
HUMILITY, NOT DOUBT: A REPLY TO ADAM KOLBER

Mary Sigler*

TABLE OF CONTENTS

I. MORAL TRUTH AND MORAL HUMILITY	159
II. MORAL DISAGREEMENT AND MORAL TRUTH	162
III. FREE WILL AND MORAL RESPONSIBILITY	164
IV. PORTFOLIO OF BELIEFS	166
V. CONCLUSION.....	167

In *Punishment and Moral Risk*,¹ Professor Adam Kolber argues provocatively that reasonable retributivists lack the requisite certainty to justify the infliction of criminal punishment. Specifically, he maintains that the familiar “beyond a reasonable doubt” (“BARD”) standard for a criminal conviction sets a high bar for certainty that should also inform our confidence levels about the morality of punishment. To meet this “justificatory standard of proof,” according to Kolber, retributivists must have very high confidence in a series of supporting propositions that constitute the retributive justification—that people generally have free will; that the infliction of punishment by the state is an appropriate response to wrongdoing; and that we can realistically analyze a person’s background to determine what punishment is deserved.² In addition to these general propositions, Kolber contends, the retributivist must also be highly confident that a specific defendant’s conduct was wrongful; that law enforcement and judicial proceedings in the defendant’s case were generally just;

* Lincoln Professor of Law & Ethics, Arizona State University Sandra Day O’Connor College of Law.

1. Adam J. Kolber, *Punishment and Moral Risk*, 2018 U. ILL. L. REV. 487.

2. *Id.* at 489–90. For convenience, I have abbreviated Professor Kolber’s list while still aiming to capture the relevant propositions.

and that the defendant's punishment is proportional and not excessive relative to the benefits of inflicting it.³

Kolber's project serves as a sensible reminder about the need for moral humility in the domain of criminal justice. As Jeffrie Murphy has observed, "If we can overcome our comforting self-deceptions about our own wisdom and goodness, we will surely realize how limited are our cognitive powers to read the heart of another and how great is our potential for evil"⁴ But whereas Murphy sought to counter the tendency for arrogance and smugness to infect our institutions of punishment—a too-ready willingness to adopt an us-them mentality toward offenders—Kolber aims to undermine retributivism altogether by miring it in doubt and uncertainty. In what follows, I take issue with several aspects of Kolber's argument, focusing especially on the nature of moral truth and moral responsibility. In particular, I argue that Kolber has made a kind of category mistake by attempting to apply mathematical confidence levels to various moral propositions, that he misunderstands the significance of controversy in evaluating moral truth, and that he misjudges the relevance of the free-will determinism debate for moral responsibility. Finally, in response to Kolber's suggestion that retributivism might be saved from obsolescence by combining it with consequentialism in a "portfolio of beliefs,"⁵ I argue that this concedes too much to the consequentialist worldview. A duly humble retributivist need not succumb to the allure of Kolber's dubious algorithms.

I. MORAL TRUTH AND MORAL HUMILITY

To establish what retributivists are up against, Kolber outlines some basic propositions that he takes to be definitive of retributive theory. The list is noteworthy for, among other things, its mix of empirical and moral claims.⁶ Among the empirical claims are that individuals have free will and that a specific defendant committed the offending conduct. In both of these cases, the fact of the matter is, at least in theory, ascertainable. Put another way, it should be possible to prove (or at least falsify) these claims—to establish through observation and testing (*e.g.*, with advanced neuroimaging) the scientific fact of whether human beings as such possess the contra-causal capacity to choose their behav-

3. *Id.* at 490.

4. JEFFRIE G. MURPHY, PUNISHMENT AND THE MORAL EMOTIONS 204 (2012).

5. Kolber, *supra* note 1, at 491.

6. It is also noteworthy for including at least two propositions that are by no means essential to mainline retributivism—the need for "whole-life" character determinations and a retributive cost-benefit analysis. First, retributivists generally do not defend strong versions of character retributivism, which would require this kind of deep assessment of a person's nature before assigning punishment. *See, e.g.*, MICHAEL MOORE, PLACING BLAME 549 (1997) (noting that "only the choice theory captures" the requisite kind of moral responsibility); *see also* Mary Sigler, *Mercy, Clemency, and the Case of Karla Faye Tucker*, 4 OHIO ST. J. CRIM. L. 455, 465–67 (2007) (discussing the distinction between choice and character retributivism). Second, as Kolber seems to recognize, retributivism, which is typically situated in a deontological framework, does not require cost-benefit analysis. *See* Kolber, *supra* note 1, at 530.

ior and to determine (*e.g.*, by consulting a contemporaneous video recording) whether, as a historical matter, a person engaged in certain acts. To the extent that these are empirical claims,⁷ subject to testing and confirmation, they lend themselves to probabilistic calculation—a determination of the likelihood that repetition of the tests will yield the same results.

