A REPLY TO CRITICS OF *PUNISHMENT AND FREEDOM*

*Alan Brudner**

I'll begin by thanking the editors of the *New Criminal Law Review* for providing a forum for this exchange, Professor Markus Dubber for organizing it, and the five commentators on *Punishment and Freedom*—Professors Thom Brooks, Shai Lavi, Alan Norrie, Alice Ristroph, and Mariana Valverde—for engaging with the ideas presented in the book.[1] I'll respond to their comments according to the alphabetical order of their surnames.

I

Professor Brooks voices broad agreement with the theoretical approach to punishment taken in *Punishment and Freedom* (henceforth PF) but suggests two refinements drawn from the work of the British Idealists, T. H. Green and James Seth. One suggestion would modify legal retributivism to justify punishing "crimes" not involving an interference with free choice, such as breaching a conservation statute, speeding, and drug trafficking. The other would acknowledge a place for deterrence and rehabilitation within a retributive theory of punishment.

In PF, there is a chapter devoted to so-called public welfare offenses, within which category I would include breaching a conservation statute, speeding, and drug trafficking. Although I argue that, with a qualification mentioned below, these offenses fall outside a legal-retributivist account of punishment, the full theory of justified penal force includes them.

*Faculty of Law and Department of Political Science, University of Toronto.

1. Alan Brudner, Punishment and Freedom (2009), hereinafter PF.

New Criminal Law Review, Vol. 14, Number 3, pps 495–518. ISSN 1933-4192, electronic ISSN 1933-4206. DOI: 10.1525/nclr.2011.14.3.495.

The retributivism pertaining to willful interferences with free choice is not offered as a complete account of legitimate penal force; rather it is supplemented by an instrumentalist theory of penal laws that regulate the pursuit of satisfactions or that protect the material conditions of living autonomously. So, unless I misunderstand him, the refinement that Brooks urges is already incorporated.

Perhaps, however, Brooks is recommending that I bring public welfare laws into the legal-retributivist justification of punishment instead of treating their enforcement as constrained by principles uniquely applicable to them. To a limited extent, this suggestion is also already incorporated. I argue that public welfare offenses can also be crimes, the punishment for which is deserved, if they involve knowing (so ignorance of law would negate the crime, but not the regulatory offense) breaches of laws aimed at preserving the material conditions (agency goods) of living autonomously; for the lawbreaker, having then denied law's authority on which rights depend, cannot complain if his own right against coercion is infringed.[2] The requirement of a risk to agency goods embodies the retributivist principle of proportionality; for without it people would be liable to imprisonment for knowingly parking illegally, yet not (on the principle I propose) for negligently polluting a water supply. Drug trafficking falls into the subcategory of public welfare offenses that endanger the conditions of living autonomously; hence, the knowing commission of this offense is properly criminalized. However, contravening a statute protecting greenery cannot be criminal, in my view, for this action falls into the subcategory of breaches of welfare laws that serve a social preference. Criminalizing the knowing contravention of such a law would violate the proportionality principle.

I would resist, however, any amendment that would blur the distinction between crimes to which retributivist constraints apply and public welfare offenses to which other autonomy constraints apply. The fact that Brooks speaks of "criminalizing" ordinary traffic offenses and says that actions such as speeding and appropriating a tree from a public forest also violate rights indicates that he has such an elision in mind. If so—if Brooks is suggesting that I assimilate public welfare offenses into the category of crimes and justify punishing them by the legal-retributivist account of punishment—I will have to respectfully demur. The legal-retributivist legitimation of punishment works only when the actor's vulnerability to coercion

2. PF, at 81, 177.

follows logically from what he or she chose to do. For reasons I set out, that vulnerability follows from knowingly interfering with another's agency and from knowingly breaking the law where the broken law protects an agency good; but it does not follow from knowingly performing the action that a law prohibits where the action involves no interference or attempted interference with agency.[3] In that case, no logical circuit connects the actor's deed to his liability to coercion, and so penal force must be justified by another route—that of constrained instrumentalism.

Brooks's second suggestion is that I make room for deterrence and rehabilitation within legal retributivism by viewing these forward-looking functions of punishment as protecting *ex ante* the rights that retribution vindicates *ex post*. Legal retributivism would be strengthened, he thinks, if it were "flexible" rather than "pure."

My response is that I already accommodate the prospective ends that punishment might serve to the extent that they can be incorporated compatibly with the integrity of legal retributivism. The retributivism in PF is pure only in the sense that the vindication of rights against right-deniers is the sole justifying reason for punishing true crimes. But of course, after that justification has finished determining who deserves to be punished, for what specifically, and at what ordinal level, there is plenty of discretionary room in sentencing for fixing the exact measure of punishment with a view to deterrence and rehabilitation. Legal retributivism does not rule out these ends; it only insists that they be pursued within the self-imposition constraints generated from its own way of justifying punishment. Viewing the forward-looking functions of punishment as protecting rights rather than as satisfying interests cannot ease the stringency of these constraints, for subordinating one person to the rights of others converts those rights into particular interests. Furthermore, although my retributivism is pure, my account of penal force is not purely retributivist. As mentioned, the entire sphere of public welfare offenses lies outside the legal-retributivist account of punishment; and in this sphere, penal force is justified (subject to autonomy-respecting constraints) by the aim of preventing harm to the agency goods to which autonomous citizens are entitled.

In sum, therefore, I believe I have already incorporated Professor Brooks's suggestions to the extent that they are compatible with the theoretical framework on which we agree.

