
ARTICLES

INTUITIONS OF JUSTICE: IMPLICATIONS FOR CRIMINAL LAW AND JUSTICE POLICY

PAUL H. ROBINSON* AND JOHN M. DARLEY†

ABSTRACT

Recent social science research suggests that many if not most judgments about criminal liability and punishment for serious wrongdoing are intuitional rather than reasoned. Further, such intuitions of justice are nuanced and widely shared, even though they concern matters that seem quite complex and subjective. While people may debate the source of these intuitions, it seems clear that, whatever their source, it must be one that is insulated from the influence of much of human experience because, if it were not, one would see differences in intuitions reflecting the vast differences in human existence across demographics and societies.

This Article explores the serious implications of this reality for criminal law and criminal policy. For example, it may be unrealistic to expect the government to “reeducate” people away from their unhealthy interest in punishing serious wrongdoing, as is urged by some reformers, for it seems unlikely that the shared intuition that serious wrongdoing should be punished can be changed through social engineering, at least not through methods short of coercive indoctrination that liberal democracies would find unacceptable. Second, a criminal justice system that adopts rules that predictably and regularly fail to do justice or that regularly do

* Colin S. Diver Professor of Law, University of Pennsylvania. The authors thank Robert Kurzban for his useful comments and Scott A. Kaplan for his valuable research assistance.

† Dorman T. Warren Professor of Psychology, Princeton University.

injustice will inevitably be widely seen as failing in a mission thought important, even foundational, by the community, unless the system's unjust operation can be hidden, something that would be hard to do without breaching notions of press freedom and government transparency to which liberal democracies aspire. Finally, an understanding of the nature of people's intuitions of justice can provide more effective strategies for changing them. For example, it appears that legal and social reformers would do better not to fight people's shared intuitions of justice, but rather to harness them in service of their reform programs.

TABLE OF CONTENTS

I. THE NATURE OF JUDGMENTS ABOUT JUSTICE	3
A. JUDGMENTS ABOUT JUSTICE AS INTUITIONAL	4
B. INTUITIONS OF JUSTICE AS A HUMAN UNIVERSAL.....	8
II. IMPLICATIONS FOR THE ABOLITION OF PUNISHMENT	11
III. IMPLICATIONS FOR DISTRIBUTING PUNISHMENT IN A WAY THAT CONFLICTS WITH SHARED INTUITIONS OF JUSTICE.....	18
A. THE UTILITY OF DESERT.....	18
1. Harnessing the Power and Efficiency of Normative Social Influence: Stigmatization.....	19
2. Avoiding Vigilantism	21
3. Avoiding the Resistance and Subversion Produced by a System that Is Seen as Not Doing Justice	23
4. Shaping Societal Norms: The Persuasive Power of the Law	28
5. Gaining Compliance in Borderline Cases.....	29
6. Conclusion	31
B. CRITICISMS OF EMPIRICAL DESERT AS A DISTRIBUTIVE PRINCIPLE	32
1. Notions of Desert as Hopelessly Vague	32
2. Hopeless Disagreement as to Notions of Desert.....	36
3. Utility of Desert as Based Upon the Assumption that People Intuitively Assess Punishment According to Desert When in Fact They Look to Deterrence or Incapacitation.....	38
4. Empirical Desert as Inevitably Draconian	41
5. Empirical Desert as Potentially Immoral.....	43
C. DEVIATIONS FROM EMPIRICAL DESERT	45

1. Promoting Justice in a Complex World	46
2. Sacrificing Justice to Promote Other Interests	47
D. STRATEGIES FOR AVOIDING DEVIATIONS FROM SHARED INTUITIONS	49
1. Criminal Justice Reforms.....	49
2. Employing Civil Rather than Criminal Processes	50
IV. IMPLICATIONS FOR CHANGING INTUITIONS OF JUSTICE.....	51
A. PROPOSALS TO CHANGE INTUITIONS OF WRONGDOING AND BLAMEWORTHINESS	52
B. DIFFICULTIES IN CHANGING INTUITIONS	53
C. WHAT INTUITIONS OF JUSTICE CAN BE CHANGED?	57
D. HOW CAN INTUITIONS OF JUSTICE BE CHANGED?	60
1. Public Education to Manipulate the Analogy to Core Wrongs.....	60
2. Internalization of Group Beliefs	63
V. CONCLUSION.....	66

I. THE NATURE OF JUDGMENTS ABOUT JUSTICE

It is common for us to think that our judgments about deserved punishment for wrongdoing are entirely reasoned, like most judgments that we make in our daily lives: How fast should I drive on this stretch of road? How long should I heat this sandwich in the microwave? How carefully should I take notes on this lecture? These judgments, we may assume, are the product of reasoning based upon life experience and education. But social science evidence suggests that judgments about justice, especially for violations that might be called the core of criminal wrongdoing, are more the product of intuition than reasoning. Their intuitional nature means, among other things, that they are judgments quickly arrived at, even by people with little education or life experience, that they frequently are held with strong feelings of certainty, and that the reasons we reach such judgments with such certainty are generally inaccessible to us.¹

It is also common to assume that our intuitions of justice provide only vague judgments and that those judgments vary considerably among people. Yet, again, the common wisdom does not match the reality. Social science research demonstrates that people's intuitions of justice are quite nuanced and that, for the punishment of serious wrongdoing, our intuitions

1. See discussion *infra* Part I.A.

are widely shared across societies and demographics.²

As Parts II through IV demonstrate, these characteristics of our judgments about deserved punishment have important implications for criminal law and justice policy.

A. JUDGMENTS ABOUT JUSTICE AS INTUITIONAL

Preliminarily, we explain what we mean when we claim that judgments about deserved punishment for serious wrongdoing are “intuitions” or “intuitive judgments.” The primary claim is that humans have a system that produces intuitive judgments that is importantly different from the system that produces the deliberate operations of reasoned judgments. Since the 1970s, a group of judgment and decision-making psychologists have documented that people frequently produce their decisions via heuristic processes, in contrast to the conscious, deliberative processes that we think of as reasoning processes.³ Putting this another way, it is useful to distinguish between decisions arrived at by reasoning and decisions with similar content but arrived at via intuitive processes.

One of the earlier demonstrations of intuitive judgments was made by Amos Tversky and Daniel Kahneman.⁴ They showed that people, including statisticians who specialize in performing and teaching similar types of calculations, often gave quickly expressed intuitive solutions to statistical problems. Further, these intuitive judgments were frequently wrong, in the sense that they importantly deviated from the demonstrably correct calculational results. Yet further, the statisticians often relied on the results of these flawed intuitions not only when making quick, off-the-cuff judgments, but also when making consequential decisions about the conduct of research.

There followed three decades of research on “heuristics and biases” that demonstrated that people frequently use intuitive, short-cut methods to come to decisions and then go on to act on those decisions. In 2001, Robin Hogarth gave a thorough review of the domains of these intuitive

2. See generally Paul H. Robinson & Robert Kurzban, *Concordance and Conflict in Intuitions of Justice*, 91 MINN. L. REV. 1829 (2007) [hereinafter Robinson & Kurzban, *Concordance*] (presenting research regarding laypersons’ intuitions of justice). See also discussion *infra* Part I.B.