In this context, Kolber's invocation of the BARD standard seems apt. With appropriate testing, we should be able to identify quantifiable degrees of confidence in our decision-making. As in the case of whether a particular defendant committed a criminal offense, there is a fact of the matter: He either did, or did not, engage in the relevant conduct, and he either did or did not possess the requisite level of culpability. In practice, it can be extremely difficult to determine whether the necessary elements of an offense have been proven, but we can readily imagine the kind of evidence that would warrant a high degree of confidence in our judgment. We may be inclined to doubt an alibi provided by a family member, for example, and to credit (probably over-credit) the testimony of a disinterested eye-witness (our confidence in eye-witness testimony has been dramatically shaken by recent studies establishing its pervasive unreliability⁸). Video recordings, DNA, and voluntary confessions (all else being equal) generally constitute the gold standard of evidence and are likely to inspire higher degrees of confidence.

The problem with Kolber's indiscriminate list of retributive propositions is that most of the entries on the list represent moral, rather than empirical, claims. And moral claims differ from empirical claims precisely in that they are not testable or otherwise susceptible to proof or falsification.⁹ Instead, moral belief (to the extent that it is critically examined) is generally a product of argument and reflection, not proof. By assimilating moral and empirical claims, Kolber attempts to apply a quantitative standard of proof to moral claims, which can neither be reliably measured nor empirically proved.

To see this, consider Kolber's first proposition—that suffering is an appropriate response to wrongdoing. This is indeed a definitive claim of retributivism and, as Kolber notes, many people find the idea “barbaric.”¹⁰ Now what would it mean to make a case for (or against) this belief? Proponents of retributivism are likely to point to other strongly held commitments with which the

7. Existing technology does not actually allow us to resolve the question of free will, but at least some neuroscientists are confident that their increasing ability to expose the mechanistic nature of human behavior will undermine our conceptions of responsibility. *See, e.g.*, Joshua Greene & Jonathan Cohen, *For the Law, Neuroscience Changes Nothing and Everything*, 359 *PHIL. TRANSACTIONS ROYAL SOC'Y LONDON B* 1775, 1775 (2004). I address the moral dimension of the free will-determinism debate—its relationship to moral and criminal responsibility—below.

8. *See, e.g.*, Thomas Albright & Jef Rakoff, *Eyewitnesses Aren't as Reliable as You Might Think*, *WASH. POST* (Jan. 30, 2015), https://www.washingtonpost.com/opinions/eyewitnesses-arent-as-reliable-as-you-might-think/2015/01/30/fe1bc26c-7a74-11e4-9a27-6fdb612b88_story.html?utm_term=.b8514ff5ac5a.

9. This is true unless one is a conventionalist about morality, holding the view that whatever people happen to believe just is the substance of morality. *See MOORE, supra* note 6, at 12–30 (defining conventional morality as the view that popular opinion determines the truth of moral propositions).

10. Kolber, *supra* note 1, at 499.

retributive idea coheres. Thus, for example, Herbert Morris argued that the infliction of hard treatment on responsible wrongdoers follows from a commitment to fairness, likening punishment to paying a debt or withdrawing an unfair advantage.¹¹ Other retributivists (as well as Morris) invoke variations on the Kantian imperative that requires treating a wrongdoer as a moral agent—in particular, holding him accountable for his choices and actions.¹² As a result, they argue, an offender actually has a right to be punished—that is, to be taken seriously as a moral agent.¹³ One can certainly disagree with these lines of argument, but what would it mean to disprove them? How might we show that fairness does not require settling the offender's moral debt? Or prove that moral agency does not justify moral accountability? This strategy is doomed because, as moral propositions, they may be right or wrong, but there is no mechanism for proving or disproving them. Instead, the conscientious moral investigator will listen to and engage with the arguments and counterarguments and, after sustained reflection, identify what he takes to be the strongest and most compelling moral position in light of other values and commitments he holds.

