

Proportionality and Punishment

*Jim Staihar**

ABSTRACT: In the literature on the justification of punishment, unfair advantage theories of punitive desert are prevalent. Several variants have been defended. Although they differ in details, each assumes that a criminal would obtain some unfair advantage or, in other words, an illicit benefit unless he were punished. Criminals deserve to be punished because punishing them is necessary to remove this illicit benefit. In spite of the efforts to defend this sort of theory, none proposed so far in the literature provides a plausible account of the proportionality of punitive desert. This Essay defends a novel unfair advantage theory of punitive desert that is the first to account plausibly for the proportionality of punitive desert. Because this theory explains why and how much criminals deserve to be punished, it should play an important deontological role in limiting the severity of punishments that the state is morally permitted to impose on offenders.

I. INTRODUCTION.....	1210
II. PROPORTIONALITY AND PRIOR UNFAIR ADVANTAGE THEORIES.....	1212
A. FIRST VARIANT.....	1212
B. SECOND VARIANT.....	1213
C. THIRD VARIANT.....	1214
D. FOURTH VARIANT.....	1214
III. PROPORTIONALITY AND A NEW UNFAIR ADVANTAGE THEORY....	1216
A. THE THEORY.....	1216
B. PROPORTIONALITY.....	1222

* Assistant Professor, Department of Accounting and Information Assurance, Robert H. Smith School of Business, University of Maryland; Associate Director, Center for the Study of Business Ethics, Regulation, and Crime at the University of Maryland. For financial support while working on this Essay, I thank the University of Maryland, the University of Chicago Law School, Princeton University's Law and Public Affairs Program, and the University of Michigan. I am very grateful to Elizabeth Anderson, Jeannine Bell, Stephen Darwall, Anthony Dillof, Allan Gibbard, Thomas Green, Stephen Loeb, Stephen Macedo, Richard McAdams, Ralf Michaels, Martha Nussbaum, and Gideon Rosen for very helpful discussion of ideas related to this Essay.

IV. THREE ILLUSTRATIVE EXAMPLES	1223
A. <i>A COOPERATIVE CRIMINAL WHO COMMITS A MODERATELY SERIOUS OFFENSE</i>	1223
B. <i>A DEFIANT CRIMINAL WHO COMMITS A MODERATELY SERIOUS OFFENSE</i>	1224
C. <i>AN EXTREMELY SERIOUS CRIME</i>	1226
V. CRITICAL DISCUSSION	1227
A. <i>AN OBJECTION</i>	1227
B. <i>FIRST RESPONSE</i>	1228
C. <i>SECOND RESPONSE</i>	1229
D. <i>THIRD RESPONSE</i>	1229
E. <i>FOURTH RESPONSE</i>	1229
VI. THREE CONCLUDING REMARKS	1230
A. <i>AVOIDING OVERCRIMINALIZATION</i>	1230
B. <i>THE PROHIBITION AGAINST PUNISHING THE INNOCENT</i>	1231
C. <i>THE EXPRESSIVE CONTENT OF PUNISHMENT</i>	1231

I. INTRODUCTION

Suppose the essence of a crime consists in manifesting insufficient concern for the rights of others in committing a criminal act.¹ The manifestation of such insufficient concern entails a degree of recklessness: an awareness of an unjustifiable risk that one's criminal act will violate the rights of others.² Assume also that a more serious crime manifests a worse deficiency in the offender's concern for the interests of others, and the harmful results of a crime, if any, do not bear directly on its seriousness.³ Finally, suppose we are obligated not to commit crimes because of a more fundamental obligation not to risk unjustifiably causing others the sorts of harms that their rights protect against.

1. See Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Culpability*, 88 CALIF. L. REV. 931, 931 (2000); see also LARRY ALEXANDER & KIMBERLY KESSLER FERZAN WITH CONTRIBUTIONS BY STEPHEN J. MORSE, *CRIME AND CULPABILITY: A THEORY OF CRIMINAL LAW* 23-31 (2009).

2. See Alexander, *supra* note 1, at 931; see also MODEL PENAL CODE § 2.02(2)(c) (1985); ALEXANDER & FERZAN, *supra* note 1, at 23-31. A crime committed with recklessness could also be committed with a variety of other more culpable mental states, such as knowledge and purpose. See Alexander, *supra* note 1, at 931; see also MODEL PENAL CODE § 2.02(2)(a)-(b) (1985); ALEXANDER & FERZAN, *supra* note 1, at 31-41. I presume, though, that a mens rea of mere negligence or strict liability is not sufficient to generate criminal liability. See, e.g., ALEXANDER & FERZAN, *supra* note 1, at 69-85.

3. See, e.g., ALEXANDER & FERZAN, *supra* note 1, at 171-96 (arguing that the harmful results of a crime do not affect the culpability of the offender); Joel Feinberg, *Equal Punishment for Failed Attempts: Some Bad but Instructive Arguments Against It*, 37 ARIZ. L. REV. 117, 118-20, 132 (1995) (arguing the same).

Given these suppositions, which I take for granted for the purposes of this Essay, we would usually assume that someone who commits a crime without any exculpatory defenses deserves to be punished. More precisely, we would assume the criminal deserves to be punished in the negative sense that the state would not violate his rights by punishing him against his will.⁴ Our assumption, though, about the punitive desert of the offender stands in need of justification. For punishing someone involves intentionally harming him, and people typically have a right not to be intentionally harmed against their will.

Explaining why criminals deserve to be punished is especially challenging because an explanation should account for the proportionality of punitive desert. A plausible theory of punitive desert should explain why people deserve to be punished more severely for committing more serious crimes, and a plausible theory should offer a reasonable account of the absolute severity of the punishments that criminals deserve. The former aspect of punitive desert constitutes its ordinal ranking. The latter is its cardinal measure.

In the literature on the justification of punishment, unfair advantage theories of punitive desert are the most prevalent.⁵ Several variants have been defended.⁶ Although they differ in details, each assumes that a criminal would obtain some unfair advantage or, in other words, an illicit benefit unless he were punished. Criminals deserve to be punished because punishing them is necessary to remove this illicit benefit.⁷

In spite of the efforts to defend this sort of theory, none proposed so far in the literature provide a plausible account of the proportionality of punitive desert.⁸ This Essay defends a novel unfair advantage theory of punitive desert

4. See J.L.A. Garcia, *Two Concepts of Desert*, 5 LAW & PHIL. 219, 219–23 (1986) (expounding this negative sense of punitive desert). On the negative sense of punitive desert, the fact that a criminal deserves to be punished means that he has forfeited his right not to be punished. Cf. Christopher Heath Wellman, *The Rights Forfeiture Theory of Punishment*, 122 ETHICS 371, 373–74, 376–91 (2012) (arguing against several objections to the claim that criminals forfeit their right not to be punished).

5. Cf. DAVID BOONIN, *THE PROBLEM OF PUNISHMENT* 120 (2008) (stating that “the fairness-based approach is arguably the preeminent form of retributivism in the current literature”).

6. See, e.g., GEORGE SHER, *DESERT* 69–90 (1987); Richard Dagger, *Playing Fair with Punishment*, 103 ETHICS 473, 474–77, 488 (1993); Herbert Morris, *Persons and Punishment*, 52 MONIST 475, 475–80 (1968); Jeffrie G. Murphy, *Marxism and Retribution*, 2 PHIL. & PUB. AFF. 217, 228–29 (1973).

7. The fact that a person lacks a right to a benefit obtained wrongfully is reflected in the principle that a person may not profit from his own wrong. See, e.g., *Riggs v. Palmer*, 22 N.E. 188, 190 (N.Y. 1889) (stating that “[n]o one shall be permitted to profit by his own fraud, or to take advantage of his own wrong, or to found any claim upon his own iniquity, or to acquire property by his own crime”).

