



Not Justice: Prison as a Moral Failure

Luke Maring¹ 

Accepted: 10 February 2023

© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Lisa Tessman (2016: 164) recounts the case of a Jewish mother, running from Nazis, who faced a terrible choice. She could (a) drown her infant, or (b) accept the virtual certainty that her baby's cries would doom the refugee group she was fleeing with. Given those options, (b) is worse. If the whole group is discovered, many will die, including the infant. Still, preemptively drowning a baby—indeed one's own baby—is a terrible act.

To make sense of cases like this, Tessman turns to the concept of *moral failure*: an act that violates an important moral value such that the loss cannot be fully redeemed by gains elsewhere. The mother, through no fault of her own, faced a situation in which all her options were moral failures. Prioritizing the group is the all-things-considered best choice; but in this case, all-things-considered best is still terrible.

I will argue that criminal justice needs a moral framework like Tessman's. It needs to make sense of the fact that some options are genuinely better than others, without letting honorifics like 'justice' obscure the equally important fact the best available option might be terrible. More specifically, I will make three arguments that when punishment involves prison, it is a moral failure. §1 contends that, realistically, sending someone to prison is to make them vulnerable to intolerable cruelty. Whether wrongdoers deserve broadly humane forms of punishment or not, making them vulnerable to intolerable cruelty is to wrong them. §2 points out that whereas many have sought to justify punishment as a form of communication or expression, we send awful messages when we make people vulnerable by sending them to prison. §3 ups the ante yet further, arguing that even when the risk of intolerable cruelty is *not* realized, prison is an assault on moral agency.

If the arguments in §1-§3 succeed, they show that prison is a moral failure, not something that deserves to be called justice. §4 explores the implications that has for justifying prison: how, and in what sense, *can* prison be justified? §4 contends that prison can be justified only in the sense that preemptively drowning an infant can be justified. It must be a last resort, used only when all better options have run out. Practically speaking, this means that prison will be justified only after our societies

✉ Luke Maring
Luke.maring@nau.edu

¹ Northern Arizona University, Flagstaff, AZ, USA

have invested in schools, effective anti-poverty programs, Restorative Justice, and other measures that can preempt the need to incarcerate people.

1 Prison is a Moral Failure Part 1: Intolerable Cruelty

Prisons house a wide range of cruelty, from egregious violations to grating indignities. Examples of the egregious include the psychological torture of solitary confinement, assault, and rape. To illustrate the grating: phone services at Angola penitentiary in Louisiana interrupt conversation every few minutes—needlessly and in humiliating fashion—to announce that one is talking to an “offender.”

Smaller indignities can add up, over time, and exact a considerable toll. But for the sake of straightforward argument, this paper will focus on egregious wrongs. *Intolerable penal cruelty* is a significant harm, perpetrated within a prison system, that constitutes a serious affront to human dignity. Physical and sexual assault, severe emotional abuse, and solitary confinement are paradigm examples. But while these acts of commission are paradigm, acts somewhat nearer to omission can be intolerably cruel too. When people are crammed into overcrowded and unsanitary cells, lack even basic healthcare, live in fear, and are forced to subsist on materials that don’t deserve the name ‘food,’ simply allowing people to remain in such conditions is a significant harm that constitutes a serious affront to human dignity.

How common is intolerable penal cruelty? There is, to my knowledge, no even remotely comprehensive database that compiles cases on a global scale. (Though the incomplete, haphazard lists we do have are alarmingly large.¹) This paper will focus on prisons in the U.S. and the U.K., where data is somewhat better. But here too it is hard to find good estimates, for many of the same reasons it is hard to gauge the frequency of police misconduct—these institutions typically have little oversight and no incentive to tell on themselves. Commenting on U.S. prisons, Weill and Haney (2017: 287) explain that

prison officials often perceive independent researchers as “dangerous” because they “do not owe allegiance to staff, inmates, nor administrators” (Trulson et al., 2004, p. 456) and fear that empirical data may reflect poorly on the prison (Dilulio, 1987) or that adverse conditions will be documented and reported by researchers without contextualizing them (Trulson et al., 2004). Because prison gatekeepers strictly control access to the facilities, the methods researchers can employ, and the topics they are permitted to study (Berg, 2001), research is often restricted to mining existing institutional data (Trulson et al., 2004). Unfortunately, here, too, data are quite limited.

There is, in addition to these institutional barriers, a further obstacle to data collection. As a report on rape in U.S. prisons explains, survey “participants tend to

¹ Amnesty International’s list does not purport to be comprehensive—or even close to comprehensive. But Searching “prison abuse” on Amnesty International’s website returns one hundred eighty-one pages of incidents. See: <https://www.amnesty.org/en/search/prison%20abuse/>.

underreport behaviors that are perceived to be against society's norms ... [and that] invade privacy (Gaes & Goldberg 2004: 3). The point, here, is that our best estimates of intolerable cruelty are almost certainly low.

And yet, the unrealistically low numbers in our best reports are far too high. Struckman-Johnson et al. (1996) found that 20% of survey respondents in a prison system in the Midwestern U.S. "had been pressured or forced at least once to have sexual contact against their will while incarcerated." Struckman-Johnson et al. (2000), which also surveyed prisons in the American Midwest, found that "21% of the inmates had experienced at least one episode of pressured or forced sexual contact." Wolf et al. (2007: 595; emphasis mine) surveyed 14 prisons in a mid-Atlantic state and found that "[o]ver a 6-month period, 20% of inmates experienced some form of physical violence." Summarizing studies of violence in U.K. prisons, Liebling (2004) writes,

The results suggest that between 7 and 30 per cent of male prisoners in closed prisons report having been assaulted; between 3 and 13 per cent report a sexual attack; and between 26 and 49 per cent report threats of violence ... A detailed self-report study ... in four prisons found that 19 per cent and 26 per cent of adult males had been assaulted and/or threatened with violence in the last month, respectively, and 30 per cent and 44 per cent of young offenders had been assaulted and/or threatened with violence in the last month, respectively (Kimmitt and O'Donnell, 2002: 29-30).

It is implausible that the moral horrors in these studies are one-offs—sudden, violently cruel episodes that occur in the context of an otherwise well-functioning prison bureaucracy. Prisoner abuse, like extreme police misconduct, generally occurs within a larger institutional *culture* of abuse (Weill and Haney 2017). The moral horrors in these studies are likely just the tip of the iceberg.