3. See Daniel Kahneman, *A Perspective on Judgment and Choice: Mapping Bounded Rationality*, 58 AM. PSYCHOLOGIST 697, 710 (2003) [hereinafter Kahneman, *Judgment and Choice*].

4. See generally Amos Tversky & Daniel Kahneman, *Belief in the Law of Small Numbers*, 76 PSYCHOL. BULL. 105 (1971) (discussing people’s intuitions about random sampling).

decisions,⁵ which led Tversky and Kahneman to suggest that these “intuitive judgments occupy a position—perhaps corresponding to evolutionary history—between the automatic operations of perception and the deliberate operations of reasoning.”⁶ The intuitive system involves processes that parallel those of the perceptions system, in that they “are typically fast, automatic, effortless, associative, implicit (not available to introspection), and often emotionally charged; they are also governed by habit and are therefore difficult to control or modify.”⁷ The processes of the reasoning system “are slower, serial, effortful, [and] more likely to be consciously monitored and deliberately controlled; they are also relatively flexible and potentially rule governed.”⁸

The processes of the intuitive system are summarized in Figure 1.⁹ As it shows, those processes are identical to those of the perception system, a fact that suggests why intuitions, like perceptions, are often taken by “the intuitor” as just as correct as all of us, functioning in the mode of perceiver, take our perceptions to be. Thus, there are a set of intuition-produced decisions, choices, and problem solutions that are experienced as summaries of the ways the world is, because the person having the intuitions is unaware of the complex, potentially incorrect, cognitive processes that produced them.

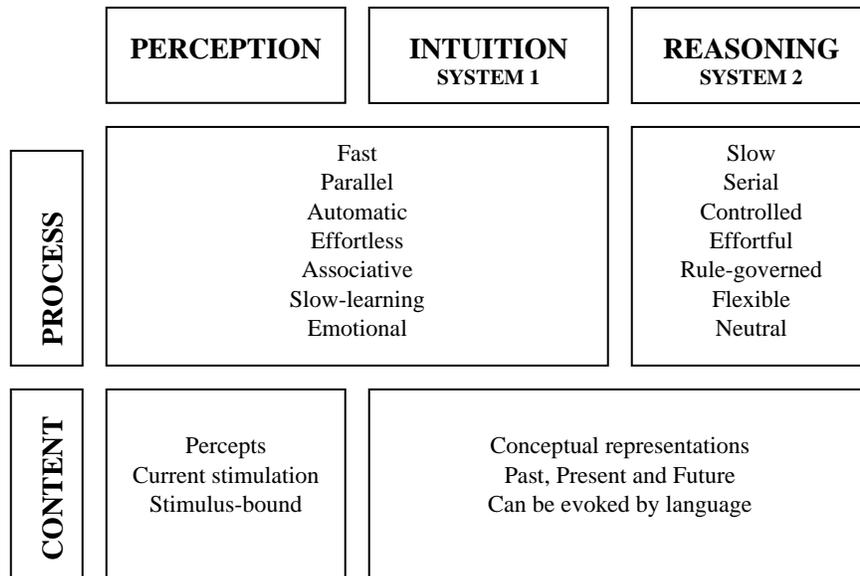
5. See generally ROBIN M. HOGARTH, *EDUCATING INTUITION* (2001).

6. See Kahneman, *Judgment and Choice*, *supra* note 3, at 697.

7. *Id.* at 698.

8. *Id.*

9. *Id.* at 698 fig.1.

FIGURE 1. Process and Content in Two Cognitive Systems¹⁰

As Kahneman explains, the perceptual system and the intuitive system “generate impressions of the attributes of objects of perception and thought. These impressions are neither voluntary nor verbally explicit.”¹¹ He then goes on to suggest that the assignment of degrees of “badness” to various stimulus situations is an intuitive system assessment, a suggestion which is quite relevant to our conceptualization of the intuitive judgments of blameworthiness of harm doers.

Some attributes, which Tversky and Kahneman (1983) called natural assessments, are routinely and automatically registered by the perceptual system or by . . . [the intuitive system] without intention or effort. . . .

The evaluation of stimuli as good or bad is a particularly important natural assessment. The evidence, both behavioral . . . and neuro-physiological . . . , is consistent with the idea that the assessment of whether objects are good (and should be approached) or bad (and should

10. This figure is reprinted with permission from Daniel Kahneman. See Kahneman, *Judgment and Choice*, *supra* note 3.

11. *Id.* at 699 (emphasis omitted).

be avoided) is carried out quickly and efficiently by specialized neural circuitry.¹²

In Jonathan Haidt's work on "moral dumbfounding," people report strong intuitions about things that are morally wrong, such as consensual, nonreproductive incest, but are unable to provide a principled explanation for their judgments.¹³ Similarly, Marc Hauser, Liane Young, and Fiery Cushman looked at judgments of morally permissible actions using the trolley problem,¹⁴ in which some number of people can be saved from being killed by a runaway trolley if some action is taken (or not taken), resulting in the death of a smaller number of people, often a single individual.¹⁵ People are asked what choice should be made in these difficult situations, and asked to explain their reasoning. While subjects commonly have strong and clear views on the proper result, they also commonly are unable to offer an explanation for their conclusions. For example, in one variation of the trolley problem, the subject can avoid the death of five people on the track by (a) pushing a bystander on the track whose body and backpack will jam the trolley's wheels, stopping it, but killing the person, or by (b) throwing a switch to divert the trolley to a side track where it will kill one person. Despite the fact that the results of the two actions are identical, 89 percent of the subjects considered the latter action moral but only 11 percent judged the former to be moral.¹⁶ More interesting for our purposes, 70 percent of the subjects could give no plausible explanation for their judgment.¹⁷ This mirrors Haidt's results concerning incest.¹⁸

Based on his review of the existing social science literature, Haidt

12. *Id.* at 701 (emphasis omitted).

13. See Jonathan Haidt, *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, 108 PSYCHOL. REV. 814, 814 (2001) [hereinafter Haidt, *Emotional Dog*].

14. For an example of a "trolley problem," see Philippa Foot, *The Problem of Abortion and the Doctrine of the Double Effect*, 5 OXFORD REV. 5, 8 (1967) (describing the problem of a driver of a runaway tram who can steer the tram either onto a track with five men working or onto a track with one man working).

15. Marc Hauser, Liane Young & Fiery Cushman, *Reviving Rawls' Linguistic Analogy, in MORAL PSYCHOLOGY AND BIOLOGY* (W. Sinnott-Armstrong ed., forthcoming), available at <http://www.wjh.harvard.edu/~mnkylab/publications/recent/ReviveRawlsChpt.pdf> (last visited Nov. 1, 2007).

16. *Id.*

17. See *id.* Even the 30 percent who gave what the authors classed as a "sufficient justification" may have been making purely ex post attempts at explanation rather than reporting reasoning they used in reaching their conclusion. The authors used what they called an "extremely liberal criterion": "A sufficient justification was one that correctly identified any factual difference between the two scenarios and claimed the difference to be the basis of moral judgment." *Id.*

18. See Haidt, *Emotional Dog*, *supra* note 13.

concludes that moral judgments derive from “quick, automatic evaluations (intuitions).”¹⁹ Similarly, Hauser concludes that “much of our knowledge of morality is . . . intuitive, based on unconscious and inaccessible principles”²⁰ In sum, there is at least some degree of consensus that many moral judgments are made by a deeply intuitive system. If such judgments were the product of a set of principles of morality learned from others, it would seem to be a straightforward matter to derive the “wrongness” of acts from these principles, just as mathematical inferences can be made from a set of axioms and subsequently explained with reference to them. Moral dumbfounding and related effects in the psychological literature suggest that this is not how these judgments are made.

