

# The Norm Shift Theory of Punishment\*

*Gideon Yaffe*

---

The philosophy of punishment's focus on the question of justification has left the question of definition neglected. This article explains why there is a need for necessary and sufficient conditions for punishment and offers a new account. Under the theory proposed, to inflict a punishment is to make fewer things permissible for another to do. Since not every such restriction is punishment, an account is offered of the additional conditions needing to be met. One implication of the resulting theory is that some prominent cases in which the question of definition needed to be answered were wrongly decided.

---

## I. INTRODUCTION

For some time, the philosophical literature about punishment has been focused, primarily, on the question of justification. What, if anything, makes it permissible, or even obligatory, for the state, or some other actor, to punish? Advocates of the familiar positions in the philosophy of punishment—retributivism, expressivism, deterrence theory, rehabilitationism—say why it is permissible, if it is, for some people, sometimes, to punish other people. They do this often by saying what it is that makes punishing a positively good thing (e.g., it gives people what they deserve, or cuts down on crime, or expresses shared condemnation of wrongdoing, or makes offenders better). Sometimes justificatory efforts proceed through appeal not to the good but to the right; they proceed by saying how punishing conforms

\* I am grateful to Doug Husak and Adam Kolber for their very careful comments on multiple drafts of this article. Thanks also to Kristen Bell, Michael Bratman, Issa Kohler-Hausmann, Daniel Markovits, Judith Resnik, and Scott Shapiro for their sustained engagement with these ideas and invaluable comments on drafts. Audiences at Fordham Law School and Yale Law School helped to improve the ideas presented here. And I benefitted enormously from additional conversations about this material with Bruce Ackerman, Stephen Darwall, Debbie Denno, Samantha Godwin, Shelly Kagan, Arthur Lau, Youngjae Lee, Tracey Meares, Sam Preston, Mike Wishnie, John Witt, Taisu Zhang, and Ben Zipursky.

*Ethics*, volume 132, number 2, January 2022.

© 2022 The University of Chicago. All rights reserved. Published by The University of Chicago Press. <https://doi.org/10.1086/716874>

to moral principles dictating how we are allowed, or required, to live (e.g., perhaps we may or ought to shun wrongdoing, or stand with victims). Even abolitionists about state punishment have been focused on the question of justification, although, unlike the other theorists, they believe that it has no acceptable answer.

The focus on the question of justification has left another question neglected: the question of definition. What are the necessary and sufficient conditions under which what one person does to another rightly bears the label “punishment”? When we focus on the variety of ways in which government imposes unwanted burdens on citizens, we find that an answer to the question of definition is badly needed. It is frequently the case that the conditions that must be met for it to be permissible for the government to do something to someone vary with the punitive status of that which it wishes to do. The government cannot permissibly punish the excused, for instance, but can permissibly quarantine them. The reason: quarantine is not punishment, and excuses shield only from punitive treatment. So, it’s no argument against quarantine that it’s not your fault you are sick, whereas if quarantine were punishment, this would be a good argument against imposing it. An answer to the question of definition, then, is essential to those tasked with determining whether the government has the right to impose a particular unwanted burden on a citizen.

Answering the neglected question of definition is the primary aim of this article. What is proposed is that a punishment is a distinctive limitation not on what another person can do but on what another person is allowed to do. To punish is to inflict a normative change. As we will see, what is especially in need of explanation is what is distinctive about those limitations on what a person can do that are properly labeled as punishments. In what follows, this simple idea is developed and its wide-ranging implications are explained.

## II. THE IMPORTANCE OF BEING PUNISHMENT

### A. *The Definitional Stop and the Punitive Restrictions*

The question of definition was not always neglected. Before H. L. A. Hart’s famous essay, “Prologomenon to the Principles of Punishment,”<sup>1</sup> theorists frequently preceded their answer to the question of justification with an account of the necessary and sufficient conditions under which what one person does to another is a punishment. After all, the thought was, if you want to know whether and why it’s permissible for people to punish each other, you first need to know what treatments of people by each other you are inquiring about.

1. H. L. A. Hart, “Prologomenon to the Principles of Punishment,” in *Punishment and Responsibility: Essays in the Philosophy of Law* (Oxford: Oxford University Press, 2008), 1–27.

So, we have this from Thomas Hobbes: “A PUNISHMENT, is an Evil inflicted by publique Authority, on him that hath done, or omitted that which is Judged by the same Authority to be a Transgression of the Law; to the end that the will of men may thereby the better be disposed to obedience.”<sup>2</sup> And, a few hundred years later, we have this from Anthony Flew: “I *propose* . . . that we take as parts of the meaning of ‘punishment’, in the primary sense, at least five elements. *First*, it must be an evil, an unpleasantness, to the victim. . . . *Second*, it must (at least be supposed to) be for an offence. . . . *Third*, it must (at least be supposed to) be of the offender. . . . *Fourth*, it must be the work of personal agencies. . . . *Fifth*, in a standard case punishment has to (be at least supposed to) be imposed by virtue of some special authority, conferred through or by the institutions against the laws or rules of which the offence has been committed.”<sup>3</sup> These definitions exclude a great deal. Notably, treatments of those who are known not to have committed any past offense are not punishments under either definition. Nor are private acts of vengeance, since they are not done with even the color of any form of authorization, such as public authorization. And for both Hobbes and Flew these exclusions are put to work. No one will deny that “punishment” of the innocent and private acts of vengeance against the guilty are both impermissible. But one could explain why punishments are permissible without also explaining why various nonpunishments, such as these, are impermissible. A theory of punishment’s justification, that is, does not aspire to be a full moral theory, allowing classification of any act as permissible or impermissible. We would not demand that Hobbes or Flew explain why we are not allowed to steal other people’s things; stealing is not punishment, and so its moral features are not among those that a theory of punishment’s justification need explain. And, similarly, one might think, it is too much to demand that Hobbes or Flew explain why we are not permitted to “punish” the innocent or engage in private acts of vengeance. Since these are not punishments, the thought is, their impermissibility is not something that an answer to the question of justification need explain.

Hart’s discussion of what he calls “the definitional stop,” however, seemed to many to show there to be something fishy about the work that Hobbes and Flew, and others, hope to do with their answers to the question of definition. Whether you call it “punishment” or not, if something that one person does to another shares all the features of punishments that are appealed to in justifying them and, like punishment, lacks features that would undermine that justification, then it, too, should be justified for the same reason that punishments are. If it is not, in fact, justified, then

2. Thomas Hobbes, *Leviathan*, ed. Richard Tuck (Cambridge: Cambridge University Press, 1996), 2.28.

3. Antony Flew, “The Justification of Punishment,” *Journal of Philosophy* 29 (1954): 291–307, 293–94.

that shows that undisclosed moral work is being done by the features of it that disqualify it from the label “punishment,” that is, it shows that the proposed answer to the question of justification is incomplete. Say punishments are claimed to be justified because they “dispose men to obedience.” It follows that if private acts of vengeance similarly dispose men, then they, too, are justified. Since they are not, the advocate of this Hobbesian answer to the question of justification must explain this. The explanation must identify a role that public authorization plays in justifying punishment, since that’s the feature that distinguishes punishments from private acts of vengeance. To note that private acts of vengeance are not punishments does no argumentative work.

One lesson one might take is that if you have an adequate answer to the question of justification, you do not need an answer to the question of definition. Imagine that all those things that we do to others that have qualities XYZ are, thanks to having these qualities, morally permissible. If the justified punishments have these qualities, then we have explained why they are permissible, and so we have answered the question of justification. This is true even if there is at most imperfect overlap between those treatments of others that are XYZ and those treatments that are punishments.

From this viewpoint, the question of definition is the plaything of those who care about the boundaries of ordinary concepts even when they are morally irrelevant. Insofar as the primary question for theorists is the question of justification, then, it is no surprise that philosophers, in the wake of Hart, mostly stopped trying to identify the necessary and sufficient conditions of punishment.

This, however, is a mistake. The question of definition matters politically and morally independently of the question of justification. State punishments are subject to a variety of restrictions that do not apply to the state imposition of various nonpunitive burdens. For this reason, the question of definition is sometimes fiercely litigated. Consider an example: forced sex offender registration.