To make matters even worse, the commissive, attention-grabbing acts that tend to show up in studies do not exhaust the scope of intolerable penal cruelty. Many incarcerated people—who have not been assaulted and have nothing actionable to report—endure conditions well beneath any reasonable standard of human dignity. I will, therefore, proceed on a realistic assumption: intolerable penal cruelty is widespread and common.

This means, unfortunately, that imprisoning someone is to impose terrible, undeserved risks upon her. Whether wrongdoers deserve broadly humane forms of punishment or not, making them vulnerable to intolerable cruelty is a clear moral failure. Even if incarceration is practically necessary—suppose the wrongdoer cannot currently be trusted in civil society and we have nowhere else to put them—it would be dishonest to paper over the fact that we are wronging them.

But what about prison reform? As it stands, sending someone to prison is to impose terrible, undeserved risks upon her. But there is no necessary, conceptual, or strictly logical link between prison and cruelty. So doesn't this argument just point out the (already widely recognized) need for prison reform?

In one sense, it does. *If* prisons suddenly become humane sites of rehabilitation, the argument in this section collapses. But, this section will argue, the forces

pulling our prisons towards cruelty are deeply embedded, not things we can readily legislate our way around. There is simply no escaping the fact that prisons are, and will always be, bureaucracies run by actual human animals.² Human animals display all sorts of moral frailties, especially when they are organized into a bureaucratic institution that sets out to harm (perceived) wrongdoers. The rest of this section makes three arguments, each of which identifies a different fact about human animals, or about prison bureaucracies, that inclines our prisons towards cruelty.

1.1 The Dynamics of Bureaucracy

In the wake of Abu Ghraib, David Luban (2005) urges sober realism in the debate about torture. Torture advocates sometimes employ fictionalized ticking time bomb cases: we have apprehended a terrorist, know that he has planted a bomb, know that he will reveal its location only under torture, and thus face an artificially simplified choice. We can torture him, or we can sacrifice scores of innocents to our scruples.

Luban rejects the case. With a nod to Wittgenstein, he (2005: 1141) calls the fictionalized ticking-time bomb scenario a picture that bewitches. It invites us to think about torture in small-scale terms: may we torture *this* person *now*? The decision we actually need to make reverberates well beyond *this* and *now*. We need to decide whether it is a good idea to entrust a bureaucratic agency—run by human animals—with the power to torture. That decision, Luban (2005: 1449–52; emphasis mine) explains, requires us to unflinchingly catalogue the ways torture bureaucracies can, and do, go wrong.

The fiction [assumes] that the interrogator operates only under the strictest supervision, in a chain of command where his every move gets vetted and controlled by [his] superiors The trouble is that this assumption flies in the face of everything we know about how organizations work. The basic rule in every bureaucratic organization is that operational details and the guilty knowledge that goes with them get pushed down the chain of command as far as possible [And there is also] a point of social psychology. Simply stated, it is this: *we judge right and wrong against the baseline of whatever we have come to consider “normal” behavior, and if the norm shifts in the direction of violence, we will come to tolerate and accept violence as a normal response.* The psychological mechanisms for this re-normalization have been studied for more than half a century, and by now they are reasonably well understood.... Abu Ghraib is not a few bad apples—it is the apple tree.

Torture bureaucracies and prison bureaucracies are not perfect analogues. But the point I am making here does not require them to be. The point is that both harm

² Philosophical terms like ‘moral agent’, or even ‘person,’ have taken on a Kantian flavor: they foreground our rationality. I need a term that foregrounds our susceptibility to environmental influences, our tendencies towards tribalism, and an array of other eminently human failings that are liable to get lost in the background when we talk about ‘moral agents.’ Hence, ‘human animals.’

people, both are infamous for intolerable cruelty, and—most importantly—both are beset by similar institutional dynamics. If it is naïve to assume that the interrogator will always operate under strict supervision, it is similarly naïve to assume that penal officers will always be closely watched. If we should expect careerist administrators in a torture bureaucracy to maintain a veneer of plausible deniability, we should expect careerist administrators in a prison bureaucracy to do about the same thing for about same reasons. If the organizational dynamics in a torture bureaucracy shield well-intentioned bureaucrats from “guilty knowledge,” the organizational dynamics in a prison bureaucracy can pull the very same trick. And for the most alarming parallel: if the astonishing human capacity to embrace normative shifts towards violence is bad news for torture bureaucracies, it isn’t great news for prison bureaucracies either.

The last parallel—that human animals readily embrace normative shifts towards violence—is so alarming because it means that a few ‘bad apples’ can initiate a much larger institutional swing towards cruelty. Realistically, any large-scale practice dedicated to harming people will contain, here borrowing from Waldron (2005: 1716-7), a few “enthusiasts who are prepared to ‘push the envelope,’ trespassing into territory that goes beyond what is legally permitted.” Waldron aims that comment at torture bureaucracies; the fact that it can be so readily understood as a comment about prison bureaucracies is the point. Over time, an actively cruel minority can shift the normative expectations of a whole institution—or, at least, it can shift expectations in the institutional sectors tasked with the hands-on business of meting out harm.

The tendency for practices of torture to spiral out of control and produce moral horrors is so obvious that many who deny an absolute moral prohibition support an absolute legal prohibition. They hold that torture may in theory be justifiable, but that human animals simply cannot be trusted to run a torture bureaucracy.³ A similar in-practice skepticism is relatively hard to find in traditional philosophy of punishment.⁴ I think it should be taken more seriously: actual prison systems—across the globe and across history—have produced abuses *no less shocking* than the ones we saw in Abu Ghraib. If there is a principled difference between torture and prison, that difference is not that whereas torture bureaucracies are inherently unruly and prone to the morally indefensible, prison bureaucracies can generally be trusted to dispense justice. The difference, if there is one, stems from the fact that whereas torture is not the only way to gather intelligence (outside fictionalized ticking time-bomb scenarios), something prison-like is the only way to isolate people who cannot currently be trusted in civil society. Torture and prison are both moral failures; but there may be better, more realistic grounds for arguing that prison is morally akin to preemptively drowning the infant. That’s not much of a moral commendation, but it would prejudice inquiry at the outset to assume that there *must* be a better one.⁵

³ See Miller (2005) for example.

⁴ But see Wolfendale (2020) for an exception.