8. Given the criticisms of prior unfair advantage theories of punitive desert, some question whether such theories warrant the widespread attention they receive. Cf. MATT MATRAVERS, *JUSTICE AND PUNISHMENT: THE RATIONALE OF COERCION* 53 (2000) (“It may seem . . . that

that is the first to account plausibly for the proportionality of punitive desert. Because this theory explains why and how much criminals deserve to be punished, it should play an important deontological role in limiting the severity of punishments that the state is morally permitted to impose on offenders.⁹

II. PROPORTIONALITY AND PRIOR UNFAIR ADVANTAGE THEORIES

To account for the proportionality of punitive desert, an unfair advantage theory must identify an illicit benefit, which unpunished criminals would retain, that satisfies two conditions. First, the benefit must be greater for committing more serious crimes. Unless the benefit satisfies this first condition, the theory will not account for the ordinal ranking of punitive desert. That is, the theory will not explain why criminals deserve to be punished more severely for committing more serious offenses.

Second, the absolute magnitude of the illicit benefits must be proportional to the absolute severity of the punishments that criminals deserve. Unless this is so, the theory will not explain the cardinal measure of punitive desert. That is, the theory will not account for the absolute severity of deserved punishments.

Arguably, prior unfair advantage theories fail to satisfy either condition. As a consequence, they cannot account for either the ordinal ranking or the cardinal measure of punitive desert. To see why, consider four variants of an unfair advantage theory that have been proposed in the literature, widely discussed, and roundly criticized.

A. FIRST VARIANT

According to one theory, when a person commits a crime, he obtains an intrinsic benefit of extra freedom from exercising an option that is morally prohibited.¹⁰ The stronger the moral prohibition against committing the crime, the greater the benefit to the criminal who commits it. According to

disproportionate space is accorded to this [unfair advantage] theory given that it is widely thought to have been discredited.”)

9. I assume an undeserved punishment would violate an offender's rights, which are deontological side constraints on how the state is morally permitted to treat the offender. See ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 30–33 (1974) (defending a conception of rights as “side constraints”). Thus, a punishment's being deserved is a necessary condition of its being all things considered justified. However, a punishment's being deserved is not a sufficient condition of its being justified all told. To be justified all things considered, a deserved punishment must also promote some sufficiently valuable aims that outweigh its costs. Cf. H.L.A. HART, *Prolegomenon to the Principles of Punishment*, in PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW 1, 8–11 (2d ed. 2008) (discussing the role of a “General Justifying Aim” in a theory of the justification of punishment); Douglas N. Husak, *Why Punish the Deserving?*, 26 NOÛS 447, 459–62 (1992) (arguing that the state is all things considered justified in imposing a deserved punishment on a criminal only if the punishment results in a sufficiently valuable reduction in future crimes).

10. See SHER, *supra* note 6, at 81–82.

the theory, an offender deserves to be punished in proportion to the intrinsic value of the extra freedom that he obtained from committing his crime.

In critically assessing the merits of this first theory, it is not clear that obtaining any extra freedom from acting immorally is an intrinsic benefit.¹¹ As a consequence, it is not clear that this first theory can explain why criminals deserve any punishment at all. To illustrate, suppose someone commits an unsuccessful attempt at murder. Say the offender tries but fails to kill a co-worker with whom he has been competing for a promotion. The offender derived no pleasure from the attempt or any material advantages. On the face of it, such a failed attempt seems to make the attempted murderer no better off in any respect even though he deserves to be punished quite severely for committing it.

Moreover, even if the freedom one obtains from acting immorally is an intrinsic benefit, there seems no reason to believe that the greater the value of the freedom, the more immoral the act. The attempted murderer we have just described seems no better off in any respect than a criminal who commits a less serious, but more desired crime, such as the successful theft of a valuable commodity.¹² Hence, this first variant of an unfair advantage theory cannot seem to account for the most basic ordinal ranking of punitive desert.

B. SECOND VARIANT

According to a second view, it is burdensome for people to comply with the criminal law in general. For any person, there are at least some crimes he desires to commit. For example, even if someone desires not to commit murder, he might still desire to steal a valuable commodity. Hence, when someone commits a crime, he obtains the illicit benefit of freedom from complying with the criminal law in general.¹³

One problem with this second theory is that the putative benefit of freedom from complying with the criminal law in general is the same for committing any offense.¹⁴ No matter how serious the crime someone commits, he obtains the same freedom from complying with the criminal law in general. Hence, this second variant of an unfair advantage theory cannot explain why offenders obtain greater illicit benefits from committing more serious crimes. So this second theory also cannot explain why criminals deserve to be punished more severely for committing more serious crimes.

11. See BOONIN, *supra* note 5, at 127–29.

12. Relative to the successful theft of a valuable commodity, an unsuccessful attempt at murder is a more serious crime because it manifests a worse deficiency in the offender's concern for the rights of others.

13. See Dagger, *supra* note 6, at 480–84.

14. See, e.g., BOONIN, *supra* note 5, at 124–25; Richard W. Burgh, *Do the Guilty Deserve Punishment?*, 79 J. PHIL. 193, 206 (1982); Jean Hampton, *The Retributive Idea*, in FORGIVENESS AND MERCY 111, 115 (Jeffrie G. Murphy & Jean Hampton eds., 1988).

C. THIRD VARIANT

According to a third theory, the average citizen finds it burdensome to restrain himself from committing any particular type of crime. When someone commits a crime, he obtains the freedom from the burden of restraining himself from committing that particular type of offense. The greater the burden of self-restraint that the average citizen incurs by abstaining from committing a particular type of crime, the greater the benefit to the criminal who commits that specific type of offense.¹⁵

Arguably, this third theory cannot account adequately for the ordinal ranking of punitive desert because the average citizen often finds it less burdensome to restrain himself from committing more serious crimes. For example, the average citizen generally has a stronger desire to steal a valuable commodity than to commit the more serious crime of murder. So the average citizen incurs a greater burden of self-restraint from choosing not to commit a theft than from choosing not to commit a murder. Hence, this third variant of an unfair advantage theory unacceptably entails that offenders deserve less punishment for committing more serious crimes.¹⁶

D. FOURTH VARIANT

According to a fourth, more sophisticated view, we are to imagine a hypothetical situation in which the state auctions off licenses to commit particular crimes with impunity.¹⁷ In the auction, there are two types of bidders with limited resources to bid. One type of bidder consists in aspiring criminals who aim to purchase licenses in order to use them to commit offenses with impunity. The other type of bidder consists in protective associations that aim to purchase licenses in order to keep them out of the hands of those who would use them to commit crimes with impunity. When someone commits a crime, he obtains an illicit benefit whose magnitude is proportional to the price that a license to commit the offense with impunity would fetch in the hypothetical auction. The higher the price, the greater the benefit to the offender.

This fourth theory faces at least two difficulties in accounting for the proportionality of punitive desert. First, it is not obvious that anyone in the hypothetical auction would bid for a license to commit certain types of crimes, such as offenses no one desires to commit. For example, consider again an unsuccessful attempt at murder from which the criminal derives no experiential or material gains. Given the budget constraints of the bidders and the low value they would presumably place on a license to commit an unsuccessful attempt with impunity, it is not obvious that such a license would

15. See Murphy, *supra* note 6, at 228.

16. See Burgh, *supra* note 14, at 209-10.

17. See MICHAEL DAVIS, TO MAKE THE PUNISHMENT FIT THE CRIME: ESSAYS IN THE THEORY OF CRIMINAL JUSTICE 83-86, 238-42 (1992).

fetch any price in the hypothetical auction even though unsuccessful attempts can be quite serious offenses.¹⁸ So this fourth theory has difficulty explaining why some offenders obtain any illicit benefits from committing quite serious crimes. As a consequence, this fourth theory has difficulty explaining why some offenders deserve any punishment at all.

Second, even if every license to commit each type of crime were to fetch some price in the hypothetical auction, licenses to commit more serious crimes might still fetch lower prices than licenses to commit less serious crimes. For even if the state were to auction off fewer licenses to commit more serious offenses, the average citizen still has a stronger desire to commit less serious crimes over more serious crimes, as we have discussed. For example, the average aspiring criminal would be willing to bid more for a license to commit a successful theft of a valuable commodity than a license to commit an unsuccessful attempt at murder.

