Punishment and Desert

Gregg D. Caruso [*]
gcaruso@corning-cc.edu

ABSTRACT
This paper explores the relationship between punishment and desert and offers two distinct sets of reasons for rejecting the retributive justification of legal punishment – one theoretical and one practical. The first attacks the philosophical foundations of retributivism and argues that it's unclear that agents have the kind of free will and moral responsibility needed to justify it. I present stronger and weaker versions of this objection and conclude that retributive legal punishment is unjustified and the harms it causes are prima facie seriously wrong. The second objection maintains that *even if* one were to assume that wrongdoers are deserving of retributive punishment, contra concerns over free will, we should still abandon retributivism since there remain insurmountable practical difficulties that make it impossible to accurately and proportionally distribute legal punishment in accordance with desert. In particular, I present the *Misalignment Argument* and *Poor Epistemic Position Argument* and argue that, taken together, they create a powerful new challenge to retributivism called the *Retributivist Tracking Dilemma.*

Punishment involves the intentional imposition of an unpleasant penalty or deprivation for perceived wrongdoing upon a group or individual, typically meted out by an authority. Everyday examples include a parent punishing their teenager for bad behavior by taking away their cellphone privileges or a university expelling a student for plagiarism. *Legal punishment* is a specific sort of punishment; it is the intentional imposition of a penalty for conduct that is represented, either truly or falsely, as a violation of a law of the state, where the imposition of that penalty is sanctioned by the state's authority. More precisely, we can say that legal punishment consists in a person or persons deliberately harming another on behalf of the state in a way that is intended to constitute a fitting response to some offense and to give expression to the state's disapproval of that offense

[*] SUNY Corning, USA.

(Zimmerman, 2011; Boonin, 2008; Walen, 2014). According to this definition, Person A legally punishes Person B if and only if A acts on behalf of the state in such a way that (1) they harm the punishee – this could include imposing an unpleasant penalty or deprivation; (2) this harm is intended by the state; (3) this harm is believed by the state to be fitting – in particular, fitting to the fact, perhaps in conjunction with some other facts, that the punishee is associated with some legal offense; (4) they thereby express the state's disapproval both of the offense and of the offender; and (5) they thereby act in some legal official capacity (see Zimmerman, 2011: 20; Caruso, 2021a).

Within the criminal justice system one of the most prominent justifications for legal punishment, both historically and currently, is *retributivism*. The retributive justification of legal punishment maintains that, absent any excusing conditions, wrongdoers are morally responsible for their actions and *deserve* to be punished in proportion to their wrongdoing. Unlike theories of punishment that aim at deterrence, rehabilitation, or incapacitation, retributivism grounds punishment in the *blameworthiness* and *desert* of offenders. It holds that punishing wrongdoers is intrinsically good. For the retributivist, wrongdoers deserve a punitive response proportional to their wrongdoing, even if their punishment serves no further purpose. This means that the retributivist position is not reducible to consequentialist considerations nor in justifying punishment does it appeal to wider goods such as the safety of society or the moral improvement of those being punished.

This paper explores the relationship between punishment and desert and offers two distinct sets of reasons for rejecting the retributive justification of legal punishment – one theoretical and one practical. The first attacks the philosophical foundations of retributivism and argues that it's unclear that agents have the kind of free will and moral responsibility needed to justify it. I present stronger and weaker versions of this objection and conclude that retributive legal punishment is unjustified and the harms it causes are prima facie seriously wrong. The second objection maintains that *even if* one were to assume that wrongdoers are deserving of retributive punishment, contra concerns over free will, we should still abandon retributivism since there remain insurmountable practical difficulties that make it impossible to accurately and proportionally distribute legal punishment in accordance with desert. In particular, I present the *Misalignment Argument* and *Poor Epistemic Position Argument* and argue that, taken together, they create a powerful new challenge to retributivism called the

Retributivist Tracking Dilemma. In light of these theoretical and practical objections, I conclude that we should reject retributive legal punishment and abandon efforts to justify state-imposed punishment as morally deserved.

1. Retributivism

According to the retributivist justification of legal punishment, wrongdoers deserve the imposition of a penalty solely for the backward-looking reason that they have knowingly done wrong. Michael S. Moore, a leading retributivist, highlights this purely backward-looking nature of retributivism when he writes:

> [R]etributivism is the view that we ought to punish offenders because, and only because, they deserve to be punished. Punishment is justified, for a retributivist, solely by the fact that those receiving it deserve it. Punishment may deter future crime, incapacitate dangerous persons, educate citizens in the behaviour required for a civilized society, reinforce social cohesion, prevent vigilante behaviour, make victims of crime feel better, or satisfy the vengeful desires of citizens who are not themselves crime victims. Yet for the retributivist these are a happy surplus that punishment produces and form no part of what makes punishment just: for a retributivist, deserving offenders should be punished even if the punishment produces none of these other, surplus good effects. (1997: 153)

This backward-looking focus on desert is a central feature of all traditional retributive accounts of punishment (see, e.g., Kant, 1797/2017; von Hirsch, 1976, 1981, 2007, 2017; Husak, 2000; Kershnar, 2000, 2001; Berman, 2008, 2011, 2013, 2016; Walen, 2014). And it is important to emphasize that the desert invoked in retributivism (in the classical or strict sense) is *basic* in the sense that it is not in turn grounded in forward-looking reasons such as securing the safety of society or the moral improvement of criminals. Thus, for the retributivist, the claim that persons are morally responsible for their actions in the *basic desert* sense is crucial to the state's justification for giving them their just deserts in the form of punishment for violations of the state's laws.[1]

[1] Retributivists typically also hold that just punishments must be *proportional* to wrongdoing. Both the justificatory thesis and the proportionality requirement for punishments are reflected in Mitchell Berman's statement of retributivism: "A person who unjustifiably and inexcusably causes or risks harm to others or to significant social interests deserves to suffer for that choice, and he deserves to suffer in proportion to the extent to which his regard or concern for others falls short of what is properly demanded of him" (2008: 269).

Depending on how retributivists view the relationship between desert and punishment, we can identify three different varieties of the view – *weak*, *moderate*, and *strong*. *Weak retributivism* maintains that negative desert, which is what the criminal law is concerned with when it holds wrongdoers accountable, is merely necessary but not sufficient for punishment. That is, weak retributivism maintains that while desert is a necessary condition for punishment, it is not enough on its own to justify punishment – other conditions must also be met. As Alec Walen describes it, weak retributivism is the view that "wrongdoers forfeit their right not to suffer proportional punishment, but that the positive reasons for punishment must appeal to some other goods that punishment achieves, such as deterrence or incapacitation" (2014). On this view, then, the desert of the wrongdoer is a necessary condition for punishment, since it removes the protection against punishing innocent people, but it does not itself provide a positive justification for punishment – additional justifications for punishment, ones that go beyond the backward-looking desert of wrongdoers, must also be provided.

Moderate retributivism, on the other hand, maintains that negative desert is necessary and sufficient for punishment but that desert does not mandate punishment or provide an obligation to punish in all circumstances – that is, there may be other goods that outweigh punishing the deserving or giving them their just deserts (Robinson and Cahill, 2006). Leo Zaibert, while eschewing the taxonomy offered here, defends a kind of moderate retributivism when he argues:

> There are many reasons why sometimes refraining from punishing a deserving wrongdoer is more valuable than punishing him – even if one believes that there is [intrinsic] value in inflicting deserved punishment. Perhaps the most conspicuous cases are those in which the refraining is related to resource-allocation and opportunity costs… To acknowledge the existence of these cases is not to thereby *deny* the value of deserved punishment: it is simply to recognize that this value, like any value, can be – and often is – lesser than other values. (2018: 20)

Mitchell Berman also defends a form of moderate retributivism, which he calls "modest retributivism" (2016), since he maintains that negative desert grounds a justified reason to punish, but not a duty. For moderate retributivists, negative desert is sufficient to justify punishment but other values and considerations may outweigh inflicting the deserved punishment.

Lastly, *strong retributivism* maintains that desert is necessary and sufficient for punishment but it also grounds a duty to punish wrongdoers. Immanuel Kant (1797/2017: Part II: 6) is perhaps the most famous representative of this latter view, but Moore also defends a form of strong retributivism and argues, like Kant, that society has a duty to punish culpable offenders:

> We are justified in punishing because and only because offenders deserve it. Moral responsibility ("desert") in such a view is not only necessary for justified punishment, it is also sufficient. Such sufficiency of justification gives society more than merely a right to punish culpable offenders. It does this, making it not unfair to punish them, but retributivism justifies more than this. For a retributivist, the moral responsibility of an offender also gives society the duty to punish. Retributivism, in other words, is truly a theory of justice such that, if it is true, we have an obligation to set up institutions so that retribution is achieved. (1997: 91)

Strong retributivists therefore defend two distinct claims: (1) that negative desert is sufficient to justify punishing wrongdoers on the grounds that they deserve it and (2) that we have a duty to do so. Moderate retributivists, on the other hand, seek only to defend the first claim.