⁵ One might think there is another argument that puts prison morally ahead of torture: torture is necessarily degrading, but prison could be humane, at least in principle. I am not making that argument for two reasons. The first reason stems from this paper’s methodological commitments. I am committed to viewing prisons realistically, and while it is logically possible for prisons to be cruelty-free, we should

To sum up the argument so far: Prison systems are bureaucracies that aim to harm people, and they are beset by the same institutional vices that produced the moral horrors at Abu Ghraib. Inevitable lack of oversight, incentives for higher-ups to protect themselves from guilty knowledge, the nearly inevitable presence of bad apples, and the astonishing human capacity to embrace normative shifts towards violence mean that prison bureaucracies are at significant risk of intolerable cruelty.

1.2 Prisoner Vulnerability

A book recently published by the World Health Organization notes,

The deprivation of liberty means that a victim who reports the violence has no possibility of escape from the retaliation by the perpetrator. A study found that 25% of respondents who had not reported their most recent experiences of assault said that they did not believe that reporting victimization would make a difference. An additional 20% did not report an assault because they feared retaliation.⁶

This finding lends further credibility to the assumption that our best, already high estimates of intolerable cruelty underestimate the problem. But the point I want to emphasize here comes from the first sentence: prisoners are profoundly vulnerable. They are trapped within reach of their abusers, whether their abusers are other inmates or prison officers.

In a vacuum, this may seem like a readily fixable problem: prisons simply need an effective system that allows inmates to report abuse and be safely separated from their abusers. In practice, however, it can be hard, even impossible, for prison officials to protect vulnerable inmates. Even Wardens and Superintendents report failure, despite their considerable institutional power.

[F]ormer Mississippi Warden Donald Cabana said, “I’ve had to negotiate no fewer than eight hostage situations, deal with riots, et cetera.” “I couldn’t protect [the women] from being sexually preyed upon,” former New York Superintendent Elain Lord told the Commission. Former New Jersey prisoner Thomas Farrow described nighttime beatings where officers targeted certain prisoners. Ron McAndrew, former warden of a maximum security prison Florida told the Commission about “goon squads,” small groups of violent officers beyond even his control. (Gibbons and Katzenbach 2006: 415).

The last subsection argued that prison bureaucracies are beset by a range of institutional vices. This section’s point is that, as a practical matter, those very same vices

Footnote 5 (continued)

not expect that from large-scale bureaucratic institutions that are dedicated to harming people and run by human animals. Second, §3 argues that even apart from the statistical likelihood of cruelty, prison is an assault on moral agency. I think prison, like torture, is inherently degrading.

⁶ Full text available here: https://www.euro.who.int/__data/assets/pdf_file/0005/249188/Prisons-and-Health.pdf.

contribute to prisoner vulnerability. The fact that careerist administrators have an incentive to maintain plausible deniability means that reports of violence may fall on uninterested ears. The fact that institutional dynamics can shield well-intentioned administrators from “guilty knowledge” means that reports are liable to get lost in the bureaucracy. The fact that it is nearly impossible to monitor penal officers at all times means that it will often be hard to prove that accusations of abuse are true. The nearly inevitable presence of ‘bad apples’ entails a risk of “goon squads.” And the astonishing human capacity to embrace normative shifts towards violence means that intolerable penal cruelty can quickly become the status quo. And if that happens, a wealth of psychological research suggests that perfectly normal people—buffeted by the sense that their role, an authority, or their peers demand it—will become cruel.⁷

So: not only do the dynamics of a prison bureaucracy facilitate cruelty, they create conditions in which prisoners have no good way to protect themselves. With enough resources, of course, some of these problems can be ameliorated. (With better pay, more training, and better work conditions, for example, prisons might attract and retain better staff.) But, as the next subsection argues, the obstacles to funding prisons appropriately are deeply embedded.

1.3 Predictable Abuses of Punishment’s Stigma

The late Lee Atwater, a high-level Republican strategist, helped design the infamous Willie Horton advertisement. The political ad warns voters that Michael Dukakis (then running against H.W. Bush) not only opposed the death penalty, but also favored weekend passes for inmates serving life sentences. Willie Horton used his weekend pass to commit murder and rape. But what likely made the ad so politically effective was not just its appeal to public safety, it was the ad’s use of *race*. Willie Horton was Black. By displaying a disreputable-looking photo, the ad dog whistled a racist message: citizens concerned about Black criminality should vote for Bush, who, according to the ad, favored the death penalty and would be tough on crime (Saul 2018).

Accusing political opponents of being soft on crime, and then promising to be tough, is a widespread form of political posturing. It is damnably effective—for reasons related to the very definition of punishment. Joel Feinberg (1965: 400) rightly distinguishes punishment from other sorts of disincentivizing penalties by its expressive content: unlike small fines or parking tickets, true punishment “is a conventional device for the expression of attitudes of resentment and indignation.” It might sometimes be appropriate to feel and express such attitudes; at the same time, however, we need to be clear-eyed about the way a practice with that expressive power is likely to be abused. If punishment is “a symbolic way of ... expressing a kind of vindictive resentment,” (Feinberg 1965: 403) every substantially resentful public is likely to have an appetite for making the resented Other suffer in prison.

⁷ See Zimbardo (2008: chapters 12 and 13).

Shameless politicians—who we will probably never be free of—will often be willing to exploit that appetite. The resented Other may be a racial minority, as it was in the Willie Horton case. But it can also be unwelcome immigrants (of whatever race), or just people the public conceptualizes as criminal.

Unfortunately, the problem, here, may be even worse than politicians shamelessly exploiting the public's latent resentment. There is some evidence of a complicated feedback loop: politicians exploit the public's resentment by supporting harsh policies; at the same time, according to Ramirez (2013: 357), the “ups and downs of [public] punitive sentiment are driven by important political factors such as the construction of crime by political leaders.” In short, tough-on-crime posturing may *drive* the public's punitive sentiment while exploiting it. The public's appetite for harsh punishment also appears to be driven by reports of violent crime on television news (particularly local news) and by sensationalist crime dramas.⁸

How does all this connect to penal cruelty? It connects in two ways. First, recall that that intolerable penal cruelty is not just a matter of overt acts, like assault or rape. Allowing people to endure conditions beneath human dignity is intolerably cruel too. But such conditions should come as no surprise when hard-to-regulate forces like latent resentment, tough-on-crime posturing, and television programming gin up punitive sentiment. As Weill and Haney (2017: 288-289) note, “the decades-long ethos of harsh punishment has created an equally harsh day-to-day atmosphere in many correctional facilities in which the limits of ‘acceptable’ mistreatment have been expanded.”

