[26].

In "Authority," Shapiro elaborates Joseph Raz' "normal justification thesis" in a way that supports my position: the normal justification thesis holds that an authority is legitimate if agents would more often conform to the balance of preexisting reasons which applied to them by obeying the authority than by deliberating how to act themselves. Under this thesis, the authority need not outperform each individual's judgment in each case in order to be legitimately authoritative in that case; if each individual were to deliberate about the balance of reasons in

[26] Idealistic as this suggestion may seem, I believe that examples of this kind of positive organization in the criminal justice system exist. For example, the increasing legalization of marijuana, at the legislative level, is arguably due in part to the actions of prosecutors who refuse to charge people for marijuana-related crimes and of judges who refuse to sentence people who have been convicted for such crimes.

each case and only obey the authority if the authority had come to the correct outcome in that case, the instrumental value of the authority for the entire group of subjects would be undermined. Nor can each individual consider obedience to authority valuable in itself and place that value alongside the balance of reasons while deciding how to act, asserts Raz, because the authority arrives at directives by considering the balance of reasons, so by placing the value of following a given directive alongside the reasons which informed that directive, one would "double count" certain reasons.

Shapiro uses Raz' example of arbitration to demonstrate why the normal justification thesis has the features which it does; two parties enter into arbitration because they agree that doing so will produce an outcome which more fairly reflects the balance of reasons than any outcome which the parties could produce by themselves. Therefore, holds Raz, each party must relinquish their power to consider the balance of reasons and place that power entirely in the hands of the arbitrator, so that the outcome announced by the arbitrator now has the compulsory force which the underlying reasons would have had. Clearly, this arrangement only works if both parties respect the arbitrator's decision, regardless of whether they feel it accurately reflects the balance of reasons, because if both parties must agree that the decision reflects the balance of reasons, the situation collapses into the *status quo* from before the parties entered into arbitration. While each party must be willing to accept "great mistakes" in the arbitrator's reasoning, each party need not be willing to accept "clear mistakes" in that reasoning; while a great mistake is one which "deviates greatly from the balance of reasons and whose detection as an error requires the agent to deliberate on the underlying dependent reasons putatively supporting the claim," a clear mistake is one which "may not deviate substantially from the balance of reasons but wears its error on its face" (Shapiro, 36). To put this abstract distinction into concrete terms, Raz gives

the example of arbitration over the sum of two integers: if an arbitrator were to make a great mistake and decide that the sum of five and three was thirty-four, neither party could fairly dispute it, because to dispute it would be to reconsider the balance of reasons and thus to nullify the arbitration. If, however, the arbitrator were to make a clear mistake and decide that the sum of five and three was seven-and-a-half, the parties could dispute it, because the knowledge that the sum of any two integers will be an integer exists independently of the balance of reasons in the specific case. Conversely, Raz states that if the arbitrator was not truly considering the balance of reasons, as would be the case if the arbitrator was incapacitated or had been bribed, the parties need not respect the arbitrator's decision, because that decision would not truly reflect and replace the preexisting reasons.

This caveat in Raz' normal justification thesis supports my position, because while a given judge or prosecutor ought to accept that other actors in the legal system sometimes make mistakes, including "great mistakes," a judge or prosecutor need not accept that other actors will make "clear mistakes." Given that the United States' criminal justice system explicitly operates on the principle of racial non-discrimination, decisions which are motivated by racial bias constitute clear mistakes, in the Razian sense, so judges and prosecutors need not honor those decisions. For example, it is the role of legislators, not that of prosecutors, to decide which laws to implement, so, in order to respect the authority of the legislators, a prosecutor must enforce even those laws which the prosecutor feels do not best serve the interests of the people. If, however, a given law is clearly meant to preserve racial inequality, a prosecutor need not enforce it, because the prosecutor need not infringe on the authority of the legislators in order to conclude that the law does not seek to reflect legitimate reasons. One may object to this example by pointing out that there is no simple, "mathematical" way to determine whether a given law is

meant to preserve racial inequality, while there is such a way to determine whether the supposed sum of two integers could be the correct answer.

I would respond that while this objection makes a good point, it is only reason to refine the practice of engaging in role infidelity, rather than reason to entirely forswear that practice; the prosecutor from the example should not refuse to enforce a law simply because that law may have unequal impacts on different races, but if, on the opposite extreme, the law is clearly written so as to have unequal impacts, the legislators who proposed it touted it for these unequal impacts, and those legislators ran for office promising to preserve racial inequality, the prosecutor can be sufficiently certain of the racial bias behind the law that they would be justified in refusing to enforce it. Another objection to this example would be that it is the role of judges, not that of prosecutors, to determine whether a given law is meant to preserve racial inequality, and if a law is in fact meant to preserve such inequality, it is the role of judges to strike it down. My response would be that the existence of the judiciary simply means that there are two sources of authority which may either be truly deliberating about the balance of reasons or may be acting so as to preserve racial inequality; if there is sufficient reason to believe that a given judge or relevant portion of a panel of judges which fails to strike down a given law is motivated by racial bias, the prosecutor need not consider the judicial approval of the law as conclusive evidence that the law is not meant to preserve racial inequality. Admittedly, what such "sufficient reason" would consist of in the case of the judiciary poses a tougher problem than it does in the case of the legislature, because judicial opinions tend to be more firmly grounded in legal reasoning and therefore tend to be less overtly politically or racially motivated than are legislator's statements about a given law.