---

3. PF, at 38–41, 176–77.

II

Before dealing with the criticisms in Professor Lavi's thoughtful comments, I want to clear up a few misunderstandings. First, I do not argue that state punishment, properly understood and practiced, is itself freedom. That would be the sort of hyperbole that has often been mistakenly attributed to Hegel and that has exposed his theory of punishment to ridicule in some quarters, most recently by Alice Ristroph. State punishment is not itself freedom, for it coerces, and one cannot be free while being coerced. Properly theorized and constrained, however, punishment is reconcilable with the *right* to be free of coercion, because the implicit assent given to it by the thinking Agent in the criminal's shoes can be imputed to the criminal insofar as he or she has the capacity to think. However, the criminal's implicit assent does not wipe out his empirical nonconsent or the wrongful violence that goes with it. Rather, justification by implicit assent presupposes that there is a wrong in the real world requiring justification. On the empirical plane, therefore, punishment negates freedom, and the philosophic justification of punishment is far from treating this as illusion. On the contrary, the negation of freedom on the empirical plane is treated as an enduring wrong that must be reconciled with the right to freedom at the transcendental level.

Second, I would not say that the criminal law can be divided into core, periphery, and exception. Lavi sees the criminal law's core in the paradigm governed by formally free choice, its periphery in the paradigm governed by real autonomy, and the exception in the ethics-based excuses. In PF I say that the penal law is divided into a core and a periphery.[4] At the core is the unity of the formal agency and real autonomy paradigms, and at the periphery lie the ethics-based excuses. This is no mere quibble, for it would be odd to regard as peripheral to the criminal law an autonomy paradigm that first generates a public measure of the relative seriousness of crimes, a right against strict penal liability for consequences, and a right to a measure of punishment proportioned to responsibility.

Now, with regard to how the penal law theory offered in PF might be applied to the case of euthanasia, I find myself in agreement with Lavi's analysis—up to a point. I agree that the formal agency paradigm, for which free choice is foundational and not itself disposable at will, would regard

4. PF, at 16, 165–68, 254–62.

suicide as self-murder and so would not give effect to a person's consent to being euthanized.[5] I also agree that the real autonomy paradigm generates a more permissive rule, for it sees human dignity as lost once agency has sunk into biological existence and so would allow a deliberative choice to avoid that fate. However, Lavi worries that once we move to the real autonomy framework, we cannot stop it from overwhelming the formal agency paradigm it superseded. And that would bring us to the unlimited right to die for whatever reason one chooses—pain, depression, or even to prove one's freedom from the law of self-preservation.

Lavi's concern raises a question that goes to the heart of PF's presentation of the penal law as a unity of paradigms. How can superseded juridical frameworks ordered to inadequate conceptions of freedom be preserved as constituent parts of a complex system? Of course, his solution to this problem is precisely what secures Hegel's place in the pantheon of great philosophers. I try to explain Hegel's idea in PF's last chapter and do not have space to repeat the exposition here. But the basic idea is that juridical frameworks based on inadequate conceptions of freedom are cancelled insofar as they claim to be fundamental and hence exhaustive of rights and duties; but they are nonetheless preserved as instances or reflections of the most comprehensive conception, which requires them for its own validation. Each conception gives up the fundamentalist claim that led to its self-contradiction and accepts a constituent status by virtue of which it is fulfilled. In that way, the normative frameworks that once claimed to tell the whole story about justice are reduced to parts of a whole. Accordingly, what stops the principle of real autonomy from overwhelming the formal agency framework is just the logical law that a part cannot be the whole. In the whole, the expansionist tendencies of the separate frameworks are tamed.

What does this mean for a law of euthanasia? I believe it means that formalism's absolute prohibition against voluntary euthanasia and assisted suicide is cancelled, but that formalism survives to the extent of constraining

5. Lavi sees in this prohibition a counterexample to my claim that the formalist paradigm cannot distinguish between crimes on the basis of the harm they inflict, for evidently it can distinguish death—the extinguishment of liberty—from other harms. I concede this point. But I maintain that, since it equates welfare with the satisfaction of contingent preferences, formalism cannot rank bodily harm and property damage on a scale of objective seriousness. Lavi states that a more "wholesome" account of formalism could yield such a ranking, but he does not elaborate.

the reasons for which someone may consent to euthanasia with legal effect. The end of the formalist paradigm is the dignity of the freely choosing self. But that dignity comes to naught if someone is forced to live without the prospect of regaining his or her free agency. Accordingly, the autonomy framework fulfills the end of the formalist one in permitting voluntary euthanasia and assisted suicide for the sole purpose of preserving one's dignity for the whole duration of one's life.

Next, Lavi poses the problem of involuntary mercy killing. I agree with him that any acquittal or mitigation of punishment in such a case must result from an ethics-based excuse that presupposes culpability according to the legal standard. However, Lavi disputes my claim that such excuses are themselves governed by standards. In his view, the excuses constitute exceptions from all standards and rules, with the consequence that they must be granted or withheld on a purely discretionary basis, preferably by a jury unaccountable for its verdict.

I believe that involuntary mercy killing can be put on the same footing as legally provoked murder. The defense of provocation is a partial moral excuse that presupposes legal culpability, and that excuse is certainly rule-bound. Murder is reduced to manslaughter if the accused gave in to rage in circumstances that would also have enraged the character of reasonable self-composure. The reasonably self-composed character is here the standard against which the accused's anger is assessed as justifiable or not. But if justified anger can mitigate, why not justified compassion—compassion that would have strained the capacity for self-control of someone who embodied the mean between callousness and whatever one wishes to call an excess of sympathy? So, I am unpersuaded that excuses are not properly governed by standards. Certainly, the excuses make an exception from the abstract, agency-based criteria of legal culpability. However, it does not follow that they lie outside the governance of all standards, for beside the legal standard is the ethical one of the reasonably virtuous character. Nor does it follow from the rule-bound nature of excuses that no indeterminacy exists in the application of the rule. Of course, there is room for discretionary judgment in application and hence for a certain degree of arbitrariness; no one would deny that. Whereas, however, a residual indeterminacy of that sort is compatible with the just application of force, the standard-less granting or withholding of excuses by a jury's inscrutable fiat is not.

