I

In her exceptionally rewarding book, Professor Erin Kelly presents a series of arguments

against retributivism that expand in an ever-widening circle from theory to practice.

Retributivist accounts of desert and criminal punishment are, at the first, found wanting

conceptually. Kelly then argues that retributivism leads to a host of predictable injustices and

indignities, and so it *upsets* rather than *restores* justice in the wake of criminal wrongdoing.

Finally, Kelly concludes with a humane and attractive alternative to retributivism: an account

of criminal justice grounded in *harm reduction* that would reshape not only how we think

about criminal punishment but also how we enforce the criminal law in an unjust society.

Alas, I am a retributivist. I am, however, not exactly happy about this fact. And my

consternation with the view has only been exacerbated by Professor Kelly's trenchant

criticisms. Since I cannot hope to do justice to all of the arguments and themes of the book,

I want to focus on two core ones. The first of these is Kelly's powerful *Inscrutability Argument*

against moral desert. Since all versions of retributivism require agents to *deserve* to be held

responsible for their actions, if the Inscrutability Argument is sound it will upset any attempt

to develop a kinder, gentler version of retributivism, since even allegedly kinder and gentler

versions will rest on the same mistake. The second core element of Professor Kelly's book

that I want to focus on is her development of the *Just Harm Reduction* theory of punishment.

This too is highly original and interesting, and it provides a humane alternative to the

excesses of many retributivist theories.

II.

In the second chapter of her book, "Skepticism about Moral Desert," Kelly offers an

epistemic argument against moral desert serving as the justificatory basis for criminal

sanctions. More exactly, she argues that the bases for moral desert—the capacities that underwrite an individual's status as a morally responsible agent—are inscrutable given the kind of evidence that is available in criminal proceedings. As a result, we cannot reliably distinguish between which defendants have excuses and which don't. This is devastating to any reasonable system of criminal punishment, which will surely include high epistemic standards for guilt as a matter of procedural justice. Since, we cannot meet those standards, Kelly concludes that, "skeptical modesty about our ability to make these judgments calls into question appraisals of blameworthiness and desert closely associated with criminal punishment," (13). The upshot, of course, is that once we embrace modest skepticism of this sort, we'll no longer have a secure *retributivist* ground for criminal sanctions.

This is an important argument because it provides a powerful reason to doubt desert-based justifications for criminal punishment. However, I am not ultimately convinced that the Inscrutability Argument is sound. Or at least, I think that suitably spelled out, it is less damning to retributivism per se than Kelly takes it to be.

II.1. The Inscrutability Argument

Professor Kelly's Inscrutability Argument begins with the thought—a thought shared by many sensible philosophers—that in order to be deserving of praise or blame, an agent must be morally competent. Competence of this sort has two main elements. First, a competent agent is able to recognize (or grasp or understand, or perceive, or …) moral considerations as such. Second, she must be able to be moved by such considerations.[1] In other words, a morally competent agent is one who is both *sensitive* and *responsive* to the demands of morality. If, however, an agent is incapacitated in some way such that she is no longer

---

[1] In places, Kelly seems to require something more than this.

sensitive to the existence of relevant moral considerations or able comply with considerations that she recognizes as bearing on her conduct, then she is not responsible for her subsequent behavior and so, not deserving of blame when she acts badly.

A problem arises here: conditions that excuse criminal defendants are quickly outstripped by the possibility of various forms psychological incapacitation that are also intuitively excusing (or at least mitigating) for moral blame but not supported by extant positive defenses in the criminal law. As Kelly puts it, "the absence of coercion, duress, insanity, and other standard legally excusing conditions does not suffice to establish that a course of conduct was freely chosen in the sense required to justify blaming practices," (52). In other words, just because someone is not (or would not be), *by the standards accessible in the United States* (or in the common law tradition more generally), excused from otherwise criminal conduct does not entail that they are morally blameworthy. But if retributivism is true, this means that such agents—agents who clearly meet the legal standards of guilt— should not be punished.

This problem is exacerbated by the fact many instances of incapacitation are "local." That is, they are not due to some general insensitivity or unresponsiveness to morality on the part of the wrongdoer, but are triggered in specific circumstances that typically obtain for reasons that are completely outside the control of the agent. Moreover, these incapacities are frequently very hard, if not impossible to identify. Anxieties, dissociations, panics, traumas, and "morally disabling social influences" (i.e., an agent's moral ecology; cf. Vargas 2013) all can offset the moral competence of agents who otherwise appear to be (and indeed in other scenarios *are*) able to recognize and conform themselves to moral requirements.