To summarize, we are suggesting that the belief that serious wrongdoing should be punished and the culturally shared judgments of the relative blameworthiness of different acts of wrongdoing are commonly intuitive rather than reasoned judgments. This being so, these judgments come quickly to mind and are accompanied by strong feelings of certainty. The fact that these intuitions are the product of interpretative habits is obscured to the person because the processes that produce them are automatic and rapid, and leave no “mental traces.”

B. INTUITIONS OF JUSTICE AS A HUMAN UNIVERSAL

Beyond the intuitional nature of judgments about justice, a second characteristic with important implications for criminal law and justice policy is the high level of agreement about relative blameworthiness. As the following opinions demonstrate, the common wisdom among criminal law theorists has been that people’s intuitions of justice are vague and the subject of much disagreement.²¹ As one scholar put it, “[E]ven assuming retribution in distribution is appropriate, there is a classic epistemological problem. How do we know how much censure, or ‘deserved punishment,’ a particular wrongdoer absolutely deserves? God may know, but as countless sentencing exercises have shown, people’s intuitions about individual cases

19. *Id.* at 814. For a discussion of the automaticity of such judgments, see *id.* at 819–20.

20. MARC D. HAUSER, MORAL MINDS: HOW NATURE DESIGNED OUR UNIVERSAL SENSE OF RIGHT AND WRONG 125 (2006). For a recent review of relevant literature, see Cass R. Sunstein, *Moral Heuristics*, 28 BEHAV. & BRAIN SCI. 531 (2005).

21. See Paul H. Robinson, *Competing Conceptions of Modern Desert: Vengeful, Deontological, and Empirical*, 67 CAMBRIDGE L.J. (forthcoming 2008) [hereinafter Robinson, *Competing Conceptions*].

vary widely.”²² Another scholar states that:

There is . . . reason to doubt that anything like a consensus exists on the seriousness of criminal conduct. While there may be some agreement on relative levels of harm, there appears to be great variation in perceptions of the absolute magnitude of harm represented by various criminal acts, and in either the relative or absolute level of culpability represented by various criminal actors.²³

Empirical data, however, refutes these commonly held assumptions. Indeed, a variety of studies have shown that subjects share both the intuition that actors who engage in serious wrongdoing should be punished and a broad consensus about the relative blameworthiness of different types of transgression.²⁴ The disconnect between these studies and common wisdom is likely the result of a failure to distinguish between intuitions about the ordinal ranking of offenses and beliefs regarding where the cardinal endpoint of the punishment spectrum should be set. That is, people widely share intuitions about whether a given offense is more or less serious than another offense, but people and societies may disagree about what the punishment for the most serious offense will be. The decision to give an offender the death penalty or twenty years for murder, for example, is independent from the decision as to whether one is more or less

22. Michael Tonry, *Obsolescence and Immanence in Penal Theory and Policy*, 105 COLUM. L. REV. 1233, 1263 (2005).

23. John Monahan, *The Case for Prediction in the Modified Desert Model of Criminal Sentencing*, 5 INT’L J.L. & PSYCHIATRY 103, 105 (1982). Other authors have made similar arguments: [A desert theorist argues that] we can work out one single, linear ordering of crimes, from least to most “serious.” Yet that scarcely seems a credible assumption. Try, for instance, to rank the following crimes in order of their “seriousness”: attempted residential burglary, trading stock on inside information, negligent vehicular homicide, bribing a mine-safety inspector, possessing an ounce of cocaine, and burning a cross on the lawn of black newcomers to a previously all-white neighborhood. To view this motley assortment along a single dimension of “seriousness” would seem no less difficult than to perceive the inner logic behind the apocryphal Chinese encyclopedist of Jorge Luis Borges’s imagination.

David Dolinko, *Three Mistakes of Retributivism*, 39 UCLA L. REV. 1623, 1638–39 (1992). Similarly, Ernest van den Haag argues:

[Another desert theorist] appears to believe that the comparative seriousness of crimes can be determined in all cases. Not so. Comparative seriousness can be determined only for some crimes, and it does not fully determine the comparative punishment deserved. If rape is a crime and murder is a crime, rape-murder must be more serious than either. Does rape-murder deserve the sum of the punishments meted out for rape and for murder? More? Less? Even when crimes are nearly homogeneous, assigning seriousness is arbitrary: Is rape more serious than assault with a deadly weapon? Is burglary more serious than fraud when fraud does more harm? What about mishandling toxic waste? Ordinal determinations of seriousness become altogether arbitrary when the seriousness of heterogeneous crimes must be compared.

Ernest van den Haag, *Punishment: Desert and Crime Control*, 85 MICH. L. REV. 1250, 1254 (1987) (reviewing ANDREW VON HIRSCH, *PAST OR FUTURE CRIMES: DESERVEDNESS AND DANGEROUSNESS IN THE SENTENCING OF CRIMINALS* (1985)).

24. For a general discussion of these studies see Robinson & Kurzban, *Concordance*, *supra* note 2, at 1832–80.

blameworthy for committing homicide under extreme emotional distress or for the same crime committed during a botched robbery. People may vary widely in their views about the former but are surprisingly consistent in their intuitions about the latter.

Further, these intuitions on the relative blameworthiness of various actors have been shown to be both nuanced and consistent. This has been tested through varied study designs, including studies that ask the subjects (1) to put offenses or offense scenarios into one of a set of predetermined categories; (2) to rank order offenses or offense scenarios; or (3) to assign numerical values to each of a number of offenses or offense scenarios.²⁵

The studies confirm that subjects consistently differentiate between situations and that they share intuitions about how these variations affect the blameworthiness of the offender.²⁶ This was seen in a number of ways. First, punishment was uniformly imposed by subjects for serious wrongdoing.²⁷ Second, incremental changes in facts produce predictably significant changes in punishment.²⁸ Finally, subjects demonstrate a high degree of accord about the relative amount of punishment that is deserved for different offenses.²⁹ Perhaps surprisingly, cross-cultural studies have shown the same consistency as domestic studies and have been consistent with domestic studies in the ordinal rankings derived.³⁰ These results demonstrate the consistency of the ordinal ranking of different cases not only within cultures but also among them.³¹ Commentators looking at the results of these studies have come to conclusions that, “it is apparent that there was considerable agreement as to the amount of official punishment appropriate to each act” and that there was “general agreement in . . . [relative rankings] across all countries.”³² In Part III.B.2, we will look more closely at this issue and review one recent study in particular that demonstrates the astounding level of agreement on nuanced differences.

One might speculate about the possible explanations for this surprising

25. *Id.* at 1837–40.

26. *See id.*

27. *See id.* at 1848–54.

28. *See id.*

29. *See id.* at 1854–65.

30. *See id.* at 1857–58, 1862–65.