In what follows, I will limit my discussion to moderate and strong varieties of retributivism and leave weak retributivism aside. I will do so because, first, most leading retributivists defend one of these stronger forms of retributivism and it is my desire to address the dominant view, not a subordinate view held by few. Second, weak retributivism is considered by many retributivists to be "too weak to guide the criminal law" and as amounting to nothing more than "desert-free consequentialism side constrained by negative desert" (Alexander, Ferzan, and Morse, 2009: 7). In fact, some theorists simply define retributivism in a way that excludes *weak retributivism* from consideration altogether. Mitchell Berman, for example, maintains that the "core retributivist thesis" is that, "[t]he goodness or rightness of satisfying a wrongdoer's negative desert morally justifies [i.e., is sufficient for] the infliction of criminal punishment, without regard for any further good consequences that might be realized as a contingent result of satisfying the wrongdoer's desert" (2016: 4). Lastly, the weight the criminal law gives desert and the way retributivism is practically implemented in the law (especially in the United States) indicate that the desert of offenders is typically seen as sufficient for punishment (see Caruso, 2021a: 7-9).

For these reasons, I will take as my target the claim that the desert of offenders provides sufficient grounds for punishment and that we are therefore

justified in sometimes punishing wrongdoers for no purpose other than to see the guilty get what they deserve. Since this core claim is held in common among all moderate and strong varieties of retributivism, I will henceforth drop the moderate/strong distinction and focus instead on this shared feature.

2. Rejecting Retributivism: Skeptical and Epistemic Arguments

My first objection to the retributive justification of legal punishment is that it's not at all clear that agents have the kind of free will and moral responsibility needed to justify it. *Free will skepticism*, for example, constitutes a family of views that deny that human beings have the control in action – that is, the free will – required for moral responsibility in the basic desert sense (see, e.g., Pereboom, 2001, 2014; Strawson, 1986; Levy, 2011; Waller, 2011; Caruso, 2012, 2021a; Caruso in Dennett and Caruso, 2021; Caruso and Pereboom, 2022).[2] In the past, the standard argument for free will skepticism was based on the notion of *determinism* – the thesis that facts about the remote past in conjunction with the laws of nature entail that there is only one unique future. *Hard determinists* argued that determinism is true and incompatible with free will and basic desert moral responsibility – either because it precludes the *ability to do otherwise* (leeway incompatibilism) or because it is inconsistent with one's being the ultimate or appropriate source of action (source incompatibilism).

More recently, however, a number of contemporary philosophers have presented arguments against basic desert moral responsibility that are agnostic about determinism – e.g., Pereboom (2001, 2014), G. Strawson (1986), Smilansky (2000), Levy (2011), Waller (2011, 2015), and myself (2012, 2021a). Most argue that while determinism is incompatible with free will and basic desert moral responsibility, so too is indeterminism, especially the variety posited by quantum mechanics. Others argue that regardless of the causal structure of the

[2] For an agent to be morally responsible for an action in the *basic desert* sense is for them to deserve the harm or pain of blame or punishment *just because* they acted wrongly, given that they were aware or should have been aware that the action was wrong. Understood this way, free will is a kind of power or ability an agent must possess in order to justify certain kinds of desert-based judgments, attitudes, or treatments–such as resentment, indignation, moral anger, and retributive punishment–in response to decisions or actions that the agent performed or failed to perform. These reactions would be justified on purely backward-looking grounds–that is what makes them *basic*–and would not appeal to consequentialist or forward-looking considerations, such as future protection, future reconciliation, or future moral formation (see Pereboom, 2001, 2014; Caruso and Pereboom, 2022).

universe, we lack free will and moral responsibility because free will is incompatible with the pervasiveness of *luck*. Others (still) argue that free will and ultimate moral responsibility are incoherent concepts, since to be free in the sense required for ultimate moral responsibility, we would have to be *causa sui* ("cause of oneself") and this is impossible. What all these arguments have in common, and what they share with classical hard determinism, is the thesis that what we do and the way we are is ultimately the result of factors beyond our control and because of this we are never morally responsible for our actions in the basic desert sense.

My own route to free will skepticism presents arguments that target the three leading rival views – event-causal libertarianism, agent-causal libertarianism, and compatibilism – and then claims the skeptical position is the only defensible position that remains standing. Against the view that free will is compatible with the causal determination of our actions by natural factors beyond our control, I argue that there is no relevant difference between this prospect and our actions being causally determined by manipulators (see Pereboom 2001, 2014; Caruso 2012, 2021a; Caruso and Pereboom 2022). Against event causal libertarianism, I object (among other things) that on such accounts agents are left unable to *settle* whether a decision occurs and hence cannot have the control required for moral responsibility (see Pereboom 2001, 2014; Caruso 2012, 2021a; Caruso and Pereboom 2022). I further maintain that non-causal accounts of free will suffer from the same problem (see Pereboom 2001, 2014). While agent-causal libertarianism could, in theory, supply this sort of control, I argue that it cannot be reconciled with our best philosophical and scientific theories about the world and faces additional problems accounting for mental causation (Caruso 2012, 2021a; Caruso and Pereboom 2022). Since this exhausts the options for views on which we have the sort of free will at issue, I conclude that free will skepticism is the only remaining position.[3]

If this conclusion is correct, as I believe it is, it undermines the retributive justification for punishment since it does away with the idea of *basic desert*.

[3] This is only a rough outline of my case for *hard-incompatibilism*, which maintains that the sort of free will required for basic desert moral responsibility is incompatible with *both* causal determination by factors beyond the agent's control *and* with the kind of indeterminacy in action required by the most plausible versions of libertarianism (see Pereboom 2001, 2014). For a fuller defense of these arguments, see Caruso (2012, 2021a), Caruso and Pereboom (2022), Pereboom (2001, 2014), Pereboom and Caruso (2018).

I call this objection the *Skeptical Argument* and it maintains that we should reject retributivism in light of the philosophical arguments against free will and basic desert moral responsibility. It contends that free will skepticism is the only reasonable position to adopt when it comes to the traditional problem of free will, since the other leading positions fail to preserve the control in action required for basic desert moral responsibility, and that this undermines the retributivist notion that wrongdoers deserve to be punished in the backward-looking, non-instrumental sense required. The justification of retributivism depends on the assumption that criminals are (or at least can be) deserving of blame and punishment in the basic desert sense for their criminal behavior. Yet, free will skepticism maintains that *no one* is ever deserving of blame and punishment in the basic desert sense for any of their actions. Hence, free will skepticism entails that retributive punishment cannot be justified, and thus retributivism should be rejected.[4]

But what if one is not totally convinced by the arguments for free will skepticism? Well, I maintain that *even in the face of uncertainty about the existence of free will*, it remains unclear whether retributive punishment is justified. This is because the burden of proof lies on those who want to inflict harm on others to provide good justification for such harm.[5] This means that retributivists who want to justify legal punishment on the assumption that agents are free and morally responsible (and hence *justly deserve* to suffer for the wrongs they have done) must justify that assumption. And they must justify that assumption in a way that meets a high epistemic standard of proof since the harms caused in the case of legal punishment are often quite severe. It is not enough to simply point to the mere possibility that agents possess libertarian or compatibilist free will. Nor is it enough to say that the skeptical arguments against free will and

[4] Critics of free will skepticism sometimes object that rejecting basic desert moral responsibility would have damaging consequences for morality, the law, society, personal relationships, and our sense of meaning in life (see, e.g., Smilansky 2005, 2011, 2017; Lemos 2016, 2018; Corrado 2018, 2019, 2021; Kennedy 2021, Sifferd 2021, Walen 2021, Zaibert 2021, Donelson 2022, McCormick 2022, Levy 2022). For responses to these and other practical objections, see Waller (2011, 2015), Pereboom (2001, 2014, 2016, 2017, 2021, 2022), Caruso (2017, 2021a, 2021b, 2021c, 2022), Pereboom and Caruso (2018), Caruso and Pereboom (2022), Caruso in Dennett and Caruso (2021).

[5] For how this burden of proof applies to retributivism, see Pereboom (2001), Vilhauer (2009, 2012, 2015), Shaw (2014, 2021), Corrado (2018), Caruso (2020, 2021a), and Jeppsson (2021). For work on legal and non-legal burden of proof, in general, see Hansen et al. (2019), Pigliucci and Boudry (2014), Walton (2014), Rhode (2017).

basic desert moral responsibility fail to be conclusive. Rather, a positive and convincing case must be made that agents are in fact morally responsible in the basic desert sense, since it is the backward-looking desert of agents that retributivists take to justify the harm caused by legal punishment.

This brings me to my second argument against retributivism, the so-called *Epistemic Argument*, which I first developed in Caruso (2020, 2021). Versions of this argument have also been developed and defended by Derk Pereboom (2001), Benjamin Vilhauer (2009, 2012, 2015), Elizabeth Shaw (2014, 2021), Michael Corrado (2018), and Sofia Jeppsson (2021), but my version of the argument can be summarized as follows:

(1) Legal punishment intentionally inflicts harms on individuals and the justification for such harms must meet a high epistemic standard. If it is significantly probable that one's justification for harming another is unsound, then, prima facie, that behavior is seriously wrong.

(2) The retributive justification for legal punishment assumes that agents are morally responsible in the basic desert sense and hence justly deserve to suffer for the wrongs they have done in a backward-looking, non-consequentialist sense (appropriately qualified and under the constraint of proportionality).

(3) If the justification for the assumption that agents are morally responsible in the basic desert sense and hence justly deserve to suffer for the wrongs they have done does not meet the high epistemic standard specified in (1), then retributive legal punishment is prima facie seriously wrong.