Second, the measures necessary to protect people in custody from intolerable penal cruelty will be *expensive*. To improve baseline prison conditions, we need substantial investments in prison healthcare, rebuilding and remodeling prisons with an eye to comfort and privacy, well-run educational and recreational programs, and an ongoing investment to provide better food. To curtail overt abuse, Gibbons and Katzenbach (2006: 408-408) recommend investing to recruit, train, and retain more qualified prison officers and establishing a fully independent government body with the resources to inspect *all* prisons and jails. The problem, of course, is that spending public funds to benefit people the public resents is not generally a winning political platform.

1.4 Putting the Pieces Together

Prison bureaucracies are beset by the very same institutional vices that produced the horrors at Abu Ghraib—including, notably, the human capacity to embrace normative shifts towards violence. That is problem enough, but it is compounded by the fact that large-scale practices dedicated to harming people are liable to contain enthusiasts who are prepared, even eager, to push the envelope. At the same time, prisoners are profoundly vulnerable, which not only means they cannot protect themselves, it may make them appealing targets to enthusiasts. And so long as

⁸ See, for example, Romer et al. (2003) and Baranauskas & Drakulich (2018).

our societies are marked by resentment, and so long as a variety of hard-to-regulate forces gin up punitive sentiment in the public, we should expect (a) intolerably cruel prison conditions and (b) limited public support for spending to fix our prisons. This is far from a comprehensive catalogue of the forces pulling our prisons towards cruelty. But it is enough to show that intolerable cruelty is not akin to software glitch, readily fixable by installing the latest update.

It is, therefore, worth entertaining the thought that our prison systems will *never* achieve a stable, secure state of moral enlightenment. Even the relatively rare, relatively humane prison bureaucracy deserves a suspicious eye: there is no law of nature that stops it getting worse. It will become cruel if the wrong people get institutional power, if key people succumb to the wrong institutional incentives, if the wrong social movements harness anger inspired by bleak social conditions and rise to prominence, or if politicians withdraw prison funding. If the last decade or so of politics has taught us anything, it is that bureaucratic institutions can readily become worse. That lesson arguably goes double for bureaucratic institutions dedicated to harming people.

So: Sending someone to a U.S. or U.K. prison is—for reasons that will be very difficult to legislate our way around—to impose terrible, undeserved risks upon her. It is a moral failure.

2 Prison is a Moral Failure Part 2: Defective Communication

As noted above, Feinberg (1965) rightly distinguishes punishment from other disincentivizing penalties by its expressive content. That contribution paved the way for a raft of justificatory theories. *Communicative* theories argue that punishment is justified because it is necessary to communicate with the wrongdoer. *Expressive* theories argue that punishment is justified because it is necessary to signal our normative commitments to society at large.

These theories are unambiguously correct about one thing: punishment has meaning, both for the wrongdoer who suffers and for the society watching on. What that meaning turns out to be, however, is fraught—particularly if we are realistic about the institutions that do the ‘speaking.’ What do we communicate to wrongdoers when our prisons are sites of intolerable cruelty? What, exactly, is the message we express to society at large?

2.1 What does Punishment Communicate?

Antony Duff (2001) is the seminal contemporary justification of punishment as communication. At its core is the distinction between exclusionary and inclusionary punishments. Pure deterrence, for Duff, is a paradigm case of exclusion. Pure deterrence, he (2001: 79) writes,

addresses [wrongdoers], not as members of the normative community of citizens, but as threatening outsiders against whom the community must protect itself. It implicitly excludes them from membership of the citizen community by longer addressing them in the terms of that community's values.

Inclusionary punishment, by contrast, respects the wrongdoer as a community member. Rather than stooping to use the language of brute threat, an inclusionary punishment speaks in the language of the community's values. It treats the wrongdoer as still being one of us.

But if wrongdoers in custody are victims of, or witnesses to, intolerable cruelty, communication goes awry. Things can go wrong in two main ways. First, wrongdoers may come to understand incarceration as an exclusionary punishment. Rather than being addressed in the putatively humane values of the community, they (or others who share their station) would be singled out for unconscionable treatment—whether that treatment consists in overt acts or a callous indifference to terrible conditions.

The second way things can go wrong might be even worse. If wrongdoers understand their punishment as inclusionary—they come to believe that they *are* being addressed in the language of their community's values—they are liable to draw grim conclusions about the substance of their community's values. Telling wrongdoers that intolerable cruelty is compatible with our values is already a striking failure of moral communication.⁹ But it is compounded by the fact that penal institutions generally employ the language of justice. When wrongdoers understand their punishment as inclusionary, then, we are not only inviting them to draw grim conclusions about our values, we are inviting them to draw grim conclusions about justice itself.

The upshot is that penal cruelty makes incarceration a highly defective form of communication. We tell wrongdoers that they have no place among us, that their community's values are inhumane, or that Thrasymachus was right about justice itself.

2.2 What does Punishment Express?

Bill Wringe argues that punishment is justified because it speaks to society at large, not to the wrongdoer suffering punishment. He (2016: 35) writes,

the purpose of the harsh treatment that punishment involves is for a society to communicate to its members that certain norms are in force and transgressions against them are viewed seriously.

But is that what we tell society when prisons are sites of intolerable cruelty? Things can again go wrong in two main ways. First, citizens may hold fast to a conception of our values that rules out intolerable cruelty. But if they do, they will face pressure to morally 'downgrade' wrongdoers in custody. If one does *not* see punishment

⁹ By communicative lights, it is a failure even if our values actually are compatible with cruelty. Communicative justifications would have us reinforce *good* values.

as a moral failure, and one thinks it would be wrong to treat any of Us so cruelly, mentally filing wrongdoers into the category of Other is the only move left. The fact that penal institutions use the language of justice again makes matters worse. It may make it harder for citizens to see punishment as a moral failure—and, therefore, encourage them to make the only move that's left.

The second way things can go wrong may, once again, be even worse. With the foregoing arguments already in place, I can be brief: when prisons are sites of intolerable cruelty, we invite society at large to draw grim conclusions about the substance of our values. Arguably, since penal institutions use the language of justice, we invite society to draw Thrasymachian conclusions about justice itself.

Expressive and communicative accounts are correct that punishment has meaning. But if we are realistic about carceral bureaucracies, the messages we communicate and express are often terrible. Punishment is therefore a failure in a second way. Not only is making people vulnerable to intolerable cruelty a failure, that failure sets the stage for another: when the risk of penal cruelty is realized, incarceration sends the wrong moral messages.