31. For a discussion of why this may be the case, see Paul H. Robinson, Robert Kurzban & Owen D. Jones, *The Origins of Shared Intuitions of Justice*, 60 VAND. L. REV. (forthcoming 2007).

32. Robinson & Kurzban, *Concordance*, *supra* note 2, at 1864 (quoting GRAEME NEWMAN, *COMPARATIVE DEVIANCE: PERCEPTION AND LAW IN SIX CULTURES* 140, 141 (1976)).

situation.³³ People might share intuitions of justice because those intuitions are the product of an evolutionarily developed mechanism that predisposes each human toward acquiring these intuitions. This explanation is consistent with the many characteristics that the intuitive system shares with the perceptual system, discussed above. One also might speculate, however, that people share intuitions of justice because such views are efficient norms for persons in a group to hold, or perhaps because of some combination of evolutionary and social influence—perhaps each human is predisposed by an evolutionarily developed mechanism to adopt the intuitions of justice of the group into which he is born and socialized, and groups tend toward such intuitions of justice as an efficient norm.

Whatever the source of intuitions of justice, it is beyond the normal influence of culture or demographic. If it were not so insulated, one would see differences in intuitions of justice among different demographics and cultures. This insulation means that there may be serious limits on whether and how social engineers can manipulate intuitions of justice, at least those intuitions of justice about core wrongdoing upon which there is broad agreement.

Therefore, these findings regarding the nature of intuitions of justice have serious implications for a variety of criminal justice debates that focus on substantial alterations of criminal justice systems, including the abolition of punishment, the distribution of punishment according to principles that conflict with shared intuitions of justice, and programs to change people's intuitions about what constitutes serious wrongdoing and about how much it should be punished, discussed in Parts II through IV, respectively. It is not our view that intuitions of justice are absolutely unalterable. On the contrary, in Part IV.D we use lessons learned here to show how intuitions of justice, especially those out from the core, might most effectively be altered. Given the breadth and stability of intuitions of justice, however, some alterations may require tactics that cannot be supported by liberal democratic societies.

II. IMPLICATIONS FOR THE ABOLITION OF PUNISHMENT

We begin by discussing the “abolitionist proposal.” We take this to be the proposal that any punishment component included in treatments assigned to criminal offenders is probably primitive, certainly misguided, and should have no part in the criminal justice system. This is meant to ban

33. For further discussion of possible explanations for the wide agreement on core intuitions of justice see Robinson, Kurzban & Jones, *supra* note 31.

punishments assigned to achieve retributive, just desert purposes and perhaps others. The movement to abolish punishment reflects these views:³⁴

At its core, criminal law is still based on the same repressive assumptions as the Inquisition from which it originates. From the beginning it has been seen to create problems instead of solving them. A penal reaction after the fact is not preventive but desocialises an ever-increasing number of people. Therefore it would be better to abolish penal means of coercion, and to replace them by more reparative means. This briefly is the abolitionist message.³⁵

In the abolitionists' view, what then should replace punishment?

First, we should think of "conflicts," or "troubles," rather than "crimes"; abolitionists offer a critique not only of punishment but of the concept of crime as that which requires or merits punishment.

Second, we should look for reparation, restoration, and reconciliation, not for retribution and punishment ("pain-delivery"); instead of inflicting penal pain on wrongdoers, as retribution or for deterrence, we should seek negotiated reparations for those who have been harmed, the restoration of relationships between the parties to conflicts, their reconciliation with each other and with the community.³⁶

The most popular modern form of the abolitionist movement is the "restorative justice movement." While it uses the term "justice" in its description, its platform is explicitly to reject the use of retributive punishment components, which most people would see as critical components to the practice of doing justice.³⁷ One of its originators, John

34. For examples of abolitionist arguments see ABOLITIONISM: TOWARDS A NON-REPRESSIVE APPROACH TO CRIME (Herman Bianchi & René van Swaaningen eds., 1986) [hereinafter ABOLITIONISM]; HERMAN BIANCHI, JUSTICE AS SANCTUARY: TOWARD A NEW SYSTEM OF CRIME CONTROL (1994); NILS CHRISTIE, LIMITS TO PAIN (1981); WILLEM DE HAAN, THE POLITICS OF REDRESS: CRIME, PUNISHMENT AND PENAL ABOLITION (1990); THOMAS MATHIESEN, THE POLITICS OF ABOLITION (1974); THOMAS MATHIESEN, PRISON ON TRIAL: A CRITICAL ASSESSMENT (1990); Nils Christie, *Conflicts as Property*, 17 BRIT. J. CRIMINOLOGY 1 (1977) ("civilizing" our responses to crime); Louk Hulsman, *The Abolitionist Case: Alternative Crime Policies*, 25 ISR. L. REV. 681 (1991); Louk H.C. Hulsman, *Critical Criminology and the Concept of Crime*, 10 CONTEMP. CRISES 63 (1986).

35. René van Swaaningen, *What is Abolitionism?*, in ABOLITIONISM, *supra* note 34, at 9.

36. R. A. Duff, *Penal Communications: Recent Work in the Philosophy of Punishment*, 20 CRIME & JUST. 1, 69–70 (1996) (explaining the values and ideas of the abolitionist movement). The passage continues by also noting a preference for "informal, participatory methods of resolving conflicts." *Id.* at 70.

37. See Paul H. Robinson, *The Virtues of Restorative Processes, the Vices of "Restorative Justice"*, 2003 UTAH L. REV. 375, 377–78 (2003) [hereinafter Robinson, *Virtues*]. As we will discuss later, however, not all "restoratists" support the abolition of punishment. See *id.* at 375–78 (explaining how many advocates of "restorative processes" do not support the antijustice agenda of the restorative justice movement).

Braithwaite, explains:

Restorative justice is a process of bringing together the individuals who have been affected by an offense and having them agree on how to repair the harm caused by the crime. The purpose is to restore victims, restore offenders, and restore communities in a way that all stakeholders can agree is just. One value of restorative justice is that we should be reluctant to resort to punishment. . . . It involves rejection of a justice that balances the hurt of the crime with proportionately hurtful punishment.³⁸

Contrary to Braithwaite's claim, the "eye for an eye" notion of doing justice to which he refers is rejected by most modern desert theorists, who take full account not just of the hurt of the crime but the characteristics and circumstances of the offender, and whose central distributive criteria is not the victim's hurt but rather the offender's blameworthiness.³⁹ The pure form of restorative justice pressed by Braithwaite, however, is one that rejects the notion of punishment on any grounds.⁴⁰

We suggest that this is not a proposal that societies will or should adopt, for two reasons. First, a majority of individuals in the society would strenuously resist the abolition of punishment because the impulse to punish serious wrongdoing is deeply ingrained. Second, having punishment available to administer to norm violators greatly reduces the frequency of norm violations that occur. We do not assert that the lengthy prison terms that the criminal justice system currently imposes on offenders, coupled with the "prisonization"⁴¹ that is inflicted on offenders in prisons, is necessary to reduce the frequency of criminal violations that exist in our society. We do say that the absence of punishment mechanisms in a society would lead to a set of violations sufficient to threaten the existence of the society.