In 1994, the US Congress passed “Megan’s Law.” The law required each state to pass its own law under which those who had been convicted of sex crimes were required to register with local law enforcement and to provide their addresses. States were then required to make that information accessible to the public through paper fliers or, later, through the internet. A 2006 update, “The Adam Walsh Act,” specified additional requirements for the registries. For instance, it required states’ registration laws to require sex offenders to appear in person to register and to be fingerprinted.<sup>4</sup>

In 2006, Alaska’s “Megan’s Law” was challenged by a man who had been convicted of a sex crime in 1984, imprisoned for six years, and released in 1990, years before the passage of either the federal Megan’s

4. The current law is 34 U.S.C.A. § 2090.

Law or Alaska's law, which was passed to comply with the federal mandate. The case, *Smith v. Doe* (538 U.S. 84), came before the Supreme Court. Doe (as he was called for the purposes of the case) challenged Alaska's law on ex post facto grounds. The US Constitution bars ex post facto laws: laws imposing a burden today on people for behavior that preceded the passage of the law. A line of cases established that the burdens in question had to be punishments for this exclusion to apply.<sup>5</sup> The Constitution does not bar, for instance, laws that impose taxes on income earned prior to the law's passage, nor does it bar laws that impose quarantine on people who contracted contagious illnesses prior to the law's passage. The ex post facto clauses of the Constitution, that is, only bar inflicting punishments unthreatened at the time of the offense for which they are issued. The question, then, before the court in *Smith v. Doe* was whether forced sex offender registration is punishment. To decide the case in a principled way, then, the court needed an answer to the question of definition.

It is possible to view the case in a way that bypasses the question of definition. We might just ask whether it is morally justified to force sex offenders to register without ever engaging with the question of whether forced registration is punishment. Many factors, including but not limited to the inefficacy of registration in reducing sexual violence and the inevitable expression of unchecked community condemnation and stigma that registration involves, are relevant to the ethical question. But it is important to see that the court does not—and, arguably, should not, given its institutional role in comparison to that of the legislature—engage with Doe's complaint in this way. Instead, the issue before the court concerned neither the moral justification of forced registration nor even the justification of punishment as a practice. Say abolitionists are right: we are unjustified in adopting a system of punishment and virtually always unjustified in punishing in specific cases. Still, it would be worse for a state that has unjustifiably adopted a system of punishment, and is contemplating what would be an unjustified punishment in a particular instance, to draw no distinction between ex post facto punishments and those that are threatened at the time of the offense. The ex post facto restriction, that is, is weighty even if forced registration is independently an ethical travesty, and it is weighty whether or not punishment is justified. The question of whether this weighty restriction burdens the government, which just is the question of definition, is important independently of the question of justification.

The point applies not just to the ex post facto restriction but to a wide variety of distinct restrictions on government action that apply if what the government is contemplating is punishment and frequently

5. *Johannessen v. United States*, 225 U.S. 227, 32 S. Ct. 613, 56 L. Ed. 1066 (1912). (“[The Ex Post Facto Clause’s] prohibition is confined to laws respecting criminal punishments, and has no relation to retrospective legislation of any other description.”)

do not apply if the government has something nonpunitive it would like to do. Consider, for instance, the principle of backward-looking proportionality according to which what we do to someone in light of a past act must be proportional to the severity of wrongdoing involved in it. Say that the six years Doe served for his crime was perfectly proportional to the wrongdoing involved in the offense, not a day too much or too little. Then, we would need to know whether forced registration is punishment to determine if the government was in conformity with the proportionality principle in forcing Doe to register. If forced registration is punishment, then the government cannot permissibly impose it on Doe for his offense, for the total punishment—incarceration plus forced registration—would thereby be disproportionately harsh; if not, then there is no reason to deny that Doe was given a proportional punishment since the only thing that enters into the calculation is the prison term he served.<sup>6</sup>

In fact, there are a suite of restrictions with which the government must comply if what it aims to impose is a punitive burden and which it might permissibly violate if the burden in question is not a punishment. The government cannot permissibly punish excused behavior, although it can do other unwanted things to people in light of their excused behavior: consider again quarantine of someone who got sick through no fault of her own. The government cannot punish disproportionately not just to wrongdoing, as just noted, but also to culpability for the past offense. The government cannot place someone in jeopardy of a punishment for an offense when it has previously placed that person in jeopardy of punishment for that very offense; by contrast, it can place someone in jeopardy of unwanted nonpunitive treatment after having previously placed her in jeopardy of punishment. The government cannot punish someone for status or for any state or event that does not include a voluntary act or omission. The government cannot punish for future behavior, even if certain

6. Proportionality turns not just on the definition of punishment but also on its severity. Whether Doe's prison term was proportional to his offense depends on how severe a punishment that prison term was (as well as how severe an offense he committed). Debates about how punishment severity is to be measured, including debates about the relevance and weight in those calculations of the subjective experience of the punished, are distinct from debates about the question of definition. Consider an analogy: One might think that the severity of a case of the flu is in part a function of patient age—the older the patient, the more severe the case. But one need not think that patient age is even part of any necessary or sufficient condition for having the flu. Similarly, answers to the question of definition do not constrain accounts of punishment severity. Rather, they set a limit on when such measures should be used in assessing the severity of that which the government does to someone; such measures are to be used only when what the government inflicts is punishment. Discussion of punishment severity has received useful discussion in recent years. As a start, see Adam Kolber, "The Subjective Experience of Punishment," *Columbia Law Review* 109 (2009): 182–236.

beyond a reasonable doubt that the person will engage in that behavior. And so on; this list is not exhaustive.

Call this suite of necessary conditions for the permissibility of government punishment—restrictions that do not, as a set, restrain other government impositions of unwanted burdens—“the punitive restrictions.” The punitive restrictions include the ex post facto restriction, the principles of proportionality to culpability and wrongdoing, the principle of excuse, double jeopardy, and so on. The punitive restrictions apply to the government’s act of imposing burden P on person D in light of C if P is a punishment of D for C. P’s punitive status is sufficient for the applicability of the punitive restrictions. In addition, however, the punitive status of P is often not just sufficient for the applicability of particular punitive restrictions but necessary. The constitutional ex post facto restriction applies, for instance, only if P is a punishment. It is true that some of the punitive restrictions also apply to government efforts to impose nonpunitive burdens on people. There are proportionality constraints on the imposition of tort damages, for instance. Still, it is frequently the case that a particular restriction applies if and only if the burden the government aims to impose is a punishment. And the full set of punitive restrictions apply if and only if P is a punishment.

The question of definition, then, is pressing. The government often has incentive to deny that what it wants to do to D is punishment—that is just what the government denied in *Smith v. Doe*. After all, if what it has in mind is not punishment, then there are at least some punitive restrictions with which it need not abide, and so it will often be allowed to inflict P on D despite the fact that it would thereby violate those restrictions. In fact, the majority in *Smith v. Doe* ruled that forced registration is not punishment and so ruled that the government need not abide by the ex post facto restriction in forcing registration.<sup>7</sup> The question is this: when is the government allowed to bypass the punitive restrictions, and when must it abide by them? This question often reduces to the question of definition: it must abide by the punitive restrictions just in case what it hopes to do is punish.

Why do the punitive restrictions, as a set, apply to punishing but not to other invasive and unpleasant things the government wants to do to people? The discussion here will not reach this question—what matters for our purposes is that they in fact apply only if what is contemplated is punishment, and not why they do. However, it is worth noting that there is no reason to expect a univocal answer to the question. Some restrictions might attach to punishment for some reasons, others for others. It seems likely, for instance, that the principle barring punishment for

7. State courts, in response to this federal ruling, established that in Alaska forced registration is punishment and so is subject to the ex post facto clause of the Alaska Constitution. See *Doe v. State*, 189 P.3d 999 (Alaska 2008).

excused conduct has its source in something about the act of holding another responsible, a genus of action of which punishing is a species. Holding another responsible is not the kind of thing that one can do permissibly unless one is sensitive to the possibility of excuse. Perhaps the proportionality principles are similar. By contrast, the ex post facto restriction probably attaches to punishing for reasons that are specific to the grounds for government legitimacy; a government that punished without abiding by this restriction acts wrongly, even if nongovernmental punishers would not. Or, at least, we should be open to this possibility. It is also possible that some punitive restrictions make sense because of contingent features of institutional design. Perhaps a double jeopardy rule makes sense because we have a system in which the government, rather than a private party, initiates the proceedings when punishment is at stake. If the government could initiate proceedings again when it fails to get the verdict it wants, this would result in abuse of power. Design a system in which the government never initiates such proceedings, and perhaps the double jeopardy restriction falls away. We should be open, anyway, to the possibility that this kind of rationale can be given for some punitive restrictions, although not for others.