3 Prison is a Moral Failure Part 3: Assault on Moral Agency

The practical identities that give our lives meaning often come with obligations. 'Parent' comes bundled with a host of caretaking obligations; 'spouse' comes bundled with obligations to support one's significant other; and 'professor' comes bundled with obligations to our students, our Department, and, perhaps, to our profession. Understanding ourselves as parents doesn't just mean that we prioritize our caretaking obligations; our estimation of ourselves comes to depend on how excellently and how reliably we fulfill them.

This intertwining of self-estimation and certain role-based obligations makes us vulnerable. Mr. Reed, an American ex-slave, describes the way his former master cruelly exploited it.

The most barbarous thing I saw with these eyes ... my older sister, she was fooling with a clock and broke it, and my old master taken her and tied a rope around her neck—just enough to keep it from choking her—and tied her up in the back yard and whipped her I don't know how long. There stood mother, there stood father, and there stood all the children and none could come to her rescue. (Guenther 2013: xxii)

The first cruelty in this scene is obvious: the master whips the older sister for "I don't know how long." But there is a second cruelty as well, one that has to be understood not in terms of straightforward physical pain, but as a harm to the mother and father in their capacity as parents. As caretakers, parents have an obligation to protect their children—especially their young, naive children—from unjust, wanton harm. The master engineered a situation in which the parents were doomed to fail as parents, in the sense that they could not fulfill that obligation.

The severity of these kinds of identity-based harms depends on two things. First, how serious is the failure relative to the identity at issue? Burning our eager kid's

favorite meal, in normal circumstances, is a small parental failure; standing by while one's child is brutally whipped, by contrast, makes it significantly harder to see oneself as a good parent. Second, to what extent does the relevant identity form a morally valuable part of one's self-understanding? Failing in a morally problematic identity—assassin, say—might cause first-personal feelings of inadequacy. But while such feelings may be *painful*, making the assassin feel them is not necessarily *cruel*. Preventing a parent from fulfilling their basic responsibilities, by contrast, is so manifestly cruel that the episode stands out in Mr. Reed's considerable experience as the most barbarous thing he has ever seen.

How do these observations bear on moral agency and prison? Few practical identities are more central to our self-understanding than 'moral agent.' Normally, of course, we are far more likely to introduce ourselves in terms of other identities—as parents, as the spouse of so-and-so, or as the friend of a friend. ("Hi, I'm a moral agent" would be a bit weird.) But the point, here, is that moral agency makes it possible to be a parent, spouse, or friend in the first place. 'Moral agent' is central to our self-understanding because it is the necessary foundation on which so many other important identities rest.

If caretaking is a basic obligation of parents, making amends is a basic obligation of moral agents. Human animals are characteristically moral agents who err. And whereas we do not harbor any normative expectation that storms or wild animals make amends after they harm us, we do expect that of people. As Linda Radzik (2003: 328; emphasis mine) points out:

To be a moral agent is not just to be the sort of creature who can judge right from wrong in general, it is to be someone who can also judge that when she has committed a wrong, she has a duty to make amends. To say that a criminal deserves to be treated as a responsible moral agent is, in part, to say that she deserves to be treated as an agent with responsibilities. *Her interest in fulfilling these responsibilities lies at the very heart of what it is to be a moral agent.*

Now for this section's upshot: The master in Mr. Reed's story makes it impossible for the mother and father to fulfill their caretaking obligations; incarceration makes it impossible for wrongdoers to make amends, at least when that would require reparative labors outside prison walls. The responsibility to protect our children from terrible, unjust harm is central to our understanding of parenthood; the responsibility to make amends is central to our understanding of moral agency. 'Parent' was an important, morally valuable part of the mother's and father's self-understanding; given that so many of our most meaningful identities depend on it, 'moral agent' is an important, morally valuable part of our self-understanding. The master engineered a situation in which the parents were doomed to fail as parents; prison dooms inmates to fail as moral agents. If the master in Mr. Reed's story acted barbarously, prison is barbarous too.

One might object that this argument overestimates the cruelty of prison: incarceration doesn't *stop* people from making amends, it simply *delays* the process. Wrongdoers are free to make amends after they are released.

The objection is (somewhat) correct, at least for those not serving life sentences. But it wholly misses the point. Suppose the master whipped the older

sister for “I don’t know how long”—and then, for whatever reason, allowed the parents to step in and come to her rescue. That the parents got a chance to rescue their daughter may be a good thing; but stopping them for so long is still a serious, identity-based harm. So too, it may be a good thing that incarcerated people get a chance to make amends after their release; but stopping them for months, or even years, is still a serious, identity-based harm.

Moreover, the objection is only somewhat correct because prison can, and does, shape people’s lives long after their release date. Consider just one axis of influence: incarceration interrupts gainful employment, prison jobs pay next to nothing, and a criminal record can make it difficult to find gainful employment in perpetuity. Even if a wrongdoer’s reparations don’t involve victim payments, someone who spends all hours looking for, and laboring at, a series of low-paying jobs has no time to make amends.

A second objection is that whereas the master had no good reason to be cruel, there are overriding reasons to keep at least some wrongdoers in prison.

It is true that the master in Mr. Reed’s story acts without adequate justification. The problem is that whipping the daughter and forcing the parents to watch on would be cruel even if that weren’t the case. Suppose that the master’s actions were, somehow, the only way to prevent something unequivocally worse. Under this supposition, we might revise our opinion of the master: he, like the fleeing Jewish mother who had to drown her infant, would be a tragic character, not a straightforwardly evil one. But intrinsically terrible actions do not shed their terribleness just because one’s back is against the proverbial wall. So while there might be overriding reasons to keep certain people in prison, that does not overturn the point: incarceration is still an assault moral agency.

To sum up the first three sections of this paper: Incarceration is a moral failure in three different senses. It makes people vulnerable to intolerable cruelty. It is liable to send the wrong moral messages. And even when the risk of intolerable cruelty is not realized, prison is an assault moral agency.

4 Justifying Prison?

This section makes two arguments, one negative and positive. The negative argument contends that a solution to what philosophers have come to call ‘the problem of punishment’ cannot justify prison. The problem of punishment, like the fictionalized ticking time-bomb scenario, is a picture that bewitches.

The positive argument embraces the idea that if incarceration *is* a moral failure, it must be justified *as* a moral failure. It can be justified only in same broad sense that preemptively drowning an infant can be justified: it must be a last resort, used only when all better options have run out. That means, a bit ironically, that the only way to justify prison is to pursue a program of prison abolitionism as far as we responsibly can.