The upshot of these facts about human psychology and behavior is this:

When a person fails morally, we might well wonder what underwrites our belief that she was competent enough to have done better. We may come to doubt, in a given case, that a moral "ought" implies "can" or, rather, that a moral "should not have acted badly" implies a "could have acted well" (67-68).

Here we see Kelly's skeptical modesty coming to the fore. After all, the explanation for any act of moral failure is underdetermined. It might have been, of course, that the person simply ignored or otherwise flouted a moral requirement they recognized to bear on their conduct. But the act itself can be equally well explained by a variety of local incapacitations that do not indicate a more general kind of irrationality or moral incompetence. And while the springs of our own actions are often opaque to us, this is seems even more true in the case of others' actions. Yet confidence about moral desert requires that we be, if not transparent to one another, fairly accessible. But we rarely seem to be, even in ordinary contexts, since even in these contexts, it's not always readily apparent whether a wrongful act was willful or due to a momentary failure that was outside the agent's control. When the appropriate standard of judgment is instead, "*beyond a reasonable doubt*," whether or not a wrongful action is an agent's own seems to be downright inscrutable.

But supposing this is true, it follows we can't be reasonably certain of an individual who has been successfully prosecuted in a criminal trial that she deserves punishment. Yet if *actual* punishment of wrongdoers is to make sense (and surely it must, one hopes)—if it is to be *justified*—it seems that we must ground it in something other than desert, since we can say with no confidence that the criminal in question does genuinely deserve to be held responsible. Professor Kelly has a story about this (which I'll come to in just a bit), but first: what can the retributivist say in response to the Inscrutability Argument?

II.2. Some Retributivist Replies

Kelly's argument is one that I've worried about a good deal elsewhere (see Coates 2016). So I'm tempted to simply accept much of it, and agree with her that moral desert, particularly of the sort that is entailed in ordinary contexts by judgments of blameworthiness is not a good basis for criminal punishment. However, I'm not sure that her argument quite secures the anti-retributivist conclusion that she's looking for.

A first problem emerges when we question whether Kelly really is correct to think that moral competence (or whatever capacity grounds moral desert) is indistinguishable from incapacitation. I am not sure that it is. To the contrary, I suspect that we have a fairly good sense of when people's actions reflect an underlying moral competence on their parts and when they don't. Of course, we're not *perfect* evaluators of agents' competence—we can and do get it wrong sometimes—but I suspect our general mind-reading/theory of mind capacities go a long way towards reliable evaluations. After all, in our moral practices, we've developed an impressively detailed list of mitigating, excusing, and exempting conditions. And we didn't do this by using a "psychological microscope" to probe the inner workings of agents' minds. Instead we just observed people's behavior, and came to realize that the observed patterns of behavior reliably indicated impairments to or incapacitations of the agents' moral competence. What this means is that although I might not immediately be able to tell whether your transgression is one you deserve blame for, after careful deliberation, which involves situating your behavior in a wider pattern of actions and learning about what you were thinking just before the action, I do think I can fairly accurately determine whether you were, at the moment of action, a competent agent who was flouting morality or someone momentarily incapacitated.

Now, I might be too optimistic here (though I'll say a bit to suggest I'm not in just a second), but Kelly could simply grant that I'm right about this. Still, she could also maintain that despite individuals' ability to determine competence in ordinary circumstances (i.e., outside of the criminal law), the criminal law is a great deal less responsive the wide variety mitigating, excusing, and exempting conditions than is moral practice. As a result, there will be many cases, perhaps, in which an agent is either less responsible for their action (than they would be otherwise) or not responsible at all for what they've done. Unfortunately however, because the criminal law is insensitive to those considerations, such agents be convicted in an apparently procedurally just way.

Here, though, the retributivist should simply deny the justice of conviction in such an inflexible system. Notice, however, that the retributivist insistence in not punishing in such a case isn't that we *can't* know that the agent is not deserving of punishment, but is instead precisely that we can know that she is not responsible. This points to a key divide between retributivists and Kelly that I want to focus on more when discussing her account of punishment. But though I'll discuss it at length later, I want to flag it here: if Kelly's argument succeeds in convincing the retributivist, the retributivist can simply embrace abolitionism—the view that the state shouldn't punish. It is surprising, then, that it is Kelly who is at pains to justify criminal punishment in a way that does not appeal to desert. In other words: it seems that if the Inscrutability Argument is sound, suitably circumspect retributivists could credibly turn the tables on Kelly for being too punitive.