To be open to pluralism about the rationales for the punitive restrictions—each might attach to punishment for a different reason—is to be open, correlatively, to the possibility that the normative pressures faced by would-be punishers are not of a single type. The problem of definition is important because those who aim to inflict something on another that is, in fact, a punishment are under normative pressure to conform to the punitive restrictions. But what is the nature of this normative pressure? Is it moral, legal, prudential, conceptual, logical, or metaphysical (or some combination thereof)? To be open to pluralism about the rationales for the punitive restrictions is to resist the call to answer this question. Perhaps to punish the excused is to commit a moral wrong. But this is consistent with the possibility that the pressures to conform to the ex post facto or the double jeopardy restrictions, or any other restrictions, are of a different sort entirely. Conformity with the punitive restrictions as a whole, then, is not rightly characterized as a demand of morality, or of legality, or of political legitimacy, or of prudence, or of any other single type. We should be open to this possibility. But it should not hamstring our efforts to answer the question of definition.

#### *B. Punishment-by-Intention versus Punishment-by-Nature*

The majority in *Smith v. Doe* uses an answer to the question of definition that was offered in earlier cases. Writing for the majority, Justice Kennedy explained that answer: “We must ascertain whether the legislature meant the statute to establish ‘civil’ proceedings. If the intention of the legislature was to impose punishment, that ends the inquiry. If, however, the

intention was to enact a regulatory scheme that is civil and nonpunitive, we must further examine whether the statutory scheme is so punitive either in purpose or effect as to negate the State's intention to deem it 'civil.'<sup>8</sup> So, Kennedy is identifying two overlapping but distinct sufficient conditions for punishment, what we might call "punishment-by-intention" and "punishment-by-nature." First, if the government intends that P should bear the label of punishment, then it does bear that label. The government can disavow that P is a punishment-by-intention by simply forming the intention to deny it the label of "punishment." The majority in *Smith v. Doe* held that forced registration is not a punishment-by-intention: the legislature did not intend for the label to attach.

This leaves the second category of punishment-by-nature. These are treatments that are "so punitive in purpose or effect" that the label attaches, despite the government's desire for it not to. The creation of this category by the courts serves to establish judicial oversight over the question of whether the government must abide by the punitive restrictions. The government has limited discretion over the question of whether what it wants to do is punishment. If it oversteps its bounds by refusing the label to some form of treatment that is too "punitive in purpose or effect," then the judiciary will force it to comply with the punitive restrictions. By stipulating the existence of the category of punishment-by-nature, the judiciary is freed from simply having to defer to the legislature's judgment as to whether to comply or bypass the punitive restrictions.

But although the desire for judicial oversight on this matter is laudable, we should also hope that the judiciary's decision on such a matter would fall out of the application of some independently specifiable criterion for punishment. That is, what we would like is for the judiciary to force the government to comply with the punitive restrictions thanks to the fact that what the government is contemplating doing is actually punishment, despite its desire not to label it that way. Held to this standard, what the court supplies is clearly inadequate. After all, to answer the question of whether P is too "punitive in purpose or effect," we must first determine if it is at all punitive in purpose or effect. And to determine that, we need to know when intended or actual effects are punishments, rather than nonpunishments. So, while the court reserves the right to label a treatment punishment, it provides us with no guidance about when that label is correctly applied, or when incorrectly. It does not answer the question of the necessary and sufficient conditions for punishment-by-nature but offers a circular criterion instead.

Notice, also, that the other category of punishment, punishment-by-intention, is no more illuminatingly drawn. To say that from the judiciary's point of view a government-imposed burden will be treated as punishment

8. *Smith v. Doe*, 538 U.S. 84 (2003) at 1146.

if it is intended by the legislature to bear that label is not to say that the legislature cannot mistakenly attach it. Say the legislature announces that it will bestow a birthday gift of \$100 on every citizen and that these gifts will be punishments for having been born. A fiscally conscious citizen objects: he was fully excused for having been born, and since the government cannot permissibly punish excused behavior, he argues, it cannot give him a gift on his birthday. Since the gifts are punishments-by-intention, this argument ought to succeed, as a matter of law. The government cannot label something punishment and then refuse to abide by the punitive restrictions. But what seems odd about this is that the government's contemplated behavior simply is not punishment, whether or not it wants to label it as such, and so it seems peculiar to require the government to abide by the punitive restrictions. There may be treatments the government can transform into punishments simply by intending they bear the label, but \$100 birthday gifts do not seem, as an intuitive matter, to be apt for such a transformation.

It is possible that when the government intends that something be a punishment which is not a punishment-by-nature the punitive restrictions apply only because of a separate principle requiring coherence of government behavior. The problem, we might think, with giving the \$100 to those who are excused for being born is merely that the government cannot both intend to punish and ignore excuses. It could avoid incoherence in such a case either by granting excuses or by just withdrawing the label of "punishment" from what it wants to do; either route to compliance is equally good. By contrast, when it wants to inflict a punishment-by-nature, its only route to permissibility is through compliance with the punitive restrictions. For this reason, punishment-by-nature is the more interesting category. It places constraints on government action that the government cannot bypass at will. In light of this, my focus from here forward will be on punishment-by-nature. The court is onto something in identifying this category, even if its description of the category's boundaries is circular and so unilluminating. What it is onto is the idea that there are some things that just are punishments even if the government does not intend the label to attach. There are some things such that, if you are doing them to someone else, you are punishing, even if you do not think you are and do not intend to be. When that's the case, the punitive restrictions apply and must be respected. The hard philosophical problem is to specify the necessary and sufficient conditions for membership in this class. The goal in the remainder of this article is to solve this problem.

### *C. Setting Aside the Legalistic Approach*

To solve the problem, it's important to set aside a certain attractive, legalistic way of thinking about it, one that runs directly counter, as it happens,

to the perspective adopted by the Supreme Court in *Smith v. Doe*. According to this legalistic approach, what punishment is is simply a matter of positive law. If the law says that the government cannot punish *ex post facto*, for instance, then the government obeys the law when it inflicts P on D in violation of this rule if P is a punishment according to the law. So, the only relevant question is what the law says a punishment is. From this point of view, the question of definition has to be tackled in the usual way that questions about legal definitions have to be tackled: through examination of legally authoritative texts and interpretations. To be sure, some legal systems, maybe even America's, will simply incorporate an ordinary usage of the term into the law. In such systems, we will need to investigate the necessary and sufficient conditions of the relevant extralegal concept. But in other systems, according to this legalistic perspective, we will not need to do that. It's all just a matter of what the law says.

The problem with this legalistic approach is that there are good reasons for thinking that punishment is a concept that resists arbitrary legal stipulation. The law can stipulate definitions of many things without much restriction. It can stipulate the definition of "deadly weapon," or "corporation," for instance. But there are some concepts that are so intertwined with government legitimacy that the government is necessarily constrained in its behavior by the extralegal boundaries of the concept. For instance, imagine that the law guarantees equality in voting rights. The government cannot, then, consistent with the law, deny voting rights to one of two relevantly identical people on the grounds that under the law "equality" is defined in such a way that the denial is consistent with it. The government cannot bootstrap its way into compliance with the law by stipulating a legal definition of "equality" under which two people are being treated "equally" when one is arbitrarily denied a right granted to the other.

So while a term, for instance, like "corporation" probably just means what the law says that it means, other legal terms are necessarily extralegally defined. It is a difficult problem in analytic jurisprudence and in political philosophy to specify which concepts do not allow for stipulative definition. I do not know what makes it the case that a concept used by the law must have an extralegally defined boundary. Still, even if we cannot specify the right criterion, we can spot some very plausible examples. Concepts like "equality," "person," "right," and "liberty" all seem like plausible contenders. For the same reason, "punishment" is a plausible contender. Its definition is necessarily intertwined with these others. The legalistic way of viewing the problem of definition overlooks this. In fact, plausibly, although not definitively, this is what the Supreme Court recognized in *Smith v. Doe*. The legislature, of course, gets to decide what the law is. But the legislature does not get to decide, without limit, what punishment is. When the law uses that concept, it necessarily uses a concept with an extralegal boundary. The question is what that boundary is.

