

## CARUSO'S NEGATIVE PROJECT IN *REJECTING RETRIBUTIVISM*

RAFF DONELSON

*IIT Chicago-Kent College of Law*


In his recent book, *Rejecting Retributivism* (2021), Gregg Caruso offers a rigorous set of arguments to unseat retributivism as the dominant justificatory theory of punishment and to institute in its place his own public health-quarantine response to criminal wrongdoing. In these two major ambitions, I am largely sympathetic, but I part ways with respect to certain arguments and certain ways of framing the debate. In this brief commentary, I first summarize the key arguments with which I engage. Next, I unpack a few of the places where I register modest disagreement. Finally, I close with a broad question about the public health model and the dispositions it requires of us.

### Summary

*Rejecting Retributivism* advances six major arguments against retributivism. Recall that retributivism is a justificatory theory of punishment—that is, the theory offering conditions under which punishment is morally justified. Simply, retributivists maintain that a criminal wrongdoer is to be punished because, and to the extent that, they deserve punishment for the wrong committed in the past. Of course, all justificatory theories of punishment will grant that in some way the person to be punished deserves the punishment, but what marks retributivism as distinct is a more *sui generis* desert claim, rather than one derived from other arguments. Desert, for retributivism, does all of the justificatory work. By contrast, on a classical specific-deterrence theory of punishment—a theory claiming that punishment is justified because, and to the extent that, bestowing the punishment on the offender will deter that offender from reoffending—the

---

**Contact:** Raff Donelson <[rdonelson1@kentlaw.iit.edu](mailto:rdonelson1@kentlaw.iit.edu)>

 <https://orcid.org/0000-0002-7708-5458>

offender, rather than someone else, deserves punishment, but *they* deserve the punishment because that will best deter *them* from reoffending, allegedly. With this rough definition of *retributivism* borne in mind, let us return to Caruso's arguments against it.

First, he argues that we lack the kind of free will necessary to deserve criminal punishment at all. Since retributivists appeal to desert to justify punishment and there is no desert, retributivists fail to justify punishment. This is Caruso's free-will skepticism. This controversial argument stands independent of the other arguments.

Caruso's second argument assumes for the sake of argument that one isn't bowled over by the case for free-will skepticism. One is instead moved by the arguments but agnostic. According to Caruso's second argument, even if we are agnostic about the truth of free-will skepticism, we should refrain from punishing people on a retributivist desert theory. Why? Well, as he sees it, we should only punish when we have a high level of epistemic justification in the underlying facts establishing someone's blameworthiness, but, *ex hypothesi*, we are agnostic about the truth of free-will skepticism.

The next four arguments are supposed to work even if free-will skepticism does not succeed, and I begin with the third and fourth, which can be understood as a single disjunctive argument. Many criminal justice systems, including those operative in the United States, purposely ignore factors that, according to retributivists themselves, determine the degree of one's blameworthiness (e.g., facts about one's upbringing, one's mental states, etc.). According to Caruso, if those criminal justice systems continue to behave this way, they are unjustified; however, if we try to modify the systems to take account of those factors, one will find that the state lacks the epistemic resources to apprehend the relevant factors.

Fifth, Caruso claims that our judgments about desert, specifically our judgments about how much harm an agent should receive for certain 'wrongful' acts, are too uncertain and manipulable.

Sixth and finally, Caruso argues that criminal punishment, at least as envisioned by retributivists, simply fails to solve our crime problems. While my commentary focuses on areas of disagreement, I note here that this sixth argument is, by far, the most compelling for me. Criminal punishment is always only a response to crime already committed; it does not reach the source of the crime problem. In my view, so long as punishment is our biggest effort in the battle against crime, we will always fail.

In any event, after these six strikes against retributivism, Caruso has a few critical comments for other justificatory theories of punishment, particularly deterrence theories and the moral education theory of punishment. For our purposes, it is most important to recall his comments about the moral education theory. This theory claims that punishment is justified, when it is, because and

to the extent that it educates offenders about what is acceptable and unacceptable conduct. For Caruso, this theory is to be rejected in favor of 'an altogether different nonretributive alternative' (168) because (1) there are plenty of criminals for whom the education would be useless, (2) adult offenders are often insufficiently malleable to benefit from such education, and (3) punishment may do more harm than good.

## Modest Disagreement

Below, I register a few qualms with Caruso's arguments. I start with voicing those worries that I wish Caruso had anticipated and addressed. These earlier worries raise questions for which I hold no settled view, but I think there is a case to be made for disagreeing with Caruso. After these, I turn to raising concerns that I find more devastating to Caruso's arguments.

### *A Different Epistemic Argument*

First, I consider his Epistemic Argument. Like anyone else, I tend to agree that we should refrain from marshalling the resources of the state to punish unless we have strong grounds for issuing the punishment. I also think that agnosticism may well be the best response to the free-will debate. Still, I wonder if there is not a slightly different epistemic argument that one could run, which would imperil the point made by Caruso.