4.1 Bewitching Pictures

Traditional philosophy of punishment is a treasure trove. It has rightly problematized retributivism, pinpointed defects in classic deterrence theory, illuminated the expressive dimensions of punishment, and more. At the same time, however, it is important to be forthright about its limitations. The now-standard formulation of the problem of punishment, like the fictionalized ticking time bomb scenario, is a picture that bewitches.

Philosophers have largely—though not quite uniformly¹⁰—coalesced around a standard representation of the problem of punishment. David Boonin (2008: 28) puts it as follows:¹¹

[P]unishment involves not merely acts that predictably harm offenders, but acts that are carried out precisely in order to harm them [W]e must explain not only why the line between offenders and nonoffenders is morally relevant at all but, in particular, how it can be important enough to justify not merely harming those on one side of the line, but intentionally harming them.

Fictionalized ticking time-bomb cases are a bewitching picture because they depict torture in small-scale terms. They focus attention on whether we may torture *this* person *now*—and away from the problems that predictably arise when human animals run a torture bureaucracy. The traditional problem of punishment focuses attention in a similar fashion. Without countenancing the fact that prison systems are bureaucracies run by human animals and beset by a range of institutional vices, we are to focus on abstract, officially-sanctioned harms to individuals. The vices of prison bureaucracies and the frailties of the human animals who run them simply disappear from view.

A philosophical representation of a complicated problem will inevitably make a few idealizing assumptions. The worry, however, is not about idealizing assumptions in general, but about the particular assumptions at work here. They obscure something important: a tendency to spiral out of control and produce moral horrors should be highly relevant to a practice's justification. A philosophical justification of punishment which ignores that tendency—because it responds to the idealized problem of punishment—*cannot* justify the practices it is ostensibly about. §2 illustrated this problem by discussing Duff communicative theory and Wringle's expressivism; they are not alone.¹²

¹⁰ Wringle (2016: 19–41) denies that an intent to harm is definitive of punishment, suggesting that acts which foreseeably cause harm can be punishment. His objection is not, it is worth noting, that focusing on intentional harm in the abstract obscures the specific harms that predictably arise in carceral bureaucracies.

¹¹ This way of putting the problem has become so standard that it has found its way into the Stanford Encyclopedia Philosophy. Duff and Hoskins (2017) write, "How can a practice that not only burdens those subjected to it but aims to burden them, and which conveys society's condemnation, be justified?" Benn (1967), Sayre-McCord (2001), Gert, Radzik, and Hand (2004), and many others put the problem similarly.

¹² Victor Tadros (2011: 292; emphasis in original) describes his view as follows: "My aim ... was to show that general deterrence provides a central part of the justification of punishment It is true

So while traditional philosophy of punishment is a treasure trove, it cannot justify prison. That is a problem because prison has long been the default sentence for most serious crimes, and because that state of affairs shows no imminent signs of change. Prison looms large in public consciousness, there is an enormous amount of physical infrastructure dedicated to it, and vast legal and bureaucratic systems take it as given. If we hope to justify actual practices of punishment, therefore, it will be useful to have a debate with a different starting point, one that incorporates harsh realities rather than idealizing them away. If the arguments in §1, §2, and §3 are near the mark, that debate should begin on the assumption that prison bureaucracies cannot be trusted. We should conceptualize imprisonment along the same lines as preemptively drowning an infant: as a moral failure.

4.2 Justifying Incarceration as a Moral Failure

Understanding incarceration as a moral failure means giving up on the idea that it is noble or somehow constitutes justice. In what sense, then, can it be justified? As Tessman points out, the fact that one must choose between failures does not mean that all choices are equally wrong—drowning the infant is terrible, but it is less terrible than giving away the whole group’s location. So, roughly, justifying a failure is a matter of minimizing the wrong we do.

Less roughly, ‘fully’ justifying a moral failure comes down to two conditions.¹³

1. Take all reasonable steps to *preempt* the need for the terrible act.
2. If preemptory measures fail, use *less terrible alternatives* (if there are any).

If the fleeing mother could have preempted the need to drown her baby by joining a faster-moving group of refugees—here illustrating condition 1—drowning would not be ‘fully’ justified, even as a moral failure. ‘Fully’ justifying a moral failure is about minimizing awfulness, and she wouldn’t have done that. And when preemptory measures fail—here illustrating condition 2—we should seek out less terrible alternatives. Suppose the fleeing mother could either drown her infant or render the infant permanently quiet by means of profoundly painful, but not fatal, procedure. I am assuming that both of these options constitute moral failures—they are both

Footnote 12 (continued)

that offenders would be harmed as a means to the good of others. But doing this is consistent with the best understanding of the *means principle* and its limits. Harming offenders as a means is justified as it involves enforcing the duties that offenders have to protect their victims and others from future harms.” If we are evaluating theories in light of what we should realistically expect from actual prison systems, two things stand out. First, Tadros’s book cites *no evidence* that actual penal systems achieve general deterrence. That might be because there isn’t much to cite. “The big problem for deterrence theories,” Brooks (2012: 42) observes, with wonderful, almost comic bluntness, “is that punishment does not appear to have much, if any, confirmed deterrent effect.” Second, Tadros suggests that penal systems are entitled to use wrongdoers, provided that they stay within the boundaries set by his means principle—apparently envisioning penal systems that will be judicious about harming the people in their care.

¹³ The scare quotes are a reminder that fulfilling these conditions does not turn a moral failure into anything that is aptly labeled a success. The aim is just to minimize failure.

terrible. But drowning is plainly worse. If the mother could perform the procedure, but opts for drowning instead, she would be choosing a greater failure over a lesser one.

Applying all this to prison, condition 1 requires preemptory measures. So: If extreme poverty predicts crime, prison will not be ‘fully’ justified until society installs an effective social safety net. If educational attainment reduces criminality, prison will not be ‘fully’ justified until schools are excellent and readily accessible. If better healthcare (including mental healthcare) reduces crime, prison will not be ‘fully’ justified until citizens can get the treatments they need. And so on. Conceptualizing incarceration as a moral failure implies that its justification depends, in large part, on whether our societies have been appropriately diligent about mitigating the environmental factors that promote crime in the first place.

I am not arguing societies must do literally everything possible to preempt the need to incapacitate via prison. Some measures that reduce criminality may infringe too much on other important values, or achieve reductions in crime that are too small to justify their substantial cost. Figuring out exactly which measures are worth it is a task much too large to complete here. (I think establishing a safety net, improving schools, and providing healthcare are worth it; others may disagree; settling that disagreement is a task for another day.) Still, I hope the larger point is clear: if incarceration is a moral failure, sending people to prison without taking reasonable steps to mitigate the environmental causes of crime is like drowning the infant when one could have joined a faster refugee group instead.