In "Authority," Shapiro also evaluates Raz' "preemption thesis," which holds that the directives of a legitimate authority not only give the subjects of those directives reasons to obey the directive, they also nullify or make irrelevant any reasons for disobedience. Raz suggests that the preemption thesis follows from the normal justification thesis, because, as described above, the social value of an authority only exists if the subjects treat the authority's directives as reflecting and replacing the balance of reasons in each case; if authoritative directives did not preempt whatever reasons each agent believes apply to them, the agents would not benefit from the authority's generally-superior ability to consider the balance of reasons. Furthermore, Raz asserts, each agent cannot treat the authoritative directive as a reason in itself and place it alongside the other reasons when making a decision, because if the authoritative directive is meant to reflect other reasons, the agent will inadvertently "double-count" some reasons. Shapiro responds to the point about "double-counting" reasons by arguing that an agent need not treat the authoritative directive as a monolithic, unitary reason; instead, the agent could consider the process which led to the directive, identify any relevant reasons which were omitted or considered improperly by the authority, and arrive at a decision which more accurately reflects the balance of reasons.

While an agent's consideration of the relevant content-dependent reasons could outperform the authority's consideration of those reasons in a given case, there is still something to be said for a content-independent reason to obey authoritative directives, especially when there are separate moral reasons to obey those directives (such as the earlier-described reasons for judges and prosecutors to honor the decisions of the legislature and of actors in the criminal justice system). However, as Shapiro asserts, there can be a content-independent reason to obey which carries a finite weight, so that each agent can respect the authority *qua* authority and still

disobey a given authoritative directive if the content-dependent reasons for disobedience are great enough. Shapiro states that "the benefit of [such accounts of authority], therefore, is that they attempt to account for [the] virtues of relying on authority without succumbing to the vice of rule-worship" (Shapiro, 50). This idea parallels Dworkin's idea of principles pointing in a specific direction with a certain weight and therefore being binding even if each principle is not determinative of the outcome in each case. In a case in which a prosecutor or judge feels that the relevant legislation or the actions of another actor in the criminal justice system have been skewed by racial bias and therefore do not properly reflect the content-dependent reasons in that case, the prosecutor or judge could consider whether the extent of the racial bias is sufficient to outweigh the content-independent reason which the prosecutor or judge has to maintain role fidelity and honor the result of the case.

In a similar vein as Raz' preemption thesis, Shapiro elaborates his own distinction between the "decision and constraint models" of authority: under the decision model (to which, Shapiro asserts, both Raz and his critics subscribe), an agent must make a decision about each action which they take, and different conceptions of authority represent different ways in which authoritative directives affect the agent's decision-making process. Raz' conception of authority holds that authoritative directives give preemptive reasons for an agent to decide to obey, while Raz' critics hold that authoritative directives give reasons to decide to obey which are content-independent but not preemptive or, under another stance, which are no reason at all to obey beyond the content-dependent reasons which already applied to the agent. Shapiro, however, contends that all of these conceptions of authority are incorrect, because authority conforms to a constraint model, under which an agent who has submitted to an authority has no choice but to obey the authority's directives going forward; the issuance of such a directive does

not change the set of reasons which apply to an agent or change the way in which the agent should consider those reasons, rather the issuance of the directive removes the reasoning process entirely and replaces it with the will to complete the directive.

Returning to the example of arbitration as the paradigm case of the exercise of authority, Shapiro asserts that authority conforms to the constraint model rather than to the decision model because whereas the decision model portrays the authority as a mediator, the constraint model portrays the authority as an arbitrator. While a mediator helps parties come to a solution which serves each party better than any solution which the parties could have agreed on without the mediator, an arbitrator lays out solutions which, given that the parties agreed to arbitration, are morally binding on the parties; a mediator is only valuable insofar as they lay out solutions which are, in each case, better than those which the parties could have come to alone, but an arbitrator is valuable because its existence leads to better outcomes overall than its nonexistence, even if, in some cases, the solutions which the arbitrator lays out are not as good as those which the parties could have come up with alone. The social value of the arbitrator comes not only from the arbitrator's ability to lay out better solutions than the parties under its purview most of the time, but also from the arbitrator's tendency to prevent the detrimental social conflict which would exist if parties had to work out problems amongst themselves.