Moreover, it is not obvious that protective associations would bid more for the latter license than the former. Given their budget constraints, protective associations would presumably submit only marginally higher bids than aspiring criminals to ensure outbidding them at the least cost. Like the other three, this fourth variant of an unfair advantage theory cannot seem to explain why offenders would generally obtain any greater benefits from committing more serious crimes.¹⁹

Although prior unfair advantage theories sometimes try but always fail to account for the ordinal ranking of punitive desert, none even attempts to explain the cardinal measure in any meaningful way. None tries to explain why the magnitude of the purported illicit benefits is proportional to the absolute severity of the punishments that criminals seem to deserve. For example, none offers any explanation for why a simple assault would provide the offender with an illicit benefit whose removal would require him to undertake a lengthy period of labor-intensive community service under reasonable conditions of incapacitation. Thus, prior variants cannot seem to account for the absolute severity of deserved punishments.

In the final analysis, I contend that prior unfair advantage theories cannot account for the proportionality of punitive desert because they try to identify an illicit benefit that criminals generally obtain from committing their offenses. But there seems no such benefit of a material, psychological, or more

18. Davis does describe some possible reasons to bid for a license to commit an unsuccessful attempt. *See id.* at 110–16. For example, when someone attempts to commit a crime, he does risk failing in his attempt and being detected. *Id.* at 110. So, insofar as a license to commit a failed attempt is cheaper than a license to commit a successful attempt, a potential criminal who cannot afford the latter license would still have an incentive to bid for the former. *Id.* at 110–16. It is still an open question, though, whether those in the hypothetical auction would bid for a license to commit a failed attempt given their budget constraints and imperfect probabilities of detection. *Id.*

19. *See* BOONIN, *supra* note 5, at 129–32; Anthony Ellis, *Punishment and the Principle of Fair Play*, 9 UTILITAS 81, 84–90 (1997).

abstract kind. Some offenses do generate material and psychological benefits, but not all, such as many unsuccessful attempts. Moreover, even if there is a more abstract benefit from committing any crime, there seems no reason to believe it is greater for committing more serious crimes or of a magnitude proportional to the absolute severity of the punishments that criminals seem to deserve.

So to develop a more plausible unfair advantage theory, I do not seek to identify another type of benefit that people might obtain by violating their obligation not to commit crimes. Instead, in developing my proposed theory, I seek to identify a type of benefit that criminals stand to obtain by violating a distinct obligation that they incur from committing their crimes: an obligation of restoration.

III. PROPORTIONALITY AND A NEW UNFAIR ADVANTAGE THEORY

A. THE THEORY

When someone commits a crime without any exculpatory defenses, he undermines his trustworthiness.²⁰ More precisely, he undermines the minimally acceptable degree of trustworthiness that we are warranted in demanding each other not to undermine. A person's minimally acceptable degree of trustworthiness consists in the conditions that are necessary for others' being justified in believing with a minimally acceptable credence that he is not disposed to commit crimes.

An unexcused crime undermines the offender's minimally acceptable degree of trustworthiness because it is sufficiently strong evidence of a sufficiently bad standing deficiency in the offender's concern for the rights of others, such that he lacks a sufficiently reliable character trait not to commit crimes.²¹ Hence, for a range of comparably serious crimes and situations, the offender poses an unacceptably high risk of committing any such crimes over a significant run of such situations.²²

At least three empirical findings support this negative inference about the character of an unexcused criminal. First, recidivism rates are high—a

20. See, e.g., Susan Dimock, *Retributivism and Trust*, 16 *LAW & PHIL.* 37, 53 (1997); David A. Hoekema, *Trust and Obey: Toward a New Theory of Punishment*, 25 *ISR. L. REV.* 332, 345 (1991); Jim Staihar, *Punishment as a Costly Signal of Reform*, 110 *J. PHIL.* 282, 282–84 (2013). Because trustworthiness comes in degrees, even repeat offenders undermine their trustworthiness to an additional degree by committing their later crimes.

21. Cf. Richard B. Brandt, *Blameworthiness and Obligation*, in *ESSAYS IN MORAL PHILOSOPHY* 3, 14, 16–17, 32 (A.I. Melden ed., 1958) (arguing that if someone is blameworthy for performing an act, then his performing the act warrants our inferring a defect in his character).

22. Some crimes under some descriptions are literally unrepeatable, such as matricide or patricide. But even these undermine an offender's trustworthiness, for an unexcused crime is strong evidence of a disposition to commit a broader range of crimes comparable in seriousness.

high percentage of offenders commit multiple crimes at different times.²³ Second, there is a high rate of versatility among criminals. In other words, there is a low rate of specialization among criminals—a high percentage of offenders commit multiple types of crimes.²⁴ Third, studies indicate that a high percentage of criminals have an especially bad, stable, and broad deficiency in their concern for the interests of others.²⁵ Because crimes essentially manifest insufficient concern for the rights of others, the common presence of such a standing deficiency in the character of unexcused criminals should not seem surprising.

Assuming that a criminal undermines his trustworthiness by committing his offense, I argue that the criminal unacceptably risks causing others certain especially significant harms unless he restores his trustworthiness.²⁶ For example, an untrustworthy criminal unacceptably risks committing a range of other offenses in the future. In addition, an untrustworthy criminal also unacceptably risks causing others to incur certain costs of insecurity, which would constitute harms to others even if the criminal were never to commit another offense.²⁷

23. See MICHAEL R. GOTTFREDSON & TRAVIS HIRSCHI, *A GENERAL THEORY OF CRIME* 107–08, 177, 230–31, 253 (1990) (noting the high stability of criminals' dispositions to commit crimes, and citing numerous research studies in support); PATRICK A. LANGAN & DAVID J. LEVIN, *BUREAU OF JUSTICE STATISTICS SPECIAL REPORT: RECIDIVISM OF PRISONERS RELEASED IN 1994*, at 1 (2002), available at <http://www.bjs.gov/content/pub/pdf/rpr94.pdf> (showing that of over 270,000 prisoners released in 15 U.S. states in 1994, “67.5% were rearrested for a new offense,” and “46.9% were reconvicted for a new crime” within three years of their release).

24. See, e.g., GOTTFREDSON & HIRSCHI, *supra* note 23, at 91–94, 256, 266 (discussing the high rate of versatility among criminals, and noting numerous research studies in support); Chester L. Britt, *Versatility*, in *THE GENERALITY OF DEVIANCE* 173, 173 (Travis Hirschi & Michael R. Gottfredson eds., 1994) (same); Alex Piquero et al., *Onset Age and Offense Specialization*, 36 *J. RES. CRIME & DELINQ.* 275, 275–76 (1999) (stating that “[r]esearchers investigating the sequencing of offense types over time in criminal offending have generally found that offenders exhibit some specialization amid a great deal of versatility”); Leonore M.J. Simon, *Do Criminal Offenders Specialize in Crime Types?*, 6 *APPLIED & PREVENTIVE PSYCHOL.* 35, 37 (1997) (noting the high rate of versatility even among white collar criminals, sex offenders, and those who commit crimes of domestic violence).

25. See, e.g., GOTTFREDSON & HIRSCHI, *supra* note 23, at 89 (stating that criminals “tend to be [acutely] self-centered, indifferent, or insensitive to the suffering and needs of others”); Joan McCord, *Understanding Motivations: Considering Altruism and Aggression*, in 3 *FACTS, FRAMEWORKS, AND FORECASTS* 115, 126 (Joan McCord ed., 1992) (stating that “crime is a consequence of motives to injure others or to benefit oneself without a proper regard to the welfare of others”); Joshua D. Miller & Donald Lynam, *Structural Models of Personality and Their Relation to Antisocial Behavior: A Meta-Analytic Review*, 39 *CRIMINOLOGY* 765, 780 (2001) (writing that “[i]ndividuals who commit crimes tend to be hostile, self-centered, spiteful, jealous, and indifferent to others”); cf. ALEXANDER & FERZAN, *supra* note 1, at 23–31 (arguing that the mens rea of a crime should involve manifesting insufficient concern for the interests of others); Alexander, *supra* note 1, at 944–45 (same).

26. See Staihar, *supra* note 20, at 283.

27. Cf. THOMAS HOBBS, *LEVIATHAN* 86–90 (Richard Tuck ed., Cambridge Univ. Press rev. ed. 1996) (1651) (describing vividly the costs of insecurity that people rationally must incur in response to being justified in believing that others are disposed to engage in acts of aggression).

Three costs of insecurity seem especially salient. First, others might rationally need to invest in costly precautionary measures to protect themselves from an untrustworthy offender. For example, they might need to engage in costly monitoring of the criminal and to invest in costly protective services when interacting with him is unavoidable.