(4) The justification for the claim that agents are morally responsible in the basic desert sense provided by both libertarians and compatibilists face powerful and unresolved objections and as a result fall far short of the high epistemic bar needed to justify such harms.

(5) Hence, retributive legal punishment is unjustified and the harms it causes are prima facie seriously wrong.

Note that the Epistemic Argument requires only a weaker notion of skepticism than the one defended in the Skeptical Argument, namely one that holds that the justification for believing that agents are morally responsible in the basic desert sense is too weak to justify the intentional suffering caused by retributive legal punishment. Unlike the Skeptical Argument, which aims to establish that the no

one is ever morally responsible for their actions in the basic desert sense (since who we are and what we do is ultimately the result of factors beyond our control), the Epistemic Argument does *not* require the refutation of libertarian and compatibilist accounts of free will. Instead, it simply needs to raise sufficient doubt that they succeed.

Premise (1) places the burden of proof on those who want to justify legal punishment, since the harms caused in this case are often quite severe – including the loss of liberty, deprivation, and in some cases even death. Given the gravity of these harms, the justification for legal punishment must meet a high epistemic standard. If it is significantly probable that one's justification for harming another is unsound, then, prima facie, that behavior is seriously wrong (see Pereboom, 2001, Vilhauer, 2009).

Support for premise (1) can be found both in the law and everyday practice. As Michael Corrado writes:

> The notion of a burden of proof comes to us from the adversarial courtroom, where it guides the presentation of evidence. In both criminal and civil cases the defendant is presumed not guilty or not liable, and it is up to the accuser to persuade the finder of fact. The only difference between the two cases lies in the measure of the burden that must be carried, which depends upon the seriousness of the outcome. When all that is at issue is the allocation of a loss that can be measured in financial terms, the accuser needs only to prove the defendant's fault by a preponderance of the evidence, but where the defendant's very life or freedom is at stake the burden is considerably higher: the prosecutor must prove beyond a reasonable doubt. (2017: 1)

Our ordinary everyday practices also place the burden of proof on those who knowingly and intentionally cause harm to others. In fact, even in cases where harm is *foreseeable* but not intended, we often demand a high level of justification. Let us say a newspaper receives a tip on a story that will likely cause great harm to a public figure, potentially sinking their career. In such circumstances, good journalistic standards demand that the story be independently verified and properly vetted before it is run. If the newspaper were to run the story without properly vetting it, and later discover that the tip came from an organization who seeks to undermine the public's trust in the media, we would rightly condemn

the newspaper for not applying a higher epistemic standard.[6] Things are even clearer when the harm caused is intentional.

In the case of legal punishment where the severity of harm is beyond question, I maintain that we should place the highest burden possible upon the state. If the state is going to punish someone for first-degree murder, then the epistemic bar that needs to be reached is guilt beyond a reasonable doubt. But does this burden of proof carry over to theoretical debates – for example, the debate over free will and moral responsibility? Here I follow Pigliucci and Maarten (2014) as well as Corrado (2017: 3) in distinguishing between *evidential* burden of proof, which comes into play only when there is no cost associated with a wrong answer, and *prudential* burden of proof, which comes into play precisely when there are significant costs associated with a wrong answer. As Corrado applies the distinction to theoretical matters:

> [I]n a purely philosophical contest where nothing of a practical nature hangs on the outcome it is the evidential burden of proof that is required, and the standard of proof must be "by a preponderance of the evidence": whoever simply has the better evidence must win. On the other hand, if something practical does depend on the outcome of the philosophical debate, then what would matter is the prudential burden. The costs on either side would determine the allocation of the burden and the standard by which satisfaction of the burden is to be measured. (2017: 3)

I contend that given the practical importance of moral responsibility to legal punishment, and given the gravity of harm caused by legal punishment (to the individuals punished as well as those family and friends who depend upon the imprisoned for income, love, support, and/or parenting), the proper epistemic standard to adopt is the prudential burden of proof beyond a reasonable doubt.

[6] The *epistemic encroachment literature* helps support this point since proponents of the view argue that the level of justification needed for a belief changes given the stakes at play in a given context (see Fantl and McGrath 2002, 2007; Ganson 2019; Stanley 2005). According to the thesis of pragmatic encroachment (a term coined by Jonathan Kvanvig), the pragmatic "encroaches" on the epistemic–i.e., practical considerations such as the potential costs of action on *p* if *p* is false can make a genuine epistemic difference. As a result, two subjects in different practical circumstances can differ with respect to whether they are epistemically justified in believing that *p* even though they are the same with respect to all truth-relevant factors, such as the quantity and quality of their evidence for and against *p*, the reliability of the methods they rely on in forming their attitudes toward *p*, etc. Sometimes more evidence is needed to be epistemically justified in believing as the stakes get higher and the odds longer (Ganson 2019; see also Fantl and McGrath 2002, 2007).

When premise (1) is combined with (2), which is simply a statement of the retributivist justification for legal punishment, we get the requirement that retributivists must justify their core assumption – that is, that agents are free and morally responsible in the basic desert sense and hence justly deserve to suffer for the wrongs they have done. While this demand for justification is reasonable given the strength of (1), many retributivists simply deny or ignore it. And those libertarian and compatibilist accounts that do try to justify the assumption of free will, fail to overcome the high epistemic burden of proof needed to justify retributive harm. This is because they tend to be either scientifically implausible (as in the case of agent causation), empirically unwarranted (as in the case of event causal libertarianism), beg the question (as in the case of Strawson and other forms of compatibilism), or end up "changing the subject" (as in the case of Dennett and others).[7]

Agent-causal libertarians, for instance, are willing to embrace mysterious and God-like powers and abilities to preserve free will and basic desert moral responsibility. Roderick Chisholm, for example, famously argued: "If we are responsible, and if what I have been trying to say is true, then we have a prerogative which some would attribute only to God: each of us, when we really act, is a prime mover unmoved" (1964: 32). Naturalistically minded event-causal libertarians have the advantage of avoiding miraculous sui generis kinds of causal powers, but when it comes to providing the epistemic justification needed to ground retributive punishment, they too fall far short. Consider, for instance, the prominent event-causal libertarian accounts of Robert Kane (1996), Mark Balaguer (2009), and Al Mele (2006, 2017). None of these philosophers claim to have provided reason to believe that their accounts are true rather than false or that the necessary empirical requirements posited on their respective accounts actually obtain. Rather, they all settle for the much weaker claim that their theories are *consistent* with our best scientific theories and have *not yet* been

[7] For a fuller defense of these claims, see Caruso (2021a) and Waller (2011, 2015). My goal here is not to address or refute all versions of compatibilism or libertarianism–that would be an altogether different paper. Instead, for the Epistemic Argument to succeed, one only needs to raise sufficient doubt that these defenses of free will and basic desert moral responsibility succeed. Sufficient doubt, however, is fairly easy to raise. In fact, expert disagreement may be enough on its own–after all, the free will debate has been waging for over two thousand years and no clear consensus has been reached. That said, I briefly sketch below some additional concerns. For more on these and other objections to libertarianism and compatibilism, see Pereboom (2001, 2014), Waller (2011, 2015), Levy (2011), Caruso (2012, 2021a).

ruled out. In fact, Mele remains agnostic about his very own event-causal account because, as he puts it, "I do not know of strong evidence that human brains work as they would need to work if a theoretically attractive event-causal libertarian view is true" (Mele, 2017: 205). Fellow event-causal libertarian John Lemos likewise admits: "It is my view that at this point in time we simply don't have sufficient experimental/empirical evidence nor sufficient metaphysical nor logical nor intuitive evidence to establish that libertarian free will exists" (2018: 6). Given this, it would be moral malpractice to continue to retributively punish wrongdoers on the assumption that they are libertarian free agents.

Compatibilists, on the other hand, are in a slightly better positions but the lack of epistemic justification comes at a different point than that of libertarianism. The epistemic challenge facing libertarianism is to justify the claim that we *actually* possess the powers and abilities posited by such accounts. There is no equivalent debate with regard to compatibilism. All parties agree, including skeptics, that we have the abilities discussed by most leading compatibilist accounts – including reasons-responsiveness, voluntariness, the capacity to act in accordance with moral reasons, and the like. The question instead is whether such abilities are *enough* to justify basic desert moral responsibility and, along with it, retributive harm in the case of legal punishment. It is here that insufficient epistemic justification is provided. And that's because extant compatibilist accounts still face powerful and unresolved objections – such as the manipulation argument (Mele, 2006), Pereboom's four-case argument (2001, 2014), van Inwagen's consequence argument (1983), Galen Strawson's basic argument (1986, 1994), Fischer's no-forking-paths argument (1994), Levy's luck pincer (2011), and so on. As a result, they fail to meet the prudential burden of proof. For a retributivist to assume that compatibilism is true and proven beyond a reasonable doubt is to beg the question against incompatibilists and permit unjustified harm.

Given that both libertarian and compatibilist accounts fall far short of the high epistemic bar needed to justify retributive punishment, we should conclude that retributive legal punishment is unjustified and the harms it causes are prima facie seriously wrong. We should therefore refrain from intentionally harming wrongdoers on the philosophically questionable assumption that they deserve it and that such punishment is intrinsically good.

