

Research Report No. 32/2012

Victor's Justice: The Next Best Moral Theory of Criminal Punishment?

François Tanguay-Renaud

Osgoode Hall Law School of York University, FTanguay-Renaud@osgoode.yorku.ca

Follow this and additional works at: <http://digitalcommons.osgoode.yorku.ca/clpe>

Recommended Citation

Tanguay-Renaud, François, "Victor's Justice: The Next Best Moral Theory of Criminal Punishment?" (2012). *Comparative Research in Law & Political Economy. Research Paper No. 32/2012.*
<http://digitalcommons.osgoode.yorku.ca/clpe/35>

This Article is brought to you for free and open access by the Research Papers, Working Papers, Conference Papers at Osgoode Digital Commons. It has been accepted for inclusion in Comparative Research in Law & Political Economy by an authorized administrator of Osgoode Digital Commons.



OSGOODE

OSGOODE HALL LAW SCHOOL
YORK UNIVERSITY

OSGOODE HALL LAW SCHOOL

Comparative Research in Law & Political Economy

RESEARCH PAPER SERIES

Research Paper No. 32/2012

Victor's Justice: The Next Best Moral Theory of Criminal Punishment?

François Tanguay-Renaud

Editors:

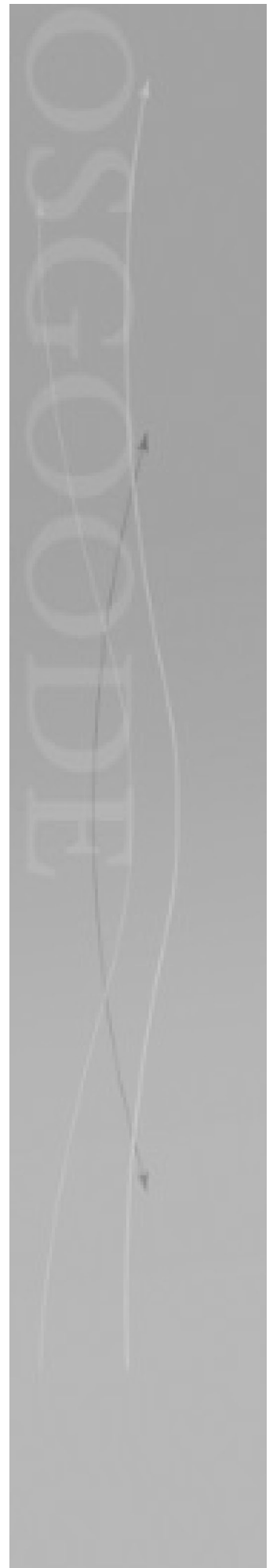
Peer Zumbansen (Osgoode Hall Law School, Toronto, Director,
Comparative Research in Law and Political Economy)

John W. Cioffi (University of California at Riverside)

Leanne Footman (Osgoode Hall Law School, Toronto, Production Editor)



Comparative Research in
Law & Political Economy



Osgoode CLPE Research Paper 32/2012

Vol. 08 No. 07 (2012)

François Tanguay-Renaud

Victor's Justice: The Next Best Moral Theory of Criminal Punishment?

Abstract: In this essay, I address one methodological aspect of Victor Tadros's *The Ends of Harm*--namely, the moral character of the theory of criminal punishment it defends. First, I offer a brief reconstruction of this dimension of the argument, highlighting some of its distinctive strengths while drawing attention to particular inconsistencies. I then argue that Tadros ought to refrain from developing this approach in terms of an overly narrow understanding of the morality of harming as fully unified and reconciled under the lone heading of justice. In a final and most critical section, I offer arguments for why this reconciliatory commitment, further constrained by a misplaced emphasis on corrective justice, generates major problems for his general deterrence account of the core justification of criminal punishment.

Key words: moral justification of punishment; state punishment; criminal law; justice; corrective justice; general deterrence; incommensurability

VICTOR'S JUSTICE: THE NEXT BEST MORAL THEORY OF CRIMINAL PUNISHMENT?

François Tanguay-Renaud*

Forthcoming: Law and Philosophy (2012)

Victor Tadros's *The Ends of Harm (TEH)*¹ is exceptionally rich in provocative novel insights. In this essay, I primarily focus on one aspect of the book's overall methodology, namely, the resolutely moral character of the theory of criminal punishment it defends. In section I, I offer a brief reconstruction of this dimension of the argument, emphasising some of its distinctive strengths while drawing attention to particular inconsistencies. In section II, I contend, more critically, that Tadros ought to refrain from developing this approach in terms of an overly narrow understanding of the morality of harming as fully unified and reconciled under the lone heading of justice. As I argue in section III, this theoretical commitment, constrained by a surprising emphasis on corrective justice, generates significant problems for his account of the core justification of punishment.

I. A MORAL THEORY OF CRIMINAL JUSTICE

Tadros spends most of *TEH* developing an argument for what is first and foremost a moral, as opposed to a political, theory of criminal punishment. Thus, he implicitly rejects the oft-encountered claim that sound argument about criminal law and punishment must remain confined to the realm of state and law—or, more broadly, to distinctively political theorizing.²

His argument builds on the assumption that paradigmatic crimes involve culpable wrongdoing that causes harm to one or more individual victims. If murders, rapes, assaults, burglaries, frauds, and torture are punishable crimes, it is primarily because they harm

* Associate Professor at Osgoode Hall Law School, and Member of the Graduate Faculty of the Department of Philosophy, York University, Toronto. Thanks are owed to Mohamad Al-Hakim, Vincent Chiao, and Hamish Stewart for discussing and challenging many ideas contained in this essay in the context of a reading group focusing on *The Ends of Harm*. I would also like to thank Vincent Chiao, Kimberly Ferzan and Victor Tadros for their written comments on an earlier version.

¹ Victor Tadros, *The Ends of Harm: The Moral Foundations of Criminal Law* (Oxford: Oxford University Press, 2011).

² Cf. Arthur Ripstein, *Force and Freedom: Kant's Legal and Political Philosophy* (Cambridge, Massachusetts: Harvard University Press, 2009), ch.10; Malcolm Thorburn, "Constitutionalism and the Limits of the Criminal Law", in R.A. Duff, Lindsay Farmer, S.E. Marshall, Massimo Renzo, and Victor Tadros (eds.), *The Structures of the Criminal Law* (Oxford: Oxford University Press, 2012), pp. 85-105, at pp. 87-88, 97-101; George P. Fletcher, *The Grammar of Criminal Law: American, Comparative, and International: Volume One: Foundations* (Oxford: Oxford University Press, 2007), p. 154.

individuals, in violation of our duty not to harm others. And like other duties that ground his overall theory—such as the duty not to use people as means, and to rescue them from harm—the duty not to harm others is fundamentally moral, in the sense of being pre-political. The criminal law merely builds on it, specifies it, and coordinates its enforcement. Thus, what *TEH* seeks to offer is a justificatory story about criminal law and punishment that starts from first principles, and resists any temptation to run ahead hurriedly to a later ‘political’ point, artificially segregated from the deep and complex reality of morality, broadly understood.

