2018

Punishment and Moral Risk

Adam J. Kolber
Brooklyn Law School, adam.kolber@brooklaw.edu

PUNISHMENT AND MORAL RISK

*Adam J. Kolber**

For every interesting moral question, we should have at least some doubt that we know the right answer. Legal theorists ignore this moral uncertainty at their peril. To take one important example, for retributivists to inflict punishment, they must believe not only that a defendant is guilty but that all other prerequisites for deserved punishment are satisfied as well. They must believe offenders have free will, even though philosophers have debated the topic for centuries. They must believe offenders can be punished proportionally, even though no one has convincingly determined how to assess proportionality. And they must believe it appropriate to make offenders suffer in response to the suffering they caused, even though some find this view barbaric.

These retributivist commitments, along with several others, are clearly controversial. One would be hard-pressed to believe a single one—let alone the conjunction—with the 95% or 99% confidence frequently attributed to the beyond-a-reasonable-doubt standard used to assess guilt. Reasonable retributivists, I argue, face too much uncertainty to justify punishment under the standard of proof they would likely set for themselves. Consequentialists, by contrast, are less vulnerable to this challenge. They can accept greater risk when punishing because they face countervailing risk by failing to adequately punish.

More generally, I argue that we hold not just beliefs but "portfolios of beliefs" that can exacerbate or hedge moral risks. These portfolios sometimes do a better job of explaining our moral and legal views than existing theories, and I show how "epistemic hybrid" theories that combine retributivism and consequentialism can avoid the inconsistencies facing current hybrid theories. We are not necessarily retributivists or consequentialists but, say, 60% retributivists or 90% consequentialists. Portfolios of beliefs can also help us understand other areas of law and morality, such as the perplexities of threshold deontology and the puzzles of tort law. Indeed, portfolios of beliefs may not only explain our beliefs but also offer normatively appealing alternatives to our existing theories that fail to take moral risk into account.

---

* Professor of Law, Brooklyn Law School. For helpful comments, I thank Peter Alces, Emad Atiq, Paul Davies, Adam Elga, David Enoch, Chad Flanders, Richard Frase, Stephen Galoob, Alex Guerrero, Nathan Hanna, Adil Haque, Doug Husak, Leora Katz, Ittay Nissan, Alice Ristroph, Gideon Rosen, Chelsea Rosenthal, Larry Solum, Patrick Tomlin, Ben Vilhauer, Alec Walen, and Leo Zaibert as well as participants in workshops and conferences at Georgetown Law School, NYU School of Law, and William and Mary Law School.

TABLE OF CONTENTS

I. INTRODUCTION

To reduce error in criminal cases, the United States and many other countries require proof at trial that a defendant committed every element of an offense beyond a reasonable doubt.[1] We set a high bar, perhaps requiring levels of confidence of 90%, 95%, or 99% because we would rather fail to punish the guilty than punish the innocent.[2] Hence, the presumption of innocence and the beyond-a-reasonable-doubt standard ("BARD standard") seem to reflect values consistent with the famous Blackstone ratio: better ten guilty people go free than one innocent person be punished.[3]

To believe that punishing a particular person is *morally* justified, however, we must believe much more than that he violated the elements of a statute. We must be sufficiently confident that all the requirements of just punishment are satisfied. Theorists have said a lot about the standard for finding an offender guilty beyond a reasonable doubt but have said little about what I call the *justificatory standard of proof*—namely, the level of confidence we must have to punish despite uncertainty as to whether the punishment is morally justified.

Moral uncertainty creates issues for all punishment theorists but especially for retributivists. Retributivists believe that those who commit a serious moral wrong deserve proportional suffering or punishment from the state.[4] The kind of retributivism I examine has several embedded propositions, including the following: (1) people ordinarily have free will and satisfy all other requirements for moral responsibility, (2) suffering (or punishment) is an appropriate response to wrongdoing, (3) we can adequately analyze a defendant's background history

1. *See In re* Winship, 397 U.S. 358, 361 (1970).

2. *Cf.*, LARRY LAUDAN, TRUTH, ERROR, AND CRIMINAL LAW: AN ESSAY IN LEGAL EPISTEMOLOGY 44 (2006) (stating that 90% or 95% are "commonly cited unofficial estimates" of the standard of proof for criminal convictions); Margaret Raymond, *The Problem with Innocence*, 49 CLEV. ST. L. REV. 449, 451 (2001) (exploring a hypothetical 99% beyond a reasonable doubt standard). As I later explain, however, my argument does not require precise quantification of the standard.

3. 4 WILLIAM BLACKSTONE, COMMENTARIES *358. Importantly, however, a probabilistic standard of proof does not automatically translate into a Blackstone-style ratio. *See* Michael L. DeKay, *The Difference Between Blackstone-Like Error Ratios and Probabilistic Standards of Proof*, 21 LAW & SOC. INQUIRY 95, 99–100 (1996).

4. *See, e.g.*, Douglas N. Husak, *Retribution in Criminal Theory*, 37 SAN DIEGO L. REV. 959, 972 (2000) ("[R]etributive beliefs only require that culpable wrongdoers be given their just deserts by being made to suffer (or to receive a hardship or deprivation)."); John Rawls, *Two Concepts of Rules*, 64 PHIL. REV. 3, 4–5 (1955) ("It is morally fitting that a person who does wrong should suffer in proportion to his wrongdoing. That a criminal should be punished follows from his guilt, and the severity of the appropriate punishment depends on the depravity of his act.").

in order to assess what he deserves, and (4) punishment can justifiably be imposed by the state.

Each of these beliefs is quite controversial (and they reflect only some of the many propositions retributivists believe). It's not obvious why retributivists would permit lower levels of confidence in these beliefs than they require for belief in guilt, yet it arguably strains reason to believe any one of them with 95% or 99% confidence. Moreover, retributivists must believe the significantly less likely proposition that all of them are true. In comparison, decisions about whether an offender violated all the elements of a statute look comparatively simple. At least we often agree about *how* to make such determinations. But the mystery of free will, to pick one important example, has persisted for centuries.[5]

So far, however, we have only addressed the tip of the uncertainty iceberg. To punish any particular offender, retributivists must not only believe these general background propositions, they must believe several claims about a particular offender who stands accused. Of course, they must conquer a familiar source of uncertainty to believe that (5) this offender engaged in the conduct alleged. They must also believe several additional propositions related to case-specific moral uncertainty that apply to common versions of retributivism, including: (6) the defendant's alleged conduct was actually wrongful; (7) police conduct and judicial proceedings were not so unlawful as to preclude just punishment; (8) the costs of giving this offender what he deserves do not grossly exceed the benefits; and (9) punishment should be proportional, and this offender's punishment is proportional (or at least not disproportional). These case-specific propositions may strain our confidence more than the general background propositions. To take a prominent example, no one has convincingly explained how to determine what punishments are proportional to what crimes nor, in my estimation, established that proportionality sets a coherent, desirable goal.[6] Yet any particular instance of punishment will require retributivists to be quite confident that punishment is not disproportional.

In Part II, I explain, more carefully and in more detail, why retributivists are subject to what I call the "epistemic challenge." Even if they have as much as 95% confidence in each of the foregoing nine propositions, their total confidence that the punishment of some particular offender is justified—if we assume that each proposition is independent of the others—will only be 63%.[7] A justificatory standard of proof requiring a mere 63% level of confidence is unlikely to be consistent with the values embedded in the BARD standard (which deem

5. One could likely say "millennia" rather than just "centuries," though there is some dispute about precisely who first recognized the problem of free will and when. *See, e.g.*, Susanne Bobzien, *Did Epicurus Discover the Free Will Problem?*, 19 OXFORD STUD. ANCIENT PHIL. 287, 289 (2000) ("[N]either Aristotle nor the early Stoics nor any contemporaries of Epicurus were concerned with [a concept of free decision or choice].").

6. *See* Adam J. Kolber, *Against Proportional Punishment*, 66 VAND. L. REV. 1141 (2013) [hereinafter Kolber, *Against Proportional Punishment*].

7. To calculate the probability of several *independent* events occurring, we can simply multiply the probability of each event occurring separately. In our case, to calculate the probability of nine independent propositions being true where each has a 95% probability of being the case, we raise 95% to the ninth power and get 63%. *See* R.B. Campbell, *Conditional Probability and the Product Rule*, U.N. IOWA COMPUTER SCI., http://www.cs.uni.edu/~campbell/stat/prob4.html (last visited Jan. 28, 2018).

it far better to fail to punish the guilty than to punish the innocent). Reasonable retributivists are likely to have too much uncertainty to justify punishment.

One possibility is that punishment simply cannot be justified with sufficient confidence and ought to be abolished. But consequentialists present another option. Consequentialists justify punishment not based on desert but on the good consequences that follow from punishment, like crime deterrence, the incapacitation of dangerous people, and the rehabilitation of offenders.[8] While consequentialists are also subject to epistemic challenges, they are less susceptible to the challenge I pose here. As I describe in Part III, consequentialists typically give more weight than retributivists to the risk of harming by *failing* to punish. Uncertainty that makes consequentialists hesitate to punish will often be counterbalanced by uncertainty that makes consequentialists hesitate *not* to punish. To the extent these forces tend to be more equally weighted than they are for retributivists, consequentialists will be less susceptible to the epistemic challenge.

Finally, in Part IV, I discuss hybrid theories that combine elements of retributivism and consequentialism. I argue that traditional hybrid theories tend to fail because they combine two theories that are fundamentally incompatible. There are, however, more plausible "epistemic hybrid theories" that take moral uncertainty into account and thereby reflect a mix of retributivist and consequentialist elements. Rather than thinking of ourselves as retributivists or consequentialists, perhaps we should describe ourselves more precisely as 60% retributivists or 90% consequentialists.

More generally, we hold not just beliefs but what I call "portfolios of beliefs." We believe different propositions with varying degrees of confidence. Just as one might purchase shares of several different companies and hold them in various amounts in an investment portfolio, one can hold different beliefs to varying degrees in a portfolio of beliefs. And just as stocks can interact in ways that increase or decrease risk, so too can the constituents of a portfolio of beliefs. These portfolios, I suggest (with additional examples outside of punishment in an appendix), will sometimes better explain our laws and practices than existing theories which fail to take moral uncertainty into account. Indeed, they may even offer normatively appealing alternatives.

II. THE EPISTEMIC CHALLENGE TO RETRIBUTIVISM

A. Typical Retributivism

When we punish, we act in ways ordinarily prohibited. Fines take people's money, prisons take their freedom, and executions take their lives. Without the state's imprimatur, such punishments would look like theft, wrongful imprisonment, and murder. Punishment theorists seek to explain why punishment is morally permissible (and perhaps even obligatory) despite the harms it undoubtedly causes.

8. C.L. TEN, CRIME, GUILT AND PUNISHMENT: A PHILOSOPHICAL INTRODUCTION 7–8 (1987).

Retributivists believe that punishment is morally justified when a person deserves it.[9] The details of retributivist views vary, so I will speak of a fairly common version which holds that *those who commit a serious moral wrong deserve proportional punishment from the state.*[10] I will focus on a pure deontological form of retributivism that takes desert to ordinarily provide a sufficient reason to punish without reliance on other possible goals of punishment like deterrence, incapacitation, and rehabilitation. While I surely do not address every version of retributivism, the thrust of my argument, with modest adjustments, applies to a broad range of retributivist views[11] (so long as they share the Blackstonian value that it is substantially worse to punish the undeserving than to fail to punish the deserving).

To say that the punishment of some actual person is morally justified, one must believe more than just that the defendant violated the elements of a criminal statute. After all, part of what we want to know is what sorts of behavior can justly be criminalized. If violation of a criminal statute were sufficient to morally justify punishment, our inquiry would end quickly. But surely that is not enough. Pre-Civil War laws that punished slaves for escape were unjust, even if they were legally recognized. Relying on compliance with a constitution provides no help either because constitutions are just special kinds of legislation. For retributivists to claim that punishment is just, they must believe that it is *morally* just and not just legally permitted.

Thus, I assume, the typical retributivist who purports to justify punishment in any particular case must believe in at least the nine claims in the introduction (and others could certainly be added) with sufficient certainty to proceed with punishment. Notice that our confidence in the overall proposition that some defendant deserves a particular punishment can be no stronger than our confidence in our least-confident component proposition. Indeed, each proposition creates *additional* uncertainty (barring the unrealistic case of 100% certainty). While I do not offer a precise value for the justificatory standard of proof, I make some arguments about which ranges are plausible. If, for example, retributivists have less than 50% confidence that some imposition of punishment is justified, they would be more confident that punishing is not justified than that it is.

---

9. Alec Walen, *Proof Beyond a Reasonable Doubt: A Balanced Retributive Account*, 76 LA. L. REV. 355, 361 (2015).

10. Mitchell N. Berman, *Punishment and Justification*, 118 ETHICS 258, 269 (2008) ("[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.").

11. For example, Alec Walen has argued that retributivists ought not consider consequentialist considerations until after desert has been affixed:

> A retributivist need not suppose that desert is the only reason to punish; the good that punishment can accomplish through deterrence and the incapacitation of potential criminals is, on any reasonable view, a reason to punish. The retributivist point is that these instrumental reasons to punish are relevant only once it has been determined that the person to be punished deserves punishment. Desert functions like a gate in the normative equivalent of a transistor; until it is switched on, the potentially greater normative force of the instrumental reasons is switched off.

Walen, *supra* note 9, at 426–27 (footnote omitted). I take my argument to apply to such a view as well, though the argument in the text may need adjustment to the extent Walen's use of consequentialist considerations at sentencing deviates from a pure retributivist approach.

Importantly, we can ask two different questions about how people ought to decide under conditions of moral uncertainty. One is, "what morally ought we to do?" and the other is, "what rationally ought we to do?" As for what we morally ought to do, there is a philosophical debate raging over such questions. One camp says that what a person *morally* ought to do is independent of his subjective confidence in what he ought to do.[12] Another camp says that what a person *morally* ought to do very much depends on his best probabilistic assessments of what he ought to do, as weighted by the seriousness of each choice.[13]

While my sympathies lie with the view that our moral obligations depend on our probabilistic assessments, those who do not share this view can read my claims as applying not to what retributivists *morally* ought to do but to what they *rationally* ought to do given their current values and beliefs. Even if retributivists are more confident in retributivism than, say, punishment abolitionism or consequentialism, they still need to consider the risks that their justification is wrong. For example, a person might be 99% confident that abandoned luggage does not contain a bomb. But if he believes that there is even a 1% chance the luggage will explode, he rationally ought to take precautions. In other words, it is irrational to guide our behavior only by confidently held beliefs without also considering less probable beliefs that speak to serious consequences.

Hence, rationally consistent retributivists committed to the values underlying the BARD standard will take heed of even small risks that punishment is undeserved given the serious harm of punishing the undeserving. For reasons of consistency, I claim, retributivists must either rethink their commitment to the BARD standard or rethink their commitment to retributivism. (But to the extent I speak of what retributivists rationally ought to do, I cannot say that acting contrary to my advice is immoral without treading into a heated philosophical debate.)