Second, others might rationally need to forgo pursuing some personally and socially valuable activities that would leave them too vulnerable to an untrustworthy offender. In other words, people might rationally need to reduce their activity levels in response to the criminal. Third, others might rationally experience higher levels of fear in response to the higher risk of the offender's committing crimes again.²⁸ Given the significance of these costs of insecurity to the lives of others, they would constitute harms that others have a right against incurring.

To avoid the unacceptable risk of causing others to incur the relevant harms, a criminal incurs an obligation to restore his trustworthiness expeditiously.²⁹ Given this obligation of restoration, the offender is obligated to undertake any burdens necessary to restore his trustworthiness to a minimally acceptable degree.³⁰ Assuming the criminal must undertake certain burdens to do so, the state may impose them on him as a punishment against his will without violating his rights. For unless the criminal suffers such burdens, he will obtain an illicit benefit consisting in his freedom from the burdens necessary to fulfill his obligation of restoration. In my view, this is the illicit benefit that a punishment can prevent a criminal from obtaining.

28. I assume that insofar as an offender poses a higher risk of committing further crimes, then others will rationally believe with a higher subjective probability that the offender will commit further crimes. For discussion of the concept of subjective probabilities and their relation to objective probabilities, see 2 DAVID LEWIS, *A Subjectivist's Guide to Objective Chance*, in *PHILOSOPHICAL PAPERS* 83, 83 (1987); John C. Harsanyi, *Bayesian Decision Theory, Subjective and Objective Probabilities, and Acceptance of Empirical Hypotheses*, 57 *SYNTHESE* 341, 355-58 (1983).

29. A criminal is obligated to restore his trustworthiness *expeditiously* because the longer he takes to restore it, the longer he will pose an unacceptable risk to others. For a statement of the importance of restoring trustworthiness in the more general context of reparations, see Margaret Urban Walker, *THE AQUINAS LECTURE, 2010: WHAT IS REPARATIVE JUSTICE?* 25 (2010) ("The gesture of reparations needs to model the kind of relationship between victims and responsible parties that creates a new or renewed basis of trust for the future, precisely what was lacking in the circumstances in which the wrong was done.").

30. If a criminal were to restore his trustworthiness to the baseline degree, he would fulfill the obligation of restoration he incurs from committing his offense. In other words, he would fulfill the obligation by restoring his trustworthiness to the level it was prior to his committing the crime. However, a minimally acceptable degree of trustworthiness is not necessarily as high as the baseline. To illustrate, suppose someone has lived a supererogatory life as a saint but then one day commits a crime. In this case, the criminal's baseline degree of trustworthiness was extraordinarily high, much higher than a minimally acceptable level, for we are not warranted in demanding each other to be saints. Rather we are warranted in demanding each other only to maintain a sufficiently reliable disposition not to violate the rights of others. For help clarifying this point, I thank Robert Ferguson and Stephen Macedo.

According to the main principle of my proposed theory, a criminal deserves a punishment for his crime that is proportional to the burdens he must undertake to fulfill the obligation of restoration he incurs from committing his crime.³¹ In other words, a criminal deserves a punishment that is proportional to the burdens he is obligated to undertake to restore his trustworthiness to a minimally acceptable degree.³² Once the criminal undertakes a punishment proportional to such burdens, he deserves no more punishment for his offense.³³

Now I argue that a criminal must in fact undertake some burdens to restore his trustworthiness. To restore it, the offender must signal his reform. “Such a signal³⁴ would be a directly observable property that is sufficiently strong evidence that an offender has rectified the prior deficiency in her concern for others” that she manifested in committing her crime.³⁵ Offenders must restore their trustworthiness by signaling their reform because people are unavoidably vulnerable to each other under any acceptable system of criminal justice available.³⁶ No acceptable means of deterrence or incapacitation available can adequately reduce the risks that untrustworthy criminals pose to others.³⁷ Moreover, there is no reliable way for others to

31. The punitive desert of an offender is commonly thought to depend on only backward-looking considerations. An offender’s obligation of restoration, though, is grounded in both backward- and forwarding-looking considerations. Insofar as this makes the conception of punitive desert under my proposed theory revisionary, it is a welcome revision.

32. As a corollary, a criminal does not deserve a punishment for his crime that is more severe than the burdens he must undertake to fulfill the obligation of restoration he incurs from committing the crime. In other words, a criminal does not deserve to be punished for his crime more severely than the burdens he is obligated to undertake to restore his trustworthiness to a minimally acceptable degree.

33. So after the state punishes an offender in proportion to the severity of the burdens necessary to satisfy his obligation of restoration, the state is not morally permitted to punish the offender further for the relevant offense. This is the sense in which my proposed theory of punitive desert provides a deontological side constraint limiting how much the state is morally permitted to punish a criminal. See *supra* text accompanying note 9.

34. In general, a signal is a directly observable property that is strong evidence of its bearer’s possessing another property that is not directly observable. Cf. Michael Bacharach & Diego Gambetta, *Trust in Signs*, in TRUST IN SOCIETY 148, 159 (Karen S. Cook ed., 2001).

35. Staihar, *supra* note 20, at 287. By rectifying the deficiency in her concern for others, the offender would develop a good will. As Annette Baier states, when we trust others, we are confident they have a good will toward us; therefore, “[r]easonable trust will require good grounds for such confidence in another’s good will.” Annette Baier, *Trust and Antitrust*, 96 ETHICS 231, 235 (1986).

36. See Staihar, *supra* note 20, at 283–87; cf. Murphy, *supra* note 6, at 231–32 (“[O]ur moral language presupposes . . . that we are vulnerable creatures—creatures who can harm and be harmed by each other.”).

37. See, e.g., Paul H. Robinson & John M. Darley, *Does Criminal Law Deter? A Behavioural Science Investigation*, 24 OXFORD J. LEGAL STUD. 173, 173 (2004); Paul H. Robinson & John M. Darley, *The Role of Deterrence in the Formulation of Criminal Law Rules: At Its Worst When Doing Its Best*, 91 GEO. L.J. 949, 951, 953 (2003); Staihar, *supra* note 20, at 284–85.

induce reform in a criminal, and none seems forthcoming.³⁸ Hence, there is a need for a sign from the offender himself that he has come to develop a sufficiently high degree of concern for the interests of others.³⁹

To be credible, I suggest that a sign of reform must be costly.⁴⁰ A criminal cannot signal his reform through mere costless means, such as merely apologizing for his crime or pleading a change of heart. Such “cheap talk” is not credible because criminals who do not care at all about others would be willing to convey it.⁴¹ So a credible sign of reform must be too costly for

38. See, e.g., GOTTFREDSON & HIRSCHI, *supra* note 23, at 268–69; Staihar, *supra* note 20, at 286–87. Absent an exculpatory defense, there is no form of clinical treatment that criminals can undergo which would provide others with the needed assurance of reform. In rare cases involving an exculpatory defense, some nonpunitive costless means might indeed reliably induce and, therefore, signal reform. In these cases, the offender suffers from some curable clinical condition impairing his capacity to respond appropriately to the reasons that count against committing crimes. For example, suppose someone commits a crime because he was recently brainwashed into thinking it right. There might be a course of psychiatric treatment that would reliably deprogram such a victim of coercive persuasion. See ROBERT JAY LIFTON, *THOUGHT REFORM AND THE PSYCHOLOGY OF TOTALISM: A STUDY OF “BRAINWASHING” IN CHINA* 86–151 (1989); Richard Delgado, *Ascription of Criminal States of Mind: Toward a Defense Theory for the Coercively Persuaded (“Brainwashed”) Defendant*, 63 MINN. L. REV. 1, 9 & nn.38 & 40, 30–32 (1978). Or assume a person commits a crime because he is suffering from a noticeable brain abnormality, such as a tumor. Then brain surgery might reliably remove his disposition to commit crimes. See Jeffrey M. Burns & Russell H. Swerdlow, *Right Orbitofrontal Tumor with Pedophilia Symptom and Constructional Apraxia Sign*, 60 ARCHIVES NEUROLOGY 437, 437 (2003). There are also nonpunitive ways to reduce some people’s temptation to commit crimes. For example, the state can provide them with better opportunities for education, employment, and treatment for problems such as addiction. But merely decreasing temptation in these ways is not the same as signaling reform. We are left with no assurance that the criminals have come to care more about others or that they are willing to work at these opportunities for the long term.