A government which acted as mediator would, according to Shapiro, be a Platonic government, because the officials would have their positions not because of the input of the populace, but because of their ability to make decisions which would benefit the populace, and it would be this ability to benefit the populace which would lend legitimacy to the government. By contrast, a government which acted as arbitrator would be a democratic government, because the officials would have their positions because of the input of the populace, and it would be this

input which would lend legitimacy to the government. Shapiro asserts that obeying the directives of a government acting as arbitrator is important because "deference to democratically elected authority is deference to a power-sharing arrangement that is socially necessary, empowering and fair. By disobeying, subjects are unilaterally, and hence unreasonably, setting the terms and direction of social cooperation" (Shapiro, 83). However, as the empirical claims underlying this thesis contend, the "power-sharing arrangement" which constitutes the United States criminal justice system is not "socially necessary, empowering, [or] fair" to people of color. Universal "deference to democratically elected authority" on the part of judges and prosecutors, in the form of deference to enacted laws, to the discretion of other actors in the criminal justice system, to the decisions of jurors, etc., would therefore not serve to fulfill the social needs of people of color, to empower them, or treat them fairly. If, under Shapiro's view, people ought to obey authoritative directives, even when they do not produce the substantively-best outcomes, because obedience furthers a beneficial and equitable social arrangement, but in some cases, obedience does not actually further such a social arrangement, there is only reason to obey authoritative directives when they are likely to produce the substantively-best outcomes. Furthermore, the justice system lacks the above-described beneficial characteristics largely because people of color have been excluded from participating in "setting the terms and direction of social cooperation," so if disobedience to authority would serve to make the process of crafting social cooperation more inclusive - such as in cases in which existing racial bias would further the disenfranchisement of people of color - this disobedience ought to be commendable under Shapiro's view.

**Chapter 2.4: Evaluating Unrestrained Role Infidelity**

Although neither Rawls, Raz, nor Shapiro were addressing the topic of this thesis in their respective works, they all presented stances which, though different from one another, tend to contradict the position which I defend in this thesis. Kimberley Brownlee, by contrast, presents a stance which is not only consistent with my position, but goes even further than my position does in encouraging criminal justice officials to prioritize substantively just outcomes. In the paper "Legal Obligation as a Duty of Deference," Brownlee critiques the conception of legal authority which Philip Soper presents in *A Theory of Law* and uses this critique to argue that actors in the legal system (or in the political system more generally) have a broad prerogative to step outside the boundaries of their roles. Brownlee begins the paper by comparing the voluntarist position on legal obligation, which holds that individuals freely subject themselves to political states and are therefore bound by the valid legal dictates of those states, and the non-voluntarist position, which holds that although people do not choose to subject themselves to political states, the existence of such states is beneficial and therefore the dictates of those states are binding. Brownlee quickly dismisses the voluntarist position on the grounds that it does not reflect the social or historical reality and then begins to examine what would be necessary for legitimate legal obligations to exist under the non-voluntarist position; Soper's argument, as described by Brownlee, holds that a person is obligated to obey the dictates of the political state in which they live because the existence of that state is better than its non-existence, insofar as the existence of the state more effectively promotes an individual's previously-held values than would the state's non-existence. The problem with this argument, as Brownlee points out, is that an individual, such as an anarchist, may deny that the existence of the state is overall better than its non-existence, and, conversely, an individual may deny that the existence of the state more effectively promotes the individual's own values than would the state's non-existence. Soper's position resists this

objection, however, because it holds that one must view the legislators with "good faith" and assume, unless there is strong evidence to the contrary, that the legislators are working to promote values which are, by their own lights, beneficial to the members of the state. By deferring to the legislators in this way, Soper holds, one is maintaining consistency with one's own values, because one is upholding an arrangement in which, were the roles reversed, the legislator would view the citizen with "good faith" and defer to the citizen's dictates.

The problem with Soper's semi-Kantian, deference-based view is, as Brownlee points out, that the state's dictates may not only promote values which differ from one's own values, they may promote values which directly contradict one's own values; if deference is ultimately based on maintaining consistency with one's own values, it cannot include acting so as to contradict those values. Not only do the dictates of the state fail to carry a preemptive reason for obedience (to borrow a term from Raz), they fail to carry any content-independent reason for obedience, because if a given dictate contradicts an individual's values and, under the non-voluntarist position, those values are what compel obedience to the the state's dictates, the individual has no reason at all to obey the dictate (beyond any content-dependent reasons which previously existed). It is especially important to understand the relationship between the state's authority and individuals' values in this way, Brownlee holds, because the state, unchecked by the values of individuals, could contradict the purpose of its own existence and produce results which are worse than those which would occur if the state did not exist. Brownlee illustrates this point as follows:

> Suppose, for example, the legislator enacts a law intended to render a particular group of persons, such as the disabled or the elderly, vulnerable to certain kinds of treatment such as sterilisation or non-voluntary euthanasia. Such a law could render those persons more vulnerable to certain forms of abuse and harm than they would be in a state of nature which lacks the institutional mechanisms to implement chosen practices" (Brownlee, 595).

Applying this idea to the racial bias which exists in the criminal justice system, although that bias is not a "chosen practice" is itself, it is certainly the result of chosen practices and the ways in which those practices are implemented, and it could not exist in its present form outside of the context of the criminal justice system; tragically, hate crimes against people of color have been and continue to be perpetrated by individuals and groups, but no individual or privately-organized group could amass billions of dollars and use those funds to detain and incarcerate hundreds of thousands of people[27]. If an individual judge or prosecutor's reason for obeying the dictates of the state is to facilitate the state's promotion of values which the individual already held, an individual who held the value of racial equality would have no reason at all to obey those dictates of the state which would promote racial inequality, especially if those dictates would promote a level of racial inequality greater than would be possible in a stateless world.