B. Background Moral Uncertainty

In this Section, I ask retributivists to examine their confidence in four background propositions that bear on retributivist justification.

1. Free Will and Other Challenges to Moral Responsibility

To believe that a person *deserves* punishment for his wrongful actions, he must have acted of his own free will. His actions must spring from his will in such a way that he is morally responsible for them. Indeed, our choices *feel* like they spring from within us. But the view that we have minds entirely independent of the forces of the universe has lost its plausibility. Most scholars now agree

12. *See* Elizabeth Harman, *The Irrelevance of Moral Uncertainty*, *in* 10 OXFORD STUD. METAETHICS 53, 53–57 (Russ Shafer-Landau ed., 2015); Brian Weatherson, *Running Risks Morally*, 167 PHIL. STUD. 141, 141 (2014); Elizabeth Harman, Ethics is Hard! What Follows? (Feb. 7, 2014) (unpublished manuscript), https://www.princeton.edu/~eharman/Ethics%20Is%20Hard%20020714%20For%20Web.pdf.

13. *See* TED LOCKHART, MORAL UNCERTAINTY AND ITS CONSEQUENCES (2000); Alexander A. Guerrero, *Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution*, 136 PHIL. STUD. 59 (2007); D. Moller, *Abortion and Moral Risk*, 86 PHIL. 425 (2011).

that we live in a physical universe and that our decisions result from the interaction between our brains and our environment in ways that are governed by laws of nature. In other words, our modern world view is mechanistic. The decisions we make result from the aggregate behavior of trillions of tiny particles in the universe.

What is controversial is how, if at all, this mechanistic view of the universe should affect our views about moral responsibility. Free-will skeptics argue that the kind of free will necessary to generate moral responsibility is incompatible with the mechanistic world we live in.[14] Suppose, they might argue, that the universe is deterministic, meaning that the way the universe is today depends only on the way it was at some other point in time and on nonrandom laws of nature.[15] If the universe is indeed deterministic, then our decisions can be traced, in principle, to physical events preceding our own births.[16] Since we are surely not responsible for events preceding our births and these events caused our current decisions, it is not obvious that we are in control of our decisions, nor that we can be deemed morally responsible for them. If we lack moral responsibility, the entire enterprise of retributivism crumbles.

Moreover, even if the universe is indeterministic—that is, even if our behavior depends in part on truly random behavior of subatomic particles—it is still hard to understand how we can be responsible as we are not responsible for the random behavior of subatomic particles. Despite numerous attempts by brilliant minds over many centuries, no one has demonstrated to widespread satisfaction how we can ever be morally responsible regardless of whether the universe is deterministic or indeterministic.

Joshua Greene and Jonathan Cohen have colorfully defended their skepticism about free will by imagining that a person's decisions are not simply the product of a cold, impartial universe but of an evil cadre of scientists who somehow carefully control all of the factors that lead their puppet person (conveniently named Mr. Puppet) to choose precisely as he does, when he does.[17] When Mr. Puppet subsequently commits a crime, the lead scientist who orchestrated his life testifies as follows:

> It is very simple, really. I designed him. I carefully selected every gene in his body and carefully scripted every significant event in his life so that he would become precisely what he is today. I selected his mother knowing that she would let him cry for hours and hours before picking him up. I carefully selected each of his relatives, teachers, friends, enemies, etc. and told them exactly what to say to him and how to treat him. Things generally went as planned, but not always. For example, the angry letters written to his dead father were not supposed to appear until he was fourteen, but by

14. Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 PHIL. TRANSACTIONS ROYAL SOC'Y B 1775, 1777, 1781 (2004).

15. *Cf.* JOHN MARTIN FISCHER ET AL., FOUR VIEWS ON FREE WILL 2 (2007) ("Something is deterministic if it has only one physically possible outcome."); Gary Watson, *Introduction* to FREE WILL 1, 2 (Gary Watson ed., 1982).

16. GALEN STRAWSON, FREEDOM AND BELIEF 4 (rev. ed., 2010) (describing the "stronger" notion of determinism).

17. Greene & Cohen, *supra* note 14, at 1780; *see also* Gideon Rosen, *The Case for Incompatibilism*, 64 PHIL. PHENOMENOL. RES. 699 (2002).

> the end of his thirteenth year he had already written four of them. In retrospect I think this was because of a handful of substitutions I made to his eighth chromosome. At any rate, my plans for him succeeded, as they have for 95% of the people I've designed. I assure you that the accused deserves none of the credit.[18]

Greene and Cohen believe that Mr. Puppet should not be deemed morally responsible for his actions. And since we have no more control over our genetics and life circumstances than Mr. Puppet does, Greene and Cohen believe that none of us can ever be held morally responsible.[19] Whether our actions are caused by nefarious scientists or just the interactions of physical forces in the universe, our choices seem to be dictated by forces beyond our control.

One solution is to give up on the notion of responsibility. We could decide that, even though people *seem* responsible for their actions, our intuitions about human responsibility are simply a relic of an earlier time. No doubt it would be difficult to part with that notion entirely. When interacting with loved ones and people we meet on the street, we will likely still treat them *as though* they are responsible entities. But it would be wrong to hold people responsible in serious contexts, including the criminal justice system, unless we really believe people can be morally responsible. Deep as our retributive impulses may be, it would be unjust to punish people unless we endorse those impulses as morally correct.

As to whether we have no choice but to treat humans as responsible agents, Greene and Cohen argue that, while it may be too difficult or impractical to avoid attributions of responsibility entirely, we *can* avoid them in the most important contexts:

> [M]odern physics tells us that space is curved. Nevertheless, it may be impossible for us to see the world as anything other than flatly Euclidean in our day-to-day lives. . . . Does it then follow that we are forever bound by our innate Euclidean psychology? The answer depends on the domain of life in question. In navigating the aisles of the grocery store, an intuitive, Euclidean representation of space is not only adequate, but probably inevitable. However, when we are, for example, planning the launch of a spacecraft, we can and should make use of relativistic physical principles that are less intuitive but more accurate. . . . For most day-to-day purposes it may be pointless or impossible to view ourselves or others in [a] detached sort of way. But—and this is the crucial point—it may not be pointless or impossible to adopt this perspective when one is deciding what the criminal law should be or whether a given defendant should be put to death for his crimes.[20]

There is, of course, an alternative view. Compatibilists believe we can hold people responsible even in a mechanistic universe. So long as we identify with our choices, are capable of acting rationally, or meet some similar criteria, they say, we can still be responsible:

> Beginning with Hume, the central idea of what is usually called "classical compatibilism" is the idea that we are at liberty—free—whenever our

18. Greene & Cohen, *supra* note 14, at 1780.
19. *Id.*
20. *Id.* at 1784.

> choices (or intentions) cause the actions chosen (intended). We have the power needed for responsibility, the ability, the free will, whenever our choices cause what we choose them to cause because we made those choices. This is a compatibilist sense of these terms, because the causation of actions by our choices to do those very actions is quite compatible with such choices themselves being caused by factors outside our control. On this version of compatibilism, being a causer in no way requires that one be an uncaused causer.[21]

While it is difficult for some to understand how we can be responsible for our choices when what causes us to make those choices is ultimately beyond our control, it is surely true that we do, in fact, regularly hold people responsible for their conduct. And for many, the pervasiveness of the practice weighs heavily in favor of compatibilism.

So how should reasonable people appraise the likelihood that we can act responsibly? As Benjamin Vilhauer has argued, the fact that the debate has raged for centuries and is still unsettled among professional philosophers is itself some evidence that reasonable people should have doubts.[22] It seems hubristic to cling tenaciously to any position, let alone with 100% confidence. To be 100% sure is to have no doubt at all that we have free will and to be completely unable to change one's mind in the face of contrary arguments and empirical discoveries.

In a large survey of professional philosophers, 59% either "accepted" or "leaned toward" a compatibilist view of responsibility.[23] The rest held another view, were undecided, or thought the question was insufficiently clear.[24] Interestingly, a substantial 12% of professional philosophers accepted or leaned toward the view that we lack free will.[25] While there is much debate about how our own views should be influenced by those of our peers and those we take to be more expert than ourselves,[26] we at least know that lots of thoughtful people who have considered the question disagree.

Only about 35% of those surveyed gave an answer that reflected straightforward confidence in the view that we have free will (as opposed to merely leaning in that direction).[27] If, for the sake of argument, the survey of philosophers provides a rough proxy for our levels of confidence, we might deem ourselves only 35% confident that we have free will. In that case, retributivism would be in big trouble. More conservatively, one might estimate 88% confidence that we have free will by counting all respondents who neither believed nor leaned toward the view that we lack free will.[28] Still, a mere 88% confidence,

---

21. Michael S. Moore, *Stephen Morse on the Fundamental Psycho-Legal Error*, 10 CRIM. L. & PHIL. 45, 69–70 (2016) (footnote omitted).

22. Benjamin Vilhauer, *Free Will and Reasonable Doubt*, 46 AM. PHIL. Q. 131, 136 (2009).

23. David Bourget & David J. Chalmers, *What Do Philosophers Believe?*, 170 PHIL. STUD. 465, 475–76, 494 (2014).

24. *Id.* at 494.

25. *Id.*

26. *See, e.g.*, David Christensen & Jennifer Lackey, *Introduction* to THE EPISTEMOLOGY OF DISAGREEMENT: NEW ESSAYS 1 (David Christensen & Jennifer Lackey eds., 2013); Adam Elga, *Reflection and Disagreement*, 41 NOÛS 478 (2007); David Enoch, *Not Just a Truthometer: Taking Oneself Seriously (but Not Too Seriously) in Cases of Peer Disagreement*, 119 MIND 953 (2010).

27. Bourget & Chalmers, *supra* note 23, at 475–76, 494.

28. *Id.* at 494.

as we will later see, puts substantial strain on retributivism in light of the BARD standard and the other sources of uncertainty we will discuss.

I am certainly not suggesting we can straightforwardly determine our confidence in the existence of free will by examining the results of one survey of philosophers. But given that the free will debate has raged for centuries and that you have not written a paper that has managed to sway the masses of other thoughtful people, it appears stubborn or narcissistic to hold views on the topic with a level of confidence too close to certainty.

Moreover, while concerns about free will pose the most prominent threat to moral responsibility, less famous threats lurk nearby. For example, much has been written about the tremendous role that chance plays in our lives. Even if we have free will, some doubt that we can be morally responsible in a world in which luck affects our personality and preferences, the circumstances we happen to face, and the ways in which our intended actions turn out to help or harm others.[29] Merely having the power to freely choose actions may not be enough for moral responsibility.

Similarly, Gideon Rosen (and others) have written about the threat to moral responsibility posed by moral ignorance.[30] According to Rosen, many immoral actions result from ignorance about what we ought to do, and often it's not our fault that we're ignorant.[31] While people may sometimes know that they are doing something wrong all things considered and persist anyhow, such instances are rarer than we think and potentially too difficult to reliably identify.[32]

Most threatening of all, perhaps, to be confident in moral responsibility means that you believe it will likely survive not only existing attacks but those yet to be imagined. So, with this brief summary of challenges to moral responsibility, I encourage you to write down your confidence in:

Proposition (1): Human beings ordinarily have the kind of free will (and other properties) required for moral responsibility. (Confidence: ___%)[33]

Some might claim that retributivists need not worry if they are mistaken about moral responsibility. If it turns out that we are not morally responsible, retributivists will not themselves be morally responsible for anything either, including whatever unjust punishment they subsequently cause or advocate causing. On the contrary, however, even if we are not morally responsible, good and bad states of the world still exist. And the causing of unjust punishment will

29. *See* NEIL LEVY, HARD LUCK: HOW LUCK UNDERMINES FREE WILL AND MORAL RESPONSIBILITY (2011); Daniel Statman, *Introduction* to MORAL LUCK 1–23 (Daniel Statman ed., 1993); Neil Levy, *Less Blame, Less Crime? The Practical Implications of Moral Responsibility Skepticism*, 3 J. PRAC. ETHICS (2015), http://www.jpe.ox.ac.uk/papers/less-blame-less-crime-the-practical-implications-of-moral-responsibility-skepticism; *see generally* THOMAS NAGEL, MORTAL QUESTIONS 24–38 (1979); BERNARD WILLIAMS, MORAL LUCK: PHILOSOPHICAL PAPERS 1973-1980, at 20–39 (1981).

30. *See, e.g.*, Gideon Rosen, *Culpability and Ignorance*, 103 PROC. ARISTOTELIAN SOC'Y 61 (2003); Susan Wolf, *Moral Saints*, 79 J. PHIL. 419 (1982); Michael J. Zimmerman, *Moral Responsibility and Ignorance*, 107 ETHICS 410 (1997). *But see, e.g.*, Elizabeth Harman, *Does Moral Ignorance Exculpate?*, 24 RATIO 443 (2011).

31. Gideon Rosen, *Skepticism About Moral Responsibility*, 18 ETHICS 295 (2004).

32. *Id.*

33. Most speak of free will as a requirement of moral responsibility. If you believe free will is not required for moral responsibility, then treat this proposition as simply speaking to your confidence that we ordinarily satisfy all requirements of moral responsibility.

certainly be bad—a result that one should seek to avoid, even if one is not morally responsible for it.[34]

2. *Punishment Is an Appropriate Response to Wrongdoing*

Even if we assume hereafter that we have free will and satisfy the other requirements for moral responsibility, the retributivist path is far from vindicated. For retributivists also believe people deserve to suffer or be deprived of liberty because of their wrongful actions. According to John Kleinig, "[t]he principle that the wrongdoer deserves to suffer seems to accord with our deepest intuitions concerning justice."[35] Hence, many retributivists claim that the punishment or suffering of wrongdoers is intrinsically good. As Michael Moore put it, "punishing just deserts is not a proxy for deterrent policy; it is, as any retributivist will say, a freestanding, intrinsic good that those who deserve punishment receive it, even when no other good (such as deterrence) is thereby achieved."[36]

True, many people have intuitions that when someone does something seriously wrong, it's good for the person to suffer or be deprived of liberty. Moore discussed the Russian nobleman in *The Brothers Karamazov* "who turn[ed] loose his dogs to tear apart a young boy before his mother's eyes."[37] According to Moore, we share the intuition that the nobleman should be punished even if we know that doing so will lead to no other benefits (like deterrence or incapacitation).[38] "Violations of others' moral rights," Moore argued, "should make us angry at those who flout morality."[39]

It's not clear how strong or widely shared such intuitions are or how easily they generalize to less serious wrongdoing.[40] Surely many slaveholders held the intuition that slavery was a morally appropriate feature of the natural order, revealing that our intuitions are time- and culture-bound in ways that can lead us astray. Even if we do have strong, widely shared, generalizable intuitions that people deserve to suffer for serious wrongdoing, our intuitions may be wrong. Some of our moral sentiments may "be part of our nature without being the better

34. *See* Adam J. Kolber, *Free Will as a Matter of Law*, *in* PHILOSOPHICAL FOUNDATIONS OF LAW AND NEUROSCIENCE 9, 26–27 (Dennis Patterson & Michael Pardo eds., 2016) [hereinafter Kolber, *Free Will as a Matter of Law*]. Indeed, we may have strong reasons related to moral uncertainty to reject "deflationary ethical theories" such as "nihilism, according to which no action is better than any other, as well as relativistic theories according to which no ethical theory is better than any other." *See* Jacob Ross, *Rejecting Ethical Deflationism*, 116 ETHICS 742, 742 (2006).