III

Professor Norrie makes three criticisms of PF, two of which deal with my understanding of the role of intention in the criminal law, the other challenging the remarks I make at the book's end about the relation between penal and social justice. But before employing the heavy artillery, Norrie engages in a few preliminary skirmishes, and I want to say something in response to these comments just to clear up a misunderstanding.

Norrie shares with Professor Valverde the view that, because the book is about reconciling punishment with individual autonomy, it is as much Kantian as Hegelian in inspiration. Thus Norrie writes: "In Brudner's account [of criminal law], Hegel operates alongside Kant to . . . 'complete' him by showing how the core Kantian understanding of freedom needs to be seen in a broader context of concepts, which together establish a complex unity." The suggestion is that the book embodies a position identifiable by names joined with a hyphen: Kant-Hegel (or perhaps Kegel), while Valverde claims to see in the book a "heavy dose" of Kantianism despite its avowed debt to Hegel. Behind these characterizations of my position lie two misconceptions. One is that the Kantian idea of formal freedom with which I begin organizes the core of the criminal law and is, in Norrie's words, "the core . . . understanding of freedom." The other is that the autonomy paradigm that completes the formalist one is also Kantian.

Let me emphasize that I regard my position as Hegelian through and through. Although Hegel's legal and political thought completes a development initiated by Kant, the completion bears as little resemblance to the beginning as an adult bears to an embryo. Throughout the book I distinguish my position from Kant's, and I criticize the Kantian theory of punishment as being untrue to its own requirement that punishment be consistent with freedom. For Kant, punishment is the execution of an external threat of disadvantage meant to shape the incentives of an agent assumed to be motivated by natural self-interest. That is why he calls authorized coercion "a hindering of a hindrance to freedom."[6] Contrary to what Norrie thinks, this formula is not identical in meaning to Hegel's negation of a negation of right, for hindrance (like Hobbes's "impediment") is physical whereas negation is conceptual. There is, no doubt, a retributivist

6. I. Kant, Metaphysics of Morals, Part I, 57 (M. Gregor trans., 1991).

side of Kant; however, that side stands in tension with a Hobbesian one, according to which the threat of punishment is required to make possible the surrender of a right of self-help to a public authority that institutes a civil condition.[7] It is because he views punishment as essentially an external threat that Kant argues against imposing punishment in circumstances where the threat could not have been effective to deter—that is, in a kill-or-die situation.[8]

However, the principal way in which an Hegelian theory of criminal law differs from a Kantian one is that it treats the paradigm of penal justice informed by a formal conception of freedom as a constituent element of a complex system rather than as the whole or even as the core of penal justice. For Kant, the formalist paradigm of penal justice is the whole truth about penal justice. For Hegel, it must be complemented by a paradigm ordered to real autonomy, which first generates the idea of objective harms to agency goods and of grades of responsibility for harmful outcomes. Since it is patently a mistake to see the difference between assault and assault causing bodily harm or between murder and manslaughter as peripheral to the criminal law, it is also a mistake to think (as Norrie does) that the formalist paradigm forms the core of criminal law, implying that everything else is periphery. The paradigms of real autonomy and formal freedom are mutually complementary and mutually essential. Neither is core and neither is periphery. Together they form the core of the penal law, and although the excuses rooted in the communitarian paradigm are, on this view, peripheral to the core, they too are essential to a penal law consistent with full dignity; for the human being is not only a free agent and an autonomous subject but also a determinate character whose worth is validated in relationships structured by ethical expectations. All this takes us far beyond Kant.

Norrie says that "there is [also] something essentially Kantian going on" in the paradigm of real autonomy, and Valverde agrees. But that is because they, not I, conflate Kant and Hegel. Of course, the idea of autonomy (or positive freedom) is central to Kant's practical philosophy as a whole. However, it plays a limited role in his *Rechtslehre* and a minor one

7. Id. at 124. No doubt, Hobbes and Kant require the sovereign's coercive power for different reasons: Hobbes to reconcile the individual's surrender of power with his right of self-preservation; Kant to reconcile this surrender with the injunction against unilateral obligations entailed by the duty of self-respect.

8. Id. at 60–61.

compared to the role autonomy plays in Hegel's philosophy of right. Kant thought that a rational agent could be autonomous only in the moral life—only through the inward virtue that consists in acting purely from respect for law, setting aside material self-interest. In the legal sphere, the role of the autonomous or self-legislating will is confined to that of a notional sovereign in which only some (not wage-laborers or women) participate and whose constraints on state action are unenforceable against those who legislate and execute in its name.[9] Thus, even the best ruler is a despot. Moreover, the notional sovereign gives laws regulating external action that free agents *could* legislate for themselves, but that, given their fixed natural inclination to self-love, the legal system must regard as operating as an external force.[10] That is why Kant has two theories of punishment: a retributive one for ideally self-legislating agents (*homo noumenon*), and a threat-based one for human beings as they are (*homo phaenomenon*).