In the paper "Responsibilities of Criminal Justice Officials," Brownlee takes the above position a step further, asserting that there need not be evidence of present or likely future substantive injustice in order to justify role infidelity by criminal justice officials; rather, Brownlee claims, there only needs to be a conceptual gap between the performance of one's defined role and the promotion of substantive justice. Brownlee defends this claim by elaborating the concept of a "moral role," the set of moral responsibilities and prerogatives which one shoulders in virtue of their having a specific office or social function, but which is not necessarily coextensive with the official boundaries of that office or social function. A doctor, for example, could be said to have a greater duty to help strangers who are having medical

[27] Of course, it is not as though there is a fund in which billions of tax-dollars are used explicitly for the incarceration of people of color *en masse*, but, given the empirical claims which underlie this thesis, it is reasonable to assert that the criminal justice system produces racially biased results beyond anything which an individual or private group could achieve.

emergencies, even if those emergencies occur when the doctor is not at work, than does a layperson, because the doctor's moral role includes a duty to care for the sick. Given that each office within the criminal justice system exists in order to promote the fulfillment of a certain moral role, Brownlee holds, each person who holds such an office has an obligation to fulfill the moral role specific to that office and has no *pro tanto* obligation to respect the boundaries of their official role (which Brownlee terms "office dictates"). The office of judge, for example, exists in order to promote just outcomes in cases and in order to ensure that all parties in a case receive equal treatment and ability to advocate their position before the court. An individual judge's obligation, according to Brownlee, is to strive to fulfill the above-described goals which motivate the existence of the office of judge, not to merely complete the dictates of that office. Therefore, under Brownlee's view, in contrast with mine, an actor within the criminal justice system should not engage in role infidelity when their personal moral beliefs contradict a given outcome which role fidelity would bring about; rather, an actor should engage in role infidelity when it is necessary in order to fulfill the moral role which they already possess in virtue of their office, regardless of their personal beliefs.

Criminal justice officials' moral roles and their official roles may be motivated by the same beneficial impulses, Brownlee claims, but this does not mean that the office dictates are best-designed, or even well-designed, so as to promote the fulfillment of the moral roles. One reason for this which Brownlee provides is that the roles within the criminal justice system are meant to be coordinated, so that if each of the individuals who hold the offices relevant to a given case do their jobs effectively and impartially, there will likely be a substantively just outcome. However, as both Brownlee and I have described, in many cases, at least one criminal justice official is affected by racial bias, which means that if each other official were to simply

adhere to their office dictates, the results would be skewed by racial bias. Given that the moral role of judges and prosecutors excludes the perpetuation of racial bias and, as the empirical claims underlying this thesis demonstrate, the perpetuation of such bias is consistent with a judge or prosecutor adhering to their office dictates, Brownlee would hold that a given judge or prosecutor has no moral reason at all to adhere to their office dictates *qua* office dictates.

Brownlee's stance admittedly goes further than mine, because I believe that even if racial bias and role fidelity are conceptually consistent, there must be evidence of actual bias which is likely to impact a case in order to justify role infidelity in that case. Brownlee's stance also diverges from mine in that she asserts that while some offices in the criminal justice system have more discretion built into them than others do, individuals who hold offices with little or no discretion built into them should nevertheless step outside the boundaries of their roles if they see fit. The most extreme example of this is Brownlee's claim that a prison guard should release or simply refuse to detain a convict whose sentence the prison guard feels is unjustified or likely to be detrimental to the convict's wellbeing. While a prison guard may take measures to change sentencing guidelines and to improve the conditions in prison, such as by asking to speak as an expert before Congress, submitting a brief to the Supreme Court, or simply voting for representatives who support prison reform, it would severely damage public trust in the criminal justice system if the guard were to simply let a convict go. If prison guards were to frequently make the unilateral decision to let prisoners out, the likely result would be an increase in vigilantism, because communities would not have reason to believe that offenders, once duly convicted, would actually serve out their sentences.

Despite the parts of Brownlee's argument which I feel overreach, Brownlee's points about moral roles and about the criminal justice system being like an ecosystem in which each

actor's behavior affects each other actor's duties support my argument, as does Brownlee's stance about the potential social consequences of role infidelity, which is as follows: "[the] potential costs of non-adherence for the judge and perhaps for people's confidence in criminal justice institutions must be weighed against other costs that arise through general adherence (including costs that should weaken people's confidence)" (Brownlee, 124). In this quote, Brownlee points out that while role infidelity may have costs for the individual who performs it (such as being fired or impeached) and while it may create public distrust in the criminal justice system, role infidelity, done right[28], can also serve to foster public trust in the system. If the public sees that criminal justice officials are willing to step outside their roles when necessary in order to promote a just outcome, rather than blithely following their office dictates even those dictates lead to unjust outcomes, the public may feel more protected by the justice system and feel more confident about turning to the justice system for help.