35. JOHN KLEINIG, PUNISHMENT AND DESERT 67 (1973).

36. Moore, *supra* note 21, at 61; *see also* VICTOR TADROS, THE ENDS OF HARM: THE MORAL FOUNDATIONS OF CRIMINAL LAW 60 (2011) (describing the retributivist view that "it is good or right that wrongdoers suffer not for any further benefit that their suffering might have, but for its own sake."); TEN, *supra* note 8, at 46 ("Contemporary retributivists treat the notion of desert as central to the retributive theory, punishment being justified in terms of the desert of the offender.").

37. *See* MICHAEL MOORE, PLACING BLAME: A GENERAL THEORY OF CRIMINAL LAW 163 (1997).

38. *Id.*

39. *Id.* at 164.

40. Nathan Hanna, *Retributivism Revisited*, 167 PHIL. STUD. 473, 477–78 (2014).

angels of it."[41] Indeed, many scholars deem it "barbaric"[42] to respond to wrongdoing with retribution.

Moore would say that our anger at those who flout morality "need not be tainted by cruel, sadistic, fearful, or resentful emotional accompaniments."[43] From a first-person perspective, Moore argued that if we ourselves did the nobleman's deeds, we would appropriately feel guilt and quite possibly judge that we deserve to suffer for our wrongdoing.[44] Since it is unlikely that we would have cruel and sadistic sentiments toward ourselves, the judgment that we deserve to suffer for our own wrongdoing survives common criticism that such judgments are barbaric.[45]

I suspect that many who engage in wrongful conduct lack the intuition that they deserve punishment for their bad deeds. We more readily understand the reasons why we did something wrong and are more likely to advocate non-punitive or less punitive solutions to make amends. Even when others engage in wrongful conduct, our anger at them could be channeled through non-punitive approaches like denying offenders benefits, thinking less of them, shunning them, ending friendships with them, and so on. We could even channel our sentiments into consequentialist punishment. So even if Moore is right that our bad deeds ought to make us feel guilt and deserving of punishment, reasonable people can disagree about whether wrongdoing deserves punishment as opposed to some other plausible reaction.

Consider then your confidence in proposition (2) below. Notice that for each proposition after the first, you are asked to take the prior numbered propositions as true. This step is important so that we can treat each proposition as independent of the others. For example, you might think it more likely that punishment is an appropriate response to wrongdoing if we assume that we ordinarily satisfy all of the requirements of moral responsibility. When determining your confidence in proposition (2), you should make that assumption and adjust your confidence level accordingly:

> *Proposition (2)*: Taking the prior numbered proposition as given, those who commit serious wrongs deserve to be punished (or to suffer) in response. (Confidence: ___%)

3. *It Is Possible to Adequately Assess Desert*

Retributivists typically believe that people ought to get what they deserve. The legal system, however, generally only considers a narrow period of defendants' lives and focuses almost exclusively on the conduct for which they are formally accused and convicted. Prior good deeds and non-criminal bad deeds may inform sentencing for recent offenses, but they receive limited independent

41. Barbara H. Fried, *Beyond Blame*, BOS. REV. (June 28, 2013), http://bostonreview.net/forum/barbara-fried-beyond-blame-moral-responsibility-philosophy-law.

42. *See, e.g.*, TADROS, *supra* note 36, at 61 ("Until the recent revival of retributivism, the common view was that retributivism is barbaric in treating the suffering of human beings as good.").

43. MOORE, *supra* note 37, at 164.

44. *Id.*

45. *Id.*

consideration. We pay even less attention to the good or bad things that have already happened to defendants that may or may not have been deserved.

Our narrow focus on recent crimes is convenient from a consequentialist perspective because it would be extraordinarily expensive to investigate and prove all the desert-related facts about a defendant's past. But the principles underlying retributivism arguably require a more global perspective.

Under what is sometimes called the "whole-life" view of desert,[46] to assess what a person deserves, we need to consider all of his good and bad deeds *and* all of his life circumstances since birth. For example, while some people may have engaged in lots of bad acts, they may have also suffered greatly already—so much so, perhaps, that punishing them for recent criminal behavior may make their total desert less appropriate from a whole-life perspective than if the person were simply left alone.[47]

Similarly, some may have engaged in so many good deeds that have never been rewarded that their well-being would best match their whole-life desert if we *refused* to punish them for recent criminal conduct. Indeed, people could bank up opportunities to commit crimes for which they would deserve no punishment. Get-out-of-jail-free cards would be disastrous from a consequentialist perspective but hard to escape from a whole-life retributivist perspective.

And, of course, actually creating a system to monitor whole-life desert would be difficult or impossible. Even defendants themselves may know little about their lives as young children. Steps to monitor people's deeds and suffering might well create new evils of intense surveillance and privacy invasion. Moreover, promoting whole-life desert more generally would seem to require (potentially) objectionable redistribution. People sometimes obtain or lose property by chance, and any system that brings people's life circumstances into line with their moral desert could have radical implications for our treatment of private property and free markets.[48]

Because of the difficulties of the whole-life view, retributivists might prefer the time-period approach largely reflected in our current punishment practices. We only give offenders what they deserve for some usually recent criminal conduct with limited regard for their prior deeds or suffering. The problem with the time-period approach is that there is little to speak in favor of it aside from convenience. If desert makes more sense on a whole-life view, failing to consider a defendant's whole life will frequently lead us to over- or under-punish. And retributivists often emphasize that it is never permissible to purposely, knowingly, or recklessly over-punish.[49] But deliberately failing to examine data

46. Gertrude Ezorsky, *The Ethics of Punishment*, *in* PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT xi, xxvi (Gertrude Ezorsky ed., 1972).

47. Another consideration in favor of the whole-life view is that without it, it is difficult to make sense of the common policy of giving credit for time served in pre-trial detention. Why reduce someone's punishment for time spent unpunished in pre-trial detention? Under the whole-life view, we can treat the suffering in pre-trial detention as a consideration that counts against additional suffering post-conviction. *See* Kolber, *Against Proportional Punishment*, *supra* note 6, at 1170 n.93.

48. *Id.*

49. *See, e.g.*, LARRY ALEXANDER & KIMBERLY KESSLER FERZAN, CRIME AND CULPABILITY: A THEORY OF CRIMINAL LAW 102 n.33 (2009).

that bears on whole-life desert seems to constitute precisely the kind of deliberate ignorance that retributivists reject. Even if you think the time-period view of desert is correct despite its sometimes arbitrary nature, a reasonable person should hold that view with healthy uncertainty.

I think the problem here is alone sufficient to defeat retributivism.[50] But as I barely scratch the surface of the potential arguments and objections, my goal here is simply to recognize doubt about the following retributivist proposition:

> *Proposition (3)*: Taking the prior numbered propositions as given, it is possible and sufficiently practical to assess the relevant background history of a defendant's deeds and life circumstances in order to assess what he deserves. (Confidence: ___%)

4. *The State Can Justly Impose Punishment*

Even if people are morally responsible and even if it is good for them to suffer according to some desert metric of appropriate duration, it is not obvious that such suffering is justly imposed *by the state*. Maybe we should be pleased when thieves get struck by lightning, but that doesn't necessarily mean that the state is justified in creating a system designed to dispense bad things to bad people. As Doug Husak has argued, even if desert is a necessary condition of punishment, it is probably not sufficient.[51]

For now, put aside one set of reasons desert might be insufficient: namely, punishing some particular offender might be too expensive, subject to mistake, and prone to abuse by politicians, judges, police officers, prison guards, and others.[52] The magnitude of these concerns will vary with the circumstances of particular defendants. But an overarching background concern is that desert is simply too weak of a goal to compel the citizenry to pay for it—or at least too weak to set up whatever minimal set of punishment institutions retributivists think necessary.

Ordinarily, we are not obliged as citizens to ensure that people get the good things they deserve. "When we say that an especially hard-working self-employed farmer deserves to succeed, or that a person of fine moral character deserves to fare well, we typically do not mean that anyone is obligated to take steps to provide what is deserved."[53] Hence Victor Tadros worries that, by expending substantial resources to deliver offenders' just deserts, the state coerces citizens to pursue goals they may not share:

> [S]tate action utilizes citizens' resources, resources which they would otherwise be permitted to use in pursuit of their private goals, for the ends that it sets itself. This amounts to the use of the labour of citizens to pursue the good. In order to justify some state project that uses resources generated

50. I discuss these matters more fully in Adam J. Kolber, The Time Period Challenge to Retributivism (unpublished manuscript) (on file with author).

51. Douglas Husak, *Holistic Retributivism*, 88. CALIF. L. REV. 991, 994 (2000); Douglas Husak, *Why Punish the Deserving?*, 26 NOÛS 447, 450 (1992) [hereinafter Husak, *Why Punish the Deserving?*].

52. Husak, *Why Punish the Deserving?*, *supra* note 51, at 450–51.

53. GEORGE SHER, DESERT 5 (1987).

by citizens, it must be shown that it is permissible to coerce citizens to work in order to pursue that project.[54]

Of course, in a representative democracy, citizens vote to create institutions which impose punishment. But it is not obvious that they have voted to create institutions to inflict *retributive* punishment. More importantly, voting results are not equivalent to moral verdicts. For example, even if a majority of voters decided to create institutions to give people their positive deserts, there would be a further question of whether such obligations are morally appropriate. Do they, perhaps, violate deontological prohibitions against using some people (including their labor) as *mere means* to further the ends of others (namely giving those others what they affirmatively deserve)? To put it another way, our worry is about moral justification, and you cannot resolve a moral matter simply by tallying votes.

If you still have no qualms about the morality of creating state institutions to inflict suffering, consider that for state punishment to be just, it is likely a prerequisite that government more generally can be just. And philosophers have debated the legitimacy of government for centuries. To the extent there is some chance that anarchists are right, then there is a chance that the state itself is unjust along with any punishment imposed by the state. With these considerations in mind, rate your confidence in:

Proposition (4): Taking the prior numbered propositions as given, state coercion generally can be just and, more particularly, the state is morally permitted to impose on the citizenry to create institutions that punish (or make suffer) those who deserve it. (Confidence: __%)

C. Total Background Moral Uncertainty

If you've been contemplating pertinent confidence levels yourself, you can now calculate your overall confidence in the four retributivist background propositions described here. To do so, just multiply your percentage confidence in each of the first four propositions. The product doesn't quite represent your overall background confidence in retributivism. There might be other background propositions we should add. Any uncertainty in such additional propositions would make your overall confidence lower. But so long as you share those four views (perhaps with minor variations), the product represents your maximum confidence that imposition of retributive punishment is ever justified.

In the next Section, we will add further propositions one must believe to deem any particular individual justly punished in retributive terms. At this point, however, you can compare your level of confidence in key background principles of retributivism to the level of confidence you believe fact finders should have when convicting an offender for violating a statute. The U.S. Constitution requires that a defendant at trial only be convicted if the fact finder (usually a jury) believes the defendant committed every element of an offense beyond a

54. TADROS, *supra* note 36, at 79.

reasonable doubt.[55] The BARD standard is high—much higher than the "preponderance of evidence" standard used in most civil contexts. It is also substantially more demanding than the clear-and-convincing-evidence standard commonly used in civil commitment hearings and punitive damage determinations.

Scholars frequently speak of the BARD standard as requiring a confidence level of 90%, 95%, or even higher.[56] Regardless of how we ought to understand reasonable doubt or explain the concept to jurors, we clearly seek to keep false convictions low. Many retributivists would balk at convicting someone with less than 90% confidence in guilt regardless of how reasonable doubt is formulated or whether the precise degree of confidence varies by case or offense. If we regularly convicted defendants with 90% confidence in their guilt, we would expect that for every ten innocent defendants, an average of one would be erroneously convicted.

Retributivists set a high burden in criminal cases because they are so eager to keep false convictions low that they are willing to tolerate many false acquittals to do so.[57] According to the famous Blackstone ratio, "[i]t is better that ten guilty persons escape than that one innocent suffer."[58] Benjamin Franklin went further, asserting it better that "a hundred guilty persons should escape than one innocent person should suffer."[59] Whatever one's preferred ratio,[60] for a given amount of punishment at issue, false convictions are worse than false acquittals.

Terms like "guilt" and "innocence" as used in the Blackstone and Franklin ratios presumably refer, at least principally, to factual errors that might be considered by a jury: for example, did the police nab the wrong person? But as I've been discussing, there are many ways punishment can be unjust from a retributivist perspective that do not involve straightforward mistakes about guilt. Whether a person was framed by an ex-lover or actually violated a statute but lacks metaphysical free will, he is *undeserving* of retributive punishment. Indeed, he is completely undeserving either way.

If your confidence in the conjunction of the four propositions falls below 50%, it seems you are no longer a retributivist (if you ever were). It would mean that in every criminal case, you would be more confident that punishment is not retributively justified than that it is. On the other hand, some may have been quite confident in the foregoing propositions. Imagine a hardcore-but-reasonable retributivist. Perhaps he will be 90% confident that we are morally responsible and 95% confident in each of the other three propositions. If so, he will have a maximum of 77% confidence that the sentence of any alleged offender to proportional punishment can be justified in retributivist terms.[61] Would that be a sufficient justificatory standard of proof? There's no need to decide yet since we're just getting started.

---

55. In re Winship, 397 U.S. 358, 361 (1970).

56. *See* sources cited *supra* note 2.

57. *See, e.g.*, Daniel Epps, *One Last Word on the Blackstone Principle*, 102 VA. L. REV. 34, 35 (2016).

58. BLACKSTONE, *supra* note 3, at *358.

59. Letter from Benjamin Franklin to Benjamin Vaughn (Mar. 14, 1785), *in* 11 THE WORKS OF BENJAMIN FRANKLIN 13 (John Bigelow ed., fed. ed. 1904).

60. *See* Alexander Volokh, n *Guilty Men*, 146 U. PENN. L. REV. 173, 174–77 (1997).

61. .90*.95*.95*.95 = 77%.

Before proceeding, however, one might ask if it is simply too difficult to quantify levels of uncertainty. I doubt it. But nothing in this Article turns on precise quantifications. Sure, it is much more elegant if you can estimate your levels of uncertainty; doing so allows us to approximate your overall confidence in retributivism.[62] But you could also answer each question with "extremely confident" or "somewhat confident." We won't be able to easily represent your overall confidence levels, but we can still generally see how your required level of confidence in moral propositions compares to your required level of confidence in factual guilt.