39. At this point, I note that my proposed theory of punitive desert is a practical one. That is, I seek to explain why and how much offenders deserve to be punished in the actual world given the natural facts that generally characterize the unavoidable conditions under which people actually live. So my theory presumes there are no extraordinary means available of obtaining epistemic access to or changing an offender’s disposition to commit crimes. Cf. JOHN RAWLS, *A THEORY OF JUSTICE* 157–61 (Harvard Univ. Press rev. ed. 1999) (developing a theory of distributive justice with a similarly practical aim and its own presuppositions about the natural facts under which it applies).

40. See Staihar, *supra* note 20, at 287–89. The concept of a costly signal has wide interdisciplinary application. See, e.g., DOUGLAS G. BAIRD ET AL., *GAME THEORY AND THE LAW* 122–58 (1994) (law); ERVING GOFFMAN, *THE PRESENTATION OF SELF IN EVERYDAY LIFE* 32, 36 (1959) (sociology); A. MICHAEL SPENCE, *MARKET SIGNALING: INFORMATIONAL TRANSFER IN HIRING AND RELATED SCREENING PROCESSES* 47–48 (1974) (economics); AMOTZ ZAHAVI & AVISHAG ZAHAVI, *THE HANDICAP PRINCIPLE: A MISSING PIECE OF DARWIN’S PUZZLE* 58–60 (Naama Zahavi-Ely & Melvin Patrick Ely trans., 1997) (biology). For an elementary game-theoretic analysis of signaling, see AVINASH DIXIT & SUSAN SKEATH, *GAMES OF STRATEGY* 41 2–16, 424–26 (1999). For an application of signaling theory specifically to trust, see Bacharach & Gambetta, *supra* note 34, at 159–62.

41. For analyses of the credibility conditions on cheap talk, see, e.g., DIXIT & SKEATH, *supra* note 40, at 425–26; Joseph Farrell, *Meaning and Credibility in Cheap-Talk Games*, 5 GAMES & ECON. BEHAV. 514, 514–15 (1993); Joseph Farrell & Matthew Rabin, *Cheap Talk*, 10 J. ECON. PERSP. 103, 103–04 (1996); Robert Stalnaker, *Saying and Meaning, Cheap Talk and Credibility*, in *GAME THEORY AND PRAGMATICS* 83, 83–100 (Anton Benz et al. eds., 2006).

criminals who have not rectified the revealed deficiency in their concern for others.⁴² In other words, a credible sign of reform must be too costly for criminals who have not come to care sufficiently about the interests of others by developing a sufficiently benevolent character. Thus, to demonstrate reform, I assume a criminal must send others a costly signal that he has developed a sufficiently benevolent character.⁴³ “In general, benevolence is a trust-warranting property that is inconsistent with the kind of insufficient concern and impulsiveness typical of criminals.”⁴⁴

To demonstrate the development of a highly benevolent character, the criminal can signal that she has acted with a sufficiently high degree of benevolence for a sufficiently long time after committing her crime. To demonstrate that she has acted with such benevolence, she can sacrifice some of her sufficiently important personal interests for a sufficiently long time for the sake of benefiting others. To make such a sacrifice for others, the criminal can engage in labor-intensive community service, and she can do so under reasonable conditions of incapacitation to mitigate the costs of insecurity that others rationally must incur during the interim. The more service she performs for the sake of benefiting others, the stronger it will serve as evidence that she has rectified the prior deficiency in her concern for others.⁴⁵

Hence, to fulfill his obligation of restoration, and thereby restore his trustworthiness to a minimally acceptable degree, a criminal must undertake some burdens.⁴⁶ The offender deserves to be punished in proportion to those burdens because unless he suffers in proportion to them, he will obtain an

42. This is the “nonpooling condition” on the credibility of a costly signal of reform. Staihar, *supra* note 20, at 289; cf. Bacharach & Gambetta, *supra* note 34, at 160.

43. See Staihar, *supra* note 20, at 287.

44. Staihar, *supra* note 20, at 287; see Bacharach & Gambetta, *supra* note 34, at 154 (noting benevolence as a trust-warranting property).

45. See Staihar, *supra* note 20, at 287; cf. KEALLY MCBRIDE, PUNISHMENT AND POLITICAL ORDER 136 (2007) (reporting that “prisoners who work in prisons are 24 percent less likely to return to prison after release”); LINDA RADZIK, MAKING AMENDS: ATONEMENT IN MORALITY, LAW, AND POLITICS 99 (2009) (writing that “[t]he greater the sacrifice that is required to make the reparation payment, the more evidence we have that the wrongdoer is remorseful for her past action”).

46. Like other trust building or maintaining processes, the process of a criminal’s restoring his trustworthiness has a multi-layered “inferential structure.” Cf. Bacharach & Gambetta, *supra* note 34, at 162. In addition to undertaking the required burdens, other steps might also be necessary to restore a criminal’s trustworthiness, such as apologizing for the crime and compensating any victims. The criminal might need to undergo some form of therapy and take steps to eliminate aspects of his situation that pressure him to commit crimes, such as unemployment, corrupting social influences, and problems of addiction. Much will depend on the specifics of the case. However, because these other steps are not necessarily burdensome for the criminal, they need not be part of his punishment properly understood.

illicit benefit consisting in his freedom from the burdens necessary to fulfill the obligation of restoration he incurs from committing his crime.

B. PROPORTIONALITY

At this point, I hope to have provided a reasonable explanation of why criminals deserve some punishment. Now I argue that my proposed theory also yields a plausible account of the proportionality of punitive desert. Regarding the ordinal ranking, more serious crimes manifest a worse deficiency in the offender's concern for the interests of others. As a consequence, more serious crimes undermine an offender's trustworthiness to a worse degree.⁴⁷

It seems natural to assume then that criminals who commit more serious offenses must do more to restore their trustworthiness. They must sacrifice more of their personal interests for a longer time for the sake of benefiting others. They must undertake a longer period of labor-intensive community service. In short, they must undertake a more severe burden to signal their reform. Thus, someone who commits a more serious crime deserves a more severe punishment because unless he suffers a more severe punishment, he will obtain the greater illicit benefit of freedom from the more severe burden necessary to fulfill his obligation of restoration.⁴⁸

47. In arguing that more serious crimes, such as murder, undermine an offender's trustworthiness to a worse degree than less serious crimes, such as theft, I am assuming the more serious crimes are committed with no exculpatory defenses. An extremely serious crime, though, such as murder, might not undermine an offender's trustworthiness at the time of assessment if it were committed with a range of excuses. The relevant excuses block the otherwise justified inference from the fact that someone performed a criminal act to his having a standing motivational defect in his character at the time of assessment. Examples of such excuses include, among others, involuntary intoxication, provocation, hypnosis, somnambulism, and temporary insanity. Proponents of such a character theory of excuses are numerous. See, e.g., DAVID HUME, *A TREATISE OF HUMAN NATURE* 412, 477-78 (L. A. Selby-Bigge & P.H. Nidditch eds., 2d ed. 1978) (1738); Michael D. Bayles, *Character, Purpose, and Criminal Responsibility*, 1 *LAW & PHIL.* 5, 7 (1982); Michael D. Bayles, *Hume on Blame and Excuse*, 2 *HUME STUD.* 17, 21 (1976); R.B. Brandt, *A Motivational Theory of Excuses in the Criminal Law*, in 27 *CRIMINAL JUSTICE* 165, 165-66 (J. Roland Pennock & John W. Chapman eds., 1985); Richard B. Brandt, *A Utilitarian Theory of Excuses*, 78 *PHIL. REV.* 337, 354 (1969); George Vuoso, *Background, Responsibility, and Excuse*, 96 *YALE L.J.* 1661, 1682-85 (1987); cf. T.M. SCANLON, *WHAT WE OWE TO EACH OTHER* 277-79 (1998) (suggesting that some considerations constitute excuses because they "sever the connection between the action or attitude and the agent's judgments and character"); P.F. STRAWSON, *Freedom and Resentment*, in *FREEDOM AND RESENTMENT AND OTHER ESSAYS* 1, 8 (1974) (stating that "[w]e shall not feel resentment against the man he is for the action done by the man he is not; or at least we shall feel less"); Gary Watson, *Responsibility and the Limits of Evil: Variations on a Strawsonian Theme*, in *RESPONSIBILITY, CHARACTER, AND THE EMOTIONS: NEW ESSAYS IN MORAL PSYCHOLOGY* 256, 260 (Ferdinand Schoeman ed., 1987) (describing excuses that, according to Strawson, "present the other . . . as acting uncharacteristically due to extraordinary circumstances").