Even autonomy's role in giving laws is limited in Kant's *Rechtslehre.* Because Kant assumes the fixed reality of a freely choosing animal (rational nature) inclined toward selfish advantage and vulnerable to external causes, he sees no imperative of justice to embody positive freedom—freedom from vulnerability to external causes—in positive law. Justice is completed by the sum of the conditions under which the negative freedom of choice of each is rendered compatible with the negative freedom of choice of all[11]; it includes no positive entitlement vis-à-vis the sovereign to the conditions of an autonomous life. For Kant, freedom from vulnerability to external causes is the chimerical goal of an endless, inward struggle to free oneself from the natural conditions of human existence.[12] There is no right to such a freedom for Kant, and so there is nothing in his *Rechtslehre* (as there is in Hegel's) about the injustice of strict penal liability for consequences, nor is there a demand that the right to self-determination be actualized against contingency in general by the state's police power.[13]

9. Id. at 126, 130.

10. Id. at 57–58.

11. Id. at 56.

12. I. Kant, Critique of Practical Reason 32–33 (L. W. Beck trans., 1956).

13. Georg Wilhelm Friedrich Hegel, Philosophy of Right, para. 115, 117–18, 120, 230 (T. M. Knox trans., 1967). It is thus mind-boggling that Norrie would say that "demetaphysicalizing" Kant's autonomy is deradicalizing it. On the contrary, it is taking autonomy out of the inward sphere and making it legally practical. Norrie claims that the formal freedom with which I begin is Kantian rather than Hegelian because Hegel sees a moment of embodiment in

For Kant, then, the legal sphere assumes the heteronomy of agents, and it is a contingent matter that the agent autonomously makes the law the determining ground of its choice; the legal order places no reliance on this.[14] For Hegel, the truth is just the reverse. Because Hegel thinks that the agent's animal inclination to selfishness is potentially overcome in a permission for reasonable self-love won back from the political community, he regards law as a limitation the thinking being is naturally inclined to accept; and it is a contingent matter that this or that individual is moved to obey by the threat of adverse consequences. That is why Hegel criticizes Kant for viewing law as an external restriction of liberty[15]; it is why Hegel has *only* a retributivist theory of punishment; and it is why he has a fully developed legal and welfare system ordered to real autonomy, whereas Kant leaves autonomy to the moral life. Norrie's conflating of Kant and Hegel obliterates these gaping differences.

Now we come to Norrie's main criticisms. First, he argues that my position on the mens rea required for criminal liability leads to the result that, if their story is believed, the defendants in the *DPP v. Morgan* ought to be acquitted of rape even though their belief in the victim's consent was unreasonable.[16] That is indeed the conclusion to which my position leads. However, Norrie thinks that this result respects the defendants' formal freedom at the expense of the victim's real autonomy, for the victim "has not had sex under conditions that give her the chance to choose her own ends." This result also, he says, protects the accused's formal freedom while ignoring community norms and expectations. So I have here focused only on formal freedom, whereas the full story about the penal law is supposed to synthesize formal freedom, real autonomy, and communal belonging. Norrie writes, "in placing so limited a conception of freedom at the core [sic] of the criminal law, Brudner's account in a

property. But this betrays a misunderstanding of Hegel. Hegel himself says that the standpoint from which he begins is a free will abstracted from all determination (Philosophy of Right, para. 35). The movement toward embodiment in property reveals the inadequacy of that starting point because it reveals the necessity of something not taken into account in the formalist foundation. So Hegel is critical of formalism (showing its equivocations), whereas Kant is not. I try to bring out this critical side in my discussion of the limits of formalism in PF at 55–58.

14. Kant, supra note 6, at 56–57.

15. Hegel, supra note 13, at para. 29.

16. DPP v. Morgan [1976] AC 182.

case like *Morgan* is stuck with a form of freedom that he himself concedes is radically incomplete."

My response is that a synthesis must be distinguished from a soup. In a soup, diverse ingredients are thrown into one pot and mingled in a way that makes their distinct identities disappear. In a synthesis (as in a well-prepared meal), the identities of the constituent elements are preserved because they form distinct, albeit mutually complementary parts of a whole. Norrie's complaint is that not all the conceptions of freedom operative in the penal law are in play at once with respect to the same question (here, culpability for sexual assault). Perhaps he could explain how this would be possible. If concern for the victim's autonomy determined her assailant's liability to punishment, how would the accused's right to be punished only for interferences he chose be respected? In the account of criminal law I give, each conception of freedom is individually respected because each pertains to a separate question of liability. The issue of criminal liability for a wrong is decided by the formalist paradigm alone. So, if the defendants in *Morgan* did not choose to have nonconsensual sex, then no denial of rights can be imputed to their choice; and if no denial of rights can be imputed to their choice, then they have not implicitly assented to the infringement of their own. Punishment would then be external violence inconsistent with the inviolability of the agent. We are not "stuck" with an incomplete form of freedom; rather, this is the conception of freedom that properly pertains to the question of liability to punishment of any severity.[17]

But that is not the end of the matter. The victim's right to choose her sex partner is not affected by the court's decision whether or not her assailant is punishable for sexual assault; for even if the assailant is acquitted of that crime, he will be liable in tort to his victim on the basis of negligence. So that is where the victim's sexual choice is protected—in private law. Norrie claims that on my account, the wrong of sexual assault is avoided by virtue of a simple mistake of fact regarding consent. But that is a misrepresentation of what I say. The wrong of sexual assault is committed by nothing

17. Perhaps Norrie is unclear about why, in Hegel's philosophy of law, incomplete conceptions of freedom get to organize a legal paradigm within the system of paradigms. The reason (as I try to explain in ch. 9) is that each is necessary but insufficient on its own for a penal law consistent with full freedom; and each instantiates the overarching form of mutual recognition that holds the paradigms together as parts of a whole. From the fact that a conception of freedom cannot be the fundamental end of the penal law, it does not follow that it cannot order a part of the penal law.

more than a voluntary sexual touching in the circumstance of the passive agent's nonconsent.[18] Knowledge of nonconsent is inessential to the wrong. Accordingly, what is avoided by a mistaken belief in consent is criminal liability for the wrong, not the wrong, and so the victim has recourse to a civil remedy. Keeping the realization of the victim's right separate from the assailant's criminal liability preserves the distinction between public punishment and private revenge. Norrie's mingling of the two issues has the opposite effect.