More importantly, if it really were impossible to approximate our levels of uncertainty, it might strengthen the epistemic challenge to retributivism. People convicted of crimes could appropriately ask how confident we are that they deserve their assigned punishments. Replying that we don't know or that such levels of confidence cannot be expressed seems to weaken the case that the imposition is appropriate.

D. Case-Specific Factual Uncertainty

We now turn from uncertainty about general retributivist propositions to uncertainty that varies by case. One proposition retributivists must believe in some particular case is that the accused engaged in the conduct alleged. We need specific facts to answer this in real life, but for now, consider the minimum confidence we must have that the accused engaged in alleged conduct in order to justly convict:

Proposition (5): Taking the prior numbered propositions as given, this offender engaged in the conduct alleged.[63] (Required Minimum Confidence ___%)

Of course, for many people, the minimum required confidence will depend on how they interpret the BARD standard. There is, however, no general agreement as to what degree of confidence is required to convict beyond a reasonable doubt. There is also no general agreement that the BARD standard can or should be understood quantitatively.[64] Courts are reluctant to quantify reasonable doubt.

62. Larry Solum has argued that the propositions I offer are not independent of each other. Larry Solum, *Kolber on Punishment and Moral Risk*, LEGAL THEORY BLOG (Jan. 11, 2017), http://lsolum.typepad.com/legaltheory/2017/01/kolber-on-punishment-and-moral-risk.html. Were it true, we would be unable to simply multiply probabilities in the way that I do. Contra Solum, however, I ensure independence by instructing readers to provide their levels of confidence *taking prior propositions as given*. Arriving at these confidence levels is, of course, quite challenging and somewhat artificial, but there is surely considerable independence among the substantive topics I discuss. For example, many of those who believe it barbaric to make offenders suffer nevertheless believe in free will. The same sort of independence will often be found in the fact-specific propositions in the next Section.

63. In some cases, offenders may be innocent of conduct charged but have committed other wrongs of which they have never been formally accused. I believe most retributivists would deem it unjust to punish for acts that have not been brought to a formal tribunal. Those who would support rough justice, however, can respond to this proposition, if the particular defendant under consideration has committed other comparable wrongful conduct, by giving their confidence in the belief that such rough justice is morally appropriate.

64. *Compare* Ronald J. Allen & Alex Stein, *Evidence, Probability, and the Burden of Proof*, 55 ARIZ. L. REV. 557, 560 (2013) (arguing that fact finders do and should continue to "decide cases predominantly by applying the relative plausibility criterion guided by inference to the best explanation, rather than by using mathematical probability"), *with* Peter Tillers & Jonathan Gottfried, *Case Comment*—United States v. Copeland, *369 F. Supp. 2d 275 (E.D.N.Y. 2005): A Collateral Attack on the Legal Maxim That Proof Beyond a Reasonable*

As one court put it, "[t]he idea of reasonable doubt is not susceptible to quantification; it is inherently qualitative."[65] Alternatively, perhaps we fear quantification because it implies that innocent people are sometimes punished:

> *Any* specification of a degree of belief necessary for a finding of guilt (such as 95 percent confidence) involves an *explicit* admission that wrongful convictions will inevitably occur. For instance, if jurors could somehow discover that they had a confidence of 95 percent in the guilt of the accused, this would generally suggest that as many as one in every twenty innocent defendants will be wrongly convicted.[66]

Some scholars, such as Ronald Allen and Alex Stein, believe that mathematical approaches to reasonable doubt lead to paradoxes and do not reflect the kinds of reasoning actually used in court.[67] Rather, they claim, court cases present opportunities for storytelling in which fact finders seek to determine the relative plausibility of each side of a dispute.[68] The view represents a marked departure from the traditional way in which the criminal law standard of proof is understood.

Fortunately, the debate about quantifying the BARD standard is largely irrelevant for our purposes. I do take retributivists to believe that it is much worse to punish an innocent person than to fail to punish a guilty person. But other than that, nothing I say relates to how we tell jurors to assess guilt. I am asking about the perspective not of a juror in a courtroom but of a theorist providing moral justification. The theorist is being asked to consider how confident we are that a particular offender engaged in particular conduct. How we instruct jurors to decide cases is a separate issue.

E. Case-Specific Moral Uncertainty

Even if we were sure that necessary background conditions were satisfied and that some defendant engaged in the conduct alleged, we would still have to cope with moral uncertainty that varies on a case-by-case basis.

Doubt Is Unquantifiable?, 5 LAW, PROBABILITY, & RISK 135, 135 (2006) (arguing that the "usual reasons for unquantifiability of reasonable doubt are unsatisfactory").

65. Massachusetts v. Sullivan, 482 N.E.2d 1198, 1200 (Mass. App. Ct. 1985); *see* McCullough v. State, 657 P.2d 1157, 1159 (Nev. 1983) ("The concept of reasonable doubt is inherently qualitative. Any attempt to quantify it may impermissibly lower the prosecution's burden of proof, and is likely to confuse rather than clarify.").

66. LAUDAN, *supra* note 2, at 45.

67. Allen & Stein, *supra* note 64, at 560.

68. According to Allen:

> Notwithstanding the rhetoric about the government having to "bear the laboring oar" in criminal cases, and that the defendant may remain passive and make the government prove its case, the probability of an acquittal goes down dramatically if the defendant does not take the stand. And for very good reason. Everyone in the courtroom knows that at least one person present—the defendant—knows what he was doing on the day in question, and the failure to testify about it leads inexorably to the conclusion that the government's plausible story of guilt is true. This generalizes to the defendant having to produce a plausible story in opposition to a plausible story produced by the government (notwithstanding the rhetoric of the self-incrimination clause of the fifth amendment to the US Constitution, and its exegesis by case law).

Ronald J. Allen, *No Plausible Alternative to a Plausible Story of Guilt as the Rule of Decision in Criminal Cases in* PRUEBA Y ESTÁNDARES DE PRUEBA EN EL DERECHO (PROOF AND STANDARDS OF PROOF IN THE LAW) (Juan Cruz & Larry Laudan eds., 2010).

1. *The Defendant's Conduct Was Actually Wrongful*

Mere violation of a law is insufficient to morally justify punishment. Retributivists must believe that a particular offender engaged in immoral (and not just illegal) behavior. For example, even though we used to punish those who engaged in adult consensual sodomy,[69] and many countries still do,[70] retributive punishment for such conduct is nevertheless unjustified.

In some cases, there will be little doubt that conduct was wrongful, as when a person violently guns down bystanders in a convenience store. There are entire categories of crimes, however, in which reasonable people disagree about whether they prohibit wrongful conduct, such as drug possession, gambling, prostitution, and insider trading. Hence, the proposition that a defendant's alleged conduct was actually wrongful requires a case-specific inquiry. While retributivists could argue that we should have fewer unjust laws, even a retributivist dreamland would presumably still be saddled with some residual moral uncertainty about the wrongfulness of conduct.

Retributivists might argue that a person who violates a duly enacted statute in a democracy has, in some sense, flouted the will of the people. So even if a person did not engage in wrongful behavior, the retributivist need not worry, so long as a criminal statute was violated. For example, assume it was widely but erroneously considered morally reprehensible to possess alcohol during Prohibition. Might one nevertheless think it appropriate to punish those who possessed alcohol when doing so was prohibited by law? A retributivist might think so. The person who violated the law took a rule of the game, so to speak, and failed to follow it. Others may have relied on him to follow the law to their detriment or may have constrained their own behavior in ways that make it unfair for the offender to go entirely without punishment. So *flouting* the law might be wrongful. But the person we're imagining wasn't charged merely with flouting the law. He was charged with alcohol possession. Alcohol possession, by hypothesis, is not morally problematic, and it cannot be made so by the passage of legislation.

True, a conviction for possessing alcohol could be viewed as rough justice for flouting a duly enacted law. But notice that the wrongfulness of mere flouting will generally be less than the perceived wrongfulness of violating some statute. Criminal sentencing takes offenders to have *not only* flouted laws but *also* to have engaged in independently wrongful conduct. Hence we are likely to punish mere flouting too severely. Moreover, the seriousness of flouting may depend on factors such as the extent to which people relied on the defendant's good behavior to their detriment or how unfair it is that others had to constrain their

69. Rebecca Onion, *A 1964 Document Tallying Penalties for Sodomy and Fornication Across the United States*, SLATE (Feb. 25, 2015, 11:44 AM), http://www.slate.com/blogs/the_vault/2015/02/25/history_of_sex_based_offenses_penalties_for_sodomy_fornication_adultery.html; *see also 12 States Still Ban Sodomy a Decade After Court Ruling*, USA TODAY (Apr. 21, 2014, 6:42 PM), https://www.usatoday.com/story/news/nation/2014/04/21/12-states-ban-sodomy-a-decade-after-court-ruling/7981025/.

70. Max Bearak & Darla Cameron, *Here Are the 10 Countries Where Homosexuality May Be Punished by Death*, WASH. POST (June 16, 2016), https://www.washingtonpost.com/news/worldviews/wp/2016/06/13/here-are-the-10-countries-where-homosexuality-may-be-punished-by-death-2/?utm_term=.310b9e4f93cc.

behavior. So the wrongfulness of mere flouting will often be distinct from the perceived wrongfulness of the substantive offense.

The bottom line, then, is that retributivists need to be confident that a particular defendant engaged in morally wrongful conduct. How confident? Patrick Tomlin has argued that if you accept the BARD standard out of concern for the intrinsic bad or wrongness of punishing people who shouldn't be, we should apply the same standard to the question of just criminalization.[71] If so, reasonable retributivists would likely be insufficiently confident to punish many common crimes, and using mere "flouting" as a basis to punish substantially increases the risk that punishment will be disproportional (a topic we examine later).

Since I cannot ask you to consider your confidence that some particular offender's conduct was wrongful without giving you the details of a particular case, you might consider your minimum required confidence in the following proposition:

Proposition (6): Taking the prior numbered propositions as given, the offender engaged in morally wrongful conduct. (Required Minimum Confidence: ___ %)

2. *Police Conduct and Judicial Proceedings Were Sufficiently Just*

Many retributivists care not only that a particular person engaged in past wrongful conduct but also require that he be arrested and punished under just laws and procedures. Among them, one might require that (1) a duly enacted law proscribed the defendant's conduct and was made public before the conduct occurred, (2) the offense definition is consistent with constitutional requirements, (3) the defendant was properly arrested and charged and had advice of counsel, (4) the defendant was given the opportunity to proceed to a trial by a jury of peers, (5) no illegal evidence was used to prosecute the defendant, (6) the defendant was not forced to testify, (7) the defendant had proper opportunities to appeal, and so on.

To be sure, some retributivists might disregard these sorts of requirements. They might believe that punishment is morally justified so long as a person genuinely engaged in serious wrongful behavior. Other requirements, they say, may be established by law but needn't be present for punishment to be morally justified.

I suspect, however, that most retributivists believe a person's punishment cannot be morally justified if it is not also legally justified. So if a thief engages in serious wrongful conduct, such retributivists will say that he ought not be punished if he is denied the right to counsel or convicted by evidence that should have been excluded under the Fourth Amendment. Such retributivists will sometimes have substantial doubts as to whether a particular defendant's conviction was lawful and will always have at least residual doubts.

---

71. Patrick Tomlin, *Extending the Golden Thread? Criminalisation and the Presumption of Innocence*, 21 J. POL. PHIL. 44, 44 (2013).

You can now consider your minimum required confidence in proposition (7) below. (If you believe that punishment would be just even in the face of atrocious police conduct and unfair judicial proceedings, then treat proposition (7) as asking about your confidence in that belief.)

> *Proposition (7)*: Taking the prior numbered propositions as given, police conduct and judicial proceedings were of sufficient quality that just punishment has not been precluded. (Required Minimum Confidence: ___ %)

3. *The Value of Giving This Offender What He Deserves Does Not Grossly Exceed the Costs*

In proposition (4), we considered the possibility that desert provides too weak of a reason to compel the citizenry to pay for punishment—or, at least, too weak to set up whatever minimal set of punishment institutions retributivists think necessary. Now consider the possibility that even with minimal punishment institutions in place, it may be unjust to punish some particular defendant because the costs are simply too high.

As Doug Husak has argued, punishment is "tremendously expensive, subject to grave error, and susceptible to enormous abuse."[72] Since we are considering a deontologically minded version of retributivism, there is no strict requirement that the benefits of punishment exceed the costs. Perhaps a punishment can still be just even when its costs moderately exceed its benefits. But when its costs dramatically exceed its benefits, justice may preclude incarceration, particularly when those resources would be better deployed feeding the hungry and healing the sick. You can now determine your minimum required confidence in the following proposition:

> *Proposition (8)*: Taking the prior numbered propositions as given, the costs of punishing this offender do not so outweigh the benefits of giving him what he deserves as to make punishing him unjust. (Required Minimum Confidence: ___ %)

4. *Proportionality Is a Worthy Goal and This Sentence is Proportional*

We end our current inquiry by examining a feature of retributivism that may generate the most uncertainty of all. Namely, even if all the foregoing propositions are true, they say nothing about *how much* we are permitted to punish. We cannot impose *any* punishment unless we have confidence in some proposition that authorizes a particular amount of punishment.

In the eyes of retributivists, having done something sufficiently wrongful in the past makes a person deserve *proportional* punishment.[73] Upon closer examination, however, proportional punishment provides a far less attractive ideal

---

72. Husak, *Why Punish the Deserving?*, *supra* note 51, at 450.

73. *See, e.g.*, Mary Sigler, *Private Prisons, Public Functions, and the Meaning of Punishment*, 38 FLA. ST. U. L. REV. 149, 164 (2010) ("Retributivism . . . is centrally concerned with the imposition of punishment in proportion to an offender's moral desert."). There is some debate as to whether punishment (or suffering) should be proportional to culpability or blameworthiness or the seriousness of an offense, but the basic formula stays the same. Kolber, *Against Proportional Punishment*, *supra* note 6, at 1143.

than many think.[74] One natural way of understanding the amount of punishment we inflict is to consider the amount of suffering our punishments cause. At sentencing, however, we pay little attention to the amount offenders are likely to suffer, and we rarely correct sentences when actual suffering deviates from what was expected.[75] Though prisoner suffering varies considerably, judges generally ignore this fact or, if they do consider it, they do so surreptitiously.[76] So, if punishment depends on suffering, we don't care enough about proportionality to make the amount of punishment fit the crime.

Of course, one solution to this problem is simply to do more to take individual subjective experience into account. But when we combine proportionality with individualized consideration of suffering, proportionality looks even less attractive. Many people would be upset if we imposed a shorter prison term on a wealthy socialite than an equally blameworthy hardened gang member, even though the socialite would quite plausibly experience more total suffering.[77] In other words, not only do we fail to inflict proportional suffering but, when examined closely, we are *averse* to proportional suffering.