Brownlee addresses several objections to her view, and even though that view is bolder than mine, the responses which she offers to these objections serve to support my argument. One objection which Brownlee addresses is the "democratic procedures" objection, which holds that each office within the criminal justice system is situated within a democratic system, so non-adherence to one's office dictates on the part of the individual who holds that office undermines the democratic system. Brownlee's response is that non-adherence by a criminal justice official is no more damaging to the democratic system than is civil disobedience by ordinary citizens, which, although some would condemn it as undemocratic, is actually pro-democratic, because it allows those deprived of the opportunity to inform democratic

[28] Brownlee acknowledges that role infidelity has the potential to degrade public trust in the criminal justice system and accordingly she lays out different ways in which an official can engage in role infidelity while minimizing any such damage to public trust. These suggestions are varied based on the amount of discretion built into the different official roles and the amount of publicity which they attract.

decisions to voice their opinions about those decisions. Applying this response to racial bias in particular, disproportionate incarceration and subjection to other punishments hinders people of color's ability to participate in democratic processes. More broadly, equal protection under the law is a democratic value, so discrimination within the criminal justice system is undemocratic. Therefore, by engaging in role infidelity when necessary to prevent racially biased outcomes, a judge or prosecutor would be protecting, not denigrating, the democratic procedures which govern them. This is so not only because preventing racially biased outcomes is inherently pro-democratic, but because the individuals and communities who would otherwise have been disenfranchised (literally and/or figuratively) by those outcomes would be better able to participate in democratic processes if they are not subjected to mistreatment by the criminal justice system. Furthermore, the ability of communities of color disproportionately impacted by the criminal justice system to participate in democratic processes could help change the law and the policies which govern the criminal justice system, so that role infidelity becomes less necessary over time. This point relates to Brownlee's response to the "valuable institutions" objection.

The "valuable institutions" objection holds that non-adherence by officials is likely to undermine and threaten the institutions within which the officials operate (in this case, the criminal justice system), thereby threatening not only the unjust, but the just and beneficial applications of those institutions. As Brownlee responds, role infidelity may make an institution more valuable, because it may contribute to a shift in the norms of the institution over time, so that eventually, officials will not need to engage in role infidelity in order to promote the most just possible outcomes.

Role infidelity, however, does not necessarily produce the most substantively just outcome in each case, as one objection to this portion of my argument would point out; a judge, for example, could sentence certain defendants more severely than others based on the judge's personal beliefs about which defendants are deserving of more punishment, and these personal beliefs could lead to the judge sentencing people of color more harshly than white people, through implicit or explicit bias. In such a case, the judge would avoid introducing racial bias and therefore preserve greater substantive justice by remaining faithful to their role, rather than by allowing their own beliefs to motivate role infidelity. This problem is brought into greater relief when considering potential cases in which an individual actor's personal beliefs would lead them not to bias the outcome in a given case, but to bias the democratic system which affects future cases. A Bureau of Elections official could, for example, deliberately influence an election in favor of a racist candidate, which would not only affect the current population, but could have long-term effects if the candidate, once in office, further undermined the democratic processes so as to entrench themself and their political allies in power.

This line of reasoning mirrors what Brownlee terms the "incompetent official[29]" objection, which holds that some officials, especially those incompetent at making moral judgments of their own, may more effectively fulfill their moral role by adhering to their office dictates than by trying to determine the right course of action in each case. Brownlee's response to this objection is that while some officials are incompetent, others are competent, and an official competent to fulfill their moral role better by acting on their own moral judgements, rather than by adhering to their office dictates, ought to do so, and should not fulfill their moral

[29] In the examples which I present above Brownlee's point, the hypothetical officials behave maliciously, not merely incompetently, but for the purposes of this argument, "incompetent" means "tending to make moral decisions with less accuracy than would be achieved by maintaining role fidelity" whether that tendency results from lack of aptitude, implicit bias, or explicit bias.

role suboptimally simply out of fairness or equality with their incompetent peers. I present Brownlee's response to the "incompetent official" objection so as to give a thorough representation of her position, but I admittedly do not think that this response achieves its purpose; the "incompetent official" objection gets its force from the fact that people tend to overestimate their own competence at making moral decisions. The corollary of this fact, that a given official who believes themself competent to engage in role infidelity may in fact be incompetent, is the reason which the objection presents for why each official should maintain role fidelity, and it is also, arguably, the reason why roles exist in the criminal justice system to begin with.

Another problem with Brownlee's response to the "incompetent official" objection, one which also challenges the objection itself, is that an official's competence or incompetence may be irrelevant to the question of the whether the official is able to promote a better outcome by engaging in role infidelity than by maintaining role fidelity. This is the case because there are issues in the criminal justice system about which competent people can disagree. For example, competent judges can disagree about whether the Constitution protects the right to abortion and the right of the individual to own guns, about whether affirmative action is justified, etc. There may be one side of either of these issues which the balance of reasons better supports, but given that a competent official could support the other side, an official who engaged in role infidelity based on the feeling that they "know better" about the issue would be abusing their power and undermining the criminal justice system. Brownlee, however, provides against this problem by including the caveat that in "non-morally-difficult-circumstances," meaning circumstances which do not challenge the official to engage in tough moral deliberation, role infidelity should be limited to "occasions where an official *clearly* will better act as the reasons that apply to [them]

would have [them] act when [they] attend to those reasons directly and not to the rules and orders governing [them]" (Brownlee, 128, original emphasis). Beyond establishing that an official should defer to their office dictates unless there is clear reason not to, which makes Brownlee's view easier to accept, this quote also suggests that an official has greater latitude to engage in role infidelity when there is a moral dilemma involved than when there is not.