And deep and complex morality is, with the morality of harming forming no exception. So complex, in fact, that Tadros concedes that even able philosophers will often have a hard time getting it right in the details. For example, if, as he assumes, it is generally wrong to intentionally harm someone in a way that disrespects her status as an autonomous agent, one may wonder whether some motivations for harming people may mark sufficient respect for this status to make harming permissible. Tadros thinks there are such motivations. Thus, he holds that it is permissible for the driver of an unstoppable trolley to divert it in a way that kills one person instead of the five towards which it was originally heading—insofar, he insists, as the driver diverts the trolley on the basis of the appropriate status-conferring features of the five people saved. Yet, he is unsure what specific status-conferring features the driver must have in mind when diverting the trolley, as well as how exactly these features must figure in her motivational make-up. The guidance he provides is minimal, and mostly illustrative. So, he tells us, diverting the trolley in order to kill the one, for personal glory, or to preserve those whom the driver merely understands as rare flowers or God’s artworks would not pass muster. Remarkably, though, instead of seeking to provide a deeper explanation, he then quickly concedes that questions about permissible motivations and value-conferring features of human beings are notoriously difficult and subject to profound disagreements. He even asserts that these disagreements may be such as to require “criminal law and moral judgement” to “come apart sharply.” Unlike morality, he continues, “The criminal law is not in the business of investigating the full motivational structure that guided the defendant’s decision in deciding to do what he did.”³

Coming from someone who promises a rather continuous story between the ordinary morality of harming and its instantiation in the institutional context of the criminal law, this suggestion comes as somewhat of a surprise. If deep motivational make-ups are part of the structure of moral wrongs like murder and assault, and criminal law is concerned with the condemnation and punishment of some such wrongs, should Tadros not strive to provide a clearer moral picture? Indeed, isn’t such a picture necessary if grave moral errors of judgment, condemnation, and punishment at the hands of the criminal law are to be avoided? In *TEH*, Tadros sidesteps the challenge by taking hasty refuge in

³ Tadros, *The Ends of Harm*, *supra* note 1, at p. 166.

institutional considerations: of neutrality of the state in the face of deep disagreements, of individual privacy in the face of the might of the law, and of conflicting functions of criminal justice including not only condemnation of wrongdoing, but optimal deterrence. Yet, the nature of his project begs the question of whether there is a more basic, pre-institutional moral story about the proper limits of the relationship between a person's deep motivations and wrongdoing. Unfortunately for Tadros, his failure to investigate this possibility, and his readiness to justify criminal law in terms of separate institutional considerations, expose him to attacks from theorists who are determined to affirm that criminal law theory must have a political, as opposed to a more general moral, starting point.

I say “unfortunately” because, in spite of sporadic inconsistencies such as this one,⁴ *TEH* evidences an outstanding effort to uncover the deep objective morality in which criminal law and punishment may find their basic justification. Of course, when it comes to harm, so-called political considerations may never loom far behind. After all, the justification for the modern state is commonly thought to rest, at least in part, on its ability to shield its citizens from harm. At one point, Tadros even provocatively muses that there may be a special, state-based “duty of equality that we owe our fellow citizens [and which] gives rise to a stronger duty to punish than the duty that is based simply on the universal duty to rescue,” in which he otherwise grounds his argument in *TEH*.⁵ Still, it is universal moral duties that take the center stage in the book.

As Tadros understands them, universal duties apply both outside and inside the context of the state. They also apply to the state just as they apply to ordinary individuals—a contention that, despite what some politically-leaning theorists may claim, is all but mysterious. If, as Tadros suggests, modern states essentially boil down to large groups of individuals acting in concert to fulfill certain tasks, duties on the state can simply be understood as duties that these individuals share amongst themselves.⁶ One might legitimately question whether such radical methodological individualism about the state does not leave out an important part of the moral picture. I am alluding here to the part according to which the modern state may also be a corporate moral agent in its own right, with moral duties that it may violate in ways that are irreducible to violations by its individual members.⁷ Still, according to both of these accounts, Tadros's claim that applications of universal moral duties to the state are merely particular applications—no different in kind from applications to ordinary individuals—remains highly plausible.

⁴ Tadros is also quick to help himself to institutional considerations when arguing against the permissibility of manipulatively killing wrongdoers to use their organs to cure the sick. *Ibid.*, at pp. 307-311.

⁵ *Ibid.*, at p. 311.

⁶ See *ibid.*, at p. 299.

⁷ See especially Christian List and Philip Pettit, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (Oxford: Oxford University Press, 2011), chs 1, 7, 8.

Of course, given their typically greater resources and claims, states and their institutions may have more extensive or stringent moral duties than ordinary individuals. The rule of law, with its associated requirement that the conditions under which a person may be punished be publicized *ex ante*, is a good example. On a desert island where you would assault me repetitively until I die, unless I deter you through punishment, it would no doubt be permissible for me—or a third party who arrives on the island—to punish you, even though the rule of law has not been established. Yet, “[i]n modern states, [...] where it is relatively easy to establish the rule of law, it is typically wrong to punish a person who has not violated the criminal law.”⁸ The moral importance of choice in culpable wrongdoing, and the modest costs that states must typically incur to ensure that people are provided, through prior notice, with a proper opportunity to choose to avoid punishment, militate strongly in favor of this requirement. The point is not that rule-of-law duties only apply to the state and not to ordinary individuals, such that we ought to speak of qualitatively distinct duties—perhaps even of a purely ‘political morality’—applicable to the state. As Tadros recognizes, “there may be strong reasons on the desert island to begin to develop the rule of law”—such as to ensure, when the costs are reasonable, that non-serial would-be wrongdoers are given a proper opportunity to avoid punishment.⁹ The point is, rather, that while universal moral duties may differ in scope and intensity in their various applications to ordinary individuals and states, they do not differ in kind. And here, the rule of law constitutes merely one illustrative example, alongside other universal duties such as the duty not to harm others, the duty to rescue them from harm, and so forth.

There should be no surprise here, since states, *qua* organized groups of individuals, are not all or nothing affairs, which, at a given stage of their development, mysteriously leave ordinary morality behind to enter a different normative realm. Instead, states tend to come together and evolve gradually, often with setbacks along the way—be it in respect of their *de facto* authority, their resources, the level of commitment of their individual members, or the complexity and efficacy of their institutional structures. One of the great strengths of Tadros’s approach is that it casts into sharp relief how morality can adapt itself to the circumstances and stages of development of any given state, just as it adapts itself to the circumstances and stages of development of ordinary individuals.

No doubt, political theory, understood as that particular branch of moral theory that focuses on the normative position of the state, cannot be ignored. To make complete sense of, and assess, criminal law and criminal punishment as we know them—namely, as exclusive instruments of the state—we also need to understand the reasons, including the duties, that we may have to defer and contribute specifically to *state* action in this sphere. Yet, one must strive to avoid mistaken political theoretical currents. For example, one should not assume that, because morality is sometimes the object of deep disagreements,

⁸ Tadros, *The Ends of Harm*, *supra* note 1, at p. 317.

⁹ *Ibid.*, at p. 316.

what the state decides must ultimately carry the day.¹⁰ As Tadros suggests, a helpful way of averting this kind of mistake is to conceive of the state as an instrument whose primary purpose is to assist individuals in realizing moral duties that they would not be able to fulfill, or fulfill as well, without it. When the nature or contours of such duties are controversial, the state may assist us by specifying them and their ramifications, as well as by coordinating their application, through its law. In fact, the state may often be under a rule-of-law duty to do so. At times, the state may also facilitate our conformity with duties by undertaking to fulfill them on our behalf. Still, state action is always open to moral scrutiny. The state may get things wrong, and may sometimes get them so wrong that it would be better for us to disregard completely its claimed attempts at assistance.