The abolitionists have on rare occasion addressed the empirical issue of whether it is indeed realistic to persuade people to no longer care about doing justice or punishing serious wrongdoing. For example, Braithwaite

38. John Braithwaite, *A Future Where Punishment Is Marginalized: Realistic or Utopian?*, 46 UCLA L. REV. 1727, 1743 (1999) [hereinafter Braithwaite, *Future*].

39. See Robinson, *Competing Conceptions*, *supra* note 21.

40. See Robinson, *Virtues*, *supra* note 37, at 375. One of us argued that restorative processes, such as victim-offender mediation and sentencing circles, are wonderful procedures that ought to be used much more extensively than they are today, but are not in large measure because leaders like Braithwaite use the restorative justice movement as a vehicle in what is their first priority, their antipunishment agenda. *See id.*

41. The term "prisonization" refers to the "(usually negative) changes that a person undergoes while in prison." Dena M. Gromet & John M. Darley, *Restoration and Retribution: How Including Retributive Components Affects the Acceptability of Restorative Justice Procedures*, 19 SOC. JUST. RES. 395, 395-96 (2006).

commented that:

Robinson . . . pleads in his conclusion for restorative justice to recognize and respond to the emotional need—“inherent in human nature”—that people have for retribution. Indeed I do suspect this is a desire deeply rooted in our biology that was probably once useful to the survival of peoples fighting off starvation by defending scarce assets of productive land. A person’s deeply felt desires for retribution should be addressed openly in conferences. However, the restorativist’s hope is that the conversation about the urge for retribution will result in it being transcended so that people can move on. The reason restorativists think this way is that they believe peoples’ natural retributive urges are not healthy things to perseverate upon. Moreover, in the conditions of contemporary societies, as opposed to the conditions of our biological inheritance, retribution is now a danger to our survival and flourishing. It fuels cycles of hurt begetting hurt. It is hoped that conversations that allow a space for the consideration of healing will help people to see this more clearly.⁴²

These are socially optimistic sentiments, but they are only intellectual ones. Braithwaite may be overoptimistic that his intellectual arguments about what would be good for people and society today—his hope to “help people see . . . more clearly”—can undo the evolutionary and social-function bases for the intuitions of justice that people have.

Against the weight of the evidence showing how fundamental the desire to do justice is to human nature and how difficult it is to erase it, Braithwaite has little to offer:

What is the proper punishment for a given type of crime? This question has obsessed the most distinguished philosophers of criminal justice. But it seems a silly question, for why should one assume that any punishment is the proper response to a crime? Why not assume that punishment is rarely the best way to respond to crime? The former assumption seems grounded in a failure of imagination and ignores the practices that citizens of all cultures utilize instead of punishment in responding to wrongdoing. We might do better to follow the lead of many Native American peoples who believe in putting the problem rather than the person at the center of this deliberation. The “right” punishment of the wrongdoer is rarely going to be the best solution to the problem.⁴³

Are we to believe that Native American peoples did not and do not

42. John Braithwaite, *Holism, Justice, and Atonement*, 2003 UTAH L. REV. 389, 407 (footnote omitted).

43. Braithwaite, *Future*, *supra* note 38, at 1728 (footnotes omitted).

believe in doing justice and do not regularly impose punishment for significant moral transgressions? We know of no society in which there are either no intuitions of justice or no use of punishment.⁴⁴ Indeed, as a cross-cultural review of the existing studies confirms, humans across a wide range of cultures share the intuition that serious wrongdoing should be punished.⁴⁵

Societies other than ours may have more successful systems for the control of deviance, or may have better systems of inculcating moral rules into their citizens, so that the occasions for punishments may arise less often. However, even if this is true, it does not disprove the contention that they do have and use punishment when the need arises.

Braithwaite is not the only restorativist to deny the empirical problem of making people not care about punishing serious wrongdoing:

We frequently must set aside our natural human responses (for example, jealousy, lust, rage and anger) and adopt more considered and reflective dispositions, because of the harm which they cause. So too could it be argued with intrinsic retributivism. Of course, it could be countered that as a psychological matter, our make-up is such that it is not possible to react in a manner other than to punish wrongdoers: inferring, therefore, that it is wasteful even to consider (other) justifications for punishment, since there is no point in arguing against that which cannot be curtailed. This, however, is repudiated by the sheer number of counter-examples where people bear no animosity towards those who have violated their important interests. Indeed, there have been loud calls for the abolition of punishment.⁴⁶

Again, what is being asserted here puzzles us. We think that it is well worth considering “other justifications for punishment.” We also agree that there are many examples “where people bear no animosity towards those who have violated their important interests,” although we suspect that, in those cases, the violators have either explained why their violations were not exactly intentional, provided recompense, or more likely both. This is different, however, from the claim that one can identify societies that have, as a matter of policy, eliminated punishments that serve a retributive impulse, and successfully maintained social control of wrongdoing.

There is, after all, an empirical question here. A recent study explored

44. See Robinson & Kurzban, *Concordance*, *supra* note 2, at 1852.

45. See *id.*

46. Mirko Bagaric & Kumar Amarasekara, *The Errors of Retributivism*, 24 MELB. U. L. REV. 124, 159 (2000) (emphasis omitted).

the acceptance of purely restorative justice practices in American society.⁴⁷ The researchers asked people to role play “referring judges” whose task it was to send specific offense cases to one of three alternative court systems.⁴⁸ One system was described to respondents as a purely restorative system,⁴⁹ another as a standard prison focused system, and a third that included the restorative procedures, but also allowed for punitive alternatives.⁵⁰

Results indicated that the respondents were willing to send just under 80 percent of the less serious cases to a purely restorative procedure, but as the seriousness of the crimes increased the respondents were more drawn to a restorative procedure that included the possibility of punitive measures being assigned.⁵¹ For midlevel serious crimes, about 40 percent were sent to pure restorative procedures and 40 percent to the mixed procedure.⁵² For high-seriousness cases, only about 10 percent of the cases were assigned to purely restorative procedures, and 66 percent to the mixed procedure.⁵³

Even these results may tend to exaggerate the willingness to forgo punishment. One might conclude that restorative justice was accepted as an option for the less serious cases only because the subject believed that the range of outcomes available in that process was sufficient to provide the punishment deserved. That is, a victim may see restorative mediation as deserved punishment. Whether it is or not will depend upon how strongly the particular participants at hand are committed to seeing justice done.

47. Gromet & Darley, *supra* note 41, at 395.

48. *See id.* at 400.

49. The purely restorative system described was the Victim Offender Mediation (“VOM”), explained as follows:

The VOM is a face-to-face meeting between an offender and his victim, with a facilitator (a neutral third party who has prepared both sides for the meeting beforehand) present. Sometimes, there may be multiple victims, or even members of the community or friends/relatives of the parties present.

During this meeting, the victim is allowed to ask the offender any question he or she wishes. During the meeting, the victim and offender may work out an agreement outlining what the offender must do to atone for his wrongs and make the victim whole again. The terms of the agreement may include an apology, monetary compensation, some services that the offender does for the victim, community service, and the like.

.....
Although there are varieties of possible outcomes of a VOM, imprisonment is not one of the possibilities. As an expert in restorative justice states, “[r]esponding to the hurt of the crime with the hurt of punishment is rejected, along with its corresponding value of proportionality-punishment that is proportionate to the wrong that has been done.”