While this suggestion may seem counterintuitive, because the "moral difficulty" of a situation would make it less clear what the right outcome would be, and thus would make it less clear whether role infidelity is justified, it nevertheless makes sense; the moral difficulty of a situation confers particular importance on bringing about the right outcome in that situation, so an individual should have more discretion to engage in role infidelity if it is necessary to bring about the right outcome than they should have in a less morally difficult situation. This point is vital to my argument because it helps draw the distinction between a "free-for-all" of role infidelity and the role infidelity in specific types of cases which I am encouraging. If, for example, a person of color were to sue their former employer for having fired the employee because of their race and the judge could tell that the jury was likely to find in favor of the employer, it would not be justified under my view or, I believe, Brownlee's, for the judge to suppress evidence presented by the employer or otherwise attempt to skew the case in favor of the employee, even if the employee truly had been fired because of their race. If, however, it were necessary for the judge to engage in similar role infidelity in order to prevent a person of color from being imprisoned because of their race, the judge would be justified in doing so, because the potential for false imprisonment confers far greater moral difficulty and importance on the case than does the potential for racist hiring and firing practices.

Although there are some cases in which competent individuals can disagree, there are also cases which are truly matters of competence and lack thereof. While I acknowledge the gravity of the "incompetent official" objection, I nevertheless believe that role infidelity is justified in some of these cases. One reason for this belief is that the background empirical claims of this thesis - that there is a significant amount of racial bias in both the outcomes and the procedures of the criminal justice system - would likely not be the case were it not for those with racist personal beliefs or biases engaging in role infidelity. Every role within the criminal justice system explicitly eschews racial discrimination, so for any actor in that system to allow such discrimination to affect their decision-making would be to engage in role infidelity; admittedly, the phrase "role infidelity" connotes a conscious decision to step outside of one's role and not all of the racially biased outcomes in the criminal justice system are the result of conscious decisions, but some racially biased outcomes certainly are. Therefore, the "incompetent official" objection cannot derive its strength from a counterfactual in which actors with racist personal beliefs engage in role infidelity, because that counterfactual already exists. Furthermore, role infidelity in the contexts in which I support it is justified in virtue of the fact that it is necessary in order to counteract the effects of the role infidelity already being performed by those with racist personal beliefs. This relates to Brownlee's point that what is required of an actor in order to fulfill their moral role may change based on the behavior of the other actors in the system and may therefore stray outside of the actor's office dictates.

A final reply to the "incompetent official" objection would be that we must assess the actions of those who engage in role infidelity using the concept of moral luck, which holds that it is appropriate to include certain factors beyond an individual's control in the moral assessment of that individual. Applying the concept of moral luck, there is nothing which is, in principle,

morally wrong about engaging in role infidelity, but if a given person steps outside the boundaries of their role in the criminal justice system in order to do something which is morally wrong, the state of affairs created by their decision can be morally condemned. If, for example, one judge believed in racial equality and another equally sincerely believed in white supremacy, the racial egalitarian judge would bring about a morally praiseworthy state of affirs by engaging in role infidelity while the white supremacist judge would bring about a morally condemnable state of affairs by so doing, even though each judge was, by their own lights, bringing about the most racially just outcome. This distinction between the two judges comes from the difference between axiological judgments, which concern the moral value of circumstances, and deontic judgments, which concern the moral value of actions[30]. Whereas, axiologically, a white supremacist judge using their personal beliefs to impact outcomes in criminal cases is morally condemnable, deontically, such a judge does not act wrongly by engaging in such role infidelity motivated by sincerely-held personal beliefs, as the other sections of this thesis argue. By relying on the distinction between axiological and deontic judgments, one can address the problem of people potentially using role infidelity towards bad ends without perpetrating the inconsistency of morally condemning individuals for performing the same actions which are laudable when performed by other people.

[30] Nelkin, Dana K., "Moral Luck", *The Stanford Encyclopedia of Philosophy*, 2019, Edward N. Zalta (ed.), https://plato.stanford.edu/archives/sum2019/entries/moral-luck/.

## Conclusion

In this thesis, I argue that judges and prosecutors have the right, and, in some cases, the duty to step outside the formal boundaries of their roles when it is necessary in order to counteract racial bias in the criminal justice system. I defend this conclusion in several different ways, including by arguing that if racial bias is already present in a given case, for a judge or prosecutor to remain within their role would be for them to allow the racial bias to impact the outcome of the case, which would be unacceptable.