This process reflects a coherence approach to moral inquiry and justification, which falls far short of establishing proofs for moral claims but is well suited for a pluralist society. In a political setting that features numerous, incompatible religious, philosophical, and moral doctrines, a more definitive approach to moral justification is simply unavailable.¹⁴ According to the foundationalist alternative, one begins with fixed authoritative principles from which further, more specific principles are derived.¹⁵ But in a pluralist society, such uncontested foundations are inevitably elusive. By contrast, the coherence approach does not depend on fixed moral principles; rather, moral justification “is a matter of the mutual support of many considerations, of everything fitting together in one coherent view.”¹⁶ The aim of the approach is thus not to “prove” the validity of a moral principle, but to evaluate whether it fits within the broader scheme of principles already taken—provisionally—as fixed. Whereas a foundationalist attempts to deduce his moral conclusions from authoritative premises (*e.g.*, the word of God), a coherentist recognizes that his enterprise will necessarily “involve a large element of trial and error and muddling through.”¹⁷

11. Herbert Morris, *Persons and Punishment*, in PUNISHMENT AND REHABILITATION 74, 76 (Jeffrie G. Murphy ed., 3d ed. 1995).

12. See generally CHRISTOPHER BENNETT, *THE APOLOGY RITUAL: A PHILOSOPHICAL THEORY OF PUNISHMENT* (2008); ANTONY DUFF, *PUNISHMENT, COMMUNICATION, AND COMMUNITY* (2003); MOORE, *supra* note 6.

13. See, *e.g.*, BENNETT, *supra* note 12, at 50; Morris, *supra* note 11, at 74.

14. See Mary Sigler, *Methodology of Desert*, 42 ARIZ. ST. L.J. 1173, 1182 (2011). For a more thorough discussion, see MOORE, *supra* note 6, at 108–110; JOHN RAWLS, *THEORY OF JUSTICE* 19–21 (1971) (describing the process of “reflective equilibrium”).

15. This account is based on my discussion of moral justification in Sigler, *supra* note 14, at 1182–84.

16. RAWLS, *supra* note 14, at 21.

17. Joel Feinberg, *Justice, Fairness and Rationality*, 81 YALE L.J. 1004, 1019 (1972).

Where does this leave the retributivist? A process of “muddling through” may not inspire high levels of Kolber-confidence, but it reflects the only viable process suitable to the moral domain, and it entails humility—the recognition that further argument, reflection, and experience may reveal a better answer. In the meantime, we need not doubt—or hedge against—what Ronald Dworkin calls the “face value”¹⁸ view of our propositions—that genocide is truly wrong, for example, or that wrongdoers really deserve to suffer punishment. For “any reason we think we have for abandoning a conviction is itself just another conviction, and . . . we can do no better for any claim . . . than to see whether, after the best thought we find appropriate, we think it so.”¹⁹ Absent empirical testing and definitive proof (unavailable in this domain), the only way to establish a working moral proposition is “through substantive normative arguments.”²⁰ Until we encounter a better argument, we have reason to credit the truth of our considered convictions.²¹ Accordingly, application of the BARD standard (or other quantitative metrics) represents a misapprehension of the nature of moral truth, producing a category mistake that imposes an inapt quantitative measure to gauge the soundness or strength of a moral proposition.

II. MORAL DISAGREEMENT AND MORAL TRUTH

Another Kolber strategy for undermining retributivism is to note repeatedly that many of the supporting propositions he identifies are the subject of considerable scholarly and popular disagreement. According to Kolber, because the retributive propositions are controversial, we should have less confidence in them. With respect to free will, for example, and the matter of whether individuals can appropriately be held morally responsible for criminal wrongdoing, Kolber notes that “numerous attempts by brilliant minds over many centuries” have failed to demonstrate “to widespread satisfaction how we can ever be morally responsible no matter whether the universe is deterministic or indeterministic.”²² As a result, Kolber concludes, “it appears stubborn or narcissistic to hold views on the topic with a level of confidence close to certainty.”²³ For “[t]o 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 the contrary arguments and empirical discoveries.”²⁴

18. Ronald Dworkin, *Objectivity and Truth: You Better Believe It*, 25 PHIL. & PUB. AFF. 87, 92 (1996).

19. *Id.* at 118.