So, Tadros argues, insofar as culpable wrongdoers have an enforceable duty to submit themselves to punishment, insofar as we all have a duty to see to it that they are punished as a means of deterring further wrongdoing, and to the extent that the state can perform such actions more justly and effectively than individuals, it may well be that we should defer (and, in many cases, contribute) to it doing so. Yet, the possibility always remains that a state's punishment practices are so unjust or so ineffective that "individual punishment will be [...] preferable to state punishment," because it is more just or more effective.¹¹ Tadros enjoins us to exercise caution. It will likely be very unusual for ordinary individuals' acts of punishment to be more just and more effective than states' punishments—even those of relatively unjust states—given the significantly more limited means of inquiry, publicity, sanction, etc., at individuals' disposal. Moreover, given the importance of long-term stability to a just society and the threat that widespread individual acts of punishment may pose to it, the threshold of state injustice and ineffectiveness for such acts to be justified will generally be exceptionally high. Here, one might wonder whether Tadros does not set this threshold too high: does he not underestimate that destabilization is sometimes required to remedy important state wrongdoing and foment the basis for justified rebellion? Regardless, his deeper point remains. Like inflictions of harm by ordinary individuals, state-imposed harm in punishment requires general moral justification.

II. THE GRAND AMBITIONS OF VICTOR'S JUSTICE

As morally discerning as Tadros's theoretical approach may be in *TEH*, as evidenced by the countless practical distinctions on which he is able to shed light, there is at least one general aspect in which it is not. I am referring to his tendency to articulate his argument in terms of one overarching moral metric, namely, justice. This tendency, reminiscent of grand value-monist theories that assume the pervasive possibility of correct moral answers to apparent moral dilemmas and uniquely optimal social arrangements, generates some awkward tensions in his account. For one thing, even if injustice is often wrong—such that

¹⁰ Cf. Ripstein, *Force and Freedom*, *supra* note 2, ch. 11.

¹¹ Tadros, *The Ends of Harm*, *supra* note 1, at pp. 306-307.

it is “normally wrong to breach a just regulation”¹²—not all wrongdoing is an injustice. What is wrong with rape, torture, or enslavement is normally not that such forms of conduct are unjust, at least if we adhere to the common understanding of the domain of justice as concerned with allocative moral questions, such as who is to get what and how much of it.¹³ Indeed, there are normally no questions, either of a comparative or non-comparative nature, to be asked about who should be raped, about the proper grounds of enslavement, or about how much one should be tortured. There are normally no such questions because these forms of conduct are simply inhumane and, except perhaps in some rare tragic scenarios, the mere asking of allocative questions about them denotes moral confusion or, worse, grave insensitivity.¹⁴ Here, I am not accusing Tadros of inhumanity, since he himself contends that what is fundamentally wrong with, say, rape is that it uses an individual in a way that disrespects her intrinsic value as an autonomous being.¹⁵ Rather, I am rejecting the propensity of what I will call “Victor’s justice” to conflate distinct moral issues, such as sheer inhumane treatment and morally deficient allocations.¹⁶ In fact, Victor’s justice is so capacious that one could perhaps be forgiven for mistakenly confusing it with morality writ large.

I say that one would be mistaken in straightforwardly equating the two, since, in *TEH*, Tadros is mostly concerned with duties which, although central to morality, constitute only one type of moral reasons—alongside, for example, *pro tanto* moral reasons. Still, the moral ambitions of Victor’s justice are sweeping. “[A] just world,” he insists, is really nothing less than “a world where everyone abides by their moral obligations.”¹⁷ Statements of this sort are the source of much obfuscating inflation in *TEH*. A world in which all would abide by their moral duties not to rape or torture would certainly be a more humane world and, *ceteris paribus*, a morally better world. Yet, it would not necessarily be a more just world since, as I suggested, conformity with such duties generally involves no allocative moral issues. Even John Rawls, who makes more of justice than most in his work, is not as expansive as Tadros in his views. That is, even he makes clear that there is value in

¹² *Ibid.*, at p. 323.

¹³ See e.g., Leslie Green, “The Germ of Justice”, Oxford Legal Studies Research Paper No. 60/2010 (2010), SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1703008; John Gardner, “What is Tort Law For? Part 1. The Place of Corrective Justice” *Law and Philosophy* 30 (2011): pp. 1-50 at pp. 6.

¹⁴ The rare tragic scenarios I have in mind are those in which, for example, the only possible way to prevent people being raped or enslaved is to rape or enslave others, or where the only possible way to prevent someone being tortured a lot is to torture another to a lesser extent.

¹⁵ See e.g., Tadros, *The Ends of Harm*, *supra* note 1, at p. 128.

¹⁶ See e.g., *ibid.*, at p. 105, where Tadros explicitly describes a failure to recognize someone’s status as a moral agent when harming them as an injustice.

¹⁷ *Ibid.*, at p. 291.

conceiving of justice as a distinctive virtue, despite famously considering it to be “the first virtue” of social institutions.¹⁸ Thus, for Rawls, a perfectly just society is not a society in which everyone abides by their moral duties writ large, but a society where duties of justice are strictly complied with.¹⁹ True, like Tadros’s understanding of justice, Rawls’s is exclusively deontic, and neglects the possibility that we might have reasons to be just over and beyond the call of duty. Unlike Tadros’s, though, his understanding has the advantage of leaving the door open for a somewhat more textured and nuanced understanding of morality (despite its more institutional focus). It allows for a morality in which both justice-based and non-justice-based duties, such as duties of sheer humanity, may coexist and, sometimes, even conflict irreconcilably—say, in tragic situations where the only way to be just would be inhumane, or the only way to be humane would be unjust.

While some may want to dismiss this point of disagreement as a mere conceptual quibble, it does color Tadros’s entire argument. For example, he contends that when William successfully assaults David, what is saliently wrong about his conduct is that he has harmed David unjustly.²⁰ Insofar as an assault is partly constituted by the harm it occasions, one can see why it may be tempting to understand its wrongfulness along the spectrum of injustice. After all, one may think, harm, understood as a morally significant setback to interests, might be allocated. It might be allocated by courts which treat it as a cost to be endured by a specific individual or shared by many. It may also be allocated by soldiers who impose it on enemies. Even ordinary self-defenders might allocate it when they opt to harm their assailants to avoid harm to themselves. Still, when harm is incurred as a result of human agency, and a suggestion is made that it was occasioned by wrongdoing, a question always arises about the kind, or ground, of wrongdoing at issue. Since the breach of a duty of justice is only one possible such ground, one might then query whether it is really the best way to understand the wrongfulness of assault. Should the wrongfulness of a random assault on the street, as in the case of William’s attack on David, really be distilled to an allocative defect, either primarily or at all? The fact that many assaults—understood as non-consensual intentional impositions of force—do not even result in significant harm certainly ought to give one pause.

Tadros’s uncompromising insistence on framing such issues as issues of justice is question-begging. Why persist in defending such a generic categorization even in the face of counter-intuitive implications? One way to go about answering this question is to inquire into the place of justice in his argument. What seems clear is that Tadros’s overall project is not to produce a refined treatment of the virtue of justice and its correct application. While the concept is omnipresent, it is not subject to any sustained analysis. Rather, Tadros seems intent on developing a general framework for assessing the moral permissibility and obligatoriness of deliberate inflictions of harm, both at the hands of ordinary individuals and states. While the ultimate focus of the argument is on punishment, other inflictions of

¹⁸ John Rawls, *A Theory of Justice*, Revised Edition (Oxford: Oxford University Press, 1999), at p. 3.