There is, however, the problem that not all judges and prosecutors personally believe in racial equality, and that by dint of following my suggestion to step outside of their roles, some criminal justice officials could promote the very racial bias which my thesis strives to combat. One way in which I respond to this problem is by pointing out that my thesis would not be relevant were it not for a substantial number of criminal justice officials who already violate the bounds of their roles in promoting racial bias (whether they do so deliberately or not); it is this racially biased role infidelity which makes necessary the countervailing, anti-racist role infidelity which I encourage. In similar fashion, I respond to the problem that my suggestions would harm democratic processes and public trust in the criminal justice system by arguing that by creating a system in which laws are less likely to be enforced in racially biased way and in which people of all races are likely to receive equal treatment under law, judges and prosecutors who engage in role infidelity would actually strengthen democratic processes and public trust in the criminal justice system.

A more general problem which my position faces is that under some views of judicial philosophy, namely that of Ronald Dworkin, the judicial role includes the consideration of the moral content of different outcomes in a given case, so a judge who applied such moral

consideration would not be stepping outside of their role, but fulfilling their prescribed role. I respond to this problem by arguing that Dworkin's view does not afford a large enough role to the consideration of morality, because there are some instances when even a morally-guided interpretation of the positive law, precedent, and underlying legal principles would still produce a racially biased outcome, and in those cases the judge or prosecutor must simply step outside of their role in order to avert such an outcome.

While I argue that Dworkin's view of the judicial role does not include enough moral consideration, and therefore requires judges to step outside of their roles on occasion in order to promote justice, I also assert that a significant degree of racial bias is necessary to justify role infidelity. Kimberley Brownlee's position contradicts this assertion, as she holds that so long as it is conceptually possible for an official to promote a more just outcome by following their own judgment than by following their role, that official, whether they are a Supreme Court justice, a prison guard, or anyone in-between, ought to engage in role infidelity. My position does not go as far as Brownlee's because I believe that there must be a baseline of respect for authority and for the boundaries of one's role, which respect requires that an official accept the possibility that the outcomes in some cases will not perfectly reflect justice. Therefore, under my view, one presumptively ought to maintain their role and may only step outside of their role when the likelihood and extent of racial bias in a given case reaches a certain threshold. I further demonstrate that my position is consistent with respect for authority by examining and responding to John Rawls' and Scott Shapiro's work about rules and authority, respectively.

Rawls holds that if one wants to make use of the value of a practice, one must respect the rules which define the practice, although if one is willing to abandon the value of the practice, one can simply step out of the practice entirely. This means that from a Rawlsian perspective, the

fact that the judicial and prosecutorial roles have boundaries does not in itself prevent judges and prosecutors from engaging in role infidelity; rather, it means that any judges or prosecutors who engage in role infidelity are not acting as criminal justice officials, but as vigilantes. I have no problem with my position being characterized as encouraging vigilantism, because I have shown that the type of role infidelity which I encourage is justified, whether it is vigilantism or not. However, I do not believe that "vigilantism" is an entirely accurate description of what I am recommending, because it seems that judges and prosecutors engaging in role infidelity would still be acting as representatives of the criminal justice system in a way in which true vigilantes never do. For example, if a judge were to engage in role infidelity in order to prevent a defendant from being imprisoned under a racially biased statute, it would seem inaccurate to assert that the judge's action is morally equivalent to breaking the defendant out of prison and that the only differences between the two scenarios are superficial. The distinction between the type of role infidelity which I suggest and vigilantism is not a trivial one, because the further an individual must step outside of their official role when they engage in role infidelity, the less their actions can change the boundaries of the official role for the better; if a given act of role infidelity is understood as pushing the boundaries of the judicial or prosecutorial role, the role may change over time to incorporate that action, but if the same act of role infidelity is understood as the aberrant behavior of a vigilante, no such evolution of the official role is possible.

Even if departure from one's official role does not necessarily equal vigilantism, Rawls and Shapiro could hold, there remains the problem that if I encourage anti-racist role infidelity in order to counteract racial bias in the criminal justice system, I ought also encourage role infidelity geared to counteract sexism, homophobia, socioeconomic and religious discrimination, etc. Rawls and Shapiro could assert that even if anti-racist role infidelity alone would not result

in coordination problems within the criminal justice system, the additive effects of all of these forms of role infidelity would lead to such problems. My response is that while I do not know enough about the legal history of the United States to say whether the above-listed forms of role infidelity would be warranted, I do know that anti-racist role infidelity is warranted, which is sufficient to support my position. Furthermore, if all of the various types of role infidelity described above are indeed warranted, this would not necessarily lead to coordination problems and disarray within the criminal justice system, because, as I suggest earlier in this thesis, role infidelity could lead to criminal justice officials coordinating their actions around more egalitarian processes and goals than they currently do.

Near the beginning of the thesis, I lay out empirical evidence of the tragic extent to which racial bias plagues the United States criminal justice system. I do not believe that, if implemented, the type of anti-racist role infidelity which I recommend would eliminate this racial bias, but I do believe that such role infidelity would help to combat racial bias in the criminal justice system, which, from my point of view, makes it worthwhile.

**Bibliography**

10 Facts About Rodney Reed's Case You Need to Know. (2019, October 11). *Innocence Project*. Retrieved from https://www.innocenceproject.org/10-facts-you-need-to-know-about-rodney-reed-who-is-scheduled-for-execution-on-november-20/.