3. The Retributivist Tracking Dilemma

While I take the Skeptical and Epistemic Arguments to be sufficient for the rejection of retributivism, there are also good practical reasons, independent of worries over free will and basic desert, for rejecting retributivism. This section spells out two of these practical concerns, the *Misalignment Argument* and *Poor Epistemic Position Argument*, and argues that taken together they create a powerful new challenge to retributivism called the *Retributivist Tracking Dilemma.* I begin with the Misalignment Argument.

Retributivism demands that we distribute punishment in accordance with desert, but it is not at all clear that the criminal law is capable of properly tracking the desert and blameworthiness of individuals in any reliable way. This is because criminal law is not properly designed to account for all the various factors that affect blameworthiness, and as a result the moral criteria of blameworthiness are often misaligned with the legal criteria of guilt. Erin Kelly (2018) makes a strong case for this claim in her recent book, *The Limits of Blame.* In it she takes issue with a criminal justice system that aligns legal criteria of guilt with moral criteria of blameworthiness – where *blameworthiness* is understood in the backward-looking *basic desert* sense. She argues that many incarcerated people do not meet the criteria of blameworthiness, even when they are guilty of crimes:

People who think carefully about criminal justice must address the problem that the *legal* criteria of guilt do not match familiar *moral* criteria for blame. Conditions that excuse moral failings – such as ignorance, provocation, and mental illness – have limited application in law. This demonstrates a lack of alignment between law and morality. Considerations that mitigate moral blame are often irrelevant to legal findings of criminal guilt. For example, poverty and other unjust deprivations of opportunity have no mitigating relevance in the courtroom; nor do serious mental illnesses such as sociopathy and schizophrenia. Some criminal defendants have diminished moral culpability and others should not be seen as morally blameworthy at all, yet such factors have no bearing on determinations of legal guilt. (2018: 3-4)

Kelly's argument underscores the problem of exaggerating what criminal guilt indicates, particularly when it is tied to the illusion that we know how long and what ways criminals should suffer.

According to Kelly, the notion of individual moral responsibility central to retributive punishment, "masks the systematic nature of social inequality

that is solidified by the criminal justice system, especially that found in the United States" (2018: 11). As she explains:

> A conception of responsibility that connects wrongdoers and moral desert is used to rationalize indefensible criminal-justice practices. We are encouraged to think that criminal conviction metes out verdicts of individual blameworthiness, and this judgment, in turn, functions, by way of its alleged ground in criminal guilt per se, as a basis for thorough-going social typecasting. The very point of criminal justice, so understood, is to assign moral responsibility to individual wrongdoers through findings of criminal guilt and the impositions of a stigmatizing punishment they are thought morally to deserve. (2018: 11)

The problem, however, with this conception of responsibility and desert is that, "it normalizes social injustice, narrows our moral perspective, and precludes a morally sensitive appreciation of the psychological and social adversity confronting many people who commit crimes" (Kelly, 2018: 11). In particular:

> A blaming perspective focused predominately on manifestations of *ill will* too readily overlooks the social and psychological context in which person's beliefs and attitudes are formed, and this focus distorts its moral findings. For example, when poverty and racial injustice are ignored, and the significance of mental illness, immaturity, or mental deficiency is disregarded, conclusions about the blameworthiness of many criminally guilty persons are exaggerated. When the relationships between criminal justice and social justice and between individual responsibility and collective responsibility are not thoughtfully calibrated, they become dangerously unbalanced. As a result, criminal justice institutions are permitted – and exploited – to punish without measure or shame. This is what happens in the United States. (2018: 11)

At the core of Kelly's criticism is the claim that since the criminal law fails to properly track desert and blameworthiness, given that it fails to take seriously the way blameworthiness is affected by poverty and social injustice, mental illness, addiction, and various other mitigating factors, we must reject retributivism and with it any criminal justice system that attempts to align legal criteria of guilt with moral criteria of blameworthiness.

Note, for instance, that courts have thrown out racial discrimination as a basis of challenging criminal conviction and sentencing – for the perverse reason that such discrimination is too common (Kelly, 2018: 8). It is also problematic that many of those actually caught up in our criminal legal system are among the most economically disadvantaged and the least psychologically healthy. The

American Psychological Association notes that, "Over the past four decades, the nation's get-tough-on-crime policies have packed prisons and jails to the bursting point, largely with poor, uneducated people of color, about half of whom suffer from mental health problems" (2014). Yet regarding the economically disadvantaged, "People should not be burdened with serious, harmful consequences for breaking the law when they have been deprived of a reasonable opportunity to lead a satisfactory, law-abiding life" (Kelly, 2018: 15). This is because, "Social injustice undermines legitimate law enforcement and dooms prospects for achieving justice through criminal law" (Kelly, 2018: 15). And regarding psychological health, the criminal law is "indifferent to individual capacities partly because it does not want to invite defendants to argue, as some surely would, that they are morally dense or unmoved by moral reasons, thereby leaving jurors to sort out whether such claims are true, on a case by case basis" (Kelly 2018: 38). Hence, whether it is by design or accident, the criminal law's specification of the conditions under which one is subject to legal punishment departs from morality's specification of when a person can be blamed.

Proponents of retributivism will, no doubt, point out that criminal law allows for various excusing conditions, like the insanity defense, but it's important to recognize that these defenses are limited in scope, seldom successful, and are by no means fine-grained enough to properly track the basic desert and blameworthiness of individuals in any reliable way. Mental illness, for instance, functions as an excuse only when it fits the legal definition of insanity, a highly specialized notion that typically does not include bipolar disorder, autism, Alzheimer's, brain damage due to injury, or many other forms of mental illness or diminished capacity (Kelly, 2018: 8). The problem is that the insanity defense is a legal concept, not a clinical (or medical) one (see Bada Math et al., 2015; Fingarette1966). This means that, first, "just suffering from a mental disorder is not sufficient to prove insanity" (Bada Math et al., 2015: 381). Second, the defendant has the burden of proving the defense of insanity by "clear and convincing evidence" (18 U.S Code §17(B)). And third, "[i]t is hard to determine legal insanity, and even harder to successfully defend it in court" (Bada Math et al. 2015: 381).

Consider, for example, the case of Eddie Ray Routh, who was convicted in Texas for killing two men at a shooting range, one of whom was celebrated sniper Chris Kyle (the former Navy Seal, who has the most recoded kills of any U.S. sniper). A former marine, Routh had been diagnosed with post-traumatic stress disorder (PTSD) and schizophrenia. On the day of the killing, Kyle and

his friend, Chad Littlefield, took Rough with them to the shooting range after Routh's mother asked Kyle for help in dealing with her troubled son. According to defense attorneys, Routh was under extreme mental distress and was convinced the two men (Kyle and Littlefield) would turn on him on the day of the killing. Rough's counsel sought the insanity defense but failed to convince the jury that Rough did not know his actions were wrong (King and May, 2018: 11). The district attorney, Alan Nash, won the jury over, stating, "I am tired of the proposition that if you have a mental illness, you can't be held [legally] responsible for what you do" (Dart, 2015).

This cavalier attitude toward mental illness, combined with the extremely limited scope of the insanity defense, results in a severe and troubling misalignment between the legal criteria of guilt and the moral criteria of blameworthiness. In fact, most moral responsibility theorists acknowledge that different forms of mental illness can, and often do, mitigate or fully exculpate moral blameworthiness and desert. David Shoemaker (2015), for example, has argued that people with high-function autism have "significantly mitigated accountability" in virtue of their empathetic impairments (2015: 173). Matthé Scholten (2016) has argued that it would be morally inappropriate to morally blame people suffering from mental disorders that fall within the schizophrenia spectrum. And Amee Baird, Jeanette Kennett, and Elizabeth Schier (2020) have argued that individuals with dementia are unfit for retributive punishment. If these theorists are correct, or correct about at least some of these conditions, then retributivism faces an alignment problem.

These misalignments are especially troubling when one considers how pervasive mental illness is among inmates in the United States. One recent study found that 64 percent of jail inmates, 54 percent of state prisoners, and 45 percent of federal prisoners report mental health issues (National Research Council 2014). And the mental health crisis is especially pronounced among women prisoners, with one study by the U.S. Bureau of Justice Statistics finding that 75 percent of women incarcerated in jails and prisons having mental illness. Another study found that there are currently three times more seriously mentally ill persons in jails and prisons than hospitals in the United States, with the ratio being nearly ten to one in states like Arizona and Nevada (Torrey et al., 2010). The study further found that 16 percent of the jail and prison population in the U.S. has a "serious mental illness," defined as someone diagnosed with schizophrenia, bipolar disorder, or major depression. This is extremely problematic for retributivists, since many, if not most, of these individuals likely fail to satisfy

the various moral criteria of blameworthiness (or, put more carefully, their conditions likely significantly reduce their blameworthiness), yet when it comes to the criminal legal system, they are found to satisfy the legal criteria of guilt. It is exactly this disconnect that makes retributivism practically problematic, since in the real world it is virtually impossible to accurately distribute punishment in accordance with moral desert.