Similar comments apply to those who understand punishment not as an infliction of suffering but as a deprivation of liberty.[78] We pay little attention to the amount we actually deprive offenders of liberty.[79] Suppose one person commits a computer crime from a fancy beach house in Malibu. Another commits the same crime from a federal facility where he is being held in quarantine indefinitely because, through no fault of his own, he suffers from a rare communicable disease. Suppose we want to imprison them so that they are deprived of liberty to an equal degree. To do so, we would have to make the conditions of confinement of the person held in quarantine much more restricted (or much longer in duration) than the conditions of the wealthy beachgoer. The reason is that the beachgoer starts out with lots of liberty, so his imprisonment will deprive him of a lot. The person in quarantine has limited liberty to begin with. His imprisonment may only deprive him of a bit of liberty relative to his baseline.[80] Hence, assuming they are equally blameworthy, depriving each offender to the *same degree* requires imprisoning them rather differently. But that seems quite unpopular. Many people would rather punish them in the same prison conditions for the same duration even though doing so inflicts a much greater deprivation of liberty on the wealthy beachgoer.[81]

Indeed, rather than inflicting proportional suffering or proportional deprivations of liberty, our actual incarcerative practices are more closely geared toward inflicting proportional *time periods* of incarceration. We have a duration

74. Kolber, *Against Proportional Punishment*, *supra* note 6.

75. Adam J. Kolber, *The Subjective Experience of Punishment*, 109 COLUM. L. REV. 182, 187–96 (2009) [hereinafter Kolber, *The Subjective Experience of Punishment*].

76. *Id.* at 192–96.

77. *Id.* at 230–35.

78. Adam J. Kolber, *The Comparative Nature of Punishment*, 89 B.U. L. REV. 1565, 1567 (2009) [hereinafter Kolber, *The Comparative Nature of Punishment*].

79. *See, e.g.*, *id.* at 1591–92.

80. A whole-life view of desert might ease baseline issues but would have other problems discussed in Section II.B.3.

81. Kolber, *The Comparative Nature of Punishment*, *supra* note 78.

fetish that leads us to pretend that all terms of incarceration of equal duration are equally severe.[82] Once we see that our intuitions about incarcerative severity are primarily pegged to duration, however, the appeal of proportionality drops off considerably. While amounts of suffering or of liberty deprivation are plausibly morally relevant, amounts of time—unweighted by other considerations—are not.

Even if we could somehow surmount these problems and determine that proportionality really does set a worthy goal, no one has discovered any widely accepted formula for determining amounts of proportional punishment. How much punishment is proportional to petty theft, sexual assault, or murder? Without principled answers to such questions, we should have significant doubts that retributivists can confidently determine even *roughly* proportional punishments.

Some empirical research suggests that people have widely shared intuitions of the *ordinal* severity of various crimes.[83] Petty theft warrants less punishment than sexual assault, which warrants less punishment than murder. Paul Robinson and Robert Kurzban took cross-cultural evidence of consistent ordinal rankings to "support the view that people everywhere share intuitions of justice about the relative blameworthiness of serious wrongdoing."[84]

The fact that our ordinal punishment intuitions are surprisingly consistent, however, helps little in justifying punishment given that intuitions about actual *amounts* of punishment vary dramatically.[85] For example, Robinson and Kurzban noted that:

> Average prison sentences vary widely from nation to nation. American offenders were required to serve an average of twenty-nine months after conviction in 1999. In contrast, the average offender in the Netherlands was released after five months, while Columbian offenders were not released until a startling mean of 140 months. Moreover, even within a culture, community attitudes toward punishment severity can vary over time.[86]

Moreover, since the amount of harm we do to those who don't deserve it varies with the severity of their punishment, it is cardinal punishment severity that matters most. To see why, imagine some society that punished petty theft with twenty years of imprisonment. Such a sentence would be grossly disproportional even if our imagined society punished sexual assault with sixty years of imprisonment and murder with death. We cannot remedy disproportional punishment simply by increasing punishments for *other crimes*. Hence, even if proportionality were a worthy goal, there is no established method of determining how much punishment is proportional. Indeed, we are often unsure which punishment modality (for example, incarceration or supervised release) is appropriate, let alone the appropriate duration and severity of some particular modality.

---

82. *Id.* at 1606.

83. Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829, 1856–61 (2007).

84. *Id.* at 1862.

85. *See id.* at 1880–82; Paul H. Robinson et al., *The Disutility of Injustice*, 85 N.Y.U. L. REV. 1940, 1962 (2010) ("While people tend to agree on the proper rank order of cases on the punishment continuum, at least for the kind of core harms at issue here, some people tend to be harsh in their 'sentencing,' while others are lenient.").

86. Robinson & Kurzban, *supra* note 83, at 1882.

The only way to be quite confident that retributivists are not overpunishing is to make their best assessment of proportionality and then dramatically underpunish relative to that assessment. Rapists, for example, could receive a month in prison, and murderers could receive two months. If we make punishments too short, however, we increase doubt as to proposition (8) concerning excessive costs. After all, it's not obvious that we can justify creating a substantial criminal justice apparatus so that we can give hardened criminals the very short sentences required to make us very confident that their sentences are not disproportional. And, of course, retributivists who believe that delivering proportional punishment is not merely permitted but obligatory are likely to frustrate their own desert goals if they must punish in the safe, sufficiently short zone.

Consider then your minimum confidence in proposition (9) below. To do so, you might want to envision a criminal justice system with punishments that you view as in some way realistic or loosely consistent with our current system:

Proposition (9): Taking the prior numbered propositions as given, punishment should be proportional (or at least not disproportional), and this offender's punishment is proportional (or at least not disproportional). (Required Minimum Confidence: ___ %)

F. A Staunch Retributivist's Overall Uncertainty

Imagine the staunchest, most confident retributivist you can. Suppose he is contemplating the case of a particular offender and is also quite confident that the offender engaged in the heinous crimes of which he stands accused. Despite raging debates about free will and proportionality, assume that he assigns 99% confidence to each of the nine propositions we've considered.

My reply to the staunch retributivist is simply to doubt that his confidence is reasonable. What my bite-size nuggets about major philosophical debates are meant to show is that reasonable retributivists should have significant doubts about most of the propositions I discussed. Moreover, even this staunch retributivist will be only 91% confident that the offender's punishment is justified overall from his retributivist point of view.[87] Still, I don't say much directly to the staunch retributivist because I doubt he is a reasonable retributivist.

G. A Hardcore-But-Reasonable Retributivist's Overall Uncertainty

Now return to the hardcore-but-reasonable retributivist we previously imagined. He was 90% confident we have free will and 95% confident in each of the other three background propositions. As for case-specific uncertainty, assume that, as to some particular offender, he has 95% confidence in each of the remaining propositions. So, despite his relatively high confidence as to each proposition, his doubts quickly multiply—literally. This retributivist's overall confidence that the offender's punishment is justified will be less than 60%.[88] The result will vary, of course, in different cases. But I suspect that reasonable retributivists will generally have insufficient confidence to satisfy a justificatory

87. 91%=.99^9; *see* Campbell, *supra* note 7.

88. 59.7%=.90*.95^8; *see* Campbell, *supra* note 7.

standard of proof consistent with the values that seem to underlie their commitment to the BARD standard—a topic we examine more closely in the next Part.

III. THE JUSTIFICATORY STANDARD OF PROOF

As noted in the introduction, scholars have had a lot to say about the guilt standard of proof. But they have said little about the justificatory standard of proof: the confidence required to deem a defendant's punishment morally justified. I will suggest that whatever the precise standard should be, reasonable retributivists are unlikely to satisfy a standard consistent with the values underlying their customary commitment to the BARD standard. I also argue that consequentialists are less vulnerable to the epistemic challenge than retributivists are.

A. What Should the Justificatory Standard of Proof Be?

Whether one subscribes to the Blackstone ratio, the Franklin ratio, or some other, the standard retributivist line is that it is far worse for an innocent person to be punished than for a guilty person to go free. It's hardly obvious why retributivists would treat errors of deservingness in general differently than errors of guilt in particular. Desert is what justifies punishment for retributivists, and any error that vitiates desert vitiates retributivist justification. So one view retributivists could hold is that errors of guilt and other errors of deservingness are in parity and should be treated the same. If the error parity thesis is true, other considerations that vitiate deservingness should be held to the same high standard as errors about guilt.

1. Same Standard for Guilt and Moral Uncertainty?

In recent years, some scholars have argued for a BARD standard as to particular components of retributive justification. Benjamin Vilhauer claimed we cannot retributively punish someone unless we believe he has free will beyond a reasonable doubt.[89] Nathan Hanna suggested that we can only treat the suffering of wrongdoers as intrinsically good if we believe it to be so beyond a reasonable doubt.[90] And Patrick Tomlin argued that, given certain plausible assumptions, we should only punish to the extent we are sure beyond a reasonable doubt that we are not overpunishing.[91] These views leave open the broader question of the overall justificatory standard of proof.[92]

89. Vilhauer, *supra* note 22, at 131–32.

90. Hanna, *supra* note 40, at 478–80.

91. Patrick Tomlin, *Could the Presumption of Innocence Protect the Guilty?*, 8 CRIM. L. & PHIL. 431, 432 (2014).

92. Perhaps Nathan Hanna would apply the BARD standard not just to this particular claim but to the entire question of whether punishment is justified. Hanna, *supra* note 40, at 480 (speaking of the BARD standard and noting that "[m]aybe we should hold *justifications of punishment* to a similar standard") (emphasis added).

Hyman Gross has claimed that "it would be wrong to condemn anyone to criminal punishment unless criminal punishment itself were justifiable beyond a reasonable doubt."[93] According to Gross:

> [I]t is politically barbaric to ensure as well as we can against the possibility of unjustified liability to punishment, while at the same time leaving the justification of punishment itself in a state where reasonable people may disagree. "We are sure beyond a reasonable doubt that you are liable to punishment," says the jury, but the legal system must add, "Though we are not nearly as certain that what you are liable to is justifiable."[94]

If Gross is right, it seems the conjunction of claims we have examined must, for retributivists, be true beyond a reasonable doubt.

We might bolster Gross's approach by considering the way some courts treat the defense of automatism. It is widely agreed that in order to be convicted of a crime, one must commit a voluntary act or omission. For example, in *People v. Newton*, Huey Newton was accused of shooting a police officer.[95] The facts were in dispute, but Newton claimed that at the time the officer was shot, he himself had been shot in the abdomen and was not in a conscious state.[96] His expert claimed that the kind of wound Newton received was "very likely to produce a profound reflex shock reaction[,]" and that it was "not at all uncommon for a person shot in the abdomen to lose consciousness and go into this reflex shock condition for short periods of time up to half an hour or so."[97] The court held that the jury should have considered the possibility that Newton was unconscious at the time he allegedly fired on the officer; if he had been unconscious, he could not have been responsible and would have no criminal liability.[98]

Importantly, the factual question of Newton's responsibility was subject to a BARD standard[99]—the same standard the court used to determine whether Newton was even at the crime scene.[100] It is hardly obvious why we would not apply the same BARD standard to the equally important question of whether Newton had metaphysical free will (and by extension whether any of us have free will).

---

93. HYMAN GROSS, CRIME AND PUNISHMENT: A CONCISE MORAL CRITIQUE 9–10 (2012). *But cf.* Leo Zaibert, *Of Normal Human Sympathies and Clear Consciences: Comments on Hyman Gross's* Crime and Punishment: A Concise Moral Critique, 10 CRIM. L. & PHIL. 91, 93–94 (2016); Eric A. Johnson, Book Review, NOTRE DAME PHIL. REV. (June 28, 2012), http://ndpr.nd.edu/news/31529-crime-and-punishment-a-concise-moral-critique (reviewing GROSS, *supra*).

94. GROSS, *supra* note 93, at 10.

95. 87 Cal. Rptr. 394, 394 (Cal. Ct. App. 1970).

96. *Id.* at 402.

97. *Id.* at 403.

98. *Id.* at 404, 406.

99. Some jurisdictions, however, examine automatism under an insanity defense (in which the BARD standard may not apply). *See, e.g.*, McClain v. State, 678 N.E.2d 104, 107 (Ind. 1997) ("[J]urisdictions are split between recognizing insanity and automatism as separate defenses and classifying automatism as a species of the insanity defense."); Deborah W. Denno, *Crime and Consciousness: Science and Involuntary Acts*, 87 MINN. L. REV. 269, 284–85 (2002).

100. *Newton*, 87 Cal. Rptr. at 404–06.

One might argue that the *law* does not directly examine whether people have metaphysical free will, while it does inquire into unconsciousness.[101] But even if true, the purported *legal* distinction is not one that has any obvious *moral* significance. So why would pure retributivists—those who are not relying on consequentialist considerations to assign punishment—treat uncertainty about *facts* that make a person undeserving of punishment significantly differently than they treat *moral uncertainty* that raises doubts about deservingness?

Retributivists might say: "My principal concern is about deservingness and either moral or factual errors can lead to the punishment of the undeserving. Prior to consideration of moral uncertainty, I only worried about uncertainty related to guilt. Now that I find much more uncertainty than I had previously recognized, I must take it into consideration. My prior allowance for uncertainty related to guilt must now be shared with other sources of uncertainty." Hence, retributivists might choose a justificatory standard of proof that matches what they previously required for guilt. If they previously required, say, 95% confidence in guilt, they would now require at least 95% confidence that a person's punishment is retributively justified overall.

Such a standard would be almost impossible to satisfy in any real case. It would take about 99.5% confidence in each of the nine propositions we examined to have an overall confidence level of about 95%.[102] Recall, too, that the person I called a hardcore-but-reasonable retributivist only mustered a 77% *background* level of confidence. And his less than 60% case-specific confidence is far lower still. Even the *unreasonable*, staunch retributivist with 91% confidence wouldn't satisfy a 95% justificatory standard of proof. In other words, a 95% justificatory standard would require retributivists to be much more demanding of certainty—indeed, too demanding—such that, by their own lights, they could no longer justifiably punish on retributive grounds.

2. *Attempts to Avoid the Error Parity Principle*

While a plausible case can be made for holding all errors related to justification to the same high standard as errors of guilt, perhaps a retributivist could sensibly treat them differently. For example, the harms of erroneously believing someone violated the elements of a criminal statute may be worse than the harms of erroneously believing someone has free will. Either error could land someone in prison (inappropriately by retributivist standards), but the psychological experience of prison is probably tougher on the person who never actually violated the elements of a statute. So retributivists could argue that the high level of confidence they require for factual guilt need not apply to all prerequisites of retributive punishment: some unjustified punishments are worse than others.

---

101. *See* Stephen J. Morse, *Avoiding Irrational NeuroLaw Exuberance: A Plea for Neuromodesty*, 62 MERCER L. REV. 837, 845 (2011) ("[F]ree will in the strong sense is not . . . a criterion at the doctrinal level. . . . Neuroscience may help shed light on folk-psychological excusing conditions, such as automatism or legal insanity, but the truth of determinism is not an excusing condition."). *But cf.* Adam J. Kolber, *Will There Be a Neurolaw Revolution?*, 89 IND. L.J. 807, 822 (2014) (conceding the surface appearance of the doctrine, but suggesting that the intent of the law's crafters could plausibly lead to an interpretation of criminal law that requires metaphysical free will). *See generally* Kolber, *Free Will as a Matter of Law*, *supra* note 34.