But before I get to Kelly's account of punishment, I want to question her skeptical modesty itself. I'm worried that, at least as Kelly develops the point, there is nothing "modest" about it. In fact, if it is a sound argument against moral desert, then it seems to generalize much too easily. Consider: given my current evidence, I can't rule out that any of

you are actually *yous* (i.e., other minds) and not just figments of my imagination, the products

of an evil demon, or cleverly packaged robots disguised as real people. After all, which of

these alternatives is the correct one is inscrutable to me *given* your behavior. Similarly, given

my current evidence I can't tell whether my family really loves me, or whether it's a Truman-

show situation, or whether they're all just very good at pretending to love me because they're

hoping for a big payday. Again, I could worry about this because they're behavior

underdetermines which of these hypotheses is true. Except none of us really thinks this way.

In other words, none of us is worried about an Inscrutability Argument for Skepticism about

Other Minds. But such an argument looks structurally isomorphic to the one Kelly offers

against desert.

Here, I think Kelly can (will?) sensibly point to the fact that by retributivist lights, if

we're wrong about desert, then we do defendants serious injustice by punishing them. Since

the stakes of getting it right are so high in the case of criminal punishment, then the

retributivist should be committed to being cautious in confidently ascribing

blameworthiness. But if I'm wrong about others' minds—if all you are just lines of code in

the Matrix—then not much of practical consequence is affected. As a result, I can be

warranted in my incaution in treating you as other minds. What saves Kelly's skepticism

from being a species of skepticism about other minds are the comparative practical stakes

that are implicated by being wrong in these two cases.

The retributivist, on this point, can again *agree* with Kelly. Some retributivists, of

course, will want to claim that even given the high stakes, there are just so many cases of

ordinary crimes in which no credible claims of incapacitation can be made. So for those

cases, we'll still be justified in ascribing moral responsibility, and as a result, we'll still be able

to appeal to desert as the basis for punishment. For other cases—which might be wider set

of cases that retributivists have wanted to admit—we'll have eschew punishment. But this shouldn't worry retributivists; it was, after all, a retributivist who insisted that it was better to let ten guilty persons go free than let one innocent person suffer.[2]

But even supposing this, there is further reason to worry about Kelly's skeptical argument. If it is sound, then retributivism can serve as a basis for criminal punishment because we should be skeptical of whether defendants deserve punishment given the forms of evidence that are accessible in criminal proceedings. And we should be skeptical of desert because we cannot distinguish between morally competent individuals and ones who are subject to psychological incapacitations of various forms. However, it's not just desert that entails moral competence (or normative competence more generally). Friendship, promissory obligation, the force of anti-paternalistic reasons, and many other things that are central to our lives also depend on others being morally competent. But if we should adopt a skeptical attitude towards their competence, then we will be forced to adopt a more objective attitude towards them more generally.[3] If I can't tell, for example, whether someone is morally competent or suffering from a local incapacitation, then it seems to me that that person's good puts demands on me that it wouldn't if it were clear that she were an autonomous agent. And if I must regard someone through paternalistic lenses, as I do with my two young daughters, then although I could take real joy in being in that person's company, how could I genuinely regard her as a friend, as someone with whom I can engage in joint activity? The upshot of Kelly's skepticism, then, isn't just that retributivists will need to reject the legitimacy of criminal punishment, but that *all* of us will need to worry about

---

[2] William Blackstone 1760.
[3] See P.F. Strawosn 1962. Some moral responsibility skeptics (e.g., Derk Pereboom 2001 and Tamler Sommers 2007) argue that responsibility skepticism does not undermine the things I've listed here and so does not entail that we should adopt a more objective attitude towards others. However, these skeptics maintain this balance by arguing that although moral competence is insufficient for desert, it is the basis for these other things. So their skepticism doesn't rest on

how we are to understand the interactions that structure our lives and to a large extent, give

them meaning. Skepticism about moral competence might be the right attitude to take, but it

is not a modest one.