It is important to see that, setting aside the legalistic way of thinking about the question of definition, it is not easy to answer. First answers that might occur to one are easily shown to be inadequate. For instance, one might think that what makes P a punishment-by-nature is that it is over some threshold of severity. Perhaps if what the government wants to do to D is sufficiently bad, that is, then it has in mind a punishment, whatever the government's opinion or intention. But this can't be right: there are both harsh treatments that are not punishments and minor inconveniences that are. In the first category, think of extended quarantine, or loss of livelihood due to a failure to qualify for a professional license. In the second category, of minor inconveniences that are punishments despite what the government might like them to be, imagine the government refusing to call a single hour of imprisonment, issued in response to a finding of guilt for a criminal offense, a punishment. Nor must P be a punishment if it has some feature that seems important to political authority, such as that it involves a deprivation of liberty. Civil commitment is deprivation of liberty, but it is not punishment; the same can be said for restraining orders and certain injunctions.

But as hard as the problem is, we will see next that it is solvable.

### III. THE NORM SHIFT THEORY

The aim of this section is to explain and defend the following answer to the question of definition:

*The Norm Shift Theory of Punishment:* P is a punishment-by-nature if and only if (1) P is a required meaningful reduction of a person D's normative freedom and (2) P occurs thanks to government S's exercise of its punitive power.

Although this theory draws on two technical notions, to be explicated in what follows—the notions of “a required meaningful reduction in normative freedom” and “a punitive power”—the central idea motivating it is simply stated: to punish is to restrict, in a very specific way, what it is permissible for another person to do by exercising a distinctive normative power. This section's central aim is to remove the darkness surrounding this theory by specifying what specific form of restriction is involved in punishment, and what is distinctive about the normative power that the punisher exercises in inflicting it. The next section goes on to explain some of the theory's implications, including why the government cannot deny the label of punishment to that which counts as such under the Norm Shift Theory.

#### A. *Reductions in Normative Freedom*

It is hardly news that there is a difference between “can” and “may” and between “can't” and “mustn't.” Ability and permissibility, inability and

impermissibility, are distinct. There is a corresponding difference between two kinds of freedom:

*Metaphysical Freedom:* D has the metaphysical freedom to A if and only if D has the ability and opportunity to A.

*Normative Freedom:* D has the normative freedom to A if and only if it is permissible for D to A.

One person has greater metaphysical freedom than another just in case there are more things she can do. But one person has greater normative freedom than another just in case there are more things that it is permissible for her to do.

Normative freedom and metaphysical freedom are, obviously, distinct. The able-bodied man sentenced to a prison term has less normative freedom than the paralyzed and bedridden man who has not been sentenced for any crime. It's impermissible for the prisoner to leave the prison, while the paralyzed man is permitted to go where he likes, even if it is not within his power to do so. But the prisoner might very well have greater metaphysical freedom than the other man. He can do jumping jacks, for instance.

Metaphysical freedom varies over time since both our abilities and our opportunities vary over time. But normative freedom also varies over time. Promises are one familiar mechanism through which normative freedom alters. Prior to promising to do something, the agent is normatively free to refrain from doing it. After promising, a formerly permissible form of action becomes impermissible. Consent, too, alters normative freedom. It is impermissible to do something to someone until that person consents, at which point that form of conduct becomes permissible for the recipient of consent. Punishment, of the kind the Norm Shift Theory is identifying, necessarily consists in a reduction in normative freedom. Punishment is to be identified, under the theory, with the change that occurs when the class of actions that it is permissible for a person to perform becomes smaller.

Although those who are punished are typically significantly reduced in metaphysical freedom, there are reasons to think that that is not the aspect of their condition that the state must own as punishment, whether it likes it or not. That is, there are good reasons to think that it is not in virtue of reductions in metaphysical freedom that a person is enduring a punishment-by-nature. Imagine that D1 and D2 commit a robbery together and spend the next five years in prison. But imagine that D1 could have left the prison at will. He could have blackmailed the warden, who would have bowed to the pressure and engineered D1's escape. So D1 enjoyed the metaphysical freedom to leave but chose not to exercise it; he elected, instead, to do his time. D2, by contrast, did not have this metaphysical freedom. There was nothing he could do to arrange his escape. Yet both are

punished and equally so. The state is under the same pressure to abide by the punitive restrictions with respect to D1 as it is with respect to D2; it could not bypass those restrictions through the formation of the intention not to label what it is doing to D1 and to D2 “punishment.” What follows is that the feature of D1’s and D2’s condition in virtue of which the state must comply with the punitive restrictions is not reduction in metaphysical freedom. This motivates the thought that punishment-by-nature consists in a reduction in normative freedom, rather than metaphysical freedom.

According to the Norm Shift Theory, when a person is punished, there is an expansion in what actions of his would be wrongful. The punishment consists in this normative shift. The person who commits a crime and is then incarcerated is punished by the fact that leaving a certain specified place is, for him, legally wrongful. If there were no crime of escape, it would be permissible for the incarcerated to leave prison when they please; incarceration would be kidnapping. What I suggest is that the punishment involved in incarceration is the normative change. Before being sentenced, it was permissible for D to be elsewhere; afterward, it was not permissible for him to be elsewhere. Similarly, the teenager who is grounded is punished by the fact that it is now, for her, wrongful to be away from home. Before the offense, there are a class of actions that it is permissible for the agent to perform; after the offense, some of the former members of that class are no longer included in it—they have become impermissible. The punishment is the deontic change. Therefore, the Norm Shift Theory supports the commonly voiced claim that punishment necessarily involves deprivation. But the theory specifies what exactly a person needs to be deprived of in order to be punished: normative freedom. Deprivations of other sorts—bodily sustenance, access to friends and family, even metaphysical freedom—are not themselves punishments, although they often accompany punishments. One of the burdens of this article is to defend this comparatively narrow view of the kind of deprivation that is involved in punishment.

I use the term “normative freedom” rather than “moral freedom” because there is nothing in the concept that requires that the relevant body of norms are moral norms. In fact, given that the most important forms of punishment existing today are meted out by governments acting under legal authority, what is far more important than moral freedom is legal freedom, which can, and usually does, differ from moral freedom. Consider two examples. D1 is fairly convicted and sentenced to a year in prison for a crime that he did not commit. The verdict is a well-intentioned false positive. D2 is fairly convicted and sentenced to a year in prison for a crime that he did commit, but where the legally prohibited behavior was entirely morally permissible. Perhaps he looked at a white woman. In both of these cases, as well as others, there are now a list of things that it is illegal for D1 or D2 to do that were formerly legal; notably, it is now illegal, for the next

year, for either of them to be anywhere other than within the prison walls. It is legally impermissible for them to leave. But it is not morally impermissible. How could it be? Neither D1 nor D2 did anything morally wrong, nor is there any general moral obligation to obey the law.<sup>9</sup> The only moral wrong is on the part of the punisher. Still, even though there has been no change in moral permissibility for D1 or D2, there has been a change in legal permissibility. Further, and importantly, in both cases there is a reduction in normative freedom even though there is no change in moral freedom—that is, normative freedom in which the relevant norms are moral norms. Put another way, there are as many kinds of normative freedom as there are kinds of norms. If, as seems likely, legal norms are not a species of moral norms, or vice versa, then legal freedom is a different thing from moral freedom, although both are normative freedoms. And, correlatively, there can be reductions in legal freedom—and so, potentially, punishments under the Norm Shift Theory—where there are no changes in moral freedom.

### *B. Meaningful Reductions in Normative Freedom*

It would be absurd to suggest that every reduction in normative freedom is a punishment. Promising is not self-punishment. Further, every time a new piece of criminal legislation is passed, everyone suffers a reduction in normative freedom: something that was formerly legally permissible for them to do becomes legally impermissible. But no citizen suffers punishment simply on the passage of a piece of criminal legislation.

A first feature that distinguishes punishments from other reductions in normative freedom can be illustrated through reflection on the concept of “getting away with a crime.” Say that D3 and D4 commit a robbery together and are sentenced to five years in prison. It thereby becomes impermissible for either to spend those years away from prison. Both suffer a reduction in normative freedom. But imagine that D3 evades the authorities and so never spends a day in prison, while D4 languishes there for five years. This fact—this difference in their compliance with the demand that they stay in prison—results in D3 never being punished, while D4 is punished. But there is no difference in their normative freedom. D3 spent the five years doing something that it was impermissible for him to do, namely, be away from the prison, while D4 spent the time doing the permissible thing. Therefore, D3’s reduction in normative freedom was not a punishment, while D4’s was; D3 got away with it. What distinguishes them is that while both suffered a reduction in normative freedom, only D4’s reduction in normative freedom was meaningful. More carefully:

9. Or, at least, that is the consensus view among philosophers. See, e.g., M. B. E. Smith, “Is There a Prima Facie Obligation to Obey the Law?,” *Yale Law Journal*, 82 (1973): 950–76.