¹⁹ *Ibid.*, at pp. 7-8.

²⁰ Tadros, *The Ends of Harm*, *supra* note 1, at p. 284.

harm in self-defense, armed combat, compensation, and so forth, also fall squarely within its remit. Thus, it is not too farfetched to posit that what Tadros primarily wants to demystify in *TEH* is that part of morality that bears on individual and state-based allocations of harm. In fact, his goal seems to be to elucidate it as a unified and seamlessly coherent allocative system, of which “criminal justice” only forms one systemic subpart continuous with the rest.

Many aspects of Tadros’s argument are consistent with this understanding of his project. For example, in light of the tremendous social resources and other human costs required for the development and maintenance of modern institutions of criminal law and punishment, he insists that their justification must be assessed in light of possible alternatives. “Suppose,” he ponders, “that expanding the system of compensation would make people better off than the criminal justice system, using the same resources.”²¹ Were this to be case, he suggests, institutions of criminal law and punishment would likely become impermissible as channels of harm allocation. For Tadros, this direct comparison is morally intelligible, because the imposition of harm in punishment is only one possible way of remedying harm that wrongdoers inflict on others.

No doubt, punishment is a distinctive form of remediation. Punishments aim to inflict harm, and the harm aimed at is often particularly significant. Punishments also tend to have a strong condemnatory aspect. Thus, the practice is inevitably harsh, and its justification is generally held to require outstandingly robust safeguards for those who may suffer it.²² Indeed, as Tadros stresses, we tend to think of punitive allocations of harm to innocents who have no obligation to incur it, or to wrongdoers in disproportion of their deeds, as especially unjust.²³ It is unsurprising, then, that he holds punishment to be justified only when other forms of remediation would be insufficient or unavailable, and the moment of justified pre-emptive defense against harm has passed. Still, he insists that punishment primarily constitutes a means of remedying harm that must be understood along the same moral continuum as compensation. Punishments remedy harms wrongfully inflicted by harming wrongdoers in ways that prevent further impositions of such harms on their victims. I will come back to this part of the argument in the next section. For now, my aim is simply to make plain Tadros’s overall goal of accounting for criminal punishment as one tool amongst many for allocating harm, as part of a wider attempt to systematize the whole morality of harming under the umbrella of justice.

This novel approach has important costs, one of which is what is lost, and perhaps confused, by assuming that all strands of the morality of harming are ultimately concerned

²¹ *Ibid.*, at p. 344.

²² *Ibid.*, ch. 14.

²³ See e.g., *ibid.*, at pp. 138, 314.

with, and can be reconciled under, the value of justice. What about the possibility that competing harm-related desiderata—say, humanity and justice—may, as suggested earlier, conflict in incommensurable ways? Does Tadros's unifying tendency not threaten to occult such critical conflicts? No doubt, there is a certain explanatory appeal to the tidiness of an approach according to which there is only ever one big moral puzzle to be assembled, as opposed to countless smaller ones that may not fit together neatly. Still, one cannot help but worry: does such an approach not run the risk of creating a false impression that perfect conciliation of competing values is always possible? At the broader level of social arrangements such as those of the criminal law, does it not run the risk of nurturing illusory promises of more uniquely optimal institutions and practices than are morally achievable?

I am unsure whether Tadros truly conceives of the morality of harming as a seamlessly unified system in which practical conflicts and dilemmas are only superficially apparent, and ultimately resolvable in a uniquely correct way. That said, in *TEH*, Tadros's justice-driven reconciling instincts are dominant. Consider Kant's oft-cited hypothetical scenario in which two equally-situated shipwrecked sailors struggle to survive by climbing on the sole remaining floating plank, which can only support one of them. One might think that due to a conflict of incommensurables—by hypothesis, two equally-situated lives—morality's guidance remains indeterminate as between them. Either each sailor is tragically bound to let go of the plank or act wrongly, all things considered, or each is equally permitted to use pre-emptive force against the other. Tadros denies the existence of any such irresolvable symmetrical tension, insisting that a uniquely just resolution must be possible. What morality requires, he argues, is "to find a fair mechanism" for resolving the apparent conflict, "such as tossing a coin."²⁴ One might question whether resort to chance by flipping a coin is really more just than leaving the outcome to be decided by the relative strengths bestowed on the protagonists by the fortunes of nature. For all I know, with some argument, Tadros might even be convinced that it is not, and that his egalitarian intuitions can at least sometimes be accommodated in the latter case. For example, it could be argued that luck often has as much to do with one's physical endowments as with the result of a coin toss. However, my main point is that, to persuade Tadros, the argument would likely have to be articulated in terms of the answer that justice requires, either substantively or procedurally.

I say "likely," because Tadros's elevation of justice to the rank of master moral value is perhaps not as absolute as I make it to be. Justice is certainly a touchstone of his understanding of morality, but he also concedes at various points that it may have to take a back seat if something like a deadly nuclear explosion were to result from strict adherence to it. Then again, Tadros may simply be a threshold deontologist who, barring the heavens falling in such a way, is absolutely committed to a justice-based morality. Admittedly, he does sometimes muse about the "rare" possibility of less extreme practical conflicts, involving the possibility of harming, for which there is no uniquely correct moral resolution. Such conflicts may exist when they are "either just" in themselves or, more

²⁴ Tadros, *The Ends of Harm*, *supra* note 1, at p. 216.

revealingly, when they are “in another way valuable.”²⁵ The sport of boxing may be such a case, given its instrumental entertainment value, and the intrinsic value residing in the displays of skill and strength involved. Still, one wonders whether the sport’s requirement that all boxers consent, within defined parameters, to being harmed may not allow him to recharacterize boxing as a *just* form of entertainment or as a *just* display of forceful skills—were he to be pressed on this issue in light of the current discussion.

So it is unclear just how absolutely committed Tadros is to justice as the ultimate moral metric. What seems incontrovertible, though, is that he is strongly committed to articulating his argument in broad unifying terms. So far, I have conveyed my discomfort with this approach fairly generally. I now want to lay out some concerns specifically directed at how he develops his theory of punishment in light of this commitment.

III. THE NEXT BEST JUSTIFICATION OF PUNISHMENT?

Tadros labels his theory of punishment “the duty view” because of its grounding in the duties that he understands wrongdoers to incur as a result of their wrongdoing. The theory takes roots in a self-defensive scenario akin to the following.²⁶ Suppose that Anna wrongfully hires a hit man to shoot off Blake’s right arm and leg. Tadros argues that it would be permissible for Blake to pull Anna in front of him to use her as a shield against her hit man’s bullets, insofar as this maneuver is necessary for him to avoid being seriously harmed. Blake is permitted to do so because Anna has a general duty not to harm him, and the only way for her to fulfill that duty is now to shield him from the threat she culpably created. Normally, Tadros concedes, it is impermissible intentionally to use someone else as a means to a further goal, such as self-protection—and perhaps even more starkly so if that someone is to be seriously harmed in the process. However, he does not think this categorical prohibition is exceptionless: it does not extend to cases where the person to be harmfully used already has an enforceable duty to fulfill the further goal in question, such as an enforceable duty to provide protection.