<div align="center">III.</div>

In Chapter 5, "Rethinking Punishment," Professor Kelly offers a novel theory of criminal

punishment that does not depend on wrongdoers being deserving of punishment for its

justification. She calls this account the *just harm reduction*. As she sees it:

> Harm reduction is, in fact, the most a criminal justice system can achieve. A criminal
>
> justice system cannot ensure that punishment is allocated only to people who morally
>
> deserve it, much less that all wrongdoers get their just deserts. Nor can it cure a
>
> society of violence, stealing, cheating, and other forms of antisocial behavior. In
>
> order to "solve" the problem of crime, a society would have to make a commitment
>
> that goes beyond criminal justice. It would need to address the causes of crime,
>
> which is something a criminal justice system does not do (123).

And she continues:

> A criminal justice system can accomplish just harm reduction even when it does not
>
> presuppose the moral capacity of every criminal wrongdoer to have avoided his
>
> crime, and even though it does not calibrate blame for criminal transgression in a
>
> scale of moral desert. Just harm reduction is cautious and morally modest. It
>
> attempts seriously and responsibly to respond to criminal wrongdoing without

exaggerating the responsibility of wrongdoers, and in keeping with the inability, and

in keeping with the inability of the criminal justice system to address the causes of

crime. It maintains a humane and respectful attitude towards all people, including

people who perpetrate and are convicted of crimes (123).

Here Kelly is giving us an attractive picture of criminal justice. We punish on this

view to reduce and prevent harms from being done. But we are concerned, unlike naïve

utilitarians, to make sure that our punishment is doled out in a just way, which requires we

be able to support it using public reason and appeals to sound principles of *distributive* justice.

None of this, however, commits us to those convicted of wrongdoing being the proper

objects of blame. So *condemnation* of criminals drops out of this picture. And with it, Kelly

thinks, some of the nastier aspects of retributivism.

Kelly acknowledges, of course, the potential pitfalls of an approach that veers too

close to general deterrence in its justification for criminal punishment. So I want to briefly

explore what she says to differentiate her view. I'll then turn to how retributivists might

reply.

III.1. Punishment Harm Reduction

Once one rejects moral desert (and so, retributivism) as a basis for legitimate criminal

punishment, one what grounds can one hold criminals responsible? An obvious alternative is

one that appeals to deterrence. However, deterrence-based justifications of criminal

punishment are notoriously difficult to square with some fairly intuitive principles of justice.

To accommodate our intuitions about the contours of a just system of criminal

punishment—that, e.g., it is distributed equally and proportionally—Kelly appeals to the

basic notion of fairness. We should treat like cases alike, and so we have grounds to attempt to reduce harm (by deterring it) in the same ways in all cases of wrongdoing that are relevantly similar.

This is sensible, but it leads to an immediate worry for someone who's concerned to reform an overly punitive criminal justice (like the one we currently "enjoy"), since it can inspire *leveling-up* punishments. Retributivism has an easy theoretical response to this, in that retributivism is only committed to there being reason to punish wrongdoers to the degree that they deserve it. But if what matters is not desert but harm reduction, and we're concerned to treat like cases alike, how can we avoid an upwards creep of punishments? After all, one can guarantee harm reduction (or guarantee a higher degree of harm reduction) by lengthy incarceration. So if harm reduction and equity in punishments are guiding principles of this account of punishment perhaps the most straightforward way of achieving these two aims is *over-punishing* convicted wrongdoers.

In responses to this worry, Professor Kelly claims that we will no longer be prone to these sorts of mistakes once we "reject the blaming function of punishment and … address system injustices" (147). In other words, the attitude that might lead us to over punish is not a concern for reducing harm or being fair; it is instead our commitment to the idea of moral desert and because of systemic injustices that shape our criminal justice system. Indeed Kelly suggests that the former is a major factor in the latter: "the appeal of retributive justice has given punitive criminal justice policies a false sense of legitimacy" (179). So it is the purported distortions of thinking issuing from retributivism that preclude us from responding to wrongdoing in a way that respects the dignity of both victim and offender and helps create a functioning society that is good for all. Harm reduction, by contrast, is an admirable alternative.

III.2. Some More Retributivist Replies

I find Kelly's harm reduction theory of punishment to be attractive. But I am not sure that it tells against retributivism—at least in some suitably weak form. Indeed, I think the value of harm reduction itself is a good that meaningfully weighs against giving people what they deserve precisely because we have reason to care about the good of all members of our society—a set that notably includes some who act in ways that offend against the normative standards of our society.