Let there be an action  $A$ , such that  $A$  is morally obligatory if  $x$  obtains and morally impermissible if  $x$  does not obtain. For  $A$ , we should really find out whether condition  $x$  obtains or not. But what if we are rightly agnostic, given our evidential set? We stand to err on either route, so it is not obvious which error to avoid.<sup>1</sup> If this sort of situation seems possible, let me fill in the variables to yield a different epistemic argument that challenges Caruso's own epistemic argument.

Let  $A$  be punishing someone who has committed a heinous crime, and let  $x$  be that the person is truly blameworthy. Several moral philosophers, most notably Kant, have argued that we are morally obliged to punish wrongdoers (Kant 1991: 331), and many agree that we are not morally permitted to punish when someone is morally blameless. Supposing that both of these are reasonable moral requirements, Caruso, in essence, says that we need to have lots of

---

1. In some ways, the following line of argument borrows from William James in 'The Will to Believe'. James questions why we should only worry about being confident that  $p$  when the  $\neg p$  is true. Why not also be worried about suspending belief about  $p$  when  $p$  is true? See James (1997).

evidence to ensure that we do not flout the obligation to refrain from punishing the innocent. But I do not understand why not take the flipside position: we should have lots of evidence to ensure that we do not flout the obligation to punish the blameworthy.

Perhaps the obvious out is to say that Kant and company are wrong, but I'm not as sure when I think of the most heinous criminals—for instance, Dylann Roof, who killed nine Black people at a South Carolina church in 2015 (CBS News 2017). If Roof is really blameworthy, it does seem wrong not to punish him, quibbles about free will notwithstanding.

### *A Pragmatic Worry about Free Will*

Second, I consider Caruso's free-will argument itself. My qualm does not concern the underlying metaphysics. Though Caruso does a fantastic job walking through decades of debate, making things clear to novices like myself, I simply cannot shake my own agnosticism on this score. My worry is a practical one about the *evaluative* upshot of free-will skepticism. Talking about his public health model of handling crime, Caruso writes, 'Such an approach . . . adopts a historical whole person approach that sees individuals as *byproducts of their histories and circumstances*', (154). I don't like seeing individuals as mere '*byproducts of their histories and circumstances*'. The existentialist Jean-Paul Sartre would say that seeing other people like that is a case of bad faith, a case of denying part of our nature (Sartre 1956: 47–48). Part of a person is their history and circumstance, but we also have the capacity for more, Sartre thinks (47–48). Maybe Sartre's thought is silly romanticism, but I suspect there might be more to this. If we are as Caruso suggests, little explains why human beings have any special value that could be the source of duties of respect.

To see this concern in another light, consider Caruso's discussion on whether humans really have a sort of self-determination that serves as one of those human capabilities that we should respect and foster. Humans are self-determining, for Caruso, in the same way as the helmsman in the example from Aristotle's *Politics*: the helmsman cannot control the wind, cannot make the boat glide upon dry land, but he still charts his own course (Caruso 2021: 216). Of course, there's a serious disanalogy that Caruso fails to mention. The helmsman 'charts his own journey' even though everything about the journey was to go exactly as it did well before the helmsman ever existed. Sure, the ship's movement results from determinations of the helmsman's will, but the will was determined by forces outside of it. Metaphysically speaking, there is no great difference between this situation and the one where the helmsman is made to steer in a certain way by a pirate at knifepoint. Caruso himself relies on such manipulation analogies

earlier in the book to cast aspersions on compatibilism, but later on, he claims this unimpressive form of self-determination has value. I don't see it.

### *On the Disjunctive Argument*

For the aforementioned worries, I offered what I take to be the best case against Caruso's powerful arguments. Still, those worries are offered in a more exploratory tone. The following three worries concern matters where I feel on surer footing. I begin with the disjunctive argument.

Recall that, for Caruso, contemporary criminal justice systems currently fail to consider important (mitigating) factors in assigning criminal blame. In doing so, these systems are unjustified; however, if one were to try to modify the systems to take account of those factors, one will find that the state lacks the epistemic resources to apprehend the relevant factors. I am not convinced.

Caruso rightly notes that sentencing courts should consider things like poverty, unjust deprivations of opportunity, and serious mental illness. Of course, if one were to read sentencing briefs, one would see that sentencing judges and juries hear about those factors all the time (Meixner 2022). If a court hears the mitigation and still issues a sentence that scholars or defense counsel find to be too long, that simply indicates a philosophical disagreement about the mitigating force of particular information. It does not indicate an *inability* of the criminal justice system to consider such factors.