*Meaningful Reduction in Normative Freedom:* D suffers a meaningful reduction in normative freedom if and only if there is an act A such that (i) at time t1 it is permissible for D to A; (ii) at time t2 it is impermissible for D to A; (iii) after t2, D does not A; and (iv) the fact that D does not A after t2 is explained, at least in part, by the fact that it is impermissible for him to A after t2.

Conditions (i) and (ii) suffice for a reduction in normative freedom, but when conditions (iii) and (iv) are met, that reduction is meaningful.

To see the need for condition (iv), consider another pair of examples: Unbeknownst to him, D is secretly convicted of a crime and secretly sentenced to a day of house arrest. Thanks to the sentence, it is impermissible for D to leave his home for a day. As it happens, D spends that day at home with no awareness that he is thereby doing what is required. D has not been punished, and the reason is that he has not suffered a meaningful reduction in normative freedom. But conditions (i), (ii), and (iii) are satisfied—leaving his home was permissible before the secret sentence and impermissible after, and D did not leave his home after. However, condition (iv) is not satisfied. Since D does not know of the sentence (or even of the trial) and stays at home for other reasons, his staying at home is in no way explained by the fact that it is impermissible for him to leave. Condition (iii) suffices for conformity with the shifted norm, but the conjunctions of conditions (iii) and (iv) suffice for compliance with it. For a reduction of one's normative freedom to be meaningful, one must comply with the normative changes, and not just conform to them.<sup>10</sup>

To say that where there is compliance there is conformity explained in part by the impermissibility of the alternative is not necessarily to imply that compliant agents conform for any particular reason, much less thanks to taking the impermissibility of the alternative to provide them with sufficient reason to conform. They might conform out of nothing but fear, or to please a friend, or even involuntarily, and it might still be the case that part of what explains their conformity is the impermissibility of alternative behavior. This much seems to be true: where there is compliance, there is conformity that includes among its causes at least one factor that would not have contributed to conformity as it did had the alternative not been impermissible. Say, for instance, that D spends a year in prison in

10. The example does not specify whether D has the metaphysical freedom to impermissibly leave home. If (1) he does not, and (2) part of the explanation for why he does not is that it is impermissible for him to do so, and (3) part of the explanation for his staying home is that he cannot leave, then his conformity will be explained in part by the impermissibility of nonconformity and will thereby count as an instance of compliance. This is the case when, for instance, a guard is stationed at his door and part of the reason why he stays home is that he knows he cannot evade the guard. In that case, he could be punished (assuming all the other relevant conditions are met) even though he does not know he is being punished.

part because his concerted efforts to escape are thwarted. Why are they thwarted? In part because it is impermissible for him to escape and so the state took steps to ensure that his efforts to do so would fail. In such a case, therefore, there is not just conformity but compliance and so the reduction in normative freedom that D suffers is meaningful.

*C. Required Meaningful Reductions in Normative Freedom, and Civil Liabilities*

Not every meaningful reduction in a person's normative freedom is a punishment. To see this, imagine a system in which prison terms can be served by proxy: if you are sentenced to a prison term, someone can do your time for you. Parents could serve time for their children, or vice versa, for instance. Insurance companies could sell products under which they would provide proxy prisoners for policy holders in the event of conviction. What I will now argue is that such a system is not a punishment system; nobody is punished under it, including those who lack proxies and serve their own time. Since those who serve their own time but are permitted to have a proxy do suffer a meaningful reduction in normative freedom, there is more to punishment than that. The argument for this conclusion requires several steps.

Start with an observation: Many, in our imagined proxy system, serve their sentences without suffering any reduction in normative freedom. Just as it was permissible for them to be away from the prison prior to the sentence, it is permissible for them to be away after, provided that the proxy is there. The sentence makes true a narrow scope normative proposition: If D has no proxy, or has a proxy who is away from the prison, then D is behaving impermissibly in being away from the prison. But when the proxy is in prison, all the same behaviors by D are permissible as they were prior to the sentence. Since a reduction in normative freedom requires a change in this respect, there is no reduction in normative freedom for those with proxies. So, under the Norm Shift Theory, those with proxies are not punished; they "got away with it."

A second observation: In the imagined proxy system, some people serve their own time. They have no proxy to serve it for them. Are they punished, or not? Or, to put the question in the form that matters to us here, is the government subject to the punitive restrictions here, even if the government does not intend that its sentences in this system be punishments? In answering we need to remind ourselves that those who have proxies are not punished, which implies that the government may not be subject to the punitive restrictions when it comes to them. So, there are two possibilities. The first is a univocal conception of the system: none of the sentences in this system are punishments, and so the government is not subject to the punitive restrictions with respect to any of them. Or, alternatively, it is best

conceived as a bifurcated system: the government is subject to the punitive restrictions in imposing sentence on those lacking proxies, but not on those who have them. If this bifurcated conception is correct, then the government cannot, for instance, impose sentence *ex post facto* when the defendant has no proxy, but it can when she does.

There is a decisive reason to favor the univocal conception, given some further assumptions about the legal system. In particular, the univocal conception must be the correct one if we assume that the system is one involving a commitment to equal treatment under the law. To be committed to equal treatment under the law is to be committed at least to this much: any differences in the effect an identical sentence has on two people are not differences intended in the issuing of the sentence (although these differences are, perhaps, foreseen). From the point of view of such a legal system, under the proxy system, the person who has a proxy and the person who does not are treated the same; while there are inevitably differences in how their sentences affect them, those differences are not intended. What this means is that from the point of view of such a legal system, there is no meaningful sense in which the normative world is different for the person who has a proxy in comparison to the person who does not; the normative world has been changed equivalently for the two. But for the punitive restrictions to apply asymmetrically for those who have and lack proxies is for the government to be subject to significantly different restrictions with respect to those who are, from its point of view, treated identically. But to take this position is to embrace contradiction. The person without a proxy has a right that the one with a proxy lacks: a right to be free of such treatment if the treatment violates one of the punitive restrictions. But this is a normative difference between the two, a difference that a government committed to equal treatment under the law cannot consistently recognize. Put simply, to give the same sentence is to treat the two the same; to be subject to the punitive restrictions with respect to one but not the other is to treat them differently. So, a government committed to equal treatment that adopts the proxy system must conceive of it univocally: neither the person with a proxy nor the person who lacks one is punished under such a system.

The conclusion: there is more to punishment than a meaningful reduction in normative freedom. After all, the person who has no proxy nonetheless suffers a meaningful reduction in normative freedom. But as we have just seen, this person is not punished. What feature distinguishes this person from the person who is?

The answer, I suggest, is that when the government punishes it sets a normative standard for the world that can be met only if D suffers a meaningful reduction in normative freedom; any alternative way the world could be would fall short. To punish is not just to cause a meaningful reduction in normative freedom but to require that. The sentences in the proxy system

cause a meaningful reduction in normative freedom for those who lack proxies but do not require that, since the world would not have fallen short, relative to the normative change produced by the sentence, had a proxy served the person's time. More carefully:

*Required Meaningful Reduction in Normative Freedom:* D suffers a required meaningful reduction in normative freedom if and only if  
 (i) D suffers a meaningful reduction in normative freedom and  
 (ii) the act in virtue of which D's normative freedom was reduced at once made it the case that no world in which D did not suffer a meaningful reduction in normative freedom would be as it ought to be.

The need for the concept of a required meaningful reduction in normative freedom derives from the fact that governments, and other actors, often change normative reality by making true narrow scope normative propositions—propositions of the form “if X, then it is impermissible for D to A.” Such acts meaningfully reduce normative freedom for any who face a situation in which X and comply with the resulting demand not to A. But the act of changing normative reality sets a disjunctive standard for how the world ought to be: the world is as it ought to be when either (1) not-X or (2) X but D complies. What we learn from reflection on the proxy system is that this is not the kind of normative change that is essential to punishment-by-nature.