2018 Homicide in the United States - Expanded Homicide Data Table 6. (n.d.). Retrieved from https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/tables/expanded-homicide-data-table-6.xls.

ACLU Death Penalty Campaign - Race and the Death Penalty. (n.d.). Retrieved from https://www.prisonpolicy.org/scans/aclu_dp_factsheet4.pdf.

Bingamon, B. (2019, November 15). Rodney Reed's Appeals: Key Players. *The Austin Chronicle*. Retrieved from https://www.austinchronicle.com/news/2019-11-15/rodney-reeds-appeals-key-players/.

Brownlee, Kimberley. "Legal Obligation as a Duty of Deference." *Law and Philosophy*, 27, no. 6 (November 2008): 583–97. https://doi.org/10.1007/s10982-008-9028-8.

Brownlee, Kimberley. "Responsibilities of Criminal Justice Officials." *Journal of Applied Philosophy*, 27, no. 2 (2010): 123–39. https://doi.org/10.1111/j.1468-5930.2010.00490.x.

Bureau of Prison Statistics - Inmate Race. (2020, December 5). Retrieved December 10, 2020, from https://www.bop.gov/about/statistics/statistics_inmate_race.jsp.

Coleman, Jules L. "Negative and Positive Positivism." *The Journal of Legal Studies*, 11, no. 1 (January 1982): 139–64. https://doi.org/10.1086/467696.

Cover, R. M. (1975). *Justice Accused: Antislavery and the Judicial Process*. New Haven, CT: Yale University Press.

Dworkin, R. M. (1975). The Law of the Slave Catchers. *Times Literary Supplement*.

Dworkin, R. M. (1986). *Law's Empire*. Cambridge, MA: Belknap Press of Harvard University Press.

Eberhardt, J. L. (2020). *Biased: Uncovering the Hidden Prejudice that Shapes What We See, Think, and Do*. New York, NY: Penguin Books.

*Examining the Demographic Compositions of U.S. Circuit and District Courts* (Rep.). (2020, February 13). Retrieved from https://www.americanprogress.org/issues/courts/reports/2020/02/13/480112/examining-demographic-compositions-u-s-circuit-district-courts/.

Graham, D. A. (2015, July 7). Most States Elect No Black Prosecutors. *The Atlantic*. Retrieved from

https://www.theatlantic.com/politics/archive/2015/07/american-prosecutors-are-incredible-white does-it-matter/397847/.

*In State with 27% Black Population, All Alabama Appellate Judges Are White* (Rep.). (2020, March 11). Retrieved from https://eji.org/news/all-alabama-appellate-judges-are-white-in-state-with-27-percent-black-population/.

Karakatsanis, A. (2019). *Usual Cruelty: The Complicity of Lawyers in the Criminal Injustice System*. New York, NY: The New Press.

Kramer, Matthew H. "Requirements, Reasons, and Raz: Legal Positivism and Legal Duties." *Ethics* 109, no. 2 (January 1999): 375–407. https://doi.org/10.1086/233899.

Leiter, Brian. "Legal Realism and Legal Positivism Reconsidered." *Ethics*, 111, no. 2 (January 2001): 278–301. https://doi.org/10.1086/233474.

McIntyre, A. (2019, Spring). Doctrine of Double Effect. In 1120059383 845662788 E. N. Zalta (Ed.), *Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/double-effect/.

Neily, C. (2019, October 14). There Are Way Too Many Prosecutors in the Federal Judiciary. *Slate*. Retrieved from https://slate.com/news-and-politics/2019/10/too-many-prosecutors-federal-judiciary.html.

Nelkin, Dana K., "Moral Luck," *The Stanford Encyclopedia of Philosophy*, 2019, Edward N. Zalta (ed.), https://plato.stanford.edu/archives/sum2019/entries/moral-luck/.

Racial Demographics. (n.d.). Retrieved October, 2020, from https://deathpenaltyinfo.org/death-row/overview/demographics.

Rawls, John. Two Concepts of Rules. *The Philosophical Review*, 64, no. 1 (1955): 3–32. https://doi.org/10.2307/2182230.

Raz, Joseph. "Authority and Justification." *Philosophy and Public Affairs*, 14, no. 1 (1985): 3–29. https://www.jstor.org/stable/2265235.

Raz, Joseph. *The Morality of Freedom*. Oxford University Press, 1988.

*Report to the United Nations on Racial Disparities in the U.S. Criminal Justice System* (Rep.). (2018, April 19). Retrieved from https://www.sentencingproject.org/publications/un-report-on-racial-disparities/.

Shapiro, Scott J., (2000). Authority. *Jacob Burns Institute for Advanced Legal Studies.* Available at http://dx.doi.org/10.2139/ssrn.233830.

Shelby, T. (2013). Racial Realities and Corrective Justice: A Reply to Charles Mills. *Critical Philosophy of Race, 1*(2), 145-162. doi:10.5325/critphilrace.1.2.0145

Soper, Philip. *A Theory of Law*. Cambridge, MA: Harvard University Press, 1984.

Spooner, L. (1860). *The Unconstitutionality of Slavery* [Burt Franklin Research and Source Work Series, no. 92, 1965]. Boston, MA.