Condition 2 requires that when preemptory measures fail, we must choose lesser failures over greater ones. Suppose, to illustrate, that while smart preemptory measures have reduced crime rates, morally serious crimes still occur. Fortunately, we have not yet reached the point where considerations of public safety require prison. Practices of Restorative Justice (RJ) have been implemented in over 80 countries (Van Ness 2005). No criminal justice administered by human animals will ever be perfect, but early signs are promising—at least when compared to prison. RJ appears to reduce recidivism, both victims and offenders generally report higher satisfaction with RJ than they do with conventional forms of criminal justice, and RJ costs less.¹⁴

Furthermore, RJ is less vulnerable than prison to this paper’s earlier critiques. §3 argued that prison is an assault on moral agency—making amends is a basic expectation of moral agents, and prison predictably undermines people’s ability to do that.

¹⁴ Brooks (2017) outlines these advantages. However, Brooks argues that programs of RJ should embrace the possibility of sentencing wrongdoers to prison. His argument is that much of the public sees RJ as being soft on crime, and that RJ could overcome this perception—and possibly play a bigger role in criminal justice—if it didn’t eschew harsh punishments like prison. The problem is two-fold. First, the advantages of RJ arguably *depend* on the fact that it eschews prison in favor of making amends. In a bid for public acceptance, Brooks’s proposal risks destroying much of what makes RJ valuable in the first place. Second, Brooks is overly hopeful, in my view, about transforming prisons into humane centers for rehabilitation. If prisons cannot realistically be trusted, programs of RJ would be wise to hold them at arm’s length. My argument in this section is that keeping RJ and prisons separate is important. That way, prison can truly function as the last resort it always should have been.

RJ, by stark contrast, is built precisely on the idea that wrongdoers are responsible for addressing the material and normative harm they have wrought. In typical cases of RJ, wrongdoers and victims work with trained mediators to hammer out a *reparative contract*—a list of concrete actions the wrongdoer must complete to make amends to the victim. So whereas prison is an assault on moral agency, RJ requires wrongdoers to be responsible moral agents.

The literature has not reached a consensus about how best to conceptualize the value of RJ. But the active taking of responsibility—using one’s agency to correct the prior misuse of one’s agency—is a strong contender. Danielle Sered runs a successful, nationally-recognized program of RJ called *Common Justice* in New York City. To After developing an agency-focused theory of accountability, she argues that prison doesn’t really hold wrongdoers accountable at all—at least insofar as ‘accountability’ means taking active steps to right one’s wrong.

All one has to do to be punished [by prison] is not escape. It requires neither agency nor dignity, nor does it require work.... No one in prison is required to face the human impacts of what they have done, to come face to face with the people whose lives are changed as result of their decisions, to own their responsibility for those decisions and the pain they have caused, and to do the extraordinarily hard work of answering for that pain and becoming someone who will not commit that harm again. While incarcerated, people are brutalized, but they are also systematically *protected and excused from all of those human burdens*. (Serred 2019: 91; emphasis mine)

The fact that RJ treats wrongdoers as agents responsible for cleaning up their mess, not as threats to be caged, is also significant for communication and expression. §2 argued that prison sends an array of awful messages. RJ can be better. We can communicate that “certain norms are in force that transgressions against them are viewed seriously”—to borrow Wringer’s language—by making wrongdoers complete reparations commensurate to the harm they have wrought. Sered’s observations in the quote above suggest that RJ might even be *better* than prison at expressing Wringer’s message. Taking moral norms seriously is, in a clear-enough sense, at odds with systematically protecting and excusing wrongdoers from the “human burdens” that come from breaking them. And if the goal is to respond to crime in an inclusive rather than an exclusive fashion—as Duff suggests—forcing wrongdoers to complete commensurate reparations is vastly preferable to locking them away.¹⁵

There is also reason to be hopeful that RJ, even when run by human animals, will dispense less intolerable cruelty than our prisons do. Again, RJ has been implemented in more than 80 countries. It is probably naïve to think that no one working in these programs has used their intuitional power for cruel ends. But there simply

¹⁵ To his credit, Duff is sympathetic with many of RJ’s aims and practices. However, Duff (2002) argues that RJ is just another form of punishment—not a fundamentally different way to hold wrongdoers accountable. In the near future, I hope to write a more theoretical argument that (i) Duff elides an important distinction between RJ and punishment, and that (ii) RJ achieves the aims of communicative and expressive theories better than punishment does. Here, however, my argument does not depend on distinguishing RJ from punishment, but from prison.

is no record of intolerable cruelty in RJ to rival the intolerable cruelty produced by our prisons. Admittedly, incidents of cruelty may well accumulate, especially if RJ plays a larger and larger role in criminal justice. The same realistic pessimism that undergirds my critique of prison in §1 means that I have to take seriously the possibility that programs of RJ, run by actual human animals, will dispense cruelty. All I can do now, however, is note that RJ has a *long* way to go before it rivals prison for cruelty. So even if actual programs of RJ sometimes result in moral failure, we can hope that RJ will in general constitute a lesser failure than prison. And if that is correct, sending people who were good candidates for RJ to prison is like drowning the infant when one could have opted for the procedure instead.

Now to make the connection to prison abolitionism explicit: Angela Davis (2003: 107) explains that abolitionism

would require us to imagine a constellation of alternative strategies and institutions.... In other words, we would not be looking for prisonlike substitutes for the prison... we would try to envision a continuum of alternatives to imprisonment—demilitarization of schools, revitalization of education at all levels, a health system that provides free physical and mental care to all, and a justice system based on reparation and reconciliation.

Davis envisions a society that makes prisons obsolete by (a) establishing a network of social services and (b) utilizing alternative approaches to criminal justice. But establishing a network of social services—better schools, an effective social safety net, and better healthcare—is to pursue the preemptory measures called for by my condition 1. Establishing a “justice system based on reparation”, such as RJ, is to provide the less terrible alternative that my condition 2 requires.

But whereas Davis optimistically casts the abolitionist program as a replacement for prison, I am more pessimistic. This late in the paper, I can only gesture towards my reasons. Practitioners of RJ, like Sered (2019: 133), are often explicit that RJ “will not fully replace incarceration; it is not a panacea.” In fact, RJ might be intrinsically unsuited for certain kinds of crime. In much domestic violence, for example, abusers “routinely use apologies as a recurring tool of abuse (often called the ‘apology cycle’)... The apology cycle is often a key factor in perpetuating the accumulating harm” (Brookes 2019: unpublished manuscript). The point is that RJ—with its focus on reparations and moral repair—is at risk of being coopted into the abuser’s apology cycle. And if RJ itself becomes a tool of abuse, it is reasonable to ask whether conventional methods like prison constitute the lesser moral failure.