20. *Id.* at 127.

21. *Id.* at 118.

22. Kolber, *supra* note 1, at 494. Note that this version of the free-will debate reflects a moral, rather than an empirical, disagreement. That is, it addresses the question of whether, whatever the underlying empirical reality, we have an appropriate normative basis to hold people morally accountable. I take up this issue directly in the following section.

23. *Id.* at 497.

24. *Id.*

Several problems plague this set of arguments, but because the final claim connects directly to the previous section, I shall start there. It is hard to imagine a more vivid illustration of the distinction between doubt (in Kolber's quantitative sense) and the humility I have argued is the appropriate posture for moral analysis. To the extent that Kolber-confidence is set high, in the 90-100% range—it seems to follow that one's mind should be, at best, only 10% open to alternative perspectives—or closed off entirely to additional argument or evidence. This quantification perfectly illustrates the limitations of the “confidence-doubt” approach to moral inquiry. For in the consideration and evaluation of moral beliefs, humility, as noted, does not require mathematical doubt, but it does require a mind *always* alive to the possibility of a better argument and a commitment to the reason-giving enterprise of moral justification. Kolber-confidence, by contrast, aspires to *foreclose* further consideration.

But I digress. The primary flaw in Kolber's attempt to sow doubt based on pervasive disagreement concerns the now familiar confusion about the difference between moral and empirical claims. Wide-scale disagreement about various moral propositions—whether punishment is justified, whether conduct is morally wrongful, whether humans have sufficient agency to be held morally responsible—does not, without more, provide any special insight into the truth of those propositions. What is needed is an argument, which Kolber does not provide, for why we should expect convergence in the domain of moral truth. As Kolber correctly notes, despite his penchant for citing disagreement, we cannot “straightforwardly determine our confidence” in a moral proposition by taking a poll.²⁵ But this disclaimer is much more devastating than he allows. Absent an argument about why these differences of opinion have any implications for moral truth, we cannot evaluate the significance of moral disagreement. “Whether diversity of opinion in some intellectual domain has skeptical implications depends on a further philosophical question: it has such implications only if the best account of the content of that domain explains why it should.”²⁶ As we've already seen, moral propositions, unlike empirical propositions, are generally not predicated on systematic observation and testing. Thus, as Dworkin notes, pervasive disagreement about the details of some observable physical entity—space aliens, say—might give us good reason to doubt the truth of the observational claims.²⁷ But in the domain of critical morality, we generally do not equate the consistency or popularity of a belief with moral truth; nor should we regard disagreement as evidence that it is false.

A related confusion is reflected in Kolber's attempt to discredit state-imposed retributive punishment because it may be unpopular with some citizens. He argues that even if it were right for people to get the punishment they

25. *Id.*

26. Dworkin, *supra* note 18, at 113.

27. *See id.* (“Suppose millions of people claimed to have seen unicorns but disagreed wildly about their size and shape. We would discount their evidence: if there were unicorns, and people had seen them, the actual properties of the beast would have caused more uniform reports.”).

deserve, it does not follow that the state is warranted in imposing it. Although this represents a perfectly plausible moral position, Kolber fails to make a persuasive case. He notes that we do not ordinarily view it as the business of the state to follow through even on valid desert claims. Then, following Victor Tadros,²⁸ he worries that in “expending substantial resources to deliver offenders’ just deserts, the state coerces citizens to pursue goals they may not share.”²⁹ True enough, but how is this different from the many other goals a state may pursue—war, welfare, gun control—that numerous citizens might oppose? Kolber provides no argument that retributive punishment is uniquely vulnerable to the fact that some citizens do not agree with it.