48. At this point, I remind the reader that I presume the harmful results, if any, of a crime do not affect its seriousness in the sense relevant to the offender's culpability or punitive desert. See *supra* text accompanying note 3. So because successful and failed attempts at the same crime

In addition to explaining the ordinal ranking, my proposed theory also provides a reasonable account of the cardinal measure of punitive desert. It naturally explains why criminals deserve punishments that are proportional to lengthy sentences of labor-intensive community service performed under reasonable conditions of incapacitation. For according to my theory, the absolute severity of the most severe punishment that a criminal deserves for an offense corresponds to the absolute severity of the burdens that he must undertake to fulfill the obligation of restoration he incurs from committing the offense. This account of the proportionality of punitive desert not only seems humane and reasonable, but also is more plausible than any alternative account proposed in the literature.

IV. THREE ILLUSTRATIVE EXAMPLES

A. *A COOPERATIVE CRIMINAL WHO COMMITS A MODERATELY SERIOUS OFFENSE*

Consider someone who commits a simple assault out of animosity toward his victims. The offender's crime undermined the minimally acceptable degree of trustworthiness that we are warranted in demanding him not to undermine. The offense reveals a particularly bad deficiency in his concern for the rights of others.

To restore his trustworthiness, the offender must apologize to his victims and provide compensation for any harm he caused them. But although necessary, that might not be sufficient. An apology is "cheap talk," and providing compensation need not be burdensome, especially if paid by a third party on behalf of the offender. So even if the criminal had not rectified at all the revealed deficiency in his concern for others, he could still be willing to provide an apology and compensation.

To provide a credible sign of reform, the criminal must sacrifice important personal interests for a long time for the sake of benefiting others. More specifically, for some small number n , he must engage in n years of labor-intensive community service, and he must do so under reasonably humane conditions of incapacitation, in a prison, to mitigate the risk he poses to others while the service is performed.⁴⁹

undermine an offender's trustworthiness to the same degree, they deserve equally severe punishments under my proposed theory.

49. Although not necessary, a restorative punishment could also express an apology and provide compensation to any victims. Cf. R.A. DUFF, PUNISHMENT, COMMUNICATION, AND COMMUNITY 106 (2001) (noting that a criminal's undertaking a punishment "constitutes a forceful public apology"). As I noted, though, a mere apology or compensation could be provided through nonpunitive means that are distinct from a criminal's punishment properly understood. See *supra* note 46. For example, a mere apology could be expressed through cheap talk, and compensation could be provided by a third party on behalf of the offender. Unlike an offender's obligation to compensate any victims, his obligation to restore his trustworthiness cannot be fulfilled vicariously by third parties undertaking burdens on his behalf. Thus, my proposed theory of punitive desert explains the constraint on penal substitution because a criminal can restore his

Now suppose the state sentences the offender accordingly, and he is cooperative and willing to restore his trustworthiness. He apologizes to his victims, provides any compensation owed to them, and undertakes the required punishment for the sake of helping others. He undertakes n years of labor-intensive community service for the sake of benefiting others.⁵⁰

In prison, he works to produce goods, such as clothes or medical supplies, that are distributed to those in need. He is paid a small wage for the labor, and he consents to the state's garnishing the wage and distributing it to his victims and others. After increasing his trustworthiness, he even works as a tutor for other prisoners, teaching them any special knowledge he might have to share about useful trades or subjects.

By the end of his n -year sentence, the criminal has restored his trustworthiness to a minimally acceptable degree.⁵¹ He has justified our believing with a minimally acceptable credence that he is no longer disposed to commit crimes. He has demonstrated that he has rectified the revealed deficiency in his concern for the rights of others. In short, he has signaled his reform. So he has fully fulfilled the obligation of restoration he incurred from committing his crime.

Now there is no illicit benefit that the criminal stands to obtain if he is not punished more. Therefore, he does not deserve any more punishment for his offense. At this point, the state is not morally permitted to impose any additional punishment on the offender for committing the crime. Any additional punishment would violate the offender's rights, being out of proportion to the seriousness of his crime.

B. A DEFIANT CRIMINAL WHO COMMITS A MODERATELY SERIOUS OFFENSE

Consider a person who commits the same simple assault as the previous cooperative criminal. Assuming her crime undermines her trustworthiness to the same degree, she must undertake the same n -year sentence to restore her trustworthiness to a minimally acceptable degree. However, unlike her cooperative counterpart, suppose this second offender remains persistently defiant, unapologetic, and unwilling to restore her trustworthiness to any degree. She could restore it if she chose, but she refuses. Not only is she unwilling to undertake any service for the sake of benefiting others, she refuses to undertake any service at all.

Nevertheless, this defiant criminal still incurs the same obligation of restoration as her cooperative counterpart, and she must undertake the same

trustworthiness only if he undertakes the appropriate burdens himself. Cf. DAVID LEWIS, *Do We Believe in Penal Substitution?*, in PAPER IN ETHICS AND SOCIAL PHILOSOPHY 128, 128-35 (2000).

50. I assume that if a criminal is cooperative, and the state insists on punishing him, then the state is obligated to provide him with an opportunity to undertake his deserved punishment in a restorative form.

51. For a more detailed discussion of how a state-imposed punishment could constitute a credible sign of an offender's reform, see Staihar, *supra* note 20, at 287-92.

n -year restorative sentence to fulfill it. So in response, suppose the state imposes on the defiant offender an n -year nonrestorative punishment with a severity proportional to the n -year sentence actually required to fulfill her obligation of restoration. Unlike the restorative punishment that her cooperative counterpart undertakes, this nonrestorative sentence consists merely in a burdensome form of incarceration for n years without any service to others.

Although this nonrestorative punishment does not fulfill the defiant criminal's obligation of restoration, she still deserves it for her crime. Unless she suffers the nonrestorative punishment, she will obtain an illicit benefit of freedom from the burdens required to fulfill the obligation of restoration she incurs from committing her crime. This is the illicit benefit that the n -year nonrestorative punishment prevents the defiant criminal from obtaining.

After undergoing the nonrestorative punishment, the defiant criminal does not deserve any more punishment for her crime. Because she has suffered in proportion to the burdens required to fulfill her obligation of restoration, she does not stand to obtain an illicit benefit if she is not punished more. Her nonrestorative punishment has made her no better off than she would have been if she had actually restored her trustworthiness to a minimally acceptable degree.⁵²

But although the defiant offender does not deserve any more punishment after undergoing the n -year nonrestorative sentence, it is true that she is also no more trustworthy.⁵³ Thus, the defiant offender still poses a higher risk of committing further crimes, and others still rationally stand to incur additional costs of insecurity in response to her. I leave it an open

52. Although the defiant criminal and her cooperative counterpart deserve equally severe punishments for their identical crimes, the former undermines her trustworthiness to a worse degree by being defiant in response to her crime, so, she deserves some additional punishment. Hence, other things being equal, defiant criminals do deserve more punishment than cooperative ones under my proposed view.