Furthermore, the state has a duty to protect the real autonomy of its citizens, and one of the conditions of real autonomy is the equal opportunity to pursue self-authored ends free of invidious stereotyping. The state may thus enact a separate public welfare offense of negligent sexual assault—one whose aim is to eradicate sexist assumptions about the meaning of a woman's "No" and to induce people to take care that a sexual partner is truly consenting. And it may penalize breaches with severe fines attached specifically to that offense. In this way, both the formalist and the real autonomy paradigms are given their separate due, with the result that different offenses are treated differently, as proportionality requires. It is thus untrue to say (as Norrie does) that if the formal agency paradigm exculpates, other conceptions of freedom are unavailable; they are indeed available, but elsewhere in the legal system, so as not to invade another's domain.

Norrie argues that a separate offense of negligent sexual assault is inadequate to reflect our moral judgments of "attitudinal culpability," because it would penalize with the same noncriminal penalty those who commit sexual assault with callous indifference toward whether the victim is consenting and those who do so out of stupidity or naivete. But he misses the point. The object of the public welfare offense is not to censure characters but to alter behavior and prevent sexual assaults. To punish someone as a criminal for inadvertent recklessness is to introduce the blind and subjective judgments about inward dispositions and character that I argue have no inculpatory role to play in a punishment regime consistent with the right to freedom. Norrie says that he finds "common cause" with me in this, and yet he criticizes my position for leaving out attitudinal culpability (with friends like these . . .). Moreover, he does not address my arguments against the character theory of culpability, but simply insists that a punishment regime that fails to incorporate judgments of

18. PF, at 35–37, 77–78.

character in its judgments of culpability is lacking something. Well, it is certainly without something, but it is not *lacking* anything essential to penal justice; for coercing a free agent pursuant to others' opinions about his moral character satisfies no conception of public reason. To introduce into criminal culpability the moral judgments of others appropriate to social interaction would be to undo all the normative work in reconciling punishment and freedom that the formalist paradigm does. Such a blurring of boundaries is not consistent with synthesis, as Norrie thinks. It is characteristic of soup.

Norrie's second main criticism concerns my treatment of the case of *R. v. Steane*.[19] I say that according to a legal as opposed to a moral or religious understanding of intention, Steane intended to assist the enemy because he chose to assist the enemy, and therefore the exoneration he deserves should come, not from lack of intention, but from excusatory duress. Norrie argues, however, that if Steane's real autonomy is to be respected, he must be exculpated for lack of intention, for though he formally chose to assist the enemy, he did so only to avoid his family's internment in a concentration camp; and so he chose, not an end he valued, but the lesser evil. Duress, in other words, negates real autonomy, and to respect real autonomy (Norrie thinks) one must exculpate for lack of attribution the actor whose conduct was not ascribable to an autonomous choice. In contrast, I argue that the real autonomy principle limits the scope of one's responsibility for the harmful consequences of a wrong once the fault threshold for liability to any punishment has been crossed, but that the fault threshold is the formally free choice to interfere with agency or to break the law; and Steane crossed that threshold. Norrie, however, sees this jurisdictional division as my abandoning the real autonomy principle in midstream.

So here again Norrie wants a constituent principle of the penal law to invade the domain of another principle. In *Morgan*, he wanted consideration for the victim's sexual autonomy to inculpate the accused even though the accused chose nothing implying a denial of agency rights and even though the victim's autonomy is separately vindicated in private law, where mistakes must be reasonable. In *Steane*, he wants the accused's lack of autonomy to exculpate even though he chose to break a valid law and even though a complete, virtue-based excuse is available that acknowledges his guilt before the legal norm.

19. R. v. Steane [1947] KB 997.

Observe, however, what occurs when the boundaries between paradigms are disrespected in this way. If the absence of autonomous choice exculpates from crime, then someone who chooses to steal because he prefers that evil to the humiliation of going on welfare deserves exculpation, for he chose what for him was the lesser evil rather than a self-valued end. Or indeed, anyone who commits a crime from thoughtless impulse deserves, on this view, an acquittal for lack of deliberate intention. This is not, as Norrie thinks, an argument for erecting makeshift floodgates against a tsunami of principle; for there is no logical imperative to generalize a constituent principle, while there *is* a retributivist principle against extending this one. To allow his lacking real autonomy to exculpate someone who has knowingly coerced another or broken the law is to allow a negation of right to stand because, autonomous choice or not, the accused made a free choice to which a denial of rights is imputable. The lesson here is the one taught by Plato: justice consists in each part of the whole's minding its own business and in that way supporting the work of the others. The principle of real autonomy minds its own business in governing measure of punishment and liability for public welfare offenses and in staying out of liability to punishment per se.

Norrie's third main criticism concerns the brief remarks I make at the book's end about the relation between penal and social justice. I argue that, with certain qualifications not relevant here, a penal system that punishes people for property crimes can be just even if unjust disparities of wealth exist in society. This is so because, barring circumstances of life-threatening destitution for which the law of necessity already provides, it cannot be permissible to violate possessory rights that the perfectly just polity would itself protect and that belong even to a thief. Norrie wonders how I can move from the worldless abstraction of the free will to the right to worldly possessions. He says that I, on the one hand, want to take a "rigorous Kantian line" on the abstraction of freedom from all determinations but, on the other hand, want to go "toward Hegel," who moves from abstraction to a right to worldly things.

My response is that I do not take a Kantian line at all and that I do not move toward Hegel "on the other hand," but am with Hegel from start to finish. Hegel himself begins with the negative movement of abstraction, but then explains the transition from abstract freedom to property roughly as follows. Faced with the apparent independence of the external world, the end-status involved in the capacity for free choice is self-contradictory, for

it is dependent on that from which it is a reflex. So the person is impelled to remove this contradiction by making unfree things means to its ends. Through acquisition, the person gives reality to its hitherto merely formal end-status. Now, if a perfectly just polity is one that secures all the conditions of realized end-status for everyone, then the just polity would vindicate private possessory rights against those who violate them. Therefore, it cannot be unjust for a polity that falls short of perfection to do the same.