### *On Our Shifting Desert Judgments*

In addition to claiming prematurely that the criminal justice system cannot consider mitigating evidence, Caruso also claims that we cannot in practice properly assess desert claims. As he puts it, 'There simply is no magic ledger to look to that objectively and impartially spells out a rank order of wrongs in one column and the punishment deserved for each in another' (142).

This argument is puzzling. Either we think we can answer moral questions or we cannot. If we cannot, Caruso's own project of exposing moral flaws in the retributivist framework is doomed. If we can, I fail to see why this specific rank-ordering task presents an insurmountable obstacle. That there is little or no consensus now over the proper rank-ordering or that we now disagree with humans from the distant past is of little *theoretical* consequence, unless we are given to a skepticism that would threaten Caruso too (Bambauer and Roth 2021: 1667).

And speaking of theoretical consequences, a word is in order about the theoretical versus the practical. We must keep in mind that retributivism is a

philosophical view and that sophisticated proponents of the view are not apologists for whatever nonsense emerges from courts or corrections facilities (Stewart 2018: 45; Brooks 2021; Western and Wildeman 2009). Arguments against what happens in actual practice are just irrelevant for assessing the philosophical position.

Moreover, even in talking about actual practice, I find Caruso to offer an overblown pessimism. The very discussion in the book betrays the point he tries to make. In talking about the racist crack and powder cocaine disparity in American law, Caruso likely believes that the reader finds this rank-ordering absurd too (Caruso 2021: 143–45). Many, many people believe that this is patently absurd and unjustified (Donelson 2019: 24–25); three separate federal commissions recommended getting rid of the disparity (ACLU 2016). The American Congress refuses to do so, not because the moral calculus is difficult, but because they are pandering to racist and often uninformed constituencies. This is an easy moral question, and there are likely plenty more.

### *On Moral Education and the Positive View*

As a final note, I want to say something about Caruso's own positive view. This will sound flippant but hopefully not wholly inaccurate. Caruso spends a few paragraphs unfairly maligning the moral education view only to adopt it later as his own with a new name and embedded in a larger scheme.

To ground this discussion, I reiterate Caruso's worries about the moral education view. First, Caruso says there are plenty of criminals for whom the education would be useless; second, adult offenders are often insufficiently malleable to benefit from such education; and third, punishment may do more harm than good.

Caruso voices three concerns, but the second is just a specialized version of the first, so there are only two concerns: a concern about the education not helping every wrongdoer and a concern that the alleged cure may be worse than the disease.

We can dispose of the first worry rather quickly by noting that punishment is not supposed to be a panacea to every instance of wrongdoing. Those with severe mental illness, which inhibits criminal responsibility, are not to be punished on most theories of punishment, the moral education theory included (Caruso 2021: 166–68). That does not mean that the punishment theorist thinks literally nothing should be done. Like anyone else, the moral education proponent can insist on finding nonpunitive ways of dealing with problems that punishment is not designed to fix. Maybe incapacitation and medical treatment are apt responses in some cases; maybe some other nonpunitive response is in order.

The second worry about the moral education theory is tougher to dispatch. Maybe punishment invariably or usually harms more than it helps. As Caruso notes, this is an empirical question. To go about answering this question, we would have to settle on what punishment is. If our model for investigation is the American system replete with solitary confinement (18), prison rape,<sup>2</sup> and few opportunities for self-improvement (254), I can grant that the research is unlikely to redeem that kind of punishment. But punishment need not look like that. Caruso, for instance, has some good things to say about Norwegian corrections (324). This system of corrections looks like punishment to me, insofar as it deprives people of the freedom to come and go, denies them certain livelihoods, and forces them to take on chores they may not wish to perform, and yet, it also seems, by many accounts, to effect rehabilitation (254). If this is punishment, it may well do the work of moral education, just as Caruso himself would like.

In deflecting these two worries, I suggest that moral education theory offers a partial solution to wrongdoing that must be supplemented by other interventions and that there are signs of its effectiveness. If what I've said is sound, where else can Caruso find fault with the view? I ask the question genuinely, not rhetorically. I genuinely wonder if Caruso simply offers a version of the same view, one whose branding focuses mostly on the nonpunitive supplement.

## A Final Broad Question

Caruso's book shines not only because he shows off his expertise in the arcane free-will debates but also because he displays an enviable facility with social science research. Though I am no social scientist myself, I did notice something that struck me as an error.

Caruso mentions that over a few years, Medellín, Colombia, cut its murder rate by 95 percent after some targeted infrastructure spending (254). In reading this, I recalled two things. First, for the past decade, Medellín has had higher levels of armed robbery every year (Macia 2019). Second, the lowered homicide rate likely has more to do with successful government negotiations with a major militia group than with neighborhood-level intervention (O'Hanlon 2016).