Misalignments can also occur in cases of psychopathy. Many philosophers and psychologists have argued that psychopaths' impaired capacity for empathy, diminished responses to fear-inducing stimuli, and failure to conform to social norms indicate that they are not fully responsible for their actions, yet the law continues to hold them legally responsible (see, e.g, Levy, 2007, 2014; Morse, 2008; Kennett and Fine, 2004; Kennett, 2010; Glen, Raine, and Laufer, 2011; Shoemaker, 2009, 2011; Nelkin, 2015). To the extent that these philosophers are correct that psychopaths are neither fully morally responsible nor deserving of blame and punishment, yet the law fails to recognize this, we have another case where the moral criteria of blameworthiness and the legal criteria of guilt fail to track each other.

Another example of this disconnect can be found in the rather common practice of prosecuting children as adults. Thirteen states in the U.S. have no minimum age for prosecuting a child as an adult, leaving eight-, nine-, and ten-year-old children vulnerable to extreme punishment, trauma, and abuse within adult jails and prisons. Australian law also currently allows children as young as ten to be charged with a criminal offence, falling below the average minimum age of criminal responsibility worldwide of 12.1 years. Around 600 children under 14 are locked up in Australia prison cells every year. While it may be possible for children to satisfy the legal criteria of guilt, since they tend to be rather permissive, it's not at all clear children satisfy the moral criteria of blameworthiness. Since children under 14 are especially immature and impulsive and have not yet developed mature judgment or the ability to accurately assess risks and consequences, it's unclear they have the executive functioning needed for basic desert moral responsibility (see Hirstein, Sifferd, and Fagan, 2018). If this is correct, then children are not morally responsible in the sense required for retributive punishment, yet in many states and countries they can still satisfy the legal criteria of guilt. This is problematic.

In addition to the above, there are other potential sources of misalignment as well. For instance, a growing number of theorists have argued that deeply disadvantaged backgrounds, poverty, and pervasive and systemic social

injustice can affect blameworthiness and basic desert moral responsibility but are seldom considered mitigating in the criminal law (Bazelon, 1976; Delgado, 1985; Murphy, 1973; Tonry, 2004, 2020; Heffernan and Kleining, 2000). In fact, some legal scholars have even proposed that we adopt a "social adversity" or "rotten background" defense, analogous to the insanity defense, allowing that testimony be permitted in appropriate cases concerning the influence of deep disadvantage and requiring that the jury be directed "that a defendant is not responsible if at the time of his unlawful conduct his mental or emotional processes or behavior controls were impaired to such an extent that he cannot justly be held responsible for his action" (Chief Judge Bazelon, *United States v. Brawner* (1972) 471 F.2d 969 (D.C. Cir.): 1032). Unfortunately, the criminal law does not currently permit social adversity as an affirmative defense, nor does it (generally) view deep disadvantage and adversity as mitigating of desert and punishment.[8]

Of course, philosophers have long been sympathetic to social adversity, though they have tended to pay more attention to the state's moral authority (or standing) to punish than to whether disadvantage sometimes lessens or in extreme cases negates blameworthiness (Tonry, 2020: 17). For instance, philosopher Jeffrie Murphy rejected his own retributive "benefits and burdens" theory of punishment because of the social adversity problem. A large proportion of defendants in criminal courts, he noted, are deeply disadvantaged, and cannot reasonably be said to enjoy the benefits of living in a secure, ordered society. That being so, their retributive punishment cannot be justified until "we have restructured society in such a way that criminals genuinely do correspond to the only model that will render punishment permissible – i.e., make sure that they are autonomous and that they do benefit in the requisite sense" (Murphy, 1973: 243). Retributivist Andreas von Hirsh further observed in *Doing Justice* that, "as long as a substantial segment of the population is denied adequate opportunities for a livelihood, any scheme for punishment must be morally flawed"

[8] There are, however, a few notable exceptions, such as U.S. Supreme Court Justice Elena Kegan's description of the life of Evan Miller in *Miller v. Alabama* (567 U.S. 460, 2012), the decision that declared mandatory life imprisonment without parole for offenders under age 18 unconstitutional: "[I]f ever a pathological background might have contributed to a 14-year-old's commission of a crime, it is here. Miller's stepfather physically abused him; his alcoholic and drug-addicted mother neglected him; he had been in and out of foster care as a result; and he had tried to kill himself four time, the first when he should have been in kindergarten" (567 U.S. 460 [2012]).

(1976: 149). And perhaps the most influential punishment theorist in our time, British philosopher Antony Duff, offered an ideal retributive punishment theory in *Trials and Punishments* but concluded that, "punishment is not justifiable within our present legal system; it will not be justified unless and until we have brought about deep and far-reaching social, political, legal, and moral changes in ourselves and our society" (1986: 294).

Given, then, the various ways moral and legal culpability can be misaligned, retributivism suffers from the fact that a person can easily be found criminally guilty and eligible for retributive punishment without being (fully) morally blameworthy for their criminal wrongdoing. Excuses, whether morally or legally established, indicate diminished responsibility, even for those theorists who believe in basic desert moral responsibility. Yet the legal criteria of guilt are not sufficiently sensitive to all the various excusing conditions that can diminish or remove moral blameworthiness. Yes, the law allows for certain excusing conditions (e.g., insanity, self-defense, automatism, etc.), but these are by no means sensitive enough to track moral blameworthiness in any kind of reliable way. Criminal law often dismisses ignorance, age, systemic injustice, and diminished capacities short of insanity, as legitimate excuses. For this reason, the notions of desert and moral blameworthiness will not function well as the basis of legal guilt. I therefore agree with Kelly when she concludes that, "We should abandon efforts to justify state-imposed punishment as morally deserved. We should reform the criminal justice system without aiming for moral desert and retribution" (2018: 46).

Perhaps, at this point, a retributivist could argue that these problems are more indicative of our broken criminal justice system, not retributivism *per se*, and that a properly reformed criminal justice system could better align the legal criteria of guilt with the moral criteria of basic desert and blameworthiness, thereby avoiding the previous objection. But I think there are in principle reasons to doubt this. This brings me to my second practical objection, which I call the *Poor Epistemic Position Argument* (or PEPA) (see Caruso, 2021a; see also McCay, 2019). PEPA can be summarized as follows: (a) for the state to be able to justly distribute legal punishment in accordance with desert, it needs to be in the proper epistemic position to know what an agent basically deserves, but since (b) the state is (almost) never in the proper epistemic position to know what an agent basically deserves, it follows that (c) the state is not able to justly distribute legal punishment in accordance with desert.

Retributivists, I assume, will try to resist this argument by challenge (b), the key premise. They could do this by arguing either that retributivism is no more epistemically worse off than its alternatives, or that the state *can* properly assess what wrongdoers justly deserve in all, or almost all, relevant cases. But there are serious problems with each of these responses. The problem with the first reply is that it in no way aids retributivism to point to the epistemic limitations of other theories. While it may be true that alternatives to retributive punishment, like, say, consequentialist deterrence theories, may have their own epistemic worries – for instance, knowing, ahead of time, what types of punishments will successfully deter crime – this in no way gets retributivism off the hook. Intentionally inflicting harm on wrongdoers in accordance with desert requires a way of properly assessing desert, and this remains true regardless of whether the alternatives generate their own epistemic worries. This means that it would be morally unacceptable for the retributivist to acknowledge that the state is not in a good epistemic position to know what agents truly deserve – especially given the amount of harm caused by legal punishment. Rather than saving retributivism, the first strategy essentially acknowledges that it would be practically impossible for the state to punish in accordance with desert, *without this leading to injustice,* since, due to the state's epistemic limitations, some people will inevitably end up getting more punishment than they deserve (and some less).

What, then, of the second reply, that the state can, in fact, properly track the desert of offenders? The problem here is that the state (or representatives of the state) is in a severely compromised epistemic position to know what an individual basically deserves. There are several reasons for this. First and foremost, for the state to be able to justly punish in accordance with desert, it would need to be capable of obtaining, processing, and weighing far more information than is typically available in a normal criminal trial. As argued above, the moral criteria of blameworthiness and accountability are more sensitive to the capacities and epistemic states of agents, as well as the way poverty, racism, and prior abuse can affect blameworthiness, than are the criteria of legal guilt. Since the state is epistemically limited in what it can know and what it can consider, it is virtually impossible for it to epistemically determine what an individual basically deserves (see Caruso, 2021a; McCay, 2019).

To further complicate matters, the lives of individuals are complex and the line between victim and criminal is not always clear. Research shows that violent offenders, more often than not, are victims long before they commit their

first crime (Western, 2015; Wachs and Evans, 2010; Tonry,1995, 2014). The complex reality of the lives of these offenders, as both violent perpetrator and violence victim, is one the criminal justice system is ill-prepared to acknowledge or treat. The depressing truth is that most people who go to prison have had to deal with violence all their lives. A recent Boston Re-entry Study found that half of the people interviewed had been beaten by their parents; 40% had witnessed someone being killed; 30% grew up with other family violence; and 16% reported being sexually abused. Nine out of ten of the people interviewed got in fights throughout adolescence. An additional 50% said they were seriously injured in assaults or accidents as children.