102. 99.5%^9=95.6%; *see* Campbell, *supra* note 7.

I don't think retributivists can easily make this claim. Unjustified punishment is unjustified, and if imposing it is morally impermissible, the extent of the impermissibility may be beside the point. True, we are discussing cases of acting under uncertainty. So retributivists could argue that they can tolerate a greater risk of acting impermissibly when expected harms are lower. But in my example, unjustified punishment varied in seriousness based on the amount offenders would likely suffer emotionally. Such considerations are usually ignored by retributivists at sentencing.[103] Unless retributivists are willing to consider the subjective experience of punishment more generally, this option seems off the table. Even if they did consider subjective experience more carefully, it wouldn't show that the *cause* of undeserved punishment matters in any deep way. It would simply show that more undeserved punishment is worse than less undeserved punishment.

The real problem with the error parity principle relates to considerations that are illicit for pure retributivists. Namely, we are inclined to tolerate a greater risk of error about free will than about guilt, for example, because an error about the former may still lead to the deterrence and incapacitation of dangerous people, while errors that put an innocent person in prison weaken the law's deterrent message and do nothing to incapacitate the dangerous. These, of course, are traditional consequentialist concerns unavailable to pure retributivists. (In Part IV, however, I suggest how "epistemic hybrid theories" might improve matters for retributivists.)

Even if retributivists can reject perfect-error parity, I doubt they can explain the large gulf separating their treatment of errors of guilt from errors of deservingness more generally. True, examination of guilt lends itself reasonably well to adjudication in court. But it's not clear why we devote so much of our resources to examining guilt in the face of other uncertainty. If our justificatory standard of proof is in the 80% range, the attention we place on guilt will likely shift our total uncertainty by just a few percentage points. In other words, a realistic picture of retributivist uncertainty seems to make the issue of guilt relatively unimportant in the face of so many other sources of uncertainty.

Perhaps retributivists could reply as follows: "It's not fair to compare the very high BARD standard to the standard we need for moral justification. The BARD standard is high, in part, to compensate for moral uncertainty. Precisely because there is so much uncertainty when it comes to criminal justice, we ought to be awfully certain that the accused is factually guilty." Hence, they could argue, the BARD standard should be viewed, in part, as a corrective to the problem of moral uncertainty. Imagine you lost your keys in one of nine rooms but can only search one of them: you will search that one room *very* carefully.

This response has some plausibility. Suppose we considered all the pertinent uncertainty surrounding justification *except for the likelihood of guilt* and reached a total level of confidence of 75%. Choosing a very high 99% confidence level for guilt would only reduce our total level of confidence to 74%.[104] By contrast, a comparatively low standard of factual guilt, like 80%, would leave

103. *See* Kolber, *The Subjective Experience of Punishment*, *supra* note 75, at 187–96.
104. 75*.99=74.25%; *see* Campbell, *supra* note 7.

us with a 60% overall confidence. And if we take the justificatory standard of proof to be somewhere between 60% and 74%, the choice of a BARD standard really could make a difference in justifying punishment—it wouldn't always wash out.

But it would often wash out. And matters get worse for retributivists when we consider how a real-world legal system could possibly reflect a sufficiently high justificatory standard of proof. In easily imaginable criminal justice systems, tasks will be spread among different actors. For example, in the United States, one might say that questions about metaphysical free will and the wrongfulness of statutory conduct are decided primarily by legislators while the legitimacy of police conduct, court procedures, and sentencing is decided largely by judges, and issues of guilt that make it to trial are determined largely by jurors.[105] We would plausibly have to establish minimum required levels of confidence for all or many of our key propositions because they are often decided by different people or are otherwise compartmentalized. And by setting minimum requirements, we cannot let high levels of certainty in some areas bolster low levels in others. Rather, we need to set one particular level in each area. That's certainly what we do with guilt. We don't *overtly* change the standard of proof for guilt simply because we are very confident the alleged offense addresses highly immoral conduct.

But setting minimum requirements as to each proposition would quickly lead to a low confidence in justification overall. If retributivists required each proposition to be true with 90% confidence, they would be willing to punish with only 39% confidence in the overall conjunction.[106] That seems clearly too low, as they would be more confident punishment is unjustified than justified. While many retributivists readily concede that our current practices are imperfect, it's hard to even imagine a system of retributive justice that could reduce error sufficiently.

Overall, I claim that pure retributivists do not have obvious means of distinguishing errors of guilt from other errors of deservingness in any substantial way. While my argument does not require parity in treatment between errors as to guilt and other errors bearing on justification, I argue that retributivists must have a high degree of confidence not only that those they punish violated the law but that the moral grounds for retributively punishing are correct as well. Even if the required confidence in deservingness is less than the required confidence associated with the BARD standard, the pure retributivist justification of

105. Of course, these lines are hardly clear cut. For example, the jury's power to nullify verdicts potentially allows jurors to weigh in on whether statutorily prohibited conduct was wrongful. *See, e.g.*, Rachel E. Barkow, *Recharging the Jury: The Criminal Jury's Constitutional Role in an Era of Mandatory Sentencing*, 152 U. PA. L. REV. 33, 59 (2003) ("The jury possesses the power to elaborate the governing norms underlying criminal laws from the perspective of the community and its sense of moral blameworthiness."); Michael T. Cahill, *Punishment Decisions at Conviction: Recognizing the Jury as Fault-Finder*, 2005 U. CHI. LEGAL F. 91, 100. Also, we might think that jurors weigh in on questions of wrongfulness to the extent that statutory elements represent mixed questions of law and fact. *Cf.* Youngjae Lee, *Reasonable Doubt and Moral Elements*, 105 J. CRIM. L. & CRIMINOLOGY 1, 24 (2015) (arguing that "the beyond a reasonable doubt requirement should not apply to moral or normative questions").

106. .9^9= 39%; *see* Campbell, *supra* note 7.

punishment is likely to be in grave danger once reasonable numbers are attached to propositions (1)–(9).

B. Typical Consequentialism

Consequentialists hold that the right action (or inaction) in some particular situation depends on the consequences of that action (or inaction).[107] In the punishment context, consequentialists seek to maximize good consequences of punishment such as crime deterrence, incapacitation of dangerous people, and rehabilitation of offenders relative to bad consequences such as the suffering punishment imposes on offenders and their friends and family, punishment's financial costs (especially those associated with incarceration), and the loss of pleasurable or productive opportunities offenders would have had if they weren't punished.[108]

Consequentialists must manage enormous uncertainty about empirical facts. For example, they should try to set punishments at levels that will minimize the financial and emotional costs of incarcerating someone relative to the gains from deterring crime, incapacitating dangerous offenders, and rehabilitating them. It is extremely difficult to do so with the limited real-world information currently available.

The task of resolving consequentialist uncertainty about empirical facts is a gargantuan one, and if consequentialism were in a contest with retributivism as to which has more of such uncertainty, consequentialism would almost certainly win.[109] There is, indeed, a powerful "epistemic challenge to consequentialism," it's simply not the one I focus on here.

The reason I focus on the challenge to retributivism is that the path to resolving consequentialism's *empirical* uncertainty is much clearer than the path to resolving retributivism's *moral* uncertainty. At least in principle, we know how to gather evidence and set up experiments to estimate how punishment policies will affect deterrence, incapacitation, and rehabilitation. The method to resolve age-old debates about free will and proportionality, however, is highly disputed and has been for centuries.

C. Consequentialism Is Subject to Moral Uncertainty as Well

Of course, consequentialists must wrestle with moral uncertainty as well. They must decide which consequences hold intrinsic value or disvalue. Some consequentialists focus on increasing pleasure (or other positive mental states) relative to pain (or other negative mental states).[110] Others find intrinsic value in

107. *Consequentialism*, STAN. ENCYCLOPEDIA PHIL., https://plato.stanford.edu/entries/consequentialism/ (last substantively updated Oct. 22, 2015).

108. *See* TEN, *supra* note 8, at 7–8.

109. When assessing empirical facts bearing on the justification of punishing someone, retributivists focus almost exclusively on historical facts surrounding the crime's commission. They require little, if any, of the vast information consequentialists need to make predictions. The "whole-life" view of retributivism, however, might be impractically, perhaps impossibly, mired in factual uncertainty. *See* discussion *supra* Subsection II.B.3.

110. WILL KYMLICKA, CONTEMPORARY POLITICAL PHILOSOPHY 13 (2d ed. 2002).

the satisfaction of desires relative to the frustration of desires.[111] Still others focus not on what people actually desire but on what they ought to desire.[112] And some focus on a non-uniform list of goods that hold intrinsic value.[113]

But while there are important debates about the fundamental axiological commitments of consequentialism, these debates are probably not important for present purposes.[114] Incarceration itself is bad whether framed as causing emotional suffering, frustrating offenders' life goals, depriving them of liberty, or costing the state resources that could be better spent elsewhere. Similarly, reducing crime is valuable whether framed as promoting positive experiences, satisfying desires, or increasing liberty and economic productivity.

True, decisions about what holds intrinsic value *can* have dramatic policy implications. If the harms of punishment are understood as bad subjective experiences, we need to measure and account for subjective experiences differently than we would if the only harms of punishment were understood as objective deprivations of liberty. Consequentialists must also trade off items that at least superficially appear incommensurable, like the financial costs of imprisonment and the value gained by reducing crime. They must trade off the harms of separating children from their incarcerated parents with the benefits of deterring crimes by those afraid of being separated from their children. Even with all the empirical facts in the world, such tradeoffs will involve controversial value judgments held with varying degrees of certainty. Still, while consequentialists debate these matters with each other, they have no obvious, dramatic effects on the general debate between consequentialists and retributivists about the justification of punishment.

There are, however, at least two fundamental features of consequentialism that do raise important issues of moral uncertainty for present purposes. First, consequentialists believe that the value or disvalue of some consequence is independent of the nature of the action or inaction that created it. So the value of some consequence does not depend on whether it was brought about intentionally or only as a foreseen consequence. Similarly, the value of some consequence does not depend on whether an actor caused the consequence or merely allowed it to occur.

This consequentialist tenet contrasts to some extent with conventional moral thinking. To take a classic example, consider a wartime fighter pilot who bombs an enemy ammunition facility knowing (but not intending) that he will also kill two civilians in the process. Many would deem his actions permissible while deeming impermissible the conduct of a similar pilot who, seeking just as much military advantage as the first pilot, bombs a site *intending* to kill two civilians as a means of damaging the opposing side's morale and ending the war

111. *Id.* at 14–16.

112. *Id.* at 16–20.

113. *Cf. id.* at 14.

114. One notable exception concerns those retributive-consequentialists who believe that deserved suffering (or punishment) should factor into consequentialist calculations as being intrinsically valuable, as discussed in MOORE, *supra* note 37, at 155–59 and Michael T. Cahill, *Retributive Justice in the Real World*, 85 WASH. U. L. REV. 815, 833–35 (2007). Such views will systematically advocate more punishment than views that count the harms of suffering and punishment as bad consequences.

sooner. Under this conventional view, it is sometimes permissible to knowingly kill civilians but is never permissible to kill them intentionally: We are responsible for harms we intend in a different way than we are for harms we merely fail to prevent.

By contrast, consequentialists will generally view the death of the two civilians as equally bad no matter whether each was intended or merely foreseen.[115] This feature of consequentialism is especially important in the punishment context. When comparing the harm of intentionally confining a convicted kidnapper to prison to the risk of harm that kidnapper might cause if unconfined, consequentialists focus on relative amounts of harm. They do not especially worry that incarceration is an *intentional* confinement, whereas leaving the kidnapper free to confine others might be described as a merely foreseen (or foreseeable) *allowing* of the confinement of others.

So, if we are deciding whether or not to institute a death penalty in some state, empirical research might show that the penalty would deter the killing of several people per year for every execution that occurs.[116] Holding all else equal, the typical consequentialist might support the death penalty on the ground that it leads to fewer deaths overall. The conclusion depends on a matter about which reasonable consequentialists should have some moral uncertainty—whether deaths brought about intentionally are as disvaluable as deaths brought about merely as foreseen consequences of other policies.[117]

The second fundamental consequentialist proposition that introduces moral uncertainty is that the harms of punishing one person can be justified by benefits to others. Some would consider this an unacceptable "using" of people. Indeed, those steeped in the Kantian tradition accuse consequentialists of treating offenders merely as means whereas they should be treated as ends in themselves.[118] So, in the prior example, the offender executed by the state is losing his life not because the loss benefits *him* in any way but simply because it benefits others by preventing crimes.

Nigel Walker responded to worries about using people merely as a means by noting how often we do it:

> Urban redevelopment evicts families from homes. Airports sacrifice the peace of a few for the convenience of many. Sufferers from some communicable diseases undergo irksome restrictions in the interests of public health. Soldiers are enlisted—not always voluntarily—to risk life and limb for country. The list of examples could be much longer. The short point is that anyone who condemns deterrent or precautionary sentences on the ground that they harm offenders for the sake of others must either condemn

115. Consequentialists can still distinguish between intentional and foreseen actions in certain respects. They might legally forbid intentional killings of civilians in ways they do not forbid foreseen killings. Their reasoning, however, would focus on yet further consequences, like how many people are likely to die in total under one legal regime relative to another (for example, the two legal regimes may have different effects on behavior going forward).

116. *See* Cass R. Sunstein & Adrian Vermeule, *Is Capital Punishment Morally Required? Acts, Omissions, and Life-Life Tradeoffs*, 58 STAN. L. REV. 703, 706 (2005).

117. *See* David Enoch, *Intending, Foreseeing, and the State*, 13 LEGAL THEORY 1 (2007) (challenging the moral significance of the intending/foreseeing distinction in the context of state action).

118. *Id.* at 8–9.

> many of the things that are done to the innocent, or explain why only the guilty should be immune from being "treated as means."[119]

Still, Walker considered the Kantian objection more powerful when especially severe harms such as "maiming, castration, disfigurement, and torture" were imposed on the one to benefit the many,[120] and one might argue that incarceration should be treated as one of these serious harms.

It is not obvious, however, why we would permit the use of people merely as a means for modest harms but not for severe harms. One might expect a slightly different rule: using a person would have to grow substantially more important as the harm to the person being used grows. So framed, the rule seems more congenial to consequentialists. Nevertheless, the intending-foreseeing distinction and the means-ends distinction are ones that reasonable people disagree about. I don't find these concerns insurmountable for consequentialists by any means. Still, consequentialists ought to wrestle with some moral uncertainty that their views neglect important facets of moral life.

D. Why Consequentialism Is Less Vulnerable to the Epistemic Challenge

While both retributivists and consequentialists must wrestle with moral uncertainty, there is an important asymmetry between the two. When retributivists decide whether or not to punish, they trade off the good of punishing the deserving against the harm of punishing the undeserving. And for most retributivists, it seems, it is far worse to punish a person who does not deserve it than to fail to punish a person who does. Thus, retributivists place substantial weight on the scale in favor of not punishing. As I argued in Part II, reasonable retributivists will likely find the weight so substantial that they cannot justify punishment in the face of moral uncertainty (or must dramatically reconsider their advocacy of Blackstone-like ratios).