Tadros maintains that Anna’s duty to shield Blake from her hit man’s bullets is one such enforceable duty, given the seriousness of the harm that Blake would incur were she to fail to discharge her duty, and the improbability that he could be compensated adequately *ex post facto*. To be sure, while innocent bystanders may sometimes have enforceable duties to rescue people from harm, such duties tend not to extend to shielding people from bullets. They do not extend to such heroic feats because of the considerable personal costs—including harm to self—that are likely to be incurred in the process. However, a culpable wrongdoer like Anna generally cannot complain of such costs. She

²⁵ *Ibid.*, at p. 214-215.

²⁶ See especially *ibid.*, ch. 12.

cannot complain because these costs would be incurred as part of preventing the realization of a threat that she had a duty not to create, chose to create, could have avoided creating at no moral costs to herself, and now has a derivative duty to negate—barring which, the person to whom the duty is owed will suffer serious harm, unlikely to be adequately compensated. In other words, once Anna cannot call off the wrongful hit that she culpably ordered, she is under a duty to do the next best thing—namely, to shield Blake from it—and Blake is permitted to harm her to enforce it.

Now, suppose that the hit man is successful and Blake's right arm and leg have been shot off. Tadros maintains that Anna's duty not to harm him is not extinguished as a result. Rather, her duty becomes a duty to do what is now the next best thing. What is that? Tadros's answer, which lies at the core of his theory of punishment, is rather unexpected. Suppose that Claire attacks Blake and is about to shoot off his other arm and leg. Tadros thinks that since Anna now cannot stand in the way of, or otherwise undo, the harm she occasioned, she owes it to Blake to protect him from relevantly similar harming. Thus, protecting Blake from Claire's attack is the next best thing Anna can do. It is the next best thing, since it would put Blake in as close as possible a position as the one in which he would have been had Anna initially complied with her duty. Had Anna not ordered the shooting of Blake's right arm and leg, he would only have incurred harm to his left arm and leg at Claire's hands. Thus, all else being equal, by protecting Blake from Claire, Anna ensures that Blake will only have incurred harm to one arm and one leg, almost as if she had not attacked him.

Here, Tadros again argues that were Anna to be unwilling or unable to rescue Blake in this way, the latter would be permitted to harm her to get her to do her duty to protect him. He would be permitted to do so given the enforceable character of Anna's duty to bear that degree of harm—assuming, once again, that performance of her duty is necessary to protect him from great harm and that other compensation would be either inadequate or unlikely to be forthcoming. In other words, Blake would be permitted to harmfully use Anna for protection if that were the only way to remedy, or rectify, the harm that she wrongfully caused to him. Tadros presents this explanation as the foundation of what he calls his remedial justification of punishment. Punishment, one may think, consists in the intentional infliction of harm to someone in response to wrongdoing (or, at least, to what is alleged, even if disingenuously, to constitute wrongdoing). At the core of Tadros's argument is the thought that such a practice can be justified as a means of compelling actual wrongdoers to remedy the harm they culpably occasioned to their victims.

As I said, though, the decision to ground this argument in the principle of "next best" conformity to reasons is surprising. It is surprising since, historically, champions of this principle of rationality have been staunchly opposed to its extension to punishment. Articulated in deontic terms, to mirror Tadros's own understanding, the principle stands for the proposition that when we fail to conform to a duty that we owe someone, and it is too late to reverse course and conform, we still owe that person something. We owe it to her to put her back, so far as we can, in the position in which she would have been had we conformed with our primary duty. To paraphrase Joseph Raz, we incur a secondary duty to compensate her—insofar as compensation is understood as rendering an equivalence, or

as near an equivalence as can be, to the performance of the primary duty.²⁷ Thus, when the primary duty breached is a duty not to harm others, a question always arises about how the duty violator (aka wrongdoer) can best compensate his victim for the harm she incurred. This question is a question of corrective justice—that is to say, a question about whether and how the wrongdoer should seek to “make-up” for his failure to conform to the duty he owed his victim and, thus, attempt to counteract his deed. This way of thinking about punishment fits seamlessly within Tadros’s overall justice-based approach. Still, any attempt to characterize the justification of punishment as a question of corrective justice encounters powerful objections.

The most oft-cited objection is that, unlike compensation, punishment is not something that the wrongdoer owes, because it is not something that the wrongdoer can give. In John Gardner’s words, “It is something that is inflicted upon him by others, and the norms regulating it belong, in the final analysis, to their normative position and not to his.”²⁸ This punishment-cannot-be-given (PCBG) objection is overkill, and must be considerably refined if it is to have any bite against Tadros’s argument. For example, one may conceivably give oneself punishment for breaching a self-regarding duty. Consider the case of Deepti who punishes herself for a broken self-made promise to study diligently for a failed exam by cancelling all her subsequent recreational outings. One may also give oneself punishment for breaching duties owed to others. Thus, a voluntary jump by Anna in the way of Claire’s bullets can easily be interpreted as Anna punishing herself for breaching her duty not to harm Blake. The PCBG objection may seem to have more teeth in this latter case. Even if we concede that Anna can give herself punishment by jumping in the way of Claire’s bullets, isn’t this punishment, by definition, aimed at herself, so that it cannot be given to the person to whom she owes compensation? Not so fast. Insofar as Anna punishes herself as a means of shielding Blake from Claire’s bullets, her self-punishment does give Blake something meaningful: it gives him protection against wrongful harm he would otherwise incur. So, Anna’s punitive sacrifice may be intended as, and yield, compensation for Blake.

An intuitively more promising focus might be on cases in which the wrongdoer is unwilling or unable to punish herself, so that punishment can only be inflicted on her by a third party. Such cases are the ones that Tadros ultimately wants to justify, since they are relevantly closer to state-imposed criminal punishment. In such scenarios, one might think, punishment is, by definition, not given, but forced upon the wrongdoer. Here again, though, the objection fails since the same might be said of other forms of compensation that are

²⁷ Joseph Raz, “Personal Practical Conflicts”, in Peter Baumann and Monika Betzler (eds.), *Practical Conflicts: New Philosophical Essays* (Cambridge: Cambridge University Press, 2004), pp. 172-194, at p. 191, fn 25.

²⁸ John Gardner, “Punishment and Compensation: a Comment”, in Russell Christopher (ed.), *George P Fletcher’s Essays on Criminal Law* (Oxford: Oxford University Press forthcoming), online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1837373.

typically not conceived as punitive, such as compensatory pecuniary damages in tort law. Civil defendants may be equally unwilling to pay the pecuniary damages they owe in compensation. When this happens, their payment of damages may be enforced, say, by garnishment of their earnings or bank accounts. The relevant point is that, in both cases, individuals have an enforceable duty to provide compensation in a given way, which undermines the complaint they would otherwise have about being harmfully used for this further goal. Civil defendants must pay the compensatory damages they owe, failing which they may suffer seizure of their assets. Similarly, Tadros tells us, culpable wrongdoers unwilling or unable to prevent further harm to their victims have an enforceable duty to submit to being harmfully used for their protection.

The PCBG objection is only successful, I think, if we understand it as taking issue, more modestly, with Tadros's wider claim that the protection that a wrongdoer owes his victim may be provided by the general deterrence effect of his punishment. His claim takes the following form. If Blake is permitted to harm Anna punitively by forcing her to shield him physically from Claire's bullets, he is also permitted to harm Anna in different ways that would deter Claire from attacking him altogether. For example, Blake may punish Anna in ways that will lead Claire to fear sufficiently that he would punish her in the same way, were she to attack him, that she will opt to desist from it. The permissibility of such deterrent punishment seems to follow seamlessly from the permissibility of punishment in the forced shielding scenario. Yet, Tadros's claim goes further. He contends that if Blake's punishment of Anna is not aimed at deterring a specific wrongdoer like Claire, but at anybody who, more generally, may wrongfully harm him in a similar way in the future, then he is also permitted to punish her for this reason. He may punish her as a means of *general deterrence*.