While it's easy to portray violence as a characteristic of certain people – thugs who are beyond redemption, people with no conscience – the violent offender of political debates is mostly a fiction. Violence is as much a characteristic of places and circumstances as of people (see Caruso, 2021a). Analysis of the life histories of the men and women who end up in prison indicates that violence typically arises in the context of poverty where conditions are "chaotic" and "lack informal sources of social control" (Western, 2015: 14). Look closely and you will find that there are lifetimes of trauma that fill the prison system. This situational perspective on violence diverges from the criminal justice perspective in which offenders and victims represent distinct classes of people and punishment involves the assessment of individual culpability. This divergence has led some, like Michael Tonry (2014), to argue that deep social disadvantages should be recognized as an excusing or mitigating defense in the criminal law, as well as be recognized as an appropriate basis for mitigating the severity of punishment. Views that deny this, he argues, "fail to acknowledge the existence of social science evidence on human development that makes clear that many offenders offend for reasons for which no plausible case can be made that they are morally responsible" (2014).

If retributivists want to reform the criminal justice system so as to be more sensitive to these mitigating factors, then they need to squarely face the problem of PEPA, since the state is often not in a position to discern the relevant mitigating factors, let alone properly weigh them so as to determine how much they should mitigate moral and legal culpability. And not only is the state epistemically limited in what it can know and properly track, it is also limited in terms of the time, effort, and resources it can dedicate to each individual case. Due to these practical limitations, the state often resorts to plea-bargaining and settling cases prior to a trail, which deprives poor and disadvantaged defendants from a

proper hearing of the facts or a close examination of their moral and legal guilt. In fact, in the U.S. 97 percent of federal cases and 94 percent of state cases end in plea bargains, with defendants pleading guilty in exchange for a lesser sentence. The image of a criminal justice system where defendants get their fair day in court and a fair and just sentence once a trial is concluded, is very different from the real, workaday world inhabited by prosecutors and defense lawyers in the U.S.

In recent decades, American legislators have criminalized so many behaviors that police are arresting millions of people annually – over 10 million in 2019. Taking to trial even a significant proportion of those who are charged would grind proceedings to a halt. Since low-income people are less likely to afford bail, the bulk of America's jailed population is made up of people whose incarceration stems from being poor. And studies show that people in jail are more likely to plead guilty because it's typically the fastest track to getting home (see, Dobbie, Goldin, and Yang, 2018; Gupta, Hansman, Frenchman, 2016). As a result, many poor and disadvantaged people plead guilty to crimes they may not have committed because they fear that without proper representation, they will end up with longer sentences if they risk going to trial. Retributivists need to explain how, in the imperfect world we occupy, the state – with its limited resources and poor epistemic position – is to justly distribute legal punishment in accordance with desert. Retributivists seldom address such concerns, but they are of serious practical importance.

And as if these hurdles were not enough, there's also the fact that humans are prone to a number of cognitive biases that further put representatives of the state in a poor epistemic position to accurately judge blameworthiness and desert. For instance, data clearly shows that we often exaggerate a person's actual or potential control over an event to justify our blame judgments and we will even change the threshold of how much control is required for a blame judgment (Alicke et al., 2008; Alicke, 1994, 2000, 2008; Clark et al., 2014; Clark et al., 2018; Clark, Winegard, and Baumeister, 2019; Everett et al., 2021). Studies by Mark Alicke and his associates indicate that subjects who evaluate the actions of others unfavorably and blame them as a result readily exaggerate the putative wrongdoer's causal control and the evidence that might favor it, while at the same time discounting the counterevidence (Alicke, 2000; Alicke, Rose, and Bloom, 2012). Alicke calls this tendency *blame validation*. Evidence that blaming behavior is widely subject to problems of these kinds is mounting (Nadelhoffer, 2006).

Additional studies by Cory Clark and her colleagues (2014) have shown that a key factor promoting belief in free will is a fundamental desire to blame and hold others morally responsible for their wrongful behaviors. Across five studies they found evidence that greater belief in free will is due to heightened punitive motivations. In one study, an ostensibly real classroom cheating incident led to increased free will beliefs, presumably due to heightened punitive motivations. In a second study, they found that the prevalence of immoral behavior, as measured by crime and homicide rates, predicted free will belief on a country level. Additional studies by Clark, Baumesiter, and Ditto (2017) also demonstrate that free will beliefs are motivated by a desire to punish others and to justify holding them morally responsible – i.e., we attribute *more* free will to those we wish to punish than to those we are indifferent toward. There is good reason to think, then, that our desire to blame and hold others morally responsible comes first and drives our belief in free will, rather than the other way around.

Other biases, such as implicit racial biases, can also affect judgments of desert. There's no denying that race matters in the criminal justice system. Black defendants appear to fare worse than similarly situated white defendants. For instance:

> In a study of bail-setting in Connecticut...Ian Ayres and Joel Waldfogel ([1994)] found that judges set bail at amounts that were twenty-five percent higher for black defendants than for similarly situated white defendants. In an analysis of judicial decision making under the Sentencing Reform Act of 1984, David Mustard found that federal judges imposed sentences on black Americans that were twelve percent longer than those imposed on comparable white defendants [(Mustard 2001)]. Finally, research on capital punishment shows that "killers of White victims are more likely to be sentenced to death than are killers of Black victims" and that "Black defendants are more likely than White defendants" to receive the death penalty [(Banks et al. 2006)]. (Rachlinski et al. 2009: 1196).

Implicit racial biases may account for these racially disparate outcomes in the criminal justice system. A recent study involving a large sample of trial judges drawn from around the U.S. found, for example, that "judges harbor the same kinds of implicit biases as others" and that "these biases can influence their judgments" (Rachlinski et al., 2009: 1195).

It would seem, then, that the state is epistemically and practically compromised in a number of important ways. These include being unable to properly track and weigh the extent to which prior abuse and disadvantage

should mitigate responsibility, being unable to dedicate the time, resources, and effort needed to accurately assess desert and blameworthiness in most cases, cognitive biases that cloud our judgments, and the general difficulty of knowing how addiction, diminished capacities, social injustice, and other relevant factors should affect judgments of desert.

These considerations lead me to conclude that the state is in a severely poor epistemic position to determine what an agent basically deserves. The Poor Epistemic Position Argument (PEPA) therefore provides a powerful reason for rejecting retributivism. Of course, defenders of retributivism could try to avoid PEPA by limiting the scope of considerations to something like the current criteria of legal guilt, arguing that the kinds of mitigating factors discussed above are irrelevant to desert. But this would only bring us back to the previous objection that the legal criteria of guilt are not properly aligned with the moral criteria of blameworthiness. It therefore appears that whichever path the retributivist chooses they face serious difficulties.

I therefore contend that when we combine the Misalignment Argument with the Poor Epistemic Position Argument, we end up with the following dilemma:

> **Retributive Tracking Dilemma:** If retributivists resist broadening the range of considerations by which legal guilt can be mitigated, they end up with an unacceptable misalignment between the legal criteria of guilt and the moral criteria of blameworthiness. On the other hand, if they attempt to properly align the legal criteria of guilt with the moral criteria of blameworthiness so that that former accurately tracks the various factors that mitigate basic desert moral responsibility, they will end up confronting epistemic and practical limitations that make it virtually impossible for the state to properly distribute punishment in accordance with desert. This is because the state is severely compromised and in no epistemic position to properly track and weigh the appropriately expanded set of considerations. Hence, whichever path the retributivist chooses, they end up with an unacceptable outcome *judged on their own terms*.

I contend that this dilemma cannot be avoided by simply tidying up our current legal practices. Instead, it represents an insurmountable practical problem for retributivism. As result, and independent of worries over free will and basic desert moral responsibility, we should reject the retributivist project of trying to distribute punishment in accordance with desert.

4. Conclusion

If what I've argued is correct, we have good theoretical and practical reasons for rejecting retributive legal punishment. The Skeptical Arguments maintains that we should reject retributivism in light of the philosophical arguments against free will and basic desert moral responsibility. It contends that free will skepticism is the only reasonable position to adopt when it comes to the traditional problem of free will, since the other leading positions fail to preserve the control in action required for basic desert moral responsibility, and that this undermines the retributivist notion that wrongdoers deserve to be punished in the backward-looking, non-instrumental sense required. Understanding, though, that not everyone will be convinced by the arguments against free will and basic desert, despite their strength, I then offered a second objection to the retributivism – the Epistemic Argument – that required only a weaker notion of skepticism. It argued that the justification for believing that agents are morally responsible in the basic desert sense, and hence justly deserve to suffer for the wrongs they have done, is too weak to justify the intentional suffering caused by retributive legal punishment. As a result, I concluded that retributive legal punishment is unjustified and the harms it causes are prima facie seriously wrong.

Moving on to a pair of practical objections, I then argued that independent of concerns over free will, it is philosophically problematic to impart to the state the function of intentionally harming wrongdoers in accordance with desert since it's not at all clear that the state is capable of properly tracking the desert and blameworthiness of individuals in any reliable way. I first objected, as part of the Misalignment Argument, that because criminal law is not properly designed to account for all the various factors that affect blameworthiness, the *moral criteria of blameworthiness* are often misaligned with the *legal criteria of guilt*. I then presented a closely related argument, the *Poor Epistemic Position Argument* (PEPA), which argued that since the state is seriously epistemically compromised and in a poor epistemic position to know what an agent basically deserves, it follows that the state is not able to justly distribute legal punishment in accordance with desert. I concluded that, taken together, these two practical objections create a powerful new challenge to retributivism called the *Retributivist Tracking Dilemma.*

The core problem is that criminal justice systems, including those operating in the US, ignore factors that, according to retributivists themselves, determine the degree of one's blameworthiness and desert (e.g., facts about one's

upbringing, knowledge of wrongdoing, age, diminished capacities, mental illness short of insanity, structural features of society, and so on). If these criminal justice systems continue to operate in this way, they are unjust since they are in violation of the retributivists own proportionality principle – which prohibits punishing individuals who do not deserve it, or punishing them more than they deserve. Yet, if retributivists try to modify these systems to take account of those factors, they will end up confronting the epistemic and practical limitations outline above, which will ultimately prevent them from ever fully and successfully aligning the moral criteria of blameworthiness with the legal criteria of guilt. As a result, retributivism, when practically implemented in our criminal justice systems, is bound to cause injustice (judged on its own terms).