III. FREE WILL AND MORAL RESPONSIBILITY

In distinguishing between moral and empirical claims, I suggested above that the question of whether humans as such possess free will is an empirical issue. That is, at least in theory, it should be possible to establish (or falsify) through scientific testing the claim that human behavior is the product of mechanistic forces beyond our control.³⁰ In that case, humans could be shown to lack free will. For at least some determinists, this fact (or even likelihood) is fatal to policies that seek to hold individuals morally responsible for their conduct. Alternatively, as Kolber notes, compatibilists maintain that even if determinism is true—even if our behavior is caused by forces beyond our control—people can still be held morally responsible insofar “as we identify with our choices, are capable of acting rationally, or meet some similar criteria.”³¹

The first thing to notice here is that the free-will determinism issue is distinct from the question of whether individuals should be held morally accountable for their behavior. The latter involves a moral claim (informed by a scientific claim) about the criteria for moral responsibility. At this point, I set aside the scientific debate to focus on the moral claim—that people generally should be held morally responsible for their conduct—which Kolber correctly associates with the retributive justification. To isolate the moral claim and consider

28. Kolber, *supra* note 1, at 501 (citing VICTOR TADROS, *THE ENDS OF HARM: THE MORAL FOUNDATIONS OF CRIMINAL LAW* 19 (2011)). I confess that I do not follow Tadros’s argument—that state action relies on citizens’ resources to pursue goals “that it sets itself.” First, in a democratic state, one supposes that the people themselves are at least indirectly involved in setting those goals. Moreover, he argues that the state must be able to justify that it is permissible to coerce citizens to pursue such goals. Fair enough. But it’s not clear why punishment stands in any special need of justification, as opposed to war and other state aims.

29. Kolber, *supra* note 1, at 501.

30. It is generally recognized that we do not yet have the technological capability to prove this scientific claim. See, e.g., Greene & Cohen, *supra* note 7, at 1775; Stephen J. Morse, *Brain Overclaim Syndrome and Criminal Responsibility: A Diagnostic Note*, 3 OHIO ST. J. CRIM. L. 397, 403–04 (2006).

31. Kolber, *supra* note 1, at 495.

the matter of confidence levels, I take as true that human behavior is mechanistically determined.³²

As before, the application of confidence levels to a moral claim about the threshold for moral responsibility is simply out of place. Whether individuals possess sufficient rationality (for example) to be held accountable for their conduct, criminally or otherwise, is a normative question that cannot be outsourced to science. In this case, as with other moral questions, we are bound to “consider[] every possibility as fully as we can, and not[e] its implications for the rest of what we think.”³³ In the context of moral responsibility, this includes reflecting on our experiences of both interpersonal and public accountability and considering what it would mean to forgo these forms of human interaction—both within and beyond the domain of criminal justice.

Indeed, broadening our inquiry beyond the moral context at issue is a critical feature of the coherentist methodology. Without the fixed foundations that might otherwise ground our moral judgments, we validate a moral proposition by showing that it fits within a broader scheme of judgments, principles, and background theories (*e.g.*, about the nature of persons) that we also rely on in other domains. Thus, in the case of moral agency, we should be skeptical of a theory of the person that we are not prepared to apply consistently. For if persons as such lack the autonomy to be held responsible for their criminal behavior, we might wonder why we should entrust them with such important decisions as, say, whether, when, and how to raise children. Lacking this basic capacity—on which liberal freedom is predicated—they are presumably unfit decision makers across the board.

Notably, however, at least some proponents of eliminating retributive punishment based on determinism are not prepared to abandon notions of moral agency in other contexts. As Kolber reports, two leading neuro-enthusiasts suggest that we jettison our commitment to moral agency only in the context of criminal justice:

“For most day-to-day purposes it may be pointless or impossible to view ourselves or others in [a] detached sort of way. But . . . 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.”³⁴

Neuroscientists may yet prove that we are actually automatons lacking mental states entirely, and that all of human behavior is merely the moral equivalent of a muscle spasm.³⁵ But as Morse suggests, we should have grave doubts about whether any such creature is equipped to reflect on the implica-

32. According to Kolber and others, the compatibilist position, which accepts the mechanistic account, is the dominant philosophical position. *See id.* at 111; Morse, *supra* note 30, at 402.

33. Dworkin, *supra* note 18, at 118.

34. Kolber, *supra* note 1, at 494 (quoting Greene & Cohen, *supra* note 7).

35. *See* Morse, *supra* note 30, at 402.

tions of his circumstances.³⁶ And if we are in fact devoid of all rational capacities, our (delusional) commitment to retributive theory is probably the least of our worries.