53. Although imposing deserved punishments on defiant criminals does not restore their trustworthiness, those punishments could still promote some other valuable aims, which could make them all things considered justified. In other words, imposing deserved punishments on defiant offenders could still promote some benefits that would outweigh their costs. For example, positive retributivists contend it is intrinsically good that wrongdoers suffer the punishments they deserve. See J.L. Mackie, *Morality and the Retributive Emotions*, 1 CRIM. JUST. ETHICS 3, 4 (1982); Michael S. Moore, *Justifying Retributivism*, 27 ISR. L. REV. 15, 19–20 (1993); cf. GEORGE EDWARD MOORE, PRINCIPIA ETHICA 163–64, 214–16 (1971) (claiming that the organic whole of a criminal's suffering some punishment is intrinsically better than the organic whole of his suffering no punishment). So positive retributivists could argue that it is intrinsically good to prevent defiant criminals from obtaining the illicit benefit of freedom from the burdens required to fulfill their obligation of restoration. Thus, they might contend that a defiant criminal's suffering her deserved punishment is intrinsically good precisely because it removes this illicit benefit. Less controversially, imposing a deserved punishment on a defiant criminal could have the valuable consequence of preventing some future crimes by promoting a deterrence objective. Cf. HART, *supra* note 9, at 8–11; Husak, *supra* note 9, at 459–62.

question whether the state would be morally permitted to incapacitate the offender longer on grounds of self-defense.

I stress, though, that insofar as an additional term of incapacitation would be justifiable, the additional term would not be an additional punishment properly understood. Rather it would be a term of involuntary civil commitment analogous to the quarantine of someone with an infectious dangerous disease. As such, the state would be obligated to mitigate the degree to which any longer term of incapacitation would be harmful to the defiant offender.⁵⁴

As an alternative to a longer term of involuntary civil commitment, the state might choose to grant the defiant criminal a conditional release instead. The relevant conditions might, for example, require the offender to wear an ankle bracelet or report to a parole officer on a reasonably frequent basis. Although these conditions would not eliminate the higher risk that the defiant offender poses to others, they would partially mitigate the risk. I also leave it an open question whether the offender would owe others any compensation for any additional costs of insecurity that they might rationally incur in response to her under the terms of such a conditional release.

C. AN EXTREMELY SERIOUS CRIME

Consider someone who commits an aggravated murder, which is an extremely serious crime. The murder undermines the criminal's trustworthiness to an extremely bad degree. It justifies our believing with an extremely high credence that the murderer is disposed to commit extremely serious crimes, for the murder reveals an extremely bad deficiency in his concern for the rights of others.

Given the inevitable constraints on the duration of a human life, there might be nothing that the murderer can do to restore his trustworthiness to a minimally acceptable degree. It is possible that he cannot justify our believing with a minimally acceptable credence that he has fully rectified the revealed deficiency in his concern for the rights of others. Nevertheless, he can still partially restore his trustworthiness. He can justify our believing with a range of higher credences that he has come to care increasingly more about the interests of others.

By partially restoring his trustworthiness, the murderer would mitigate the risk he poses to others. By mitigating the risk, the murderer would mitigate the costs of insecurity that others, such as prison officials and fellow inmates, rationally must incur in response to him. Thus, the murderer is obligated to restore his trustworthiness as much as he can. And to do so, he

54. A term of involuntary civil commitment is not a punishment properly understood for at least two reasons. First, the term does not express moral blame toward the subject. Second, the term is not intentionally harmful to the subject. For more discussion about the concept of punishment, see BOONIN, *supra* note 5, at 3-28; JOEL FEINBERG, *The Expressive Function of Punishment*, in *DOING & DESERVING: ESSAYS IN THE THEORY OF RESPONSIBILITY* 95, 95-118 (1970).

might need to undertake a life sentence of labor-intensive community service under reasonably humane conditions of incapacitation, in a prison.

Hence, whether cooperative or defiant, the murderer might deserve a punishment with a severity proportional to such a burdensome life sentence. Unless he were to receive such a life sentence, he would obtain an illicit benefit of freedom from the burdens required to fulfill his obligation to restore his trustworthiness as much as he can. This is the illicit benefit that a sufficiently burdensome life sentence might be needed to prevent the murderer from obtaining.

Under my proposed theory, though, neither the murderer in this case nor any other criminal would ever deserve the death penalty, for no criminal is obligated to undertake the death penalty. An offender's obligation of restoration is ultimately grounded in a more basic obligation not to risk unjustifiably causing others certain especially significant harms, including the commission of further crimes and the costs of insecurity. This more basic obligation is subject to a limit: someone is obligated to undertake the means necessary to prevent himself from posing an unjustifiable risk of harming others only if the means are not extremely worse for him than the risk he would otherwise pose to others.

Now whether cooperative or defiant, those who commit extremely serious crimes do pose an unjustifiable risk to others even while incapacitated under reasonably humane conditions in a penitentiary. For example, prison officials and fellow inmates might rationally incur additional costs of insecurity in response to any incapacitated offender.⁵⁵ However, presumably, the death penalty would be extremely worse for an incarcerated criminal than the risk he would pose to others while serving a life sentence in a reasonably well-secured penitentiary. Therefore, at most, those who commit extremely serious crimes deserve punishments proportional to life sentences of labor-intensive community service performed under reasonably humane conditions of incarceration.⁵⁶

V. CRITICAL DISCUSSION

A. AN OBJECTION

Critics might challenge the potential for a state-imposed punishment to restore an offender's trustworthiness by functioning as a costly signal of reform.⁵⁷ They might concede that labor-intensive community service could

55. See Staihar, *supra* note 20, at 284.

56. As a further consequence of the relevant limit on one's obligation not to cause others harm, criminals also do not deserve to be incapacitated under the conditions that would literally minimize the degree to which others would be vulnerable to them. These inhumane conditions, which might involve straightjackets, sensory deprivation, and solitary confinement, would also run afoul of the limit at issue.

57. The critical discussion in this section is based on Staihar, *supra* note 20, at 289–92.

be a credible sign of reform if voluntarily undertaken independently of the state. In that case, the punishment would be entirely self-imposed. But no service could demonstrate reform if the state were to force the criminal to perform it by threatening to punish him even more severely if he refused. In that case, we would have no reason to believe that the criminal undertook the service with a benevolent motive as opposed to a selfish one.

For all we know, the offender might have made the sacrifices only to avoid the immediate threat of a more severe punishment. Alternatively, the criminal might have merely mimicked the behavior of a benevolent person in order to convince others to extend him the personal benefits of trust, such as better friendships and employment in the future.⁵⁸ The possibility of mimicry could undermine the potential for community service to restore a criminal's trustworthiness even when performed independently of the state.

Assuming there is no credible way for any criminal to signal his reform, then no criminal is obligated to restore his trustworthiness. "Ought" implies 'can.'⁵⁹ Hence, my proposed theory fails to explain why any criminal deserves any punishment.

B. FIRST RESPONSE

In response, several reasons can justify our believing that a criminal undertook his service with a benevolent motive even if it was part of a state-imposed punishment. First, we might have a mechanism for detecting the motive with which the service was performed. People seem to have a mechanism for detecting intentions. As Oliver Wendell Holmes said, even a dog can distinguish between being kicked and merely tripped over.⁶⁰ Assuming people can detect intentions, they might be able to detect whether a criminal performs his service for the sake of benefiting others or for the sake of something else, like promoting his own personal interests.

For example, assume a selfish criminal and a genuinely benevolent one were to engage in the same general type of community service while incarcerated. We can expect their motivational differences to manifest themselves in their behavior and morphologically at the level of their facial expressions and body language. Other things being equal, the selfish criminal would likely express more frustration and reluctance in making the required sacrifices over time. We can also expect to observe differences in their output.

58. See, e.g., Bacharach & Gambetta, *supra* note 34, at 157 (describing the concept of a mimic).

59. See, e.g., David Copp, 'Ought' Implies 'Can,' Blameworthiness, and the Principle of Alternate Possibilities, in MORAL RESPONSIBILITY AND ALTERNATIVE POSSIBILITIES: ESSAYS ON THE IMPORTANCE OF ALTERNATIVE POSSIBILITIES 265, 271-75 (David Widerker & Michael McKenna eds., 2003); IMMANUEL KANT, *The Metaphysics of Morals*, in PRACTICAL PHILOSOPHY 353, 513 (Mary J. Gregor ed., Mary J. Gregor trans., 1996) (stating that "he must judge that he *can* do what the law tells him unconditionally that he *ought* to do").

60. OLIVER WENDELL HOLMES, *THE COMMON LAW* 2 (American Bar Ass'n ed., 2009) (1881).

Other things being equal, the benevolent criminal would likely be more productive and do better quality work for others over time.

C. SECOND RESPONSE

Labor-intensive community service can be part of a criminal's state-imposed punishment without the states literally forcing him to perform it. In prison, the state might provide the criminal with some humane alternatives, thus leaving the criminal with a choice between performing the service and taking a non-helpful way out.