As Tadros notes, the question of general deterrence most often arises in the context of state punishment. That is because general deterrence relies on punishment being public, and states are typically more effective than individuals at ensuring that punishment reaches a wider audience and that this audience believes that it has the power to punish them for future wrongdoing.²⁹ This point forms part of Tadros's wider argument that, although punishment of a culpable wrongdoer is centrally owed to those whom he has wrongfully harmed, the state is often permitted, and best placed, to act as punisher on their behalf. It may so act either with victims' consent or without it, to the extent that victims have a duty of self-respect to consent to it, or a duty to rescue others from wrongful harm by allowing the state to deter other wrongdoers through punishment (at little cost to the victims themselves).³⁰ I do not wish to say any more about these arguments, or about Tadros's more complex arguments concerning states' duties to punish wrongdoers and the precedence of their permission to do so. My aim here is merely to note that they all rest primarily on the permissibility of punishing culpable wrongdoers to deter generally further wrongful harm to their victims.

²⁹ Tadros, *The Ends of Harm*, *supra* note 1, at p. 304.

³⁰ *Ibid.*, at pp. 294-299.

Indeed, for Tadros, the protection afforded to victims by the general deterrence effect of state punishment lies at the core of the justification of this institutional practice. In most contemporary societies, he assumes, this effect plays a “powerful” and “significant” role in ensuring that people do not develop inclinations to wrong others.³¹ Without it, “many more people” would develop such inclinations, resulting in much more harmful wrongdoing.³² The problem with such an undefended empirical claim is that it leaves the reader to wonder how measurable the general deterrence effect of specific acts of punishment really is and, insofar it can be measured, how much each punishment really contributes to general deterrence. We may agree with Tadros that the practice of state punishment has at least some general deterrence effect in society. However, if we cannot at least roughly identify how much deterrence, and deterrence of whom, is generated by specific punishments of particular wrongdoers then we cannot assess whether these punishments constitute sufficient remediation for their victims. Although primarily empirical, this problem has serious ramifications for the appeal of Tadros’s theory. If, according to the core justificatory case for the punishment of wrongdoers, a punitive act is permissible only when necessary to prevent serious harm to their victims, then we need a plausible way to assess such necessity. Otherwise, we end up with a very abstract, not to say wishful, justificatory case for punishment, not unlike cases grounded in the idea of desert, which Tadros otherwise decries on similar grounds. One way to get around this problem may be to contend that it is the deterrence effect of the practice of punishment as a whole—in which, past a certain point, individual acts of punishment may or may not play a role—that ought to be the locus of assessment. However, this concession would entail that state punishments are not, as Tadros argues, to be justified as ways for *each* punished wrongdoer to correct his breach of duty, since no actual deterrence may be provided—or “given”—by his specific punishment.

Insofar as Tadros wants to resist this concession, he must answer the following question: do states have a potentially very onerous duty to develop effective systems to assess the probable deterrence effect of *each* punishment they imposed, and to refrain from punishing whenever this effect is negligible? I believe he would be reluctant to embrace such a possibility, given the considerable other social costs which, he repeatedly emphasizes, must already be expanded to sustain modern institutions of criminal punishment, and the difficulty of monitoring with precision the many subtle and long term effects of these institutions. Rather, his response in *TEH* seems to be, in part, that, although general deterrence is the only thing that could plausibly justify using significant state resources to develop and maintain institutions of state punishment, it does not provide an exclusive justification for such punishment. Once such institutions are set up, it then becomes permissible to harm wrongdoers in punishment for the further goal of leading

³¹ *Ibid.*, at p. 281.

³² *Ibid.*, at p. 282.

them to recognize the wrongfulness of their conduct.³³ Thus, recognition may pick up as a justification where deterrence is missing.

This rejoinder is problematic. First, it is unclear that institutions of state punishment can even be calibrated to ensure that most, let alone all, punitive acts generate sufficient general deterrence to discharge punished wrongdoers' specific duties of protection to their victims. If that is true, then state punishment is markedly deficient as a means of *inter partes* compensatory remediation. So, recognition may have a lot more justificatory slack to pick up than it can handle—or, at least, that Tadros thinks it can handle—since he holds that “we should not be too confident about the ability of our institutions to achieve recognition and reconciliation. Criminal justice has typically fared very poorly in that regard.”³⁴ So Tadros's recognition argument may not provide sufficient support to his general deterrence argument to prevent him from having to endorse a much more abolitionist stance about state punishment than he otherwise seems willing to do in *TEH*.

Tadros may have another partial reply up his sleeve, since he also argues that “[i]t is permissible to harm an offender to a greater degree than is necessary fully to compensate the victim for having been wronged.”³⁵ Insofar as he is right, it may be permissible to punish wrongdoers in ways that, generally, have more powerful deterrent effects than mere compensation would allow. While greater punishments will not always translate into greater general deterrence, and greater deterrence does not guarantee that all state punishments of wrongdoers will provide adequate compensation to their specific victims, cases of insufficient *inter partes* remediation could well be lower as a result. Admittedly, this reply does not answer my objection fully, but it can at least weaken it somewhat.

Unfortunately, the additional claim is also problematic for Tadros's account: it seems to represent an important shift in his argumentative strategy away from his otherwise central concern for the next-best thing that wrongdoers owe their victims. Consider his explanation. If Anna does not bear any significant cost in protecting Blake from Claire's attack—say, she happens to be wearing a bullet-proof outfit—she cannot claim that she has fulfilled her corrective duty to Blake. For she only does what her generic duty to rescue him already requires of her, independently of her wrongdoing. Until Anna incurs a cost that she would not have had a duty to incur were it not for her wrongdoing, her enforceable duty of protection remains. Yet, to whom is this extra duty of protection owed, now that, according to Tadros, Anna, having protected Blake from Claire's attack, has made him virtually as well off as he would have been had he not been wronged by her in the first place? One possibility is that Anna owes it to Blake to provide him with protection against further wrongful harms. Tadros is sympathetic to this possibility, sometimes emphasizing that his additional argument is not meant as a substitute for the otherwise

³³ *Ibid.*, at pp. 283-286.

³⁴ *Ibid.*, at pp. 285, 108.

³⁵ *Ibid.*, at p. 286.

compensatory basis of his theory. Rather, it provides a complementary explanation for why it may sometimes be permissible to punish over and above the cap of compensation. At other times, though, Tadros appears to recognize the pressure that his additional argument puts on his *inter partes* rectificatory framework. Thus, at the conference leading to this symposium, he suggested that once state punishment of a wrongdoer has fully compensated her victim, any additional deterrence benefits might permissibly be distributed more evenly amongst other potential victims of crime.

Admittedly, Tadros still emphasizes the compensatory basis of punishment. However, when analysed alongside other arguments scattered across *TEH*, his additional distributive suggestion may also be taken as one of many indications that his commitment to this basis is unstable. For example, he argues that what a culpable wrongdoer owes her victim lessens over time, as “the connection between the person as they are now and the person who committed the offence weakens.”³⁶ This line of argument suggests that the extent of burdens that a wrongdoer may permissibly suffer in punishment is much more central to Tadros’s account than actual compensation of her victim. Indeed, were punishment to be extremely costly to the wrongdoer relative to the wrongful harm she inflicted, Tadros would likely hold it to be disproportionate and, thus, impermissible, even if the victim were left uncompensated as a result. As we have seen, Tadros is also adamant that the upper cap of permissible punishment differs from, and will often exceed, compensation. *Ergo*, how central really is compensation to the justification of punishment?