Id. at 426 (references omitted).

50. *Id.* at 400.

51. *Id.* at 408 fig.2.

52. *Id.*

53. *Id.*

In one way, these results should please restorative justice advocates, for this is one of the first studies that demonstrates that people are accepting of a restorative justice system under some circumstances. The conclusion that is less palatable for them, however, is “that in order for citizens to view a restorative justice procedure as an acceptable alternative to the traditional court system for serious crimes, the procedure must allow for the option of some retributive measures.”⁵⁴

One of us has argued elsewhere that restorative processes, such as victim-offender mediation and sentencing circles, are wonderful procedures that should be used much more widely than they are today.⁵⁵ First, a good many restorative inflictions, such as hours of service to the victim, have punitive elements. Second, there is no barrier to sincere attempts to restore comity between offender and victim and between offender and society taking place in criminal proceedings. For offenses in which intuitions of justice demand retributive impositions, however, people will see justice as demanding those impositions.⁵⁶

One group of scholars, sympathetic to restorative justice practices, rejects the idea that retributive practices should be banned from procedures that seek restorative ends. Kathleen Daly states:

By framing justice aims (or principles) and practices in oppositional terms, restorative justice advocates not only do a disservice to history, they also give a restricted view of the present. They assume that restorative justice practices should exclude elements of retribution; and in rejecting an ‘attitude of hostility’, they assume that retribution as a justice principle must also be rejected.

When observing conferences, I discovered that participants engaged in a flexible incorporation of multiple justice aims, which included:

- (1) some elements of retributive justice (that is, censure for past offences);
- (2) some elements of rehabilitative justice . . . ; and
- (3) some elements of restorative justice⁵⁷

R. A. Duff introduces a recent paper this way:

My thesis can be stated quite simply. Our responses to crime should aim for ‘restoration’, for ‘restorative justice’: but the kind of restoration that

54. *Id.* at 395.

55. See Robinson, *Virtues*, *supra* note 37, at 375.

56. See *id.* at 380.

57. Kathleen Daly, *Restorative Justice: The Real Story*, 4 PUNISHMENT & SOC’Y 55, 59 (2002) (emphasis omitted).

criminal wrongdoing makes necessary is properly achieved through a process of retributive punishment. To put it the other way round, offenders should suffer retribution, punishment, for their crimes: but the essential purpose of such punishment should be to achieve restoration. To put it yet more simply, my slogan is 'Restoration through retribution'.⁵⁸

To conclude, even if one were convinced intellectually that abolition of punishment were a desirable ideal, it is a reform that could never be successfully implemented and it would be folly to try to do so. We must face the reality that human beings will demand justice for serious wrongdoing, and that the absence of a system that allows for the imposition of deserved punishment would produce intolerable consequences, such as people undertaking to do justice themselves. A fuller account of the serious consequences is detailed in Part III.

III. IMPLICATIONS FOR DISTRIBUTING PUNISHMENT IN A WAY THAT CONFLICTS WITH SHARED INTUITIONS OF JUSTICE

Abolition of punishment altogether is not likely to ever have much political support. Even when dressed up as Braithwaitian restorative justice, such reforms are not likely to succeed if they conflict with a community's shared intuitions of justice. But there exist a wide variety of reasons that have driven reformers to adopt distributions of punishment that conflict with people's shared intuitions of justice (referred to as "empirical desert"). We will examine in greater detail such deviations from empirical desert in Part III.C, but we may note here that it is not uncommon for the deviations to occur in both directions, sometimes giving less punishment than community intuitions of justice would suggest, or even more commonly, giving greater punishment, often justified by reason of general deterrence or incapacitation. We argue that deviations in either direction can have undesirable consequences and unjustified costs that can ultimately hurt rather than help effective crime control.

A. THE UTILITY OF DESERT

There is a case to be made that there is great utility in a criminal justice system that provides for a distribution of liability and punishment in concordance with the citizens' shared intuitions of justice.⁵⁹ Agreement

58. See R. A. Duff, *Restorative Punishment and Punitive Restoration*, in *RESTORATIVE JUSTICE AND THE LAW* 82, 82 (Lode Walgrave ed., 2002).

59. For an earlier, preliminary account of these arguments, see Paul H. Robinson & John M.

between the criminal code and citizens may provide utility greater than through distribution of liability and punishment in the more traditional utilitarian manner of optimizing deterrence or incapacitation. The case for this emerges when one considers why most people generally obey the law, even on occasions when law-breaking is unlikely to engage the armamentarium of arrest, conviction, and punishment that the criminal justice system has at its disposal. That is, given the generally weak deterrent threat facing people, why do the vast majority of those societal members still act in a way consistent with the law? Social scientists have two answers to give: people obey the law because (1) they regard the law as representing the principles that moral people adhere to, and they are socialized in such a fashion as to want to live up to those moral rules; and (2) if the law specifies morally proper conduct, people naturally believe that the community believes in the “righteousness of the law” and so people fear the disapproval of their social groups if they violate the law.⁶⁰ These factors are referred to as (1) “behavior produced by internalized moral standards and rules,” and (2) “compliance produced by normative social influence,” generally involving concerns for the sanctions that others will inflict when one violates the accepted rules of conduct.⁶¹ Criminal law can have influence on people’s conduct through both of these mechanisms.

1. Harnessing the Power and Efficiency of Normative Social Influence:
Stigmatization

Actors generally feel the force of a social norm as an external force impinging on them, which is not unlike the weight exerted by the general knowledge that the criminal deterrence system has a host of penalties that await transgressors.⁶² The sanctions feared for social norm transgressions, however, are generally experienced as coming from the community. People obey the social norms of their communities for a number of reasons. Violating those social norms may involve the loss of one’s standing in the community gained from one’s past achievements, losing the benefits that flow from valued relationships with others, and suffering the shunning experienced by those who have become stigmatized within their

Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453 (1997). *See also* Kent Greenawalt, *Punishment*, 74 J. CRIM. L. & CRIMINOLOGY 343, 359 (1983) (“The idea is that since people naturally think in retributive terms, they will be disenchanted and eventually less law-abiding if the law does not recognize that offenders should receive the punishment they ‘deserve.’”).

60. *See* Robinson & Darley, *The Utility of Desert*, *supra* note 59, at 468.

61. *See id.* at 468–69.

62. *See id.* at 469.

communities.⁶³ “If one is thought to have committed a crime, one may lose one’s job, ability to borrow money, ability to command trust from others, and possible business partners.”⁶⁴

Other scholars have made similar observations based on the concepts of social capital and reciprocal trust. They have noted that generalized trust depends on the willingness of individuals to adhere to their community’s social norms; moreover, trust and social capital are viewed as products of people’s compliance with these social norms. Increasingly, social scientists have come to realize the social value of having a reputation as trustworthy. Untrustworthy people cannot be counted on to do their share of the countless social exchanges that life in communities depends on. Thus, one seen as not living up to community norms signals, among other things, an unwillingness to live up to the future commitments made in exchange for benefits that others might choose to give in the present.⁶⁵ Criminal behavior, known to the community, even if not resulting in formal criminal liability, is a strong signal that a person is not one in which the community should reside trust or confidence.