But then Norrie asks: does this mean that any distribution of property is equally valid so that one has a right to be a billionaire while others starve? The answer is: of course not. But redistribution will supervene on property rights rather than define them. In Hegel's synthesis of paradigms, private property is validated through the institution of contract prior to the common good of real autonomy, but it is nonetheless subject to that good and hence to redistribution for its sake. The common good thus limits property externally; it does not shape the right itself. Accordingly, the just polity will recognize possessory rights based on the formal will's need for realization, but it will redistribute property for the sake of the real autonomy of all. In Norrie's soup, by contrast, the real autonomy of all determines what possessory right the individual has, with the result that no one's home is her own unless the common good has allotted it. But then the individual person's independent end-status is submerged in the common good, which is to turn the good into the bad.

IV

Buried beneath Professor Ristroph's polemic against metaphysics lie a few arguments one can come to grips with. One is that the kind of punishment I seek to justify is violent, and no purported reconciliation of punishment and freedom can distinguish it from violence. Another is that the thinking Agent whose notional consent is supposed to justify punishment cannot do so because, as an entity without corporeal existence, it cannot represent the actual criminal, and so its hypothetical consent cannot count as the criminal's. Finally, Ristroph urges us to return to Hobbes for a liberalism "more honest" than mine—one that faces up to the morally naked violence that punishment is.

Ristroph's claim that punishing criminals is irredeemably violent rests on a stipulated definition of violence. For her, violence "refers to exercises

of physical force against vulnerable bodies, and it is the vulnerable body that is critical, not the consenting (or nonconsenting) mind." By this definition, chopping wood is a violent act. Although nothing stops Ristroph from viewing it so, I assumed a definition of violence that, emphasizing its etymological kinship with "violate," contained the idea of nonconsent or wrongfulness. Because, as I argue, the law deems actual criminals not to be consenting to punitive force, punishment indeed involves (wrongful) violence, the justification of which must face up to, rather than erase, its wrongfulness. That is why I usually distinguish the justified violence of punishment from "naked" or "criminal" violence, although (as Ristroph observes) I occasionally slip and drop the adjectives.

Ristroph admits that it is conceptually unobjectionable to define violence as containing the idea of nonconsent; indeed, she observes that consent "make[s] the difference between violence and nonviolence" in the sexual context. But here, she says, empirical consent is required to negate violence; no hypothetical consent of a thinking Agent will do. And so she pounces on what she claims is an inconsistency between my acknowledging the necessity for empirical consent to negate wrongfulness in sexual contact and my allowing hypothetical consent to justify punitive force. I am accused of playing "fast and loose" with the notion of individual inviolability, "invoking it to preserve broad criminal liability for sexual offenders while readily abandoning it to justify state punishment as consensual."

There are several layers of confusion here. First, I do not argue that hypothetical consent to punitive force is a substitute for empirical consent, and that one may therefore disregard the criminal's actual nonconsent. In fact, I explicitly deny this and say that hypothetical consent to punitive force justifies what remains a wrong, much as a taking of property justified by necessity remains a wrong.[20] So, empirical consent to sex negates wrongfulness, whereas hypothetical consent to empirically unconsented-to punitive force justifies an abiding wrong; there is surely no contradiction in that.

Still, Ristroph may wonder why hypothetical consent to force can justify an admitted wrong in the punitive context but not in the sexual context. I suggest a reason in discussing whether necessity can ever justify an invasion of bodily autonomy.[21] Of course, a claim of necessity to justify a sexual

20. PF, at 37–38, 279.

21. PF, at 231.

assault is fanciful in the extreme; but even if one's life depended on having intercourse and no willing partners were available, a necessity claim could not succeed because, the body being inseparable from free agency, no public reason acceptable to both parties justifies violating one person's agency to save another's. In the case of punitive force, by contrast, there is (as I argue at length) a public reason to justify an abiding wrong—namely, the requirement that mutual respect for agency prove itself as the sole ground of valid right-claims through the demonstrated nemesis of the criminal's claimed right to an unlimited liberty. Ristroph may smile at this metaphysical justification of a wrong, but she may not see it as playing fast and loose with the notion of inviolability.

A further confusion in Ristroph's argument has to do with the parity she assumes between consent to sex and consent to the deprivation of one's liberty. The law allows everyone of age to consent to sex because sex is an optional end one may pursue consistently with one's freedom to pursue any end. But submission to the deprivation of one's liberty is not such an end, and so the law disallows empirical consent to punishment (much as it disallows consent to slavery) and deems the criminal not to consent whatever his outward behavior. This is exactly what Hobbes does in disallowing any empirical waiver of the right to self-preservation, to which point I'll return. But it is precisely because the law postulates a nonconsenting prisoner regardless of the empirical facts that any justification of punishment must justify what remains a (violent) wrong.

Ristroph seems to sense that her preoccupation with the conventional definition of violence leads nowhere, because she exhorts herself to "start afresh" with a question of real substance: "what are the conditions under which the state may permissibly exercise power over its citizens?" Here begin her deprecations of my thinking Agent who, stripped (she says) of all corporeal attributes, cannot coherently consent to imprisonment of the body nor plausibly claim to represent the embodied criminal in the dock. And if the thinking Agent cannot represent the criminal, then its notional consent, even supposing it were demonstrated, cannot count as the criminal's. Indeed, Ristroph argues, no consent but that of the actual criminal can legitimize force against him; and since (following Hobbes) no one can alienate her right to defend herself against violence to her body, punitive force is never legitimate. The thinking Agent is merely a fig-leaf that liberals like me use to cover up the naked (nonjustified) force applied against murderers and rapists by imprisoning them.