1. The Consequentialist See-Saw

By contrast, for consequentialists, the moral uncertainty of punishment is counterbalanced by the moral uncertainty of *failing* to punish. While consequentialists should worry to some extent about disregarding the difference between intentional and foreseen acts, *failing* to treat foreseen harms as seriously as intended harms will sometimes lead us to allow unnecessary harm. So just as there is moral risk when *ignoring* the difference between intended and foreseen actions, there is moral risk when honoring the distinction as well. Similarly, *failing* to use people as a means can also lead to preventable harm. So even if fears about actively causing harm and using people solely as a means ought to trouble consequentialists a bit, nothing compels them to be guided in any substantial way by those fears (though I will explain in Part IV what could happen if they were).

---

119. NIGEL WALKER, WHY PUNISH? 54 (1991).
120. *Id.*

Hence, consequentialists are less vulnerable to epistemic moral risks than retributivists are. For consequentialists, the risk of allowing unnecessary harm or death by foregoing punishment can be much more serious than the risk of an immoral using or an immoral failure to take intentionality into account. While consequentialists ought to have some doubts about the principles that underlie their views, so long as their overall confidence in the consequentialist scheme is strong enough, moral uncertainty need not radically shift their behavior.

2. *Consequentialist Standards of Proof*

Retributivists are also more vulnerable to the epistemic challenge than consequentialists because consequentialists are usually not committed to the values underlying the BARD standard in the way that most retributivists are.[121] For a good consequentialist, the legal standard by which we ought to convict people is itself derived from a consequentialist calculation. When BARD is so interpreted, it cannot possibly conflict with consequentialism. By contrast, I have argued that the retributivist commitment to the values underlying the BARD standard should lead most retributivists not to punish. While consequentialists also face moral uncertainty, it is counterbalanced by reasonable fears that failing to punish will allow unnecessary harms to occur.

As for the consequentialist justificatory standard of proof, most consequentialists seek to maximize good consequences relative to bad. From one perspective, such consequentialists set an impossibly high justificatory standard: they are advised to pick the *single* best action among millions of possibilities. But given limits on human memory and processing, we can often decide among just a few realistic options, and consequentialists will deem it more rational to pick options with better expected consequences than those with worse expected consequences. Since both punishing and failing to punish risk serious harms, neither is necessarily safer than the other.

3. *Reply to Claim of Double Standard*

Responding to an earlier draft of this Article, Larry Solum challenged the asymmetry of my epistemic argument. He claimed: (1) the truth of consequentialism is also uncertain and many philosophers reject it; (2) retributivists also recognize the downside of failing to punish; and (3) there are reasons to doubt the efficacy of deterrence, incapacitation, and rehabilitation and our ability to calibrate punishments to achieve those aims.[122]

Solum's first claim (that the truth of consequentialism is also in doubt) misunderstands my methodology. I do *not* make a ground-floor overall assessment of the likely truth of retributivism relative to consequentialism. Doing so would largely just rehash the age-old debate between them. Rather, I

121. Walen, *supra* note 9, at 401 (stating that consequentialism may "call for a surprisingly low [standard of proof] in criminal cases.").

122. Solum, *supra* note 62; *see also* Larry Solum, *More from Kolber on "Punishment and Moral Risk,"* LEGAL THEORY BLOG (Mar. 2, 2017), http://lsolum.typepad.com/legaltheory/2017/03/more-from-kolber-on-punishment-and-moral-risk.html.

acknowledge that people come to the theorists' table with existing views, and I examine whether reasonable theorists of both stripes can accommodate moral risk. So I readily grant retributivists and consequentialists relatively high levels of confidence in core tenets of their respective views.

As for Solum's second claim, it is true that many retributivists recognize a downside to failing to punish. My claim is that traditional deontological principles coupled with the values underlying Blackstone-like ratios typically lead retributivists to put a heavy thumb on the scale against punishing in the face of uncertainty. By contrast, once consequentialists overcome the moral uncertainty required to disavow certain core deontological views, additional moral uncertainty will tend to have the see-saw effect I described: failing to punish can be as morally risky as actually punishing.

Finally, Solum challenges the efficacy of consequentialist aims and our ability to calibrate punishment to achieve them. But *contra* Solum, consequentialists need not believe that deterrence, incapacitation, and rehabilitation are particularly efficacious. They need only believe they are the sorts of good things that *can* justify punishment. If they don't work well, then we are not justified in inflicting substantial punishment to achieve them. But no one plausibly doubts that punishment sometimes deters and prevents crime; the debate concerns how effective punishment is at the margins. Solum is surely right that consequentialists must cope with substantial *empirical uncertainty*. But as I've noted, they are *morally* uncertain about both punishing and failing to punish, so moral uncertainty need not make them as cautious about punishment as it makes retributivists.[123]

IV. PORTFOLIOS OF MORAL AND LEGAL BELIEFS

A. The Core Conflict Between Retributivism and Consequentialism

Our intuitions about punishment often have both retributivist and consequentialist components.[124] Hence many have tried to combine these sentiments into coherent hybrid theories.[125] In my view, these attempts have been unsuccessful. Retributivism and consequentialism are deeply conflicted. In their typical elaborations, each theory *excludes* the considerations of the other. Pure retributivists say we must punish proportionally without regard to deterrence, incapacitation, and rehabilitation.[126] Considering these non-desert factors, many retributivists believe, would unacceptably *use* prisoners merely as a means to benefit others. Consequentialists, by contrast, say we should focus precisely on those instrumental considerations without regard to desert. The past is in the

123. Adam J. Kolber, *Reply to Solum on "Punishment and Moral Risk,"* NEUROETHICS & LAW BLOG (Jan. 11, 2017), http://kolber.typepad.com/ethics_law_blog/2017/01/reply-to-solum-on-punishment-and-moral-risk.html.

124. *See, e.g.*, Kevin M. Carlsmith et al., *Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment*, 83 J. PERSONALITY & SOC. PSYCHOL. 284, 295–97 (2002); Kevin M. Carlsmith, *On Justifying Punishment: The Discrepancy Between Words and Actions*, 21 SOC. JUST. RES. 119, 127, 133–36 (2008).

125. *See, e.g.*, Ken Levy, *Why Retributivism Needs Consequentialism: The Rightful Place of Revenge in the Criminal Justice System*, 66 RUTGERS L. REV. 629, 635 n.18 (2014).

126. *Id.* at 645–46.

past, and so giving independent weight to desert distracts us from the consequences we really ought to care about.

True, retributivist and consequentialist considerations are often positively correlated; those who are most culpable are also likely to be the most dangerous and most in need of incapacitation and rehabilitation. But sometimes retributivist and consequentialist considerations are negatively correlated or have no obvious relationship. Suppose two offenders commit equally blameworthy rapes. Prior to conviction and sentencing, just one of these offenders develops testicular cancer that requires major surgery and drug treatments that, let us assume, dramatically reduce his ability to commit another sexual offense and his interest in doing so.[127] Since both offenders are equally blameworthy, retributivists would recommend equal punishments. Consequentialists, however, are likely to recommend less punishment for the one who is less likely to reoffend.

And such conflicts can arise in many different ways. Retributivists might view having only recently reached the age of eighteen as mitigating an offender's culpability, while a consequentialist might believe the offender is in a high-risk age group and warrants extra punishment. Retributivists might view certain signs of mental illness as mitigating culpability while consequentialists might view those same signs as evidence of future dangerousness. The bottom line is that retributivism and consequentialism often do not play well together. While they can, of course, be strung together, the challenge is to combine them in a principled way that does more than just roughly satisfy our pre-reflective punishment intuitions.

B. Traditional Hybrid Theories

There have been many attempts to bridge the gap between retributivism and consequentialism. I'll mention two leading contenders.

1. Limiting Retributivism

Norval Morris influentially defended a "limiting retributivist" view in which "desert is not a *defining* principle, but is rather a *limiting* principle."[128] In his view, "the concept of a just desert properly limits the maximum and the minimum of the sentence that may be imposed, but does not give us any more fine-tuning to the appropriate sentence than that."[129] Morris would have us sentence between the upper and lower bands using consequentialist considerations.[130]

Limiting retributivism has certainly been a popular compromise of retributivist and consequentialist approaches. But is there any principled basis for this particular blend? Consider how Morris looked approvingly at a case in which a

127. See a similar example in MICHAEL MOORE, LAW AND PSYCHIATRY 241–42 (1984).

128. Norval Morris, *Desert as a Limiting Principle*, *in* PRINCIPLED SENTENCING: READINGS ON THEORY & POLICY 180, 180 (Andrew von Hirsch & Andrew Ashworth eds., 2d ed. 1998); *see also* E. THOMAS SULLIVAN & RICHARD S. FRASE, PROPORTIONALITY PRINCIPLES IN AMERICAN LAW: CONTROLLING EXCESSIVE GOVERNMENT ACTIONS 161–62 (2009).

129. Morris, *supra* note 128, at 180.

130. *Id.* at 184.

judge sentenced each of nine white men to four-year sentences for what appear to be racially motivated attacks in London in 1958:

> This sentence was at least double the sentence normally imposed for their offenses, and was stated by the sentencing judge to be in excess of his normal sentence for such offenses, but it was within the legislatively prescribed maximum for those offenses. It was imposed expressly as an exemplary punishment, to capture public attention and to deter such behavior by a dramatic punishment. It needs no refined analysis to demonstrate that these nine offenders were selected for *unequal* treatment before the law. . . . I am arguing that if the increased penalty is within the legislatively prescribed range, then any supposed principle of equality does not prevent such a sentence from being in the appropriate case a just punishment.[131]

Morris's approach sacrifices retributivist principles by using these offenders as tools to further larger policy goals. Yet it also abandons consequentialism when the legislative boundaries are reached. Why think that legislatures set morally appropriate boundaries? More importantly, why should decisions at those boundaries dramatically shift our *method* of analysis?

Even if legislatures do set appropriate boundaries, why do the boundaries kick in suddenly and sharply?[132] Just shy of the boundary, retributivism matters not a bit. But at the boundary, no consequentialist consideration trumps the retributivist limits. If retributivism and consequentialism both offered important benefits, balancing them would not likely involve such sharp on-and-off switches. Hence, Morris would likely have no good answer to crime victims predictably harmed by rigid adherence to the proportionality boundaries. He could only offer the standard retributivist answer that proportionality limits consequentialist goals, even when enforcing those limits causes avoidable harm and even when we have no convincing theory as to how to calculate the boundaries of proportional punishment.

2. *The Separate Question Approach*

Another celebrated approach to punishment hybridity, offered most notably by H.L.A. Hart, requires us to separate our reasons for creating punishment institutions from our reasons for distributing particular punishments.[133] According to Hart, "it is perfectly consistent to assert *both* that the General Justifying Aim of the practice of punishment is its beneficial consequences *and* that the pursuit of this General Aim should be qualified or restricted out of deference to principles of Distribution which require that punishment should only be of an offender for an offence."[134] So one way to satisfy our mixed retributivist and consequentialist intuitions is to create laws and punishment institutions that seek to deter crime, incapacitate dangerous people, and rehabilitate offenders but to

131. *Id.* at 180–81.

132. *See generally* Adam J. Kolber, *Smooth and Bumpy Laws*, 102 CALIF. L. REV. 655 (2014); Adam J. Kolber, *The Bumpiness of Criminal Law*, 67 ALA. L. REV. 855 (2016); Adam J. Kolber, *Smoothing Vague Laws*, *in* VAGUENESS AND LAW: PHILOSOPHICAL AND LEGAL PERSPECTIVES 275 (Geert Keil & Ralf Poscher eds., 2016).

133. H.L.A. HART, PUNISHMENT AND RESPONSIBILITY 3–4 (1968).

134. *Id.* at 9.

limit those goals by the requirement that offenders be punished according to their desert.

I fail to see, however, how separating these questions helps us. If you're inclined to distribute punishment retributively, then this goal would seem to be part of your motivation for setting up punishment institutions in the first place. Are we to believe that there is some important order to these questions, as though the justification of punishment could depend on whether we first focus on institutions or first focus on how those institutions will subsequently achieve the offender-level goals we seek?

Hart wrote that "in relation to any social institution, after stating what general aim or value its maintenance fosters we should enquire whether there are any and if so what principles limiting the unqualified pursuit of that aim or value."[135] But it's not at all clear how to distinguish a general motivation for a practice from limits on it. Consequentialists seek to promote deterrence, but the deterrence has to be efficient. It has to be such that additional deterrence would lead to worse consequences overall. So could Hart's methodology support a purely consequentialist approach if we say that the general motivation for punishment is deterrence but, in allocating punishment, we must limit ourselves only to the amount of deterrence that leads to the best consequences?

Alternatively, if we can limit punishment to the amount promoting the "net best consequences" as part of punishment's general justifying aim, then it seems that some limiting principles can already be built into the general justifying aim. Why not put retributive limitations into the general justifying aim? Surely the nature of who you are punishing and how much they should be punished serve as part of the motivation for creating punishment institutions.

Furthermore, suppose we're constructing some aspect of the institution of punishment but only expect it to apply to a small number people. For example, perhaps we're setting up rules that only to a country's leader or top leadership. Are we focusing on an institution or the distribution of punishment? And why would it matter if we are focusing on a narrow portion of the population or a large portion? It's true that one might purport to make the general justifying aim of institutions consequentialist but distribute punishment in accordance with desert. The problem is explaining how asking separate questions helps to bridge the gap between retributivism and consequentialism that Hart himself described as "partly conflicting"[136] and "partly discrepant."[137]

Doug Husak correctly criticized the outsized influence Hart's separate-questions approach has had on punishment theory.[138] Perhaps the least criticism goes to Hart himself who self-consciously labeled the pertinent chapter of his book a "Prolegomenon to the Principles of Punishment" that was intended only

135. *Id.* at 10.

136. *Id.* at 1.

137. *Id.* at 10.

138. Douglas Husak, *A Framework for Punishment: What is the Insight of Hart's 'Prolegomenon'?*, *in* HART ON RESPONSIBILITY 91, 93, 107–08 (Christopher Pulman ed., 2014). For additional criticism of Hart's hybrid theory, see WHITLEY R.P. KAUFMAN, HONOR AND REVENGE: A THEORY OF PUNISHMENT 85 (2013) ("Hart fails to prove there is a fundamental distinction within the institution of punishment between two different 'questions,' the general aim and the limiting principle.").

to "develop[] this sense of the complexity of punishment."[139] But the distinction Hart set out to draw is so unclear that perhaps it should be no surprise that other theorists were drawn to it. They can find in it whatever they like; such is the strength of the urge to combine the two conflicting theories of retributivism and consequentialism.