Given these theoretical and practical difficulties, I conclude that we should reject retributive legal punishment and abandon efforts to justify state-imposed punishment as morally deserved.

REFERENCES

Alexander, L., K. Kessler Ferzan, S. Morse. (2009). *Crime and Culpability: A Theory of Criminal Law.* New York: Cambridge University Press.

Alicke, M.D. (1994). Evidential and extra-evidential evaluations of social conduct. *Journal of Social Behavior and Personality* 9: 591–615

Alicke, M.D. (2008). Blaming badly. *Journal of Cognition and Culture* 8: 179–186.

Alicke, M.D. (2000). Culpable control and the psychology of blame. *Psychology Bulletin* 126: 556-74.

Alicke, M.D., D. Rose, D. Bloom. (2012). Causation, norm violation and culpable control. *Journal of Philosophy* 106: 587-612.

Alicke, M.D., J. Buckingham, E. Zell, T. Davis. (2008). Culpable control and counterfactual reasoning in the psychology of blame. *Personality and Social Psychology Bulletin* 34: 1371–1381.

Bada Math, S., C. N. Kumar, and S. Moirangthem. (2015). Insanity defense: Past, present, and future. *Indian Journal of Psychological Medicine* 37(4): 381–387.

Baird, A., J. Kennett, E. Schier. (2020). Homicide and dementia: An investigation of legal, ethical, and clinical factors of Australian legal cases. *International Journal of Law and Psychiatry* 71: 1–13.

Balaguer, M. (2010). *Free Will as an Open Scientific Problem.* Cambridge: MIT Press.

Bazelon, D. L. (1976). The morality of the criminal law. *Southern California Law Review* 49: 385–405.

Berman, M. (2008). Punishment and justification. *Ethics* 18: 258–290.

Berman, M. (2011). Two kinds of retributivism. In *Philosophical Foundations of Criminal Law,* eds. R.A. Duff and S. Green, pp. 433–457. New York: Oxford University Press.

Berman, M. (2013). Rehabilitating retributivism. *Law and Philosophy* 32: 83–108.

Berman, M. (2016). Modest retributivism. In *Legal, Moral, and Metaphysical Truths: The Philosophy of Michael S. Moore*, eds. K. Kessler Ferzan and S. J. Morse, pp. 35–48. New York: Oxford University Press.

Boonin, David. (2008). *The Problem of Punishment.* Cambridge: Cambridge University Press.

Caruso, G.D. (2012). *Free Will and Consciousness: A Determinist Account of the Illusion of Free Will.* Lanham: Lexington Books.

Caruso, G.D. (2017). Free will skepticism and the question of creativity: Creativity, desert, and self-creation. *Ergo* 3(23): 591-607.

Caruso, G.D. (2020). Justice without retribution: An epistemic argument against retributive criminal punishment. *Neuroethics* 13(1): 13-28.

Caruso, G.D. (2021a). *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice.* Cambridge: Cambridge University Press.

Caruso, G.D. (2021b). Retributivism, free will skepticism, and the public health-quarantine model: Replies to Kennedy, Walen, Corrado, Sifferd, Pereboom, and Shaw. *Journal of Legal Philosophy* 46(2): 161-216.

Caruso, G.D. (2021c). Rejecting retributivism: Reply to Leo Zaibert. *The Philosopher* 109(4): 118-26.

Caruso, G.D. (2022). Preemptive incapacitation, victim's rights, desert, and respect for persons: Replies to McCormick and Donelson. *Journal of Practical Ethics.* https://journals.publishing.umich.edu/jpe/news/57/

Caruso, G.D. and S.G. Morris. (2017). Compatibilism and retributivist desert moral responsibility: On what is of central philosophical and practical importance. *Erkenntnis* 82(4): 837-55.

Caruso, G.D., and D. Pereboom (2022). *Moral Responsibility Skepticism.* Cambridge: Cambridge University Press.

Chisholm, R. (1964). Human freedom and the self. *The Lindley Lecture,* Department of Philosophy, University of Kansas; reprinted in *Free Will,* G. Watson, ed., Oxford: Oxford University Press, 1982, pp. 24-35.

Clark, C.J., J.B. Luguri, P.H. Ditto, J. Knobe, A.F. Shariff, and R.F. Baumeister. (2014). Free to punish: A motivated account of free will belief. *Journal of Personality and Social Psychology* 106(4): 501-13.

Clark, C.J., A. Shniderman, J.B. Luguri, R.F. Baumeister, P.H. Ditto. (2018). Are morally good actions ever free? *Consciousness and Cognition* 63: 161-82.

Clark, C.J., B.M. Winegard, R.F. Baumeister. (2019). Forget the folk: Moral responsibility preservation motives and other conditions for compatibilism. *Frontiers in Psychology,* February 7: https://doi.org/10.3389/fpsyg.2019.00215.

Corrado, M.L. (2018). Criminal quarantine and the burden of proof. *Philosophia.* Available at: https://doi-org.libproxy.lib.unc.edu/10.1007/s11406-018-0026-2.

Corrado, M.L. (2019). Fichte and the psychopath: Criminal justice turned upside down. In *Free Will Skepticism in Law and Society: Challenging Retributive Justice,* eds. E. Shaw, D. Pereboom, and G.D. Caruso, pp. 161-76, New York: Cambridge University Press.

Corrado, Michael L. (2021). The limits of the state's power to control crime. *Journal of Legal Philosophy* 46(2): 126-31.

Dart, T. (2015). American sniper' jurors: Defendant was using PTSD diagnosis as an excuse. *The Guardian.* Available at: www.theguardian.com/us-news/2015/feb/25/american-sniper-trial- jury-eddie-ray-routh-ptsd-excuse.

Delgado, R. (1985). Rotten social background: Should the criminal law recognize a defense of severe environmental deprivation? *Law and Inequality* 3: 9–91.

Dennett, D.C., G.D. Caruso. (2020). *Just Deserts.* Cambridge: Polity Press.

Dobbie, W., J. Goldin, C.S. Yang. (2018). The effect of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review* 108(2): 201–240.

Donelson, R. (2022). Comments on Gregg Caruso's *Rejecting Retributivism. Journal of Practical Ethics.* https://journals.publishing.umich.edu/jpe/news/56/

Duff, R.A. (1986). *Trials and Punishment.* New York: Cambridge University Press.

Everett, J.A.C., C.J. Clark, P. Meindl, J.B. Luguri, B.D. Earp, J. Graham, P.H. Ditto, A.F. Shariff. (2021). Political differences in free will belief are associated with

differences in moralization. *Journal of Personality and Social Psychology* 120(2): 461-83.

Fantl, J., M. McGrath. (2002). Evidence, pragmatics, and justification. *Philosophical Review* 111: 67-94.

Fantl, J., M. McGrath. (2007). On pragmatic encroachment in epistemology. *Philosophy and Phenomenological Research* 75(3): 558-589.

Fischer, J.M. (1994). *The Metaphysics of Free Will.* Oxford: Blackwell Publishers.

Ganson, D. (2019). Pragmatic Encroachment. In *The Routledge Encyclopedia of Philosophy. Taylor and Francis.* Retrieved 15 Oct. 2022, from https://www.rep.routledge.com/articles/thematic/pragmatic-encroachment/v-1. doi:10.4324/9780415249126-P079-1

Glenn, A.L., A. Raine, and W.S. Laufer. (2011). Is it wrong to criminalize and punish psychopaths? *Emotion Review* 3: 302–304.

Gupta, A., C. Hansman, and E. Frenchman. (2016). The heavy costs of high bail: Evidence from judge randomization. *Journal of Legal Studies* 45(2): 471–505.

Hansen, Hans V., Fred J. Kauffeld, James B. Freeman, Lilian Bermejo-Luque (eds.) (2019). *Presumptions and Burdens of Proof: An Anthology of Argumentation and the Law.* Tuscaloosa: University of Alabama Press.

Heffernan, W.C., J. Kleinig. (eds.) (2000). *From Social Justice to Criminal Justice: Poverty and the Administration of Criminal Law.* New York: Oxford University Press.

Hirstein, W., K. Sifferd, and T. Fagan. (2018). *Responsible Brain: Neuroscience, Law, and Human Culpability.* Cambridge, MA: MIT Press.

Husak, D. (2000). Holistic retributivism. *California Law Review* 88: 991–1000.

Jeppsson, S. (2021). Retributivism, justification and credence: The epistemic argument revisited. *Neuroethics* 14(2): 177-190.

Kane, R. (1996). *The Significance of Free Will,* New York: Oxford University Press.

Kant, I. (1797/2017). *The Metaphysics of Morals,* Mary Gregor, tr. Cambridge University Press.