And even if the state requires all criminals to engage in some hard labor for a specified time, criminals can invariably choose to perform even more service for others and to put even more effort into producing better quality work for them. This is true of criminals who serve their sentences both inside and outside of a formal penitentiary.

D. THIRD RESPONSE

The possibility that a selfish criminal will mimic the behavior of a truly benevolent person presumes that he believes others would willingly extend him the personal benefits of trust in the future. However, these benefits, such as better friendships and employment, are not guaranteed to anyone, especially criminals. Thus, offenders are likely to be highly uncertain about whether they would receive any personal payoffs from signaling their reform. In light of this uncertainty, a criminal's performing his community service for the sake of benefiting others could be the inference to the best explanation of what really motivated him in choosing to perform it.⁶¹

E. FOURTH RESPONSE

Suppose an offender engages in labor-intensive community service for a long time but without the right motive. He makes the sacrifices for the sake of promoting his own long-term personal interests in anticipation of receiving the personal benefits of trust sometime in the future. Even so, the offender's service can still restore his trustworthiness to some degree.

Criminals tend to be impulsive, lacking in self-control.⁶² They tend to do what would best promote their own short-term personal interests; they tend not to be the kind of people who make significant short-term sacrifices for the sake of promoting long-term gains. When a criminal performs labor-intensive community service for the sake of promoting his own long-term interests, he signals that he has developed greater self-control. He demonstrates that he no

61. See generally Gilbert H. Harman, *The Inference to the Best Explanation*, 74 PHIL. REV. 88 (1965).

62. See, e.g., ROBERT H. FRANK, *PASSIONS WITHIN REASON: THE STRATEGIC ROLE OF THE EMOTIONS* 161–62 (1988) (noting that criminals are typically impulsive); GOTTFREDSON & HIRSCHI, *supra* note 23, at 85–120 (noting that criminals typically lack self-control).

longer gives as much priority to his own short-term interests in deciding what to do.

As a result, the threat of punishment will have a greater deterrent effect on the offender, as will the other negative long-term personal effects of crime. Being less impulsive, he will be more willing to work at opportunities for employment, education, and treatment that will make him better off in the long term and ultimately reduce his temptation to commit crimes. More generally, he will be in a better position to resist promoting his own short-term personal interests at the expense of harming others.

In summary, all criminals who undertake the right kind of state-imposed punishments in the right way under the right conditions can restore their trustworthiness to at least a partial degree. Assuming criminals can restore their trustworthiness by undertaking the relevant burdens, they are obligated to do so and, therefore, deserve to be punished in proportion to such burdens under my proposed view.

VI. THREE CONCLUDING REMARKS

A. AVOIDING OVERCRIMINALIZATION

This theory of punitive desert avoids many problems of overcriminalization.⁶³ As noted earlier, it presumes the essence of a crime consists in manifesting insufficient concern for the rights of others.⁶⁴ As a result, the minimally acceptable degree of trustworthiness relevant to my theory concerns only one's willingness to violate the rights of others. It does not concern one's willingness to cause others harms of less significance, such as mere offense or hurt feelings. Hence, under my proposed theory, the state is morally permitted to criminalize only those types of acts that unjustifiably risk violating the rights of others. The state is not morally permitted to criminalize mere harms to self, acts that cause others mere offense or hurt feelings, or acts of harmless wrongdoing, if there are any.⁶⁵

To clarify, though, I do not assume that the state ought all things considered to criminalize every type of act that unjustifiably risks violating the rights of others. Although such acts make one deserving of punishment under my view, the state might still have reasons of utility or justice not to criminalize them. For even if people deserve to be punished for committing certain acts, the costs of criminalizing those acts might still outweigh the benefits.

63. See generally DOUGLAS HUSAK, *OVERCRIMINALIZATION: THE LIMITS OF THE CRIMINAL LAW* (2008) (describing the problem of overcriminalization and defending various constraints on criminalization).

64. See *supra* text accompanying note 1.

65. See generally JOEL FEINBERG, *HARM TO OTHERS* (1984); JOEL FEINBERG, *HARM TO SELF* (1986); JOEL FEINBERG, *HARMLESS WRONGDOING* (1988); JOEL FEINBERG, *OFFENSE TO OTHERS* (1985); cf. JOHN STUART MILL, *ON LIBERTY* 145–46 (David Bromwich & George Kateb eds., 2003) (1859). Thus, infidelity causes hurt feelings in intimate relationships. But insofar as infidelity does not violate anyone's rights, it should not be criminalized.

Alternatively, criminalizing such acts could result in injustices to innocent third parties.

To illustrate, consider defamation. Those who commit libel or slander do manifest insufficient concern for the rights of others. Hence, they deserve to be punished under my proposed theory. Nevertheless, the state might not be morally permitted to criminalize defamation because doing so could have an intolerable chilling effect on the general public's freedom of expression.⁶⁶ Thus, all things considered, the state might be justified in treating defamation only as a tort, not a crime. As such, defamation would generate civil liability, but not criminal liability. In the final analysis, this theory of punitive desert is part of the solution to, not the problem of, overcriminalization.

B. *THE PROHIBITION AGAINST PUNISHING THE INNOCENT*

In addition to explaining why criminals deserve to be punished, my view also explains why an innocent person, who has not committed any wrongful acts, does not deserve any punishment. In virtue of being innocent, the person has not undermined his minimally acceptable degree of trustworthiness. To undermine his trustworthiness to a sufficiently bad degree, the person must perform an act that is sufficiently strong evidence of a sufficiently bad deficiency in his concern for the rights of others.⁶⁷ Only the commission of a crime without a range of exculpatory defenses passes the required evidentiary threshold.⁶⁸

C. *THE EXPRESSIVE CONTENT OF PUNISHMENT*

Punishing criminals does more than intentionally harm them. The punishment of a criminal expresses an attitude of moral blame toward him, such as an attitude of resentment or indignation.⁶⁹ My proposed theory of punitive desert illuminates part of the content of blame that punishment can express. On a standard view, attitudes of moral blame consist at least partly in

66. See *Garrison v. Louisiana*, 379 U.S. 64, 79–80 (1964) (Black, J., concurring) (arguing that statutes criminalizing defamation of public officials are unconstitutional under the First Amendment, and noting their general chilling effect on people's freedom of expression); Jean Hampton, *Correcting Harms Versus Righting Wrongs: The Goal of Retribution*, 39 UCLA L. REV. 1659, 1700 (1992) ("For example, a liberal state will believe that even though it has an obligation to respond in a retributive fashion to libel and slander and to a wide variety of racist and sexist literature, it also has an obligation not to threaten freedom of speech, which might occur as a result of such punishment.").

67. Of course, this is not to say that people will actually trust a trustworthy person. Irrational fears and prejudice abound. Thus, others might incur the costs of insecurity even in response to an innocent person. But an innocent person is not responsible for the costs that others incur because of their own irrational beliefs or ignorance about what he has done.

68. See, e.g., GOTTREDSON & HIRSCHI, *supra* note 23, at 259–61 (noting the serious inaccuracies in the best available strategies for predicting future criminal behavior when applied before any crimes have taken place); Dimock, *supra* note 20, at 53.

69. See Feinberg, *supra* note 54, at 98–101 (discussing the conceptual relation between punishment and the expression of moral blame).

certain demands and presuppositions.⁷⁰ When others blame someone for committing a crime, I suggest they presuppose that he undermined his trustworthiness by committing the crime, and they demand him to restore his trustworthiness expeditiously by undertaking certain burdens in order to signal his reform. This demand can be constitutive of both blame and punishment.

In addition to clarifying the content of one important demand that punishment can express, my theory also explains why this demand is warranted: Unless a criminal restores his trustworthiness expeditiously by undertaking the required burdens, he will pose an unacceptable risk of causing others certain especially significant harms, including the commission of further crimes and the costs of insecurity.

70. See, e.g., STEPHEN DARWALL, *THE SECOND-PERSON STANDPOINT: MORALITY, RESPECT, AND ACCOUNTABILITY* 17 (2006); STRAWSON, *supra* note 47, at 14–15, 21–22; Watson, *supra* note 47, at 258, 263–65.