In explaining the retributivist circuit between crime and punishment, I do not say that the thinking Agent in the criminal's shoes consents to imprisonment. I say that, having acted on a right-denying principle, the criminal cannot complain if his or her own agency is interfered with.[22] What material form the interference takes is something that can be determined only by asking how the conceptual circuit between right-denial and license to coerce is best replicated in the material world; and I argue that imprisonment (as the pure curtailing of liberty) rather than corporal punishment is the mode of punishment most adequate to the retributive idea.[23] Still, Ristroph is correct to infer the criminal's implicit consent to imprisonment from this argument; for having licensed an interference with its agency, the thinking Agent in the criminal's shoes would also license the kind of interference that best replicates in the world what it has licensed in idea.

How can the thinking Agent license the imprisonment of a body it does not have? But it is Ristroph who says that the thinking Agent has no body, not I. I say that the license to interfere with the criminal's agency comes from the thinking Agent "in the criminal's shoes." I explain this image as signifying that the thinking Agent shares the criminal's human condition of finitude, and I specify the meaning of finitude as location in a particular place at a particular time.[24] Only a body can be so situated and so limited. I say that the thinking Agent must be unaware of its principal's subjective ends and have none of its own, but I never say it is noncorporeal and, indeed, say the opposite. All of Ristroph's arguments (and jokes) about the thinking Agent's unrepresentative-ness rest on this misreading.

Slipshod readings underlie Ristroph's further criticisms. She says that I fail to defend my claim that punishment is a necessary implication of crime. I do not defend this claim because it is not a claim I make. In the book's first few pages, I formulate in a still un-nuanced way the condition of just punishment within the formal agency paradigm. I say punishment is consistent with formal freedom just in case "it can be conceived as chosen by the agent . . . as the necessary implication of its actual choice to pre-empt the free choice of another."[25] Ristroph relies on this sentence but

22. PF, at 38–41, 76–77.
23. PF, at 51–55.
24. PF, at 4.
25. PF, at 5.

ignores its later elaboration, qualification, and refinement. In the end, I defend at length two claims about crime's connection with punishment: one, that the criminal's *moral vulnerability* to punishment is a necessary implication of his choice to interfere with another's agency; two, that the public authority has a duty to punish perpetrators of crimes as a general rule but not a duty to punish every single perpetrator of a crime.[26]

Ristroph claims that I equivocate about whether there is any necessity for vindicating law's authority against the criminal's implicit right-denial. I say that law's authority is *inherently* invincible "and cannot be negated by logically absurd claims making pretenses to existential validity."[27] That is why the public authority need not punish in every case of intentional wrongdoing. But I also say that if it did not punish criminals as a general rule, law's authority would be "pervasively unrealized and hence no authority at all."[28] Ristroph thinks that these two statements contradict each other. She is wrong, because there is a distinction between law's inherent invincibility and its actualized authority. Because law's authority is inherently or conceptually invincible, it is not undermined by any particular crime. But obviously, if crimes were generally permitted, law's authority would be unreal and self-contradictory as an authority. Put another way, the proposition that law's authority is invincible against this or that crime is perfectly compatible with the proposition that a generally unrealized authority is no authority.[29]

When Ristroph is not misreading what I wrote, she is caricaturing it. Thus, she writes, purporting to summarize the import of my argument:

> Freedom . . . means one thing for criminals and another for everybody else. For the criminal, freedom is formal agency and is unimpaired by incarceration or even execution. For the potential victim, freedom is real autonomy, with a right to bodily security, the opportunity for actualized choice, and certain material goods. The kind of freedom that is consistent with punishment is the same kind of freedom that is consistent with slavery, which is, again, good enough for criminals but not for the rest of us.

---

26. PF, at 38–41, 45–47, 76–77.

27. PF, at 47.

28. PF, at 47.

29. Ristroph also sees a contradiction between my arguing that an institution of punishment is necessary for law's authority and my denying that it vindicates the right of the individual victim. But that is because she does not understand the distinction between criminal law and the law of torts.

Not one of these sentences bears a relation to what I wrote. Both criminals and their potential victims have rights of formal agency and real autonomy. The criminal's right of formal agency is reflected in the requirement that he be liable to punishment only on the basis of subjective fault and only if his acts publicly manifested a right-denying choice. His right of real autonomy is reflected in the requirement that he be punished only for consequences that are connected in some degree to his choice, and that the measure of his punishment be proportioned to the degree of connection; in this way both the incidence and the measure of punishment are self-authored. I never say anything as silly as that the criminal's freedom is unimpaired by incarceration or execution. On the contrary, I say that these coercive acts are wrongs requiring justification. Moreover, the same principle against the alienability of independence that rules out voluntary servitude for all of us also rules out voluntary submission to punishment by criminals. That is why the law deems criminals not to be consenting to punishment, and it is why, as I said, any justification of punishment must be a justification of a wrong.

But perhaps I am taking the wrong tack in responding to Ristroph point by point. After all, anyone who thinks she has refuted Kant's statement that "if you kill [another], you kill yourself" by pointing to the empirical possibility of killing another without killing yourself is likely tone deaf to arguments pitched at the transcendental register. So, instead of persisting with those arguments, I'll try to show that Ristroph, though a self-professed empiricist of the most radical hue, cannot avoid the metaphysical claims she criticizes in me.

As an alternative to what she sees as Hegel's cover-up theory of punishment, Ristroph offers Hobbes's "honest" one. In Hobbes's account, according to Ristroph, human individuals are "thinking bodies" who voluntarily surrender their natural right of self-governance to a sovereign in order to escape the anxieties that hound them in an anarchic state of nature. Since they thereby authorize all acts of the sovereign, the latter cannot wrong them, and so sovereign power is legally unlimited or despotic. Nevertheless, the sovereign's subjects do not give up their right to defend themselves against violent hands being laid on them, not even if the hands belong to law enforcement agents of the sovereign, for the right to self-preservation is inalienable. Therefore, even if they have murdered or raped, they are at liberty to resist arrest and punishment. This does not mean that the sovereign has a duty not to use violence against them. On the contrary, the sovereign has the same right to preserve itself against the criminal as the criminal has

to preserve itself against the sovereign. With respect to each other, criminal and sovereign are in a state of nature and a state of war.