Formalizing this general perspective, Eric Posner has created a signaling version of social norm theory. One obeys social norms in the present to signal to others that one will behave cooperatively in future social exchanges, and being included in these exchanges is critically important for one’s own benefit.⁶⁶ In other words, one attempts to develop a reputation for normatively trustworthy behavior so that one can trade on or benefit from that reputation in the future.

It is important to note that people generally do not want to violate social norms, and if they think that legal codes map those social norms,

63. *See id.* *See also id.* at 472–75 (discussing the relationship between criminal law and societal norms).

64. *Id.* at 469.

65. *See* Jack Knight, *Social Norms and the Rule of Law: Fostering Trust in a Socially Diverse Society*, in *TRUST IN SOCIETY* 354, 354–65 (Karen S. Cook ed., 2001).

66. *See* ERIC A. POSNER, *LAW AND SOCIAL NORMS* 18–27 (2000). Notice the reliance on the motive of self-interest here, which is characteristic of the law and economics movement. Norms, unlike laws, are not published in codified structures, so it is possible to wonder how these apparently invisible social rules get learned. People commonly learn norms from the reactions of observers when a norm is either lived up to or failed. Experienced actors know how to interpret these expressions, and they alter their subsequent behaviors accordingly. Further, observers often deliver not only expressions, but also immediate rewards for living up to norms and sanctions for failing to do so. People learn that the likelihood of longer-term sanctions is generally signaled by these immediate cues of distress or disdain when social norms are violated. A good deal of the learning that takes place during childhood socialization lies in learning how to see initial signs of disapproval and modify one’s behaviors accordingly.

they obey the laws. Violators of criminal laws are stigmatized by their community, which is a successful way that the criminal justice system controls conduct. Compared to the financial and social costs of using the system of arrest, trial, conviction, and imprisonment to deter potential offenders, stigmatization can be more powerful and is essentially cost-free. This wonderfully efficient result will occur, though, only if the community views violating a legal code as violating a social norm. The criminal justice system's power to stigmatize depends on the legal codes having moral credibility in the community. The law needs to have earned a reputation for accurately representing what violations do and do not deserve moral condemnation from the community's point of view. This reputation will be undercut if liability and punishment rules deviate from a community's shared intuitions of justice. For example, in some inner-city communities, in which very high proportions of young African-American males have done prison time for actions that the community does not regard as criminal, it may be that being an "ex-convict" no longer stigmatizes the individual in that community.⁶⁷

2. Avoiding Vigilantism

Vigilantism may be the most dramatic reaction to a perceived failure of justice.⁶⁸ Vigilante justice, descriptively, refers to occasions in which groups of citizens come together to enforce rules that are not otherwise being enforced by the formal forces of the legal system. Historically, vigilante groups often sprang up in newly settled areas without real law enforcement mechanisms. Vigilante action, however, also occurs even when police and courts are available if citizens perceive that the criminal justice system is failing in its responsibilities to punish wrongful acts.

Some of these latter cases make the reader quite vividly understand

67. See generally James P. Lynch & William J. Sabol, *Assessing the Effects of Mass Incarceration on Informal Social Control in Communities*, 3 CRIMINOLOGY & PUB. POL'Y 267 (2004) (discussing evidence that suggests that incarceration may negatively impact social control in neighborhoods); Dina R. Rose & Todd R. Clear, *Incarceration, Social Capital, and Crime: Implications for Social Disorganization Theory*, 36 CRIMINOLOGY 441 (1998) (describing the impact of criminal convictions on communities and local social means of controlling crime).

68. For discussions on vigilantism, see Ray Abrahams, *What's in a Name? Some Thoughts on the Vocabulary of Vigilantism and Related Forms of 'Informal Criminal Justice,'* in INFORMAL CRIMINAL JUSTICE 25 (Dermot Feenan ed., 2002) (discussing vigilantism and "informal criminal justice" generally); Rebekah Lee and Jeremy Seekings, *Vigilantism and Popular Justice After Apartheid*, in INFORMAL CRIMINAL JUSTICE, *supra*, at 99 (examining vigilantism more specifically in the context of South Africa); Elisabeth Ayyildiz, *When Battered Woman's Syndrome Does Not Go Far Enough: The Battered Woman as Vigilante*, 4 AM. U. J. GENDER & L. 141, 146-49 (1995) (discussing vigilantism and how battered woman's syndrome fits into the concept).

the impulses that drive citizens to vigilante action. In one case, occurring in 1981 in a small town in Missouri, a man had, for some years, been terrorizing the town, raping young girls, robbing farms for antiques, and blatantly intimidating anyone who accused him of crimes or who might testify against him in the various trials in which he was charged with these crimes. The story includes a lawyer who was constantly manipulating the laws to this person's advantage. The police in this small town were probably also intimidated, and were unable to keep the man from committing crimes and intimidating people. One day, the man was standing watch out in his truck in front of the place of business of a grocer whom he had threatened with death. This proved to be bad timing on his part. At their wit's end about what to do, the townsfolk were meeting nearby, and a large number of them came out of their meeting to see the man sitting in his truck, intimidating the grocer. He was shot with several weapons and killed. The circumstances suggest that the crime had been witnessed, but nobody could be found, then or ever, to report what they had witnessed.⁶⁹

The case described above illustrates a classic vigilante response to an obviously intolerable situation. As Peter French identifies, three types of mental states drive the urge to action.⁷⁰ The first is the deep and personal feelings the citizens have of being deprived, in this case of justice.⁷¹ The second is a long-held feeling of powerlessness, or in this case the shocking powerlessness of the law to do what is just.⁷² Finally, these two feelings generate further feelings of "contempt, hatred, scorn, disdain, and loathing" for the offender.⁷³ These are powerful responses, and they motivate equally powerful actions. We would suggest, in the case in which the legal system is tolerating some actions that citizens consider morally wrong, that these cognitive/affective appraisals attach themselves differentially to the wrongdoer and the legal system. Hatred and loathing are directed toward the offender; contempt, scorn, and disdain toward the legal system that tolerates the offense. When the system punishes a citizen for some action that the community considers morally acceptable, reactions of sympathy and support are generated toward the unjustly punished person, and again, contempt, scorn, and disdain are generated toward the legal system, along with, we suspect, fear and anger.

69. See generally HARRY N. MACLEAN, IN BROAD DAYLIGHT (1988) (describing the story of the murder of Ken Rex McElroy in Skidmore, Missouri).

70. See PETER A. FRENCH, THE VIRTUES OF VENGEANCE 6 (2001).

71. *Id.*

72. *Id.*

73. *Id.*

3. Avoiding the Resistance and Subversion Produced by a System that Is Seen as Not Doing Justice

The resort to vigilantism is a dramatic reaction to the system's failure to do justice, one in which the disaffected persons move outside of the criminal justice system. A similar dynamic results, less dramatically, but more commonly, when deviations from perceived deserved punishment prompt individuals to pervert the operation of the criminal justice system itself, sometimes to produce punishment that the law normally would not provide, but sometimes to prevent punishment that the law would provide.