Tadros’s acknowledgment that inchoate wrongdoers may permissibly be punished accentuates this tension. It may be true, as he argues, that at least part of the explanation for why such wrongdoers may be punished is that they bear partial responsibility for the harms that people suffer from other completed crimes. They bear such responsibility because they are a threat to us, a threat that we must expend resources in averting, thus decreasing the resources that we have to prevent other crimes.³⁷ Yet, since, by definition, inchoate wrongs have no direct victims, it is unclear to whom punishment of such wrongdoers is owed as a next-best thing. If it is owed to society as whole—that is, to all potential victims of crime—then one is again left to wonder whether *inter partes* compensation really occupies the foundational position that Tadros otherwise claims it does in his theory.

At one point, Tadros even seems to concede inadvertently that a “stronger” justification for state punishment may be grounded in the special duty of equality that citizens owe to each other.³⁸ For all I know, this alternative argument might allow him to elude my earlier objection linked to general deterrence. That is, he might be able to argue

³⁶ *Ibid.*, at p. 308.

³⁷ *Ibid.*, at pp. 287-288, 325-326.

³⁸ *Ibid.*, at p. 311.

that even when wrongdoers' punishments add no general deterrence to what is already otherwise provided by their state's practice of punishment, they still have a duty of equality to undergo punishment just like those whose punishment has an actual deterrent effect—perhaps more as a matter of luck than anything else. However, in *TEH*, Tadros deliberately eludes this line of argument, whose focus is much less on *inter partes* corrective concerns than on the broader issue of the fair distribution of the ills (and goods) of punishment amongst citizens in society. Thus, in what follows, I take Tadros at his (admittedly ambiguous) word, disregard this troubling alternative proposal, and continue to focus on his so-called rectificatory commitment.

Indeed, Tadros's decision to frame his theory in terms of the principle of "next best" conformity to duties is also unsettling for another reason. Is protecting victims of wrongful harm from future harm really the next best thing that those who have harmed them in breach of their duty not to do so can offer? For one thing, if, as Raz suggests, doing the next best thing requires the duty violator to provide the victim with an equivalence (or as close an equivalence as can be) to the fulfillment of his primary duty—so that his breach is, in a sense, "made-up"—then it is at least plausible that certain harms have no equivalence (or even near equivalence). Raz speaks of such harms as non-compensable harms. He gives the example of hurt feelings arising from having been unable to attend one's mother's wedding overseas, as a result of a friend's breached promise to facilitate it.³⁹ We may also think in such terms of various harmful criminal wrongs. Consider rapes and the damage they can occasion to people's sense of sexual integrity and dignity as autonomous agents, assaults resulting in the permanent loss of a limb or, say, murders and the irreversible consequent loss of life for both the person killed and those who cared about her. In other words, it is at least plausible that some harms are incommensurable with anything else, and that any claims or attempts to rectify them with something as good, or nearly as good, *ex post facto* can only add insult to injury.⁴⁰

At first sight, Tadros seems sensitive to this possibility. After all, his argument for the justification of punishment overtly rests on the premise that some wrongfully-imposed harms are unlikely to be compensated adequately. The problem is that, in this context at least, Tadros seems to understand compensation in a much narrower sense than the one in which I am using the term. While I have been using it, and will continue to use it, generically to refer to the provision of an equivalence, or near equivalence, for a breach of duty, Tadros mostly uses compensation to refer to a specific currency of compensation—namely, monetary compensation. As I hinted above, he generally resorts to other terms such as rectification and remediation to capture the broader idea.⁴¹ The distinction is

³⁹ Raz, "Personal Practical Conflicts", *supra* note 27, at p. 192.

⁴⁰ Raz himself presents a strong argument for such incommensurability in: Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), at ch. 13.

⁴¹ See e.g., Tadros, *The Ends of Harm*, *supra* note 1, at pp. 277, 283, 322. To be fair, Tadros sometimes also uses compensation in its broader sense, but he does so rarely and inconsistently (see e.g. pp. 286, 288, 291).

important because Tadros seems to think that, in principle, all wrongful harms can be broadly compensated (rectified, remedied) at least in part. No doubt, his argument for this claim is clever and nicely illustrated by the case of Anna, Blake, and Claire. According to Tadros, it is almost always possible to compensate (in the broad sense) a victim of wrongful harm by preventing her from incurring again the very same wrongful harm.⁴² Can it really be denied that something can be compared with, and judged equivalent to, itself? Thus, Tadros argues that Anna compensates Blake for her hit man shooting off his right arm and leg—or, to use his terminology, rectifies this wrongful harm—by protecting him from having his other arm and leg shot off by Claire.

A key difficulty with this line of reasoning is that the wrongful harm that would have been imposed by Claire is not the same as the one imposed by Anna. It is not the same, except perhaps in terms of its abstract categorization as a harm of comparable magnitude caused by an assault. Here are some possible differences. Claire would have shot off Blake's left arm and leg, and not, like Anna, his right arm and leg. Thus, unless Blake is ambidextrous, the harm Claire would have inflicted upon him would likely have been markedly different in degree. Anna also stands in a different position to the harm she imposed—namely, she was an accomplice, while Claire would have inflicted her harm directly. Anna's harm would also have been imposed at an earlier point in time than Claire's harm. More fundamentally, though, the harm that Anna wrongfully imposes on Blake is not the same as Claire's since, quite simply, it is the harm that Anna, and not Claire, wrongfully imposes. Thus, Anna preventing Claire from inflicting similar wrongful harm on Blake is not necessarily equivalent, or nearly equivalent, to her having fulfilled her duty not to harm him in the first place. If, let us suppose, the wrongful harm Anna inflicts on Blake is of the non-compensable kind, then protecting him from similar harm will do nothing to compensate him.

Tadros's insistence that the primary justification of state punishment is its general deterrence effect only compounds the problem. Why? Because the general deterrence effect of punishment is likely best conceived as a diffuse phenomenon which, in general, does not precisely track the types of wrongful harms that are deterred. At least, this is the way in which Tadros himself seems to understand the phenomenon. Punishment, he writes, may help people see that wrongdoing is not an option and, thus, play a significant role in reducing people's inclinations to perpetrate wrongful harm.⁴³ Yet, nowhere does he suggest that the general deterrence effect of state punishment can be specifically tailored—

⁴² I say "almost," because it is much harder to argue that a victim who dies as a result of wrongdoing can still be compensated. Tadros attempts to get around this problem by arguing that the next-best thing for the wrongdoer to do is then to compensate the victim's family and those whom the victim cared about. However, he acknowledges the trickiness of such cases, especially if the victim did not care about anyone else. See *ibid.*, at p. 281.

⁴³ *Ibid.*, at pp. 281-282.

or at least always be so tailored—to deter wrongful inflictions of harm that match those for which wrongdoers are punished. In fact, it would seem farfetched to hold that criminal punishment of a wrongdoer can only be justified if it deters at least one instance of the very same kind of wrongful harm. Insofar as I know, Tadros does not hold this view, despite the fact that most of his hypothetical examples tend to emulate my comparison between Anna and Claire, and involve fairly symmetrical harmful wrongdoing. This interpretation of his position is reinforced by his assertion, explored earlier, that the extent of next-best duties of protection is set not so much by the kind and magnitude of wrongful harm that a wrongdoer actually imposed, but by the amount of harm that could permissibly have been inflicted on him to avert his original threat. Thus, even if Tadros is right that protection against future wrongful harm can compensate prior wrongful harm of the same kind and magnitude, issues of compensability still stand in the way of his broader general deterrence justification of state punishment.