The idea that we do not have a punishment unless the meaningful reduction in normative freedom that a person suffers is required bears an intuitive link to our pretheoretic conception of punishment. In our pretheoretic conception, punishment is primarily a personal burden, and at most secondarily a social one. By its nature, punishment is to be borne by the offender; if society bears the burden also (by, say, paying for prisons), it does so as a means to assuring that the offender bears it. What this implies is that if the state makes a demand that can be fully met without any change in the offender's situation, the state is not issuing a punishment. The notion of a required meaningful reduction in normative freedom captures this aspect of our pretheoretic conception of punishment.

An immediate corollary of this is that civil damage awards, such as tort damages, are not punishments. Tort damages can be paid by anyone, not just the defendant. In fact, given the ubiquity of insurance, most tort damages are not paid by defendants, but instead by insurance companies. The state has no objection to that. The defendant has not failed to do what was demanded by the state when an insurance company, or his benevolent benefactor, pays the plaintiff on his behalf. In declaring the defendant's tort liability, the state does not thereby make it the case that it would be impermissible for D to fail to pay the money; when D fails to pay, but the insurance company pays instead, D has not done anything impermissible.

What the state has done is to make true the following narrow scope normative proposition: if the plaintiff is not paid for her injury, then D has done something impermissible. This change causes some people to suffer a meaningful reduction in normative freedom. It causes this whenever the defendant pays the plaintiff personally. But it does not require such a reduction, as can be seen by the fact that both the defendant who pays and the defendant whose insurance company pays are in identical normative positions, from the law's point of view.

One implication of this is that fines are typically at most punishment-by-intention, and not punishment-by-nature. When a judge fines a defendant, the judge typically makes true a conditional: if nobody pays D's fine for him, then it is impermissible for D not to pay it.<sup>11</sup> This imposes no meaningful reduction in normative freedom on those defendants who are lucky enough to have their fines paid by others. While the sentence does meaningfully reduce the normative freedom of those who pay their own fines, the reduction is not required. This means that it is not a punishment-by-nature, under the Norm Shift Theory. But it could still be a punishment-by-intention: the government might intend that it bear the label of punishment, and so be bound by the punitive restrictions. Public outrage, then, when criminal behavior by corporations is punished by fines paid, indirectly, by shareholders and customers may derive from the sense that such corporations have not really been punished; they have been punished in name only. This thought is accommodated by the Norm Shift Theory. In such cases the corporation has not suffered a punishment-by-nature but at most a punishment-by-intention.

The concept of required meaningful reduction in normative freedom goes some way, but not the whole way, in providing an account of the criminal/civil divide. There are nonpunishments (e.g., civil injunctions) that do indeed involve required meaningful reductions in normative freedom and so must fail to be punishments for another reason. When a school district is required to educate children who live far away, and the district complies, the district does indeed suffer a meaningful reduction in its normative freedom. There is compliance with a shifted norm. Further, this is required. The world is as the court demands only when the school itself complies by bringing the students there and educating them; it could not, say, hire a private school to take them on. But what the school undergoes is not a punishment. The resources for explaining this are developed in the next subsection.

11. Interestingly, courts have sometimes ruled that insurance companies cannot be required to pay a corporation's criminal fines. See, e.g., *Philadelphia Indem. Ins. Co. v. Sabal Ins. Grp., Inc.*, 786 F. App'x 167 (11th Cir. 2019). But this does not rule out the possibility of other parties, besides the corporation, doing so.

*D. The Punitive Power*

As the case of civil injunctions illustrates, not all required meaningful reductions in normative freedom are punishments-by-nature. Contract, also, is replete with examples. Q hires D, a famous artist, to paint his portrait. Prior to the formation of the contract, it was permissible for D to refrain from painting Q's portrait; after, it is not.<sup>12</sup> D has suffered a reduction in normative freedom. Further, when D paints the portrait, fulfilling his end of the bargain, his reduction in normative freedom is meaningful. And, importantly, this meaningful reduction in normative freedom was required: the contract sets a normative standard that is not met when, for instance, D hires another painter to paint the portrait, even if it is just as good or better. This is a case where substitution will not do. But it would be ridiculous to suggest that the contract punishes D. The result: there is more to punishment than a required meaningful reduction in normative freedom.

The reason that the painter has not been punished by the contract is that the normative change that the contract makes—the shift from the permissibility to the impermissibility of refraining from painting the portrait—takes place thanks, in part, to the special entitlements of the parties to the contract to bind themselves in something akin, or perhaps identical, to the way in which people bind themselves through promises. Within certain domains, we are entitled to control what we are required to do. That's not the kind of entitlement that is exercised when, for instance, the state issues a criminal sentence and so makes it impermissible for the offender to be, for a time, anywhere outside the prison walls. The missing necessary condition of punishment, that is, is the distinctive kind of entitlement that figures in the explanation of the normative change in which punishment consists.

To make good on the idea that the distinctive feature of the required meaningful reductions in normative freedom that are punishments is their relation to a special entitlement held by the person who shifts the relevant norms, it helps to draw on the idea of a "normative power," or a power to change the facts about what people are required and permitted to do. Normative powers are powers; they are a form of might. The possessors of them can make changes in the world, albeit to the normative parts of the world. Meaningful reductions in normative freedom are punishments when they are produced through an exercise of a different normative power from that which is exercised in promise, in consent, or in property transactions. They stem from the exercise of a distinctive normative power. This, anyway, is the idea expressed by the Norm Shift Theory.

12. Some contract theorists would deny this, asserting instead that what D is required to do is disjunctive: paint or pay damages. See Daniel Markovits and Alan Schwartz, "The Myth of Efficient Breach: New Defenses of the Expectation Interest," *Virginia Law Review* 97 (2011): 1939–2008. However, even if this is so, the point remains. What is important is that the norm produced by the contract is not narrow scope, even if it is disjunctive in content.

But supporting this idea requires two things: First, it requires at least some steps toward an account of the individuation conditions for normative powers. In virtue of what is the normative power exercised in the making of promises a different normative power from that exercised in consent, or in punishment? Second, it requires a showing to the effect that the normative power exercised in punishment is distinct, given the account of normative power individuation, from other normative powers. And, further, this showing must evade the obvious worry about circularity: it cannot be that the normative power exercised when, for instance, a judge issues a sentence is distinctive in virtue of the fact that what the judge inflicts is a punishment. That would be to distinguish the power from others in a way that rendered it incapable of being used to distinguish punishments from other forms of meaningful reduction in normative freedom. Rather, there must be independent criteria for the relevant power and some reason to think that when a normative power meets those criteria, the change that is produced through its exercise is a punishment.

For convenience, the term “punitive power” will be used to refer to the distinctive power the government exercises when it produces a required meaningful reduction in normative freedom, and when that is a punishment. The burden is to specify in a noncircular way what distinguishes the punitive power from those that are exercised in the production of required meaningful reductions in normative freedom that are not punishments.

I suggest that there are two separate sufficient conditions for normative power distinction, and the punitive power is distinct from others either because one of these conditions is met or because the other is. First, the punitive power is distinct in virtue of the conditions under which it is possessed. Other normative powers are possessed under different conditions. Second, the punitive power is distinct in virtue of the entitlements that accompany it, entitlements that do not accompany the power exercised when, for instance, the government issues a civil injunction. So, I am drawing on the following two principles of normative power individuation:

*Distinction in Possession:* If normative power N1 is possessed only in circumstances X and normative power N2 is possessed only in circumstances Y, and  $X \neq Y$ , then N1 and N2 are distinct.

*Distinction in Entitlement:* If normative power N1 necessarily co-occurs with entitlement E, but normative power N2 does not necessarily co-occur with E, then N1 and N2 are distinct.

I will not be defending either of these principles of normative power individuation here. I believe them to be sufficiently plausible to warrant relying on them. However, if both of these principles turn out to be false,

the Norm Shift Theory might still survive, for there may be some alternative way to distinguish the punitive power from other powers to produce required meaningful reductions in normative freedom.

The punitive power, as will be elaborated in the remainder of this subsection, is a normative power to produce required meaningful reductions in normative freedom, but it is distinct from all other such powers either by its circumstances of possession or by its accompanying entitlements, or both. This account of the punitive power helps the Norm Shift Theory of Punishment to avoid circularity. The question of whether a particular government's normative power is a punitive power does not turn on the question of whether that which is produced through its exercise is a punishment. Rather, the explanation goes in the direction it ought: if the relevant power is possessed under the right circumstances, or accompanied by the right entitlements, then that which is produced through its exercise is a punishment, according to the Norm Shift Theory.