So what? The *so what* of this story is not about Medellín or whether infrastructure spending can help with crime. The bigger question is about how governments approach crime-fighting. They do not adopt a scientific approach; they employ an active approach. Governments usually do lots of things at once. When your city has the highest levels of murder on Earth, you negotiate with the

---

2. Prisons reported 18,884 allegations of sexual assault in 2018, 2.5 times the 7,575 reported six years earlier in 2012. See Maruschak Buehler (2021).

militias, you change out chiefs of police, you fund various initiatives—in scientific speak, you manipulate lots of variables. This approach is perhaps justified because we are incensed and impatient with wrongdoing. One thing I worry about in making the public health analogy is that in medicine, we have controlled experiments. With controlled experiments, we patiently isolate variables and come to understand which interventions work. In the world of crime and punishment, for now at least, we only have natural experiments. We are not willing to have endless trials to understand what works and doesn't. We have one natural experiment, and we hope that we didn't miss an important variable, like the 2016 deal with FARC (Revolutionary Armed Forces of Colombia) (Aljazeera 2019).

My big question for Caruso is whether he can imagine a world where we treat crime and punishment in a more patient and scientific way. The book urges so many big changes to our thinking. Caruso urges us to drop our belief in free will, to see people as byproducts of circumstance, to stop being retributive, and, most of all, to react to crime not as evil to be haphazardly attacked but as a theoretical question to be solved. Perhaps that last one is the biggest ask of all.<sup>3</sup>

## References

- ACLU. 2006. *Cracks in the System: Twenty Years of the Unjust Federal Crack Cocaine Law*, 6 <<https://www.aclu.org/other/cracks-system-20-years-unjust-federal-crack-cocaine-law>>
- Al Jazeera. 2016. 'Colombia Reaches New Peace Deal with FARC Rebels', <<https://www.aljazeera.com/news/2016/11/13/colombia-reaches-new-peace-deal-with-farc-rebels>>
- Bambauer, Jane R., and Andrea L. Roth. 2021. 'From Damage Caps to Decarceration: Extending Tort Law Safeguards to Criminal Sentencing', *Boston University Law Review*, 101: 1667
- Bender, Kathleen. 2018. 'Education Opportunities in Prison Are Key to Reducing Crime', AmericanProgress.org <<https://www.americanprogress.org/article/education-opportunities-prison-key-reducing-crime/>>
- Brooks, Thom. 2021. *Punishment: A Critical Introduction* (New York: Routledge)
- Caruso, Gregg D. 2021. *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice. Law and the Cognitive Sciences* (Cambridge: Cambridge University Press)
- CBS News. 2017. 'Dylann Roof, Charleston Church Shooter, Gets Nine Life Sentences' <<https://www.cbsnews.com/news/dylann-roof-charleston-church-shooter-gets-nine-life-sentences/>>
- Donelson, Raff. 2019. 'Legal Inconsistencies', *Tulsa Law Review*, 55: 24–25

---

3. I thank the Society for Philosophy of Agency for inviting me to take part in the author's-meets-critics session at the American Philosophical Association conference that led to this commentary. I also thank the editors at this journal for helpful feedback and Nick Gonano for excellent research assistance.



- James, William. 1997. 'The Will to Believe', in *Pragmatism: A Reader*, ed. by Louis Menard (New York: Vintage Books)
- Kant, Immanuel. 1991. *The Metaphysics of Morals*, trans by Mary J. Gregor (New York: Cambridge University Press), pp. 331
- Macia, Andrew. 2019. '11 Ways to Avoid Being Robbed in Medellín', *MedellinBuzz.com* <<http://medellinbuzz.com/medellin-safety/>>
- Maruschak, Laura M., and Emily D. Buehler. 2021. 'Survey of Sexual Victimization in Adult Correctional Facilities, 2012–2018—Statistical Tables' (Washington, DC: PREA Publications) <<https://bjs.ojp.gov/library/publications/survey-sexual-victimization-adult-correctional-facilities-2012-2018>>
- Meixner, John B., Jr. 2022. 'Modern Sentencing Mitigation', *Northwestern University Law Review*, 116
- O'Hanlon, Michael E. 2016. 'Once a Drug Den, Medellín Is on a New Path', *Brookings* <<https://www.brookings.edu/blog/order-from-chaos/2016/08/13/once-a-drug-den-medellin-is-on-a-new-path/>>
- Sartre, Jean-Paul. 1956. *Being and Nothingness: An Essay on Phenomenological Ontology*, trans. by Hazel E. Barnes (New York: Philosophical Library), pp. 47–48
- Stewart, Hamish. 2018. 'The Wrong of Mass Punishment', *Criminal Law & Philosophy*, 12: 45
- Western, Bruce, and Christopher Wildeman. 2009. 'The Black Family and Mass Incarceration', *Annals of the American Academy of Political and Social Sciences*, 621