One reason why he may not see, or acknowledge, the compensability problems to which I am referring may have to do with his understanding of the rationale for the duty not to harm others. Indeed, it seems commonsensical that we first need to know why we have a duty if we are to be able to work out what counts as its next-best satisfaction.⁴⁴ If, as Tadros seems to believe, the rationale for the duty not to harm others is best understood at a high level of abstraction, in terms of the inherent badness of reducing people's well-being through unjust means, then there may exist many ways to compensate such reductions. In other words, well-being may be such a generic concept that an exclusive focus on it can generate a false impression that equivalences, or near equivalences, can be found for all kinds of harms. Recall, in this light, Tadros's remarks that if a system of compensation—presumably, to be understood in his narrow sense of the term—could make people better off than a criminal justice system, while using the same resources, then the latter would likely be impermissible.

Here, we meet again Tadros-the-grand-unifier who seems to believe in deep and widespread commensurability of values. If he is right and, in principle, all harms have their price and can be compensated, then one cannot help but wonder: why not bite the bullet and simply become the advocate of a unified system of compensation using the most generic currency of all—namely, money? Such a system could have numerous advantages, including the maximization of victims' choices about how best to make use of adequate compensation, thus bolstering a value otherwise central to Tadros's account. Indeed, assuming that Anna owes something to Blake to make up for the wrongful harm she caused him, it is not obvious that this 'something' must be the direct protection or general deterrence that would result from her being harmed in punishment. After all, this is not how we interpret the duty she has in a typical tort case. Imagine that Blake's well-being could be sufficiently 'made up' were Anna to give him enough money to acquire cutting-edge prosthetic limbs, a round-the-clock personal helper, and a condominium unit in a sunny beach resort where, for years to come, he could evade the stresses of daily life. Could

⁴⁴ For an elaboration of this idea, see John Gardner, "What is Tort Law For?", *supra* note 13, at pp. 33, 40.

this not discharge her duty to do the next best thing? As the remarks recalled above suggest, Tadros comes very close to endorsing this position.

Three factors seem to stop him. First, he emphasizes the very high sums of money that may be required to compensate some wrongful harms, and the likelihood that many wrongdoers would be unable to pay. This argument strikes me as weak. The extraordinary sums it might take to compensate someone who has suffered a violent crime can hardly be considered a reason not to rely on a monetary compensation scheme when people are actually able to pay those sums. After all, in many jurisdictions, it is possible to sue people for intentional civil wrongs, such as the tort of battery. Moreover, the fact that some people might be unable to pay the required sums only shows that other responses to crime—like incarceration—may be needed as a backstop. It does not demonstrate that such other responses should be the default ones.

Second, Tadros sometimes seems to think that, without a minimum level of societal security from wrongful harm provided by the general deterrence effect of institutions of state punishment, no worthwhile monetary options such as those suggested above would be available to someone like Blake. A blunt way to put this point may be that no one can individually buy his way out of the state of nature. Again, I find this argument unsatisfactory. For one thing, it assumes that Blake could not use the money to relocate somewhere else—perhaps another country or a desert island—where this minimum level of security is available. It also seems to assume, without argument, that punishing an assaulter by, say, sending her to prison is necessarily the best way of advancing general deterrence. What if Anna were to pay the salary of five full-time police officers, and this were expected to have a vastly greater general deterrence effect than, say, sending her to prison? Could that not discharge her wrongdoing-derived duty to Blake? Tadros might try to argue that, unlike institutionalized general deterrence, such individualized protective remediation would not ensure the kind of stable social peace required for lasting individual flourishing. However, such an institutional, not to say statist, response would be in clear tension with the more bottom-up moral justification of punishment that he otherwise seeks to defend in *TEH*, and which assumes that punishment may sometimes be permissible in isolated cases like on desert islands and in the case of Anna and Blake.

Thirdly, Tadros emphasizes the importance of various non-compensatory aspects of state punishment: leading wrongdoers to recognize that what they did was wrong; educating people not to do wrong and, thus, preventing *them* from being saddled with onerous duties to compensate in the first place; enhancing people's overall security, and sense of security; facilitating the fulfilment of people's duties to rescue others; etc. This point is the most promising, although it reflects what I take to be the deeper problem with Tadros's theory. No doubt, insofar as the practice of state punishment serves such secondary functions, it is all the better for it. However, characterizing, as Tadros does, all such functions as secondary and framing the justificatory core of the practice as

compensatory, risks concealing different, and perhaps more particular, rationales for the duty not to harm others that may ascribe a much more primary role to some of them.

As I suggested, one plausible such rationale is that one should not do detrimental things to others which, in salient ways, cannot be made up. Thus, one should not lastingly deprive others of the usage of their limbs or other crucial organs. One should not kill them or kill their nearest and dearest. One should also not cause them enduring psychological trauma by violating their bodily integrity or sense of security in their homes and other surroundings. Tadros's spirited attempt to portray punishment as one practical segment of a unified and seamless morality of harm allocation, which, at a bottom, is no different in kind from other forms of compensation, obscures the possibility that a central function of the practice consists in responding to such non-compensable wrongfully-imposed harms. Of course, this does not entail that punishment cannot operate alongside a practice of compensation. It simply means that it cannot be equated with it. This argument should be added to the long list of other reasons for disaggregating punishment from compensatory practices, including the greater importance of culpability for the former, the many victimless crimes that are commonly thought to be punishable, as well as the general deterrence version of the PCBG objection.

Perhaps as importantly, the plausibly close relationship between punishment and non-compensable harms does not mean that justice has no role to play in the justification of the practice. It simply means that *corrective* justice may not play the pivotal role that Tadros thinks it does. As I suggested, *distributive* justice, and its concern for the appropriate overall division of goods and ills in society might have something important to contribute to the discussion. Furthermore, although Tadros rejects *retributive* conceptions of justice as vague and, often, barbaric, they can at least avoid my commensurability objection. They can avoid it, insofar as their focus is on giving *wrongdoers* what they deserve. That is, retributive justice need not require that anything be given to victims of wrongful harm, let alone anything nearly as good as them not having incurred this harm in the first place. Despite what may transpire from overly strict interpretations of retributive justice—e.g., 'an eye for an eye'—the concept of desert also need not entail any kind of commensurability between the wrongful ills incurred by victims, and those to be imposed retributively on wrongdoers. Insofar as I am right here, retributive justice, too, may have important insights to contribute to punishment theory. In other words, justice may not be as unified a value as Tadros sometimes seems to think it is,⁴⁵ let alone, as I argued in section II, the one seamlessly unifying master moral value. Thus, I would suggest, the most promising roads ahead in Tadros's valorous quest for the justification of criminal punishment lie away from Victor's justice, understood as an undifferentiated and master metric. They also, most certainly, lie away from Victor's corrective justice.

⁴⁵ Rawls himself insists on the distinction between distributive and retributive justice in *A Theory of Justice*, *supra* note 18, at pp. 276-277, while many since Aristotle (EN 1130^b30ff) also insist on the distinction between distributive and corrective justice.