Kelly, E. (2018). *The Limits of Blame: Rethinking Punishment and Responsibility.* Cambridge, MA: Harvard University Press.

Kennedy, C. (2021). Taking responsibility for criminal responsibility: Comments on *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice. Journal of Legal Philosophy* 46(2): 132-37.

Kennett, J. (2010). Reasons, emotion, and moral judgment in the psychopath. In *Responsibility and Psychopathy: Interfacing Law, Psychiatry and Philosophy,* eds. L. Malatesti and J. McMillan, pp. 243–260. New York: Oxford University Press.

Kennett, J., C. Fine. (2004). Mental impairment, moral understanding and criminal responsibility: Psychopathy and the purpose of punishment. *International Journal of Law and* Psychiatry 27(5): 425–443.

Kershnar, S. (2000). A defense of retributivism. *International Journal of Applied Philosophy* 14 (1): 97–117.

Kershnar, S. (2001). *Desert, Retribution, and Torture.* Lanham, MD: University Press of America.

Lemos, J. (2016). Moral concerns about responsibility denial and the quarantine of violent criminals. *Law and Philosophy* 35: 461-83.

Lemos, J. (2018). *A Pragmatic Approach to Libertarian Free Will.* New York: Routledge.

Levy, K. (2022). Let's not do responsibility skepticism. *Journal of Applied Philosophy.* Doi: 10.1111/japp.12623.

Levy, N.l. (2011). *Hard Luck: How Luck Undermines Free Will and Moral Responsibility.* Oxford: Oxford University Press.

Levy, N. (2007). The responsibility of the psychopath revisited. *Philosophy, Psychiatry, and Psychology* 14: 129–138.

Levy, N. (2014). Psychopaths and blame: The argument from content. *Philosophical Psychology* 27(3): 351–367.

Mele, A. (2006). *Free Will and Luck.* New York: Oxford University Press.

Mele, A. (2017). *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will.* New York: Oxford University Press.

McCay, Allan. (2019). Mitigation is difficult: A moral evaluation of a mitigation practice at sentencing. In *Free Will and the Law: Key Perspectives,* ed. Allan McCay and Michael Sevel, pp.188-209. New York: Routledge.

McCormick, K. (2022). Comments on Gregg Caruso's *Rejecting Retributivism: Free Will, Punishment, and Criminal Justice. Journal of Practical Ethics* https://journals.publishing.umich.edu/jpe/news/55/

Moore, M.S. (1997). *Placing Blame.* New York: Oxford University Press.

Morse, S.J. (2008). Psychopathy and criminal responsibility. *Neuroethics* 1: 205–212.

Murphy, J. (1973). Marxism and retribution. *Philosophy and Public Affairs* 2: 217–243.

Nadelhoffer, T. (2006). Bad acts, blameworthy agents, and intentional action: Some problems for jury impartiality. *Philosophical Explorations* 9: 203-20.

National Research Council (2014). The Growth of Incarceration in the United States: Exploring Causes and Consequences. Available at: www.nap.edu/read/18613/chapter/1

Nelkin, D.K. (2015). Psychopaths, incorrigible racists, and the face of responsibility. *Ethics* 125(2): 357–390.

Pereboom, D. (2001). *Living without Free Will.* Cambridge: Cambridge University Press.

Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life.* Oxford: Oxford University Press.

Pereboom, D. (2016). A defense of free will skepticism: Replies to commentaries by Victor Tadros, Saul Smilansky, Michael McKenna, and Alfred R. Mele. *Criminal Law and Philosophy* 11(3): 1-20.

Pereboom, D. (2017). Response to Daniel Dennett on free will skepticism. *Rivista Internazionale di Filosofia e Psicologia* 8: 259-65.

Pereboom, D. (2021). *Wrongdoing and the Moral Emotions.* Oxford: Oxford University Press.

Pereboom, D. (2022). *Free Will.* Cambridge: Cambridge University Press.

Pereboom, D., G.D. Caruso. (2018). Hard-incompatibilism existentialism: Neuroscience, punishment, and meaning in life. In *Neuroexistentialism: Meaning, Morals, and Purpose in the Age of Neuroscience*, eds. G.D. Caruso and O. Flanagan, pp.193-222. New York: Oxford University Press.

Pigliucci, M., M. Boudry. (2014). Prove it! The burden of proof game in science vs. pseudoscience disputes. *Philosophia* 42: 487–502.

Powers, M., R. Faden. (2006). *Social Justice: The Moral Foundations of Public Health and Health Policy.* New York: Oxford University Press.

Rachlinski, J., S. Johnson, A.J. Wistrich, C. Guthrie. (2009). Does unconscious racial bias affect trial judges? *Notre Dame Law Review* 84(3): 1195–1246.

Rhode, Conny. (2017). The burden of proof in philosophical persuasion dialogue. *Argumentation* 31(3): 535-554.

Robinson, P.H. M.T. Cahill. (2006). *Law Without Justice: Why Criminal Law Doesn't Give People What They Deserve.* New York: Cambridge University Press.

Scholten, M. (2016). Schizophrenia and moral responsibility: A Kantian essay. *Philosophia* 44 (1): 205–225.

Shaw, E. (2014). *Free Will, Punishment, and Criminal Responsibility*. Dissertation Thesis. Edinburgh University. Available at: https://era.ed.ac.uk/bitstream/handle/1842/9590/ Shaw2014.pdf?sequence=2&isAllowed=y.

Shoemaker, D. (2009). Responsibility and disability. *Metaphilosophy* 40(3-4): 438–461.

Shoemaker, D. (2011). Attributability, answerability, and accountability: Toward a wider theory of moral responsibility. *Ethics* 121: 602-32.

Shoemaker, D. (2015). *Responsibility from the Margins*. Oxford: Oxford University Press.

Sifferd, K. (2021). Why not 'weak' retributivism? *Journal of Legal Philosophy* 46(2): 138-43.

Smilansky, S. (2000). *Free Will and Illusion*. New York: Oxford University Press.

Smilansky, S. (2005). Free will and respect for persons. *Midwest Studies in Philosophy* 29: 248-61.

Smilansky, S. (2011). Hard determinism and punishment: A practical reductio. *Law and Philosophy* 30: 353-67.

Smilansky, S. (2017). Pereboom on punishment: Funishment, innocence, motivation, and other difficulties. *Criminal Law and Philosophy* 11: 591-603.

Stanley, Jason. (2005). *Knowledge and Practical Interests*. New York: Oxford University Press.

Strawson, G. (1986). *Freedom and Belief*. Oxford: Oxford University Press.

Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies* 75: 5-24.

Tonry, M. (1995). *Malign Neglect: Race, Crime, and Punishment in America*. New York: Oxford University Press.

Tonry, M. (2004). *Thinking About Crime: Sense and Sensibility in American Penal Culture*. New York: Oxford University Press.

Tonry, M. (2014). Can deserts be just in an unjust world? In *Liberal Criminal Theory: Essays for Andreas von Hirsch*, eds. A.P. Simester, U. Neumann, and A. Du. Bois-Pedain. Oxford: Hart.

Tonry, M. (2020). *Doing Justice, Preventing Crime.* New York: Oxford University Press. van Inwagen, P (1983). *An Essay on Free Will.* Oxford: Oxford University Press.

Vilhauer, B. (2015). Free will and the asymmetrical justification of holding morally responsible. *The Philosophical Quarterly* 65(261): 772–789.

Vilhauer, B. (2009). Free will and reasonable doubt. *American Philosophical Quarterly* 46: 131-40.

Vilhauer, B. (2012). Taking free will skepticism seriously. *Philosophical Quarterly* 62: 833-52.

Von Hirsch, A. (1976). *Doing Justice: The Choice of Punishment.* Boston: Northeastern University Press.

Von Hirsch, A. (1981). Desert and previous convictions in sentencing. *Minnesota La Review* 56 (841): 591–634.

Von Hirsch, A. (2007). The "desert" model for sentencing: Its influence, prospects, and alternatives. *Social Research* 74(2): 413–434.

Von Hirsch, A. (2017). *Deserved Criminal Sentences: An Overview.* Portland, Oregon: Hart Publishing.

Wachs, T.D., G. W. Evans. 2010. Chaos in context. In *Chaos and Its Influence on Children's Development: An Ecological Perspective*, eds. G. W. Evans and T. D. Wachs, pp.3–13. Washington, DC: American Psychological Association.

Walen, A. (2021). Determinism, compatibilism, and basic desert: A reply to Gregg Caruso. *Journal of Legal Philosophy* 46(2): 144-48.

Waller, Bruce. (2011). *Against Moral Responsibility.* Cambridge: MIT Press.

Waller, Bruce. (2015). *The Stubborn System of Moral Responsibility.* Cambridge: MIT Press.

Walton, Douglas. (2014). *Burden of Proof, Presumption, and Argumentation.* New York: Cambridge University Press.

Western, B. (2015). Lifetimes of violence in a sample of released prisoners. *The Russell Sage Foundation Journal of the Social Sciences* 1(2): 14–30.

Zaibert, L. (2018). *Rethinking Punishment.* Cambridge: Cambridge University Press.

Zaibert, L. (2021). Embracing retributivism. *The Philosopher* 109(4): 105-17.

Zimmerman, M. (2011). *The Immorality of Punishment.* New York: Broadview Press.