IV. PORTFOLIO OF BELIEFS

Finally, a few thoughts about Kolber's case for a "portfolio of beliefs."³⁷ Despite the intuitive appeal of the deontological framework—and the retributivism it underwrites—Kolber suggests that we might be better served by a variation on hybrid theories of moral justification so that we can "backstop" the uncertainty inherent in our preferred approach. This portfolio of beliefs would reflect a broader set of views in proportion to the strength of our confidence in each; together, these could inform our overall policy preferences.³⁸ Kolber considers one familiar precursor to this approach: threshold deontology.³⁹ In the classic scenario, we are to imagine that we have apprehended a terrorist, and we have reason to believe that he knows the location of a bomb that will shortly detonate in a crowded metropolis, killing thousands of innocents. The question is whether we can justify torturing the terrorist to obtain the relevant intelligence and spare the lives of thousands. From a deontological perspective, torture is categorically unjustifiable; yet it seems morally obtuse—or a form of moral scrupulosity—to fail to act to save thousands of lives.

A threshold deontologist allows that when the stakes are sufficiently high—when enough people are affected—it is possible to justify violating a categorical prohibition despite deontology's rejection of such cost-benefit calculations. According to Kolber, an epistemic threshold deontologist "could have high confidence in deontology and low confidence in consequentialism."⁴⁰ In this challenging case, he could choose the consequentialist solution from his diverse moral portfolio without thereby violating his moral commitments.

Whatever the merits of threshold deontology—and I confess that I am tempted by it—the move to a portfolio of beliefs seems like a dodge that purports to get the decision maker off the moral hook. What is lost in this strategy is the "moral remainder" that should be squarely confronted in any moral dilemma—the acknowledgement of remorse and regret that should attend such a terrible choice.⁴¹ The tragic dilemma at the center of the ticking-bomb hypo-

36. *Id.*

37. Kolber, *supra* note 1, at 529–30.

38. *Id.* at 529.

39. *Id.* at 530.

40. *Id.* at 531.

41. ROSALIND HURSTHOUSE, ON VIRTUE ETHICS 47–48 (1999). By invoking virtue ethics, I risk falling into Kolber's trap—adopting a "portfolio" that includes deontological and virtue ethical features. However, I am not proposing to backstop doubts about deontology with virtue ethics as my Plan B in a pinch. Rather, I believe a robust deontology, sometimes attributed to Kant, can make space for regret even in (tragic) cases of

thetical is that a decision either way is morally repugnant—inflicting torture or allowing the slaughter of innocents. In such a case, what is wanted is not a moral structure that discounts the significance of violating one’s moral principles. Instead, we require a framework that allows us to identify and experience the compromise to principle that such a tragic choice entails. That is, if we are put to a terrible choice, we should not come away unscathed by moral regret. Indeed, if “we are forced to do evil and we act contrary to one of the essential values we hold, the very fabric of our lives is shaken.”⁴²

In contrast, Kolber’s portfolio of beliefs supplies a hedge against uncertainty, eliminating moral conflict by shifting moral foundations. In this way, it does not so much reflect the humility that is called for in the moral domain, but rather a kind of moral cowardice.

V. CONCLUSION

The thrust of my objection to Kolber’s analysis of retributivism is that he invokes a justificatory standard that is unsuited to the moral claims that lie at the heart of retributive theory. Moral claims that are the product of conscientious reflection and engagement should be held with confidence *and* humility—the active recognition that moral justification is a reason-giving enterprise and that our moral propositions are subject to review and reconsideration in the light of stronger arguments. In the moral domain, there is no meaningful alternative.⁴³

morally justified decision making. *See generally* BARBARA HERMAN, *THE PRACTICE OF MORAL JUDGMENT* (1993) (offering a reinterpretation of Kantian ethics that emphasizes moral judgment and moral personality).

42. DANIEL STATMAN, *MORAL DILEMMAS* 19 (1995).

43. Note that consequentialism, as a comprehensive moral theory, rests on the sane uncertain foundation as all other moral theories. Apart from the specific consequentialist challenges Kolber identifies, the basic moral proposition that we should structure institutions to maximize aggregate utility or happiness is itself controversial and unsusceptible to proof.