Observe all the places in this story where metaphysics supplants empirical reality. Hobbes does not claim that the social contract is an empirical phenomenon—that the subjects of authority actually consented to being bound by it. He argues that it is a covenant the thinking body would necessarily enter for the sake of its self-preservation and that everyone must therefore regard himself as having entered. So, the thinking body is everyone's representative. Moreover, whether the criminal wants to consent to his punishment or not, he cannot do so because the thinking body in his shoes would not alienate his right of self-preservation. That right is fundamental to Hobbes's derivation of political authority and thus cannot be treated as an optional end. Finally, to avoid giving its subjects grounds for defection from the commonwealth, the sovereign must in general treat crime as a threat to its self-preservation whatever the empirical realities.

From this we may conclude that the choice Ristroph poses for us is not one between a metaphysical and an empiricist approach to punishment. Rather, it is a choice between two metaphysical approaches. In one, a thinking body submits to despotism for the sake of "commodious living" and peace of mind, trading its vulnerability to individuals roughly equal in strength to it for exposure to a Leviathan, in relation to which it remains in a lawless state of nature. In the other, a thinking but embodied agent recognizes a public authority on condition that authority reciprocally recognize the agent's right to be subject to no law or punishment to which a dignified end could not assent. Would not the thinking body grasp the self-contradictoriness of the first option and rethink its commitment to comfort above dignity?

V

I am indebted to Professor Valverde for the careful reading she has given PF and for the critical insights she has provided from a Derridean perspective. I won't address Derrida's refusal of Hegel's synthetic moves, for that is far too large a topic. Instead, I'll focus on Valverde's specific points.

I am not sure why Valverde says that justice is not directly discussed in PF, or why she mentions the lack of an index entry for "justice." The subtitle of the book is "A liberal theory of penal justice." So that is what

the book is about, and everything in the book directly discusses it. I doubt there is an index entry for World War II in Churchill's history of World War II (but I may be wrong).

Valverde claims that I sweep under the rug the distinction between law and justice. Yet in the Introduction, I discuss the inevitable gap between ideal penal justice and the positive penal law, though my understanding of the gap differs from Valverde's. I say that the unity of the model penal law is not to be expected of positive law, in which there is always a mixture of stable and ephemeral elements; and I say that the divergence between ideal and actual is greater in the penal law than in private law because of the former's greater vulnerability to political demagoguery.

However, Valverde thinks that there is a necessary gap between law and justice in a different sense. If I understand her, she follows Walter Benjamin in thinking that the generality of law precludes it from achieving justice because generality disqualifies law from paying proper respect to the concrete specificities of others. My response is that the penal law has its own way of dealing with the gap between generality and specificity. One way is through the excuses that work as complete defenses after the elements of the crime have been proved. In excusing because of necessity and duress, the criminal law first treats the accused as legally culpable by virtue of his or her choice to interfere with another's agency or property. However, the law then admits to the limitations of its agency-based criteria of culpability, which are divorced from the concreteness of the accused's situation; for it acquits despite legal culpability if the accused lived up to the expectations his fellow citizens could fairly have had of him in the situation he faced. So, here is a case where the penal law, understood broadly as encompassing the virtue-based excuses, achieves justice by recognizing the limitations of law in the narrow sense. My guess is that Benjamin's assertion of a necessary gap between law and justice presupposes an identification of law with law in the narrow sense.

Another point of criticism raised by Valverde is that the book is not as Hegelian as it might seem—that there are Kantian and even Hobbesian elements that are inconsistent with an Hegelian approach. I agree that there are Kantian elements but only to the extent that Hegel's system itself incorporates Kantian elements in the legal paradigm he calls Abstract Right. That does not make my approach Kantian, because no unreformed Kantian could accept the demotion of Kant's formalism to a subordinate element of a whole. Accordingly, one has to distinguish between incorporating Kantian

elements in a way that Hegel himself does (and that are consistent with Hegel's system) and combining or conflating Kant and Hegel. I certainly do the former, but the latter is what I accuse others of doing.

I am baffled by the suggestion that my argument for the reconciliation of punishment and freedom relies on a Hobbesian or contractarian argument. I specifically criticize what I see as the Hobbesian side of Kant—the side according to which the threat of punishment is needed to sustain the civil condition. I do not rely at all on a social contract theory of punishment. My argument is not that punishment is authorized by a social contract; it is that punishment is authorized by the criminal's choice to interfere with another's free will. I would be the last to suggest that Hegel is a social contract thinker.

Finally, Valverde takes issue with my reconciliation of retributive and restorative justice, arguing that they rest on antithetical presuppositions. Retributive justice belongs to the state, she says, whereas restorative justice belongs to subcommunities within society. Of course, antitheses, far from intimidating an Hegelian, attract his intense interest. In PF I argue that, appearances notwithstanding, retributive and restorative justice are mutually complementary. This is so because retributive justice is satisfied if the criminal *could* see his punishment as self-willed; it does not care whether he actually regards his punishment so. Thus, the reconciliation retributivism effects between punishment and freedom is metaphysical, not actual. But that is where restorative justice institutions come in—to reconcile actual perpetrators and victims to the penal process and to each other. However, restorative justice also requires that there be institutions that mete out punishment that only self-respecting ends *could* accept; for otherwise empirical individuals might be pressured into reconciling themselves to punishments and processes that do not treat persons as ends. It would seem, therefore, that the two processes can work together harmoniously.