To explain this fully, we need a description of the circumstances in which the punitive power is possessed and a description of the distinctive entitlements that accompany it. Consider the circumstances of its possession first. The punitive power is possessed only where there is a finding (perhaps accurate, perhaps not) that D committed C, where C is wrongful. Say that without finding that D did anything wrong, the government locks D in a prison. Has there been a required meaningful reduction in D's normative freedom? In one set of cases, there has been no reduction in D's normative freedom at all since the government is not exercising a normative power at all. In such a case, the government has not succeeded in making it impermissible for D to leave the prison. It may have succeeded in making it impossible for D to leave—the government has reduced D's metaphysical freedom. But it has not succeeded in producing any normative change. In another set of cases, the government does exercise a normative power—it does make it impermissible for D to leave the prison—but the normative power that it exercises cannot be the punitive power, since it is possessed independently of any finding of wrongdoing on D's behalf. We may have many reasons to complain about D's confinement in such a case, but the government's violation of the punitive restrictions provides us with no grounds for complaint.

To see the work that Distinction in Possession is doing here, consider involuntary civil commitment. When the government civilly commits D, it does exercise a normative power: it makes it impermissible for D to leave the institution in which he is committed. But the government does not possess this normative power with respect to D thanks to a finding that D has done something wrong. Rather, the government possesses this power thanks to a finding that D is a danger to himself or others, perhaps for reasons that involved no violations of the law by him. Civil commitments, even involuntary civil commitments, are not punishments despite involving a

required meaningful reduction in normative freedom thanks to the fact that the normative power exercised in their production is not the punitive power but a power that is possessed under different conditions from the conditions under which the punitive power is possessed. A similar point can be made about pretrial detention where the government's power to make it impermissible to leave jail is possessed not thanks to a finding of wrongdoing but thanks to a finding that the suspect cannot be trusted to return to court on his own.

But notice that the power exercised when issuing a civil injunction is often, although not always, possessed under the same conditions as that exercised when punishing. Often the government has the power to issue a civil injunction and thereby make it impermissible for D to refrain from conduct that it was previously permissible for him to refrain from, thanks to a finding of wrongdoing on D's part. So, Distinction in Possession will not distinguish the relevant normative powers in such cases. However, Distinction in Entitlement will. There is one entitlement that necessarily accompanies the punitive power: what I will call the "strong target-hardening" entitlement, which is absent in cases of civil injunction.

Those who have a target-hardening entitlement are entitled to make it very difficult or impossible for others to violate the norms they shift. Familiar and paradigm instances of target hardening are simple security measures, such as locking doors or setting prominent surveillance cameras, undertaken to prevent oneself from being a victim of crime. Locking someone in a prison cell is no different fundamentally from locking one's front door: in both cases, the action serves to make it more difficult for a person to do what is, for him, wrongful—namely, leaving the cell, or entering the home. In fact, much of the misery that is involved in punishment—although not all—derives not from the punishment itself, the shift in norms, but instead from the fact that the punished lives in a world of hardened targets. There's misery in being imprisoned. A small and rather outré part of the misery is that it is not pleasant for behavior that is permissible for others to be wrongful for oneself. At least, that's unpleasant for those who find it unpleasant to be unequal to and lesser than others; it is unpleasant for those who do not like to be subject to more stringent demands than other people. But most of the misery that the punished suffer comes not from the punishment, strictly speaking, but from the fact that they are denied many pleasurable things—many of the things that give life meaning, in fact, such as association with friends and family—thanks to the instruments that prevent them from doing the wrongful thing of leaving the prison. The target hardening, rather than the punishment, is the primary thing that makes time in a cell "hard" time.

It does not generally follow from the fact that it is impermissible for a person to do something that it is permissible for another to harden the target, making it difficult for the person to violate the norm. Sometimes

target hardening is permissible, and sometimes it is not. Further, it is an agent-relative permission. It is permissible for you to harden the target by locking your front door; it is not permissible for others to do this, even if doing so does indeed cut down on the frequency with which people violate the norm against entering your house without permission.

The target-hardening entitlement is that in virtue of which it is permissible to harden the target, when it is. In general, someone possesses the target-hardening entitlement with respect to another's wrongful act thanks to the fact that she bears the right kind of relation to the norm in virtue of which the conduct would be wrongful. The owner of a house is permitted to lock the front door in an effort to harden the target thanks to the fact that others are not permitted to enter without her permission. The validity of the norm against entry depends on her permission, and so she bears the right kind of relation to that norm to make it permissible for her to take steps to prevent others from violating it. Those who have no power over the validity of the norm—they cannot make it permissible for others to enter—also lack the target-hardening entitlement with respect to the norm.

Target-hardening measures are of two distinct sorts: some such measures reduce the other person's ability to violate the norm, whereas others make it more burdensome for him to exercise his abilities to violate the norm. A lock on the door is of the first sort; a surveillance camera is of the second sort. Target-hardening measures of the first sort decrease the probability of violation for the rational and irrational alike. Measures of the second sort, however, decrease the probability of violation only on the part of rational actors who aim to minimize their burdens in pursuit of their ends. Corresponding to this distinction in target-hardening measures is a distinction in target-hardening entitlements. Those who are permitted to use both kinds of target-hardening measures have a *strong target-hardening entitlement*. Those who are permitted, instead, to employ only the latter measures for hardening the target have only a *weak target-hardening entitlement*. When someone has a weak target-hardening entitlement, it is more important that they treat the potential norm violator as an autonomous, rational agent than it is that they prevent violations of the norm. By contrast, we find the strong target-hardening entitlement when it is permissible to treat potential violators as irrational, nonautonomous actors who cannot be counted on to comply through the exercise of their faculties for the appreciation of value.

When we find a power to shift norms together with a strong target-hardening entitlement with respect to the shifted norm, we find the punitive power. When that power is exercised so as to induce a required meaningful reduction in normative freedom, then that reduction in normative freedom is a punishment. This is true even if the bearer of the normative power leaves her strong target-hardening entitlement unexercised. When the parent grounds the teenager but does not lock the door and does not

take away the car keys, the parent is nonetheless exercising a punitive power, for the parent is entitled to take such steps even if she recognizes that they need not be taken or would be unwise to take.

Combining these remarks with our discussion of the circumstances in which the punitive power is possessed, we have the following:

*Punitive Power:* A normative power to produce required meaningful reductions in normative freedom is the punitive power if and only if (1) it is possessed with respect to D only given a finding that D has engaged in a wrongful act and (2) it is accompanied by a strong target-hardening entitlement with respect to the norms established through its exercise.

Contrast with the power exercised in the issuance of civil injunctions. There are two different things that the government is permitted to do to increase the probability that a defendant complies with a civil injunction: it can threaten civil penalties for failure, and it can threaten criminal punishment for failure. These are weak target-hardening measures. They might be highly coercive and effective, but their efficacy depends on the defendant's exercise of her power to direct her conduct in accordance with her conception of value. She complies because she does not wish to suffer the consequences, and not because she cannot do otherwise. The normative power exercised in the issuing of a civil injunction is not the punitive power thanks to the fact that the government is not entitled to force compliance with it, but only to strongly encourage, even to coerce, compliance with it. To issue a civil injunction is to maintain a picture of D as someone who can be trusted to guide his conduct in accordance with his conception of value. This means that the government is entitled to increase the probability of compliance only by altering the value of compliance, perhaps by magnifying the burdens associated with noncompliance. The government is not permitted to resort to force; it possesses only a weak, and not a strong, target-hardening entitlement.

Incarceration is punishment-by-nature for the following reasons: First, D suffers a reduction in normative freedom. Prior to sentence, it was permissible for D to be away from the prison; after, it is not. Second, the reduction in normative freedom is meaningful. D stays in the prison, and his staying there is in part explained by the fact that it is legally impermissible for him to leave. Third, this meaningful reduction in normative freedom is required. Things are not as they ought to be, legally speaking, if D does not suffer this meaningful reduction in normative freedom. Fourth, the power which the government exercises in bringing about the shifted norm is possessed in part because there was a finding of wrongdoing on the part of the person incarcerated. Fifth, the government enjoys a strong target-hardening entitlement with respect to the shifted norm—the

norm requiring D to stay in the prison. The government is entitled to take steps not just to discourage D's escape but to prevent D's escape through means that would be effective even if D were willing to bear the burdens of escape. It is for these last two reasons that the normative power the government exercises—the power to shift norms—is the punitive power.