But even if my pessimism is correct, and Davis’s grand political vision is unrealistic, her abolitionism is still morally and politically relevant. Pursuing it as far as we responsibly can is the only way to justify prison. After all, incarceration is a moral failure—it makes people vulnerable to intolerable cruelty, it sends the wrong moral messages, and it is an assault on moral agency. A terrible thing like that can be justified only in the sense that preemptively drowning an infant can be justified. It must be a last resort, used only when all better options have run out. And even then, it is a moral tragedy, not some broadly noble action that deserves an honorific like ‘justice.’

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

References

- Baranauskas, Andrew J., et al. 2018. Media Construction of Crime Revisited: Media types, Consumer Contexts, and Frames of Crime and Justice. *Criminology* 56 (4): 679–714.
- Berg, B. 2001. *Qualitative research methods for the social sciences*. Boston, MA: Allyn and Bacon.
- Boonin, David. 2008. *The Problem of Punishment*. NY: Cambridge University Press.
- Brooks, Thom. 2012. *Punishment*. NY: Routledge.
- Brookes, Derek. 2019. “Restorative Justice and Domestic Violence,” unpublished manuscript.
- Benn, Stanley I. 1967. “Punishment,” in *The Encyclopedia of Philosophy*, Macmillan Publishing Co., Vol. 7 & 8, p. 29.
- Davis, Angela. 2003. *Are Prisons Obsolete?* NY: Seven Stories Press.
- Dilulio, J. 1987. *Governing Prisons: A comparative study of corectional management*. New York, NY: Free Press.
- Duff, R.A. 2001. *Punishment, Communication and Community*. NY: Oxford University Press.
- Duff, R. A. 2002. “Restorative Punishment and Punitive Restoration,” in L. Walgrave (ed.) *Restorative Justice and the Law*, Cullompton, Devon: Willan.
- Duff, R. A. and Hoskins, Zachary. 2017. “Legal Punishment,” in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives.win2019/entries/legal-punishment/>.
- Feinberg, Joel. 1965. The Expressive Function of Punishment. *The Monist* 39 (3): 397–423.
- Gaes, Gerald G., and Andrew Goldberg. 2004. *Prison Rape: A Critical Review of the Literature*. Washington, DC: National Institute of Justice.
- Gert, Heather, Linda Radzik, and Michael Hand. 2004. Hampton on the Expressive Power of Punishment. *Journal of Social Philosophy* 35 (1): 79–90.
- Gibbons, J., and N. Katzenbach. 2006. *Confronting Confinement: A report of the Commission on Safety and Abuse in America’s Prisons*. New York: Vera Institute of Justice.
- Guenther, Lisa. 2013. *Solitary Confinement: Social Death and Its Afterlives*. University of Minnesota Press.
- Kimmett, E., and I. O’Donnell. 2002. Routine Victimization in Prisons. *The Howard Journal of Criminal Justice* 37: 266–279.
- Liebling, Alison. 2004. *Prisons and their Moral Performance: A Study of Values, Quality, and Prison Life*. NY: Oxford University Press.
- Luban, David. 2005. Liberalism, Torture, and the Ticking Bomb. *Virginia Law Review* 91: 1425–1461.
- Miller, Seumas. 2005. Is Torture Ever Morally Justifiable? *The International Journal of Applied Philosophy* 19 (2): 179–192.
- Ramirez, Mark D. 2013. Punitive Sentiment. *Criminology* 51 (2): 329–364.
- Radzik, L. 2003. “Do Wrongdoers Have a Right to Make Amends?” *Social Theory and Practice* 29 (2).
- Romer, D., et al. 2003. Television News and the Cultivation of Fear of Crime. *Journal of Communication* 53 (1): 88–104.
- Saul, Jennifer. 2018. “Dog whistles, Political Manipulation, and Philosophy of Language,” in Fogal, Harris, and Moss (eds.) *New Work on Speech Acts*. Oxford University Press.
- Sayre-McCord, Geoffrey. 2001. Criminal Justice and Legal Reparations as an Alternative to Punishment. *Philosophical Issues* 11: 502–529.
- Sered, Danielle. 2019. *Until We Reckon: Violence Mass Incarceration, and the Road to Repair*. NY: The New Press.
- Sered, D. 2019. *Until We Reckon*. NY: The New Press.
- Struckman-Johnson, C., et al. 1996. Sexual coercion reported by men and women in prison. *The Journal of Sex Research* 33 (1): 67–76.
- Struckman-Johnson, C., and D. Struckman-Johnson. 2000. Sexual Coercion Rates in Seven Midwestern Prison Facilities for Men. *The Prison Journal* 80 (4): 379–390.

- Tadros, V. 2011. *The Ends of Harm: The Moral Foundations of Criminal Law*. NY: Oxford University Press.
- Tessman, Lisa. 2016. *Moral Failure: On the Impossible Demands of Morality*. Oxford University Press.
- Trulson, C. R., J. W. Marquart, and J. L. Mulling. 2004. Breaking in: Gaining entry to prisons and other hard-to-access criminal justice organizations. *Journal of Criminal Justice Education* 15 (2): 451–478.
- Van Ness, Daniel W. 2005. “An Overview of Restorative Justice Around the World,” presented at the Eleventh United Nations Congress on Crime Prevention and Criminal Justice.
- Waldron, Jeremy. 2005. Torture and Positive Law: Jurisprudence for the White House. *Columbia Law Review* 105 (6): 1681–1750.
- Weill, Joanna, and Craig Haney. 2017. Mechanisms of Moral Disengagement and Prisoner Abuse. *Analyses of Social Issues and Public Policy* 17 (1): 286–318.
- Wolff, Nancy, et al. 2007. Physical Violence Inside Prisons: Rates of Victimization. *Criminal Justice and Behavior* 34 (5): 588–599.
- Wolfendale, Jessica. 2020. Prisons as a Torturous Institution. *Res Philosophica* 97 (2): 297–324.
- Wringe, William. 2016. *An Expressive Theory of Punishment*. Palgrave MacMillan UK.
- Zimbardo, Philip. 2008. *The Lucifer Effect: Understanding how Good People Turn Evil*. NY: Random House.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.