Earlier in the book, Kelly claims that "moral desert … is the view that culpable wrongdoing morally determines how we should respond" (48). And of course if this is your view of what desert is, then it will follow that the goal harm reduction cannot compete against or outweigh reasons of desert. But I confess that although some, like Kant, hold this view of desert, no plausible form of retributivism does. More sensibly, many who theorize about desert think that it is good to give people what they deserve or that there are pro tanto reasons to give people what they deserve. But by conceptualizing desert in this weaker way, it becomes clear that the aim of harm reduction can serve as a powerful guide for choices we make about how to punish criminal offenders.

But this is a two way street. Caring about reducing harm *for all the reasons Kelly adduces* helps constrain the sometimes problematic force of considerations of desert (or at least, our imperfect application of what we take those reasons to be). But moral desert serves as an important check on who we subject to punishment in the first place. It does this by providing us with a powerful and natural explanation for why scapegoating is wrong, for why we should treat like cases alike, and for why we shouldn't "level up." All of these can serve

the function of harm reduction, and yet they are intuitively problematic precisely because they reduce harm by treating some people in ways that they don't *deserve* to be treated.

As we've seen, though, Kelly has alternative explanations for why we shouldn't scapegoat, etc. even when the aim of harm reduction is being promoted. But once one has rejected moral desert—that is, once one denies that morality does not recognize desert as being reasons-giving—then it is very surprising that the true theory of punishment should have exactly the same contours it would have with respect to these issues as it would if retributivism was true. So even if Kelly alternative explanations for why harm reduction can't run amok succeed in constraining punishment in just the circumstances the retributivist would predict that we must constrain punishment, we'll still want an explanation of why *should* constrain punishment in these circumstances. As it stands, it seems like the theory on offer is structured around a set of intuitions (viz., retributivist ones) that Kelly thinks we would do well to reject.

Finally, I want to close by suggesting that our retributivism is less to blame for our current state of affairs that Kelly claims. As I said, she suggests that our commitment to the blaming function of punishment *and* systematic injustices lead to the sorry state of affairs that is our criminal justice system. She then suggests that if we weren't so concerned with assigning blame then we might be led to adopt more equitable and just criminal justice policies.

This might be. But I am initially suspicious of the order of explanation. Consider, for example, the difference in how crack cocaine was treated in the 1980s and how lawmakers are currently discussing the current "opiate *epidemic*." A generation ago the war on drugs was one waged with harsher and harsher prison sentences that were superficially justified with appeals to criminality and evil and all of the language that is very tightly connected to

retributivist sentiments. This regime devastate already marginalize communities and ruined many people's lives. In the wake of those failures, we currently see strong calls for treatment rather than criminality, and policy suggestions that are very much framed in ways that are concerned with *harm reduction.*

If this were all you knew about these two drug crises happening 35 years apart, then it would serve as evidence for Kelly's claim that our proneness to moral blame leads to grave ills that have furthered deep systemic injustices. But of course this isn't the whole story. The whole story starts (and probably ends) with the fact that America is a deeply racist society. And in particular it is racist in a way that *usurp* moral and aretaic language for its own immoral and vicious ends (think here of calls for "patience" that King criticizes in "Letter from a Birmingham Jail"). So blame and criminality served as a cudgel that those in power could wield over African-American communities in the 1980s in the midst of their suffering. But now that it is white communities who are suffering, the proposed solutions shift away from punitive ones.

In other words, it seems like the real explanation for over-punishment and for abuses of criminal justice—at least in this one case—is not our retributive natures leading us to racist policies but precisely the other way around. And my suspicion is that *any* moral ends, indeed even harm reduction, could similarly be usurped by systemic racism in criminal justice and law enforcement. Kelly, of course, is very sensitive to all this (her last chapter is very good on related points, and I didn't discuss it only because this is an "author meets critics" session and not "author meets people telling her what they learned from her book" session (though no doubt that would have value as well)).[4] But it is at odds with her repeated insistence that our commitment to moral blame lies behind the ills of our society. Kelly

---

[4] Mention the discussion of excusing and mitigating conditions in Chapter 3 here as another example.

might yet be right that we should free ourselves from retributivism and from the concepts

it's grounded in, but, alas, I suspect that this will not be enough to bring justice to our justice

system.

IV.

As I said at the outset, I am a retributivist. But Professor Kelly's book has chastened that

somewhat, though not perhaps in the ways she would prefer. Her Inscrutability Argument

genuinely gives my pause about the legitimacy of criminal punishment. And her Just Harm

Reduction theory of punishment seems, if not correct as its own theory, a nevertheless

important supplement that we should amend to retributivism.