The theory also explains, as any account of punishment-by-nature must, why corporal punishments, such as executions, are punishments. Corporal punishments are distinguished from the simple infliction of physical harm by this: it is impermissible for those who are corporally punished to defend themselves against the infliction of the relevant physical harm. This is a reduction in normative freedom. Prior to the sentence, it was permissible for D to defend himself against a certain kind of physical attack (a lethal attack in the case of a death sentence); after the sentence, it is not. Further, the reduction in normative freedom is meaningful. D complies with the relevant shifted norm when he submits to the physical attack and his submission is explained, in part, by the impermissibility of self-defense. And the meaningful reduction in normative freedom that he suffers is required: thanks to the normative changes the punisher has brought about, things would not be as they ought to be were D not to submit. And, finally, the power the government exercises in issuing the sentence is the punitive power. It is possessed in virtue of a finding of wrongdoing by D, and it is accompanied by the strong target-hardening entitlement: the government is entitled to make it the case that D submits, and does not defend himself, whether or not D sees and responds to sufficient reasons to do so. If the government seeks not just to physically harm someone in response to a finding of wrongdoing on that person's part but also to make it impermissible for that person to defend herself, and if the government claims an entitlement to force submission, to thwart efforts at self-defense, then the government seeks to punish. What it seeks to do is to inflict something that is a punishment-by-nature.

#### IV. THE PUNISHMENTS THAT RESIST DISAVOWAL

Much of the motivation for the Norm Shift Theory comes from the observation, elaborated in Section II, that there are some things that the government wants to do to people that just are punishments, whether or not the government is willing to accept that label. And so there are some things that the government is not permitted to do to people without abiding by the punitive restrictions, whether or not it is willing to accept that those restrictions apply. The Norm Shift Theory's success, then, turns on the question of whether those things that count as punishment under it resist disavowal of the label. Are all and only those things that count as punishment under the theory things that bear the label, whether or not the government wants them to?

One way to produce a full argument for an affirmative answer to this question would be to derive the punitive restrictions from the Norm Shift Theory. That is, if it could be shown that, for instance, the *ex post facto* restriction applies to the infliction of P if P is a punishment under the Norm Shift Theory—and that this is also true for the principle of excuse, the proportionality principles, and so on—then we would know that the government could not disavow the label when inflicting anything that counts as a punishment under the Norm Shift Theory. However, for reasons alluded to in Section II, I am doubtful that this full argument could be produced in support of any theory of punishment-by-nature. The problem is that there is no reason to think that all of the punitive restrictions apply because of punishment's nature considered in itself. They might apply, instead, because of the nature of the punisher, or because of the nature of the act of inflicting punishment, or for some other reason. In the absence of a full account of the source and limits of each of the punitive restrictions, a full argument in support of a theory of punishment is likely impossible to produce.

However, there is a line of thought that can give us some confidence that the Norm Shift Theory identifies the necessary and sufficient conditions of punishment-by-nature, even if it does not provide us with certainty. The theory identifies punishment with something that, by its nature, resists disavowal of the label. To see how the theory does this, consider a feature of punishment that has been recognized by every philosopher who has tried to solve the problem of definition: those who inflict punishment claim authority to do so; they make a claim of right. Answers to the question of definition always include this as a necessary condition of punishment. But prior theories also identify other necessary conditions that are entirely independent of this condition. So, for instance, when Hobbes says that punishment is “an Evil inflicted by publique Authority,” he is equating “evil” with pain—something that can be inflicted with or without “publique Authority.” There is no necessary connection between the identified intrinsic feature of punishment—it is painful, according to Hobbes—and the claim to authorization made by those who inflict it. The result is that no inference can be made from the proposition that D suffers something with the relevant intrinsic feature to the proposition that the inflictor claims authority to inflict it. Since it's the kind of thing that could be inflicted with or without authorization, someone who is suffering an “Evil” at the hands of the government needs to marshal additional evidence to show that the government is making a claim to right, much less the distinctive claim that it makes when it punishes, in contrast to when it taxes or issues a civil injunction. Further, if it is up to the government to decide whether or not it is making a claim to authorization, and, if so, what kind of claim it is making, then there is no possible way to succeed in showing that what the government is doing is punishing, even if it claims not to be. It is up to the government to decide.

Put another way, the category of punishment-by-intention swallows the category of punishment-by-nature.

To avoid this, we need a theory of punishment under which punishments have intrinsic features that necessarily imply that the government makes the distinctive claim to authorization that is involved in punishing whenever it inflicts something with those intrinsic features. That is, we need a theory under which the government cannot consistently both (1) grant that what it is inflicting has xyz intrinsic features and, at the same time, (2) deny that it claims the distinctive form of right to inflict that thing, a right which is definitive of punishment. The Norm Shift Theory fits the bill. The reason is that if the government grants that D is suffering a required meaningful reduction in normative freedom and, at the same time, grants that this is predicated on D's wrongdoing and that it is entitled to strongly harden the relevant target, then it admits that there is something in the world—the relevant kind of normative change—that could not possibly come to pass except through the exercise of the distinctive form of normative power that we are labeling “the punitive power.” There is no other way, except through the exercise of the punitive power, that a required meaningful reduction in normative freedom, predicated on wrongdoing and accompanied by a strong target-hardening entitlement, could come to pass. If the world includes such a thing, then the entity that enjoys the strong target-hardening entitlement is punishing, whether it is willing to accept the label or not. Acknowledgment by the government of the intrinsic features of what it is inflicting on D entails that what it inflicts came to pass through the exercise of a distinctive normative power, the punitive power, and so is punishment. Put another way, the only way that the government can free itself from the punitive restrictions is to deny that what D is suffering is a required meaningful reduction in normative freedom, or to deny that it has the relevant power in light of D's wrongful behavior, or to deny that it enjoys a strong target-hardening entitlement. But to deny any of these things would be to give to D something that the government fervently aims to withhold. The government, after all, wants to deny that what it is inflicting is punishment while at the same time retaining all the normative changes that are involved in punishment.

To see this point more clearly, return to the case of *Smith v. Doe*, which, we are now in a position to see, was wrongly decided. Forced registration involves a required meaningful reduction in normative freedom. It is illegal for Doe not to register. When he complies with this requirement, his reduction in normative freedom is meaningful. And the world would not be as it is legally required to be were someone to register in his place. In fact, it would not even be as it is legally required to be were someone other than Doe himself to put him on the registry; registration requires an in-person appearance at a police station. How did all of these normative facts come to be in place? Normative change can only happen through the

exercise of a normative power. So, the government made this happen through the exercise of some normative power. The question is whether the normative power that was exercised here was of the distinctive kind involved in punishment. But it must be. The government has the power to induce this normative change thanks to the performance of a wrongful act by Doe, namely, the sex crime for which he was earlier imprisoned. The government does not claim the power to make it illegal for Doe to refrain from registering for any other reason. And, further, the government has a strong target-hardening entitlement: Doe can be arrested and forced to comply with registration requirements. Therefore, the power that it exercises with respect to Doe is a punitive power.

The government can disavow the punitive status of forced registration, and so avoid the *ex post facto* restriction, only by denying one of the normative features of forced registration just enumerated. It could avoid that restriction by denying that it is illegal for Doe to fail to register, for instance. There would therefore be no reduction in normative freedom and so no punishment. But, of course, that would be to give up the game entirely. The government could also avoid the *ex post facto* restriction by, for instance, allowing that it can only use threats to induce Doe to register of the sort that can be used to enforce civil injunctions. That is, it can deny that it has a strong target-hardening entitlement and assert that its entitlement is only weak. But it would then radically decrease its compliance rates with the registration demand, even among those who committed their offenses after the threat of forced registration was made. The government will not want to make that concession. There are other ways out too. But all of them involve granting normative leeway to Doe of a kind that the government is unwilling to grant. The result: the government can have its cake only by granting that what it aims to inflict is punishment, and so that which it aims to inflict must conform to the punitive restrictions.

Any theory of punishment provides a tool with which to press back against a government that wants to deny that what it hopes to do to one of its citizens is punishment and so bypass the punitive restrictions. But any theory that takes a necessary condition of punishment to be some condition that the government has unchecked power to either put in place or take away, without otherwise affecting what it wants to do to a citizen, is a theory that recognizes only the category of punishment-by-intention. But we need a theory of punishment-by-nature. It cannot be that the government must abide by the punitive restrictions only when it wants to. No prior answer to the question of definition has given teeth to this idea. The advance of the theory of punishment presented here, I suggest, is that it does.