C. *Epistemic Hybrid Theories*

Theorists tend to pigeonhole themselves into discrete theories. Doing so hides the fact that we hold our views with varying levels of confidence. More precise short-hand descriptions of our theoretical inclinations might be "70% retributivist, 30% consequentialist" or "80% consequentialist, 10% retributivist, 10% abolitionist." We hold not just beliefs but portfolios of beliefs.

By addressing moral uncertainty with a portfolio of beliefs, we may identify hybrid theories with less internal conflict than current contenders. These "epistemic hybrids" may better capture people's actual punishment intuitions and even offer normative guidance. My comments will be very speculative primarily because there are multiple ways of modeling decisions under uncertainty,[140] some of which may be better grounded normatively or descriptively than others. I turn now to three of many possible approaches that I offer simply to open up new avenues of exploration.

1. *Value-Discounted Consequentialism*

Suppose an otherwise hardcore consequentialist is uncertain as to one particular detail: Should we treat the deserved suffering of guilty offenders as having negative value (along with most consequentialists) or positive value (along with essentially all retributivists)? Upon reflection, the consequentialist charts his own path and decides that he is 80% confident that the deserved suffering of guilty offenders has disvalue and 20% confident that it has neither positive nor negative value.

When he performs his consequentialist calculations to maximize expected value, in addition to discounting based on factual uncertainty, he also discounts the suffering of offenders based on his moral uncertainty. As a result, he sometimes reaches conclusions that differ from consequentialists who lack his moral uncertainty. All else being equal, his punishments will seem more punitive relative to other consequentialists.

We might say that our imagined, morally uncertain consequentialist holds a hybrid view. But note that he doesn't straightforwardly disagree with other consequentialists about his moral beliefs. He will assent to the proposition that the deserved suffering of guilty offenders has disvalue. He simply has less confidence in the proposition than some others.

---

139. HART, *supra* note 133, at 1, 3.

140. For example, what exactly do we hold with varying degrees of confidence: intuitions? beliefs? theories? collections of theories? And how do we compare choices across theories when the theories themselves disagree about what ought to matter to us?

Ordinarily, good consequentialists will discount future expectations by their best estimates of the likelihood those events will occur. I'm suggesting that reasonable but morally uncertain consequentialists may also discount the value of consequences they expect to occur by their confidence in the value of those consequences. If, for example, these "value-discounted consequentialists" are 90% confident they can make some consequence occur and 90% confident the consequence has ten units of positive value, then the expected value of trying to make the consequence occur is 8.1 units (=.9 * .9 * 10 units). By contrast, a consequentialist who did not take moral uncertainty into account would only discount his ability to make the consequence occur and so would value the course of action at 9 units (=.9 * 10 units). That means value-discounted consequentialists will sometimes reach different conclusions about how to proceed in the face of uncertainty than conventional consequentialists.

2. *Means-Principle-Attentive Consequentialism*

According to traditional consequentialists, the suffering of prisoners is disvaluable. If punishing a prisoner will create 100 units of disvalue (and assume that's the only relevant disvalue), punishment will not be justified unless causing that disvalue leads to at least a 100-unit increase in value across society more generally.

Now imagine a consequentialist who—like others—believes that it is okay to use people merely as a means for the benefit of others. He has, however, somewhat more serious doubts about the proposition than most. Assume that our imagined consequentialist is 90% confident that it is permissible to use people merely as a means. He engages in standard consequentialist analysis and generally suppresses his uncertainty about using people because he worries more about failing to maximize value than about using people.

When the potential wrongfulness of using someone as a means gets especially severe, however, perhaps when it comes to taking a life solely to benefit others, his inclinations shift. Though he believes there is only a 10% chance that using people as a means is impermissible, that chance is too worrisome when combined with his perception that *if* using someone's life merely as a means is impermissible, then it's a very bad thing to do. It's not that he thinks it's wrong to use people merely as a means. Rather, he thinks that even if there's only a relatively small chance that it's wrong to use people, the harm of doing so is so enormous that the small chance is enough to make him refrain.

Our imagined epistemic-hybrid consequentialist is not suffering from any obvious contradiction. Unlike limiting retributivists or separate-question hybrid theorists, this "means-principle-attentive consequentialist" is not simply stacking two conflicting theories on top of each other. Rather, he is coping with the fact that matters of morals are difficult to assess and that we must consider both our confidence in various propositions and the relative harms of making moral errors.

True, we can challenge the levels of confidence and moral values that this means-principle-attentive consequentialist assigns. Rational debate does not end

when someone comes up with these numbers. But at decision time, ignoring uncertainty is irrational (and perhaps even immoral) when the severity of a moral mistake is high enough.

3. *Consequentialism-Backstopped Retributivism*

In Part II, I argued that *reasonable* retributivists cannot rely on pure retributivism to justify punishment. But just as consequentialists may hold epistemic hybrid views, so may retributivists. Indeed, by using a "consequentialism-backstopped retributivism," they can avoid many of the concerns I raised. Consequentialism-backstopped retributivism importantly depends on what retributivists believe about the world should their retributivist justifications be mistaken.

Suppose a retributivist is rather confident that people can be morally blameworthy and deserve state-inflicted proportional punishment as a result. Still, he worries he may be wrong. Assume he is 80% confident in the retributive view, 19% confident in the consequentialist view, and 1% confident in the view that punishment should be abolished. In this case, our retributivist is quite confident that punishment is justified, but his 80% confidence in retributivism might be insufficient to justify punishment of some actual offender, given his views about the relative value of punishing the deserving versus failing to punish the undeserving.

So he asks himself: "What if I'm wrong about retributivism? What if there is no such thing as free will? And what if proportionality is bogus? And what if I'm unnecessarily afraid to use people as a means?" When he examines his beliefs, he realizes that even if his retributivist views are wrong, he has a consequentialist backstop: "At least I'll be preventing crime and incapacitating a dangerous person," he reasons. It's not that he thinks such factors ought to matter with respect to punishment. He is fairly certain they should not. But when his retributivist justification is pushed to the edges, his consequentialism-backstopped justification can make a real difference: Pure retributivism might be too risky for him, but with a consequentialist hedge, he can proceed to punish. And perhaps his epistemic hybrid theory only allows him to dispense a quantity of punishment consistent with both retributivism and consequentialism. In other words, he might achieve results like those of limiting retributivists without their theoretical contradictions.

V. CONCLUSION

Judges are human, and like the rest of us, they must wrestle with moral and legal uncertainty. Legal opinions would be much more transparent if judges directly addressed their uncertainty and explained its influence on their deliberations. Judges rarely do so, of course. Courts like to seem omniscient. Perhaps such intellectual arrogance serves important societal functions: Maybe we're supposed to be confident that justice was done, not that we've achieved the highest likelihood that justice was done.

Scholars don't have the same excuse. We don't decide cases. We are supposed to be committed to analytical rigor and transparency, values that are hard

to square with the near absence of direct discussion of moral uncertainty in scholarly exchanges. Hence, I argue, all moral and legal theorists should attend more openly to moral uncertainty.

Failing to consider moral uncertainty is particularly troublesome for retributivists. Retributivists are trying to shoot an apple off the head of a person standing far away. If everything lines up perfectly, in theory, retributivists will only hit the apple. But once retributivists recognize the risk of error and their low tolerance for it, they ought to desist. Consequentialists, by contrast, should recognize that they take on some moral risk by ignoring fundamental deontological commitments. But such moral risk need not radically alter their punishment practices so long as they consider the harms of over-punishment and under-punishment as roughly similar in disvalue.

If retributivists addressed their moral uncertainty more straightforwardly, they might self-consciously adopt a consequentialist backstop to bolster retributivism in the face of moral uncertainty. Should such epistemic-hybrid retributivists be wrong about desert and the obligation to refrain from using people merely as a means, at least they will be accomplishing some good by punishing. Similarly, some consequentialists may find that they are sometimes unwilling to treat acts of punishment the same as failures to punish, even though pure consequentialism requires it when consequences are held equal. The lapse might be explained by consequentialists' residual doubts about their denial of traditional deontological principles.

Nothing I say here diminishes the project of reducing our uncertainty about morality through careful reflection and debate. But so long as some issues remain contested by thoughtful people, we will continue to have plausible doubts about moral issues. These doubts can help structure a portfolio of beliefs. Much more work must be done to develop and evaluate portfolios of beliefs, but they offer an avenue to better understand existing legal and moral debates and to find new ways of resolving them.

VI. APPENDIX: PORTFOLIOS OF BELIEFS MORE GENERALLY

There are many big debates in legal theory: Is tort law supposed to provide corrective justice or incentivize safety? Is contract law about fulfilling promises or promoting efficient economic activity? Should we interpret statutes based only on their plain meaning or consider the intentions of legislators? Should constitutions be understood in terms of their original meaning or can their meaning change over time? As with punishment theory, these questions can be addressed with portfolios of beliefs. A tort theorist might be "60% corrective justice-oriented, 40% deterrence-oriented," and a constitutional law theorist might be "50% textualist, 50% purposivist."

Just as one can hold shares of different companies in an investment portfolio, one can hold different beliefs in varying proportions in a portfolio of beliefs. And just as stocks in an investment portfolio interact in ways that can increase or decrease total risk, so too can the constituents of a portfolio of beliefs. Hence, I suggested earlier, retributivism alone might be impotent to punish but

capable of doing so with a consequentialist backstop. In other words, our backup beliefs should sometimes influence our overall policy preferences.

In criminal law, courts and legislatures frequently repeat their commitment to both retributivist and consequentialist goals[141]—even though they conflict. Similar patterns of conflict are glossed over by other legal doctrines. Often, no single theory adequately captures our intuitions. At least sometimes, our conflicting impulses can be explained by conscious or unconscious attempts to manage uncertainty. And at least sometimes, decision-making that reflects uncertainty will be superior to decision-making that blindly disregards it.

Consider some examples of how portfolios of beliefs might be deployed by philosophers and legal scholars to address troublesome puzzles in moral theory and tort law.

A. Threshold Deontology as a Portfolio of Beliefs

The debate between retributivist and consequentialist punishment theorists is rooted in an even more fundamental debate about the nature of morality. Deontologists believe that some acts are morally impermissible even if they would lead to net beneficial consequences.[142] For example, it is morally wrong to torture an innocent person, some deontologists hold, even if doing so would prevent two or three other innocent people from themselves being tortured.[143]

But what if the consequences of following a deontological rule get even worse? What if the only way to prevent the torture of thousands of innocents is to torture one innocent person? At some number of tortured innocents, call it N, many deontologists concede that the consequences win out. Describing himself as a "threshold deontologist," Michael Moore has prominently argued:

> It just is not true that one should allow a nuclear war rather than killing or torturing an innocent person. It is not even true that one should allow the destruction of a sizable city by a terrorist nuclear device rather than kill or torture an innocent person. To prevent such extraordinary harms extreme actions seem to me to be justified.[144]

There are many interesting puzzles and questions raised by threshold deontology.[145] But one noteworthy problem is that, like retributivism, deontology is fundamentally incompatible with the sort of cost-benefit analysis that characterizes consequentialism. Larry Alexander concisely made the point:

> [D]eontology and consequentialism are incommensurable because they are fundamentally opposed conceptions of what morality is about. One sees the individual as inviolate, an end in himself, and the opposite of a resource for the betterment of the world. The other sees the individual in exactly the opposite way. The threshold deontologist would have us believe that we switch from not being resources for others to being resources for others

141. *See, e.g.*, 18 U.S.C. § 3553(a)(2) (2012) (instructing courts to give sentences that "reflect the seriousness of the offense," "provide just punishment for the offense," and promote deterrence, incapacitation, and rehabilitation); CAL. PENAL CODE §1170 (West 2017); N.Y. PENAL LAW §1.05 (McKinney 2017).

142. Larry Alexander, *Deontology at the Threshold*, 37 SAN DIEGO L. REV. 893, 894 (2000).

143. *Id.*

144. MOORE, *supra* note 37, at 719.

145. Alexander, *supra* note 142, at 896–97.

> when N is reached. When N is looked at like that, however, it seems downright implausible that the moral universe is so constituted. There may be thresholds at which new phenomena emerge, but it is quite another thing to have thresholds at which things become their opposites.[146]

I cannot hope to settle the debate about threshold deontology here. I simply note that an epistemic hybrid of deontology and consequentialism can arise from a portfolio of beliefs. An "epistemic threshold deontologist" could have high confidence in deontology and low confidence in consequentialism. Most of the time, his deontological beliefs will rule the day. But when the risks of not adopting a consequentialist approach get especially high, he opts for consequentialist solutions. Such a portfolio of beliefs would lack the internal contradictions Alexander attributes to threshold deontology yet still offer behavioral advice consistent with it.

At least on some models of epistemic threshold deontology, we might even be able to calculate a precise value of N at which a particular epistemic threshold deontologist starts choosing in ways that appear consequentialist. But details aside, our imagined hybrid theorist lives life as a deontologist until the risk-adjusted evil of doing so exceeds its risk-adjusted virtue, a pattern that superficially looks like threshold deontology but arguably has a firmer theoretical basis.

B. Tort Law as a Portfolio of Beliefs

I will briefly mention three of many ways in which portfolios of beliefs could help us understand tort law. The first concerns the ultimate goal or goals of tort law. Some would say tort law should compensate wrongful injuries, some it should deter dangerous behavior, and some it should serve as a form of insurance. Some would pluralistically choose all or a subset of these goals.[147] Each approach advocates a different legal regime. Portfolio approaches yield still further options that have largely been unexplored. A person who is 70% confident tort law should solely concern compensation and 30% confident it should solely concern optimal deterrence will advocate different results than a pluralist who seeks tort awards of full compensation with modest upward and downward adjustments in the direction of optimal deterrence.

Second, portfolios of beliefs can enlighten tort procedures and standards of proof. Like criminal law theorists, tort theorists must deal with uncertainty about both facts and values, and the risk-weighted severity of mistakes may influence their views about requisite burdens of proof. Tort law usually uses the rather low preponderance of the evidence standard, but in any particular case, the moral risks of one party losing might be substantially graver than the risks of the other side losing. Asymmetric moral uncertainty should arguably affect the way we interpret rules of evidence and procedure and may help us better guide jury decisionmaking.

---

146. *Id.* at 912.

147. *See generally* Mark A. Geistfeld, *The Coherence of Compensation-Deterrence Theory in Tort Law*, 61 DEPAUL L. REV. 383, 383 (2012) (discussing theories that blend compensation and deterrence rationales).

Finally, there is a puzzle as to why tort law focuses so much on negligence and only allows strict liability in limited circumstances.[148] After all, we want to deter injuries even when they are non-negligent. Portfolios of beliefs suggest one answer: even though tort law is not generally thought to require fault, we may be reluctant to deprive people of their property rights when they faultlessly cause injury. To the extent that we have moral uncertainty about taking the property of faultless people, we may prefer a negligence standard that generally requires fault but allows strict liability in rare contexts—as indeed we do—when the consequences of limiting tort liability to negligence are particularly serious.

148. *Id.* at 392–93.