We argue that in communities in which the liability and punishment rules regularly deviate from the consensus of the community members, those criminal law rules are diminished in their moral credibility. The criminal justice system depends on those involved in it (offenders, judges, jurors, witnesses, prosecutors, police, and others) for its operation. For the system to function effectively, these people must cooperate or, at least, acquiesce to the system's demands. Otherwise, if the system is regarded as being in conflict with justice or simply failing to do justice, this critical cooperation or acquiescence may diminish or cease to exist at all. Moreover, to the degree that these deviations from justice are frequent and morally consequential, active forces of subversion and resistance are generated in the community.⁷⁴

In general, there are two ways that the criminal justice system may deviate from the community's intuitions about appropriate criminal laws: by failing to punish actions that the community thinks are morally wrong, and by punishing actions that the community regards as morally innocent. Depending on which of these intuitions is violated, the community may respond with different specific actions. If the law "allows" actions that the community regards as condemnable, then in some cases, the community will seek to mobilize informal methods that restrict and sanction these activities, as with the vigilantism discussed above. If, on the other hand, the law criminalizes actions that the community or segments of the community

74. See, e.g., Daniel J. Bell, *Family Violence in Small Cities: An Exploratory Study*, 12 POLICE STUD.: INT'L REV. POLICE DEV. 25, 30-31 (1989) (finding empirically that police in small cities are likely to subvert the law regarding domestic violence by not reporting incidents, not arresting violent offenders, or deferring to other agencies; also reporting that police respond in such a manner because they believe domestic violence to be a family matter); Lewis R. Katz & Robert D. Sweeney, Jr., *Ohio's New Drunk Driving Law: A Halfhearted Experiment in Deterrence*, 34 CASE W. RES. L. REV. 239, 240 (1984) ("The frequency of drunk driving creates a tendency to view this criminal behavior without the condemnation warranted by its often tragic consequences. Until recently, the law's treatment of drunk drivers has been shaped by this tolerant attitude, and enforced and administered by police, prosecutors and judges affected by [leniency due to their own participation in drunk driving].").

think are morally permissible, then it is likely that those actions will continue to be practiced by people, but will “go underground.” This in turn generates other consequences: the rise of entrepreneurs who profit by facilitating people’s indulgence in those activities, venues in which those activities are practiced, pay to authorities willing to overlook the commission of these actions, and so on.

In addition to these undesirable reactions attempting to engage or escape control mechanisms for these actions, we suggest that when the criminal justice system is seen as out of tune with community sentiments, a less obvious but more common and troublesome reaction occurs in the loss of moral credibility that the justice system brings on itself. In general, these reactions can be summed up as the justice system’s loss of relevance as a guide to good conduct. Initially, the reactions may be limited to conclusions about the idiocy of the specific legal rule that offends the community’s morality, but as the apparatus of the social control agents of the government, such as police and the courts, are mobilized to enforce the senseless laws, or as those forces seem to stand by passively as moral offenses are committed, there develops a generalized contempt for the system in all its aspects, and a generalized suspicion of all of its rules.

As we argue, and others have recognized, deviations from what is perceived as deserved punishment, “empirical desert,” may affect the community’s respect for the criminal justice system.

Just as the institution of the criminal law may be brought into disrepute by the too easy attribution of criminality in situations where the label criminal is generally thought inappropriate, so also may the institution be undercut if it releases as noncriminal those society believes should be punished. This does not mean that the criminal law may not be a means of educating the public as to the conditions under which moral condemnation and punishment is inappropriate. It does mean that the results may not depart too markedly from society’s notions of justice without risking impairment of the acceptability and utility of the institution.⁷⁵

The most extreme manifestation of disregard for the criminal justice system based on its failure to condemn acts that are intuitively blameworthy is the resort of citizens to vigilantism, noted above. This may occur when core moral rules are transgressed but not punished. But less significant failures may also engender resistance and subversion, generally

75. Joseph M. Livermore & Paul E. Meehl, *The Virtues of M’Naghten*, 51 MINN. L. REV. 789, 792 (1967) (footnotes omitted).

in less dramatic ways. The people whose cooperation is necessary for the functioning of the system, as noted above, may indulge in minor acts of vigilantism, such as the jurors who convict an offender based on their perception of blameworthiness independent of the instructions they receive or in spite of the applicable legal rules. Similarly, a witness who perceives the criminal justice system as failing to do justice may be tempted to distort his or her testimony to compensate for what they see as the system's failings.⁷⁶

Empirical studies support this. Research has found that Americans are likely to obey the law when they view it as a legitimate moral authority.⁷⁷ In turn, they are likely to regard the law as a legitimate moral authority when they regard the law as being in accord with their own moral codes. As Tom Tyler concludes, “[t]he most important normative influence on compliance with the law is the person’s assessment that following the law accords with his or her sense of right and wrong”⁷⁸

As this implies, people who come to believe that the legal codes are importantly deviant from their own moral codes feel less concerned with abiding by the law. Two recent experiments provide evidence for the beginnings of this rejection process in individuals who discover contradictions between the legal codes in force and their own moral intuitions of justice. In these studies, participants read about several cases in which there was a mismatch between the laws in existence and the moral intuitions of the participants. By coincidence, both studies used a case that had attracted national attention as the example of the legal codes failing to criminalize actions that the community thought were criminal. The case involved a real world instance in which one young man dragged a young girl into a semiprivate space and raped and killed her. A friend, aware of the series of actions, took no steps to intervene or report the action to authorities that might have intervened. The pair spent the next two days gambling in casinos, and on return home, the watcher bragged to friends about the crime. Since there was no law against the watcher’s actions in the state in question, no legal action was brought against him.⁷⁹ Many respondents found the absence of a law criminalizing this conduct seriously problematic. In an earlier study of community intuitions of justice, we

76. For a more detailed discussion on the utility of desert, see Robinson, *Competing Conceptions*, *supra* note 21.

77. See TOM R. TYLER, WHY PEOPLE OBEY THE LAW 68 (1990).

78. *Id.* at 64.

79. For a narrative describing the case facts and a discussion of the issues arising in it, see PAUL H. ROBINSON, CRIMINAL LAW: CASE STUDIES AND CONTROVERSIES § 15 (2005) [hereinafter ROBINSON, CASE STUDIES].

found that respondents generally punished actors who failed in a duty to rescue a person in distress, if the intervention could be achieved without serious inconvenience or danger to the potential rescuer, as it could have been in the case described here.⁸⁰

After reading this and other intuition-violating cases, in which the legal codes criminalized actions that respondents did not consider criminal, participants rated themselves significantly less likely to cooperate with police and less likely to use the law to guide their behavior.⁸¹ More specifically, in one of the studies, participants who had read cases in which the legal system behaved in ways counter to their moral intuitions rated themselves “more likely to take steps aimed at changing the law (including replacing legislators and prosecutors and breaking the law while taking part in demonstrations), less likely to cooperate with police, more likely to join a vigilante or watch group, and less likely to use the law to guide behavior.”⁸² “Overall, participants appeared less likely to give the law the benefit of any doubt after reading cases where the law was at odds with their intuitions.”⁸³ In the second study, learning about a case in which a similar mismatch occurred on one law caused participants to report a “willingness to flout unrelated laws commonly encountered in everyday life . . . as well as willingness of mock jurors to engage in juror nullification.”⁸⁴

These studies report the beginnings of the process within individuals of coming into contempt for legal codes when they learn about thirdhand examples of the legal system failing to do justice. We suggested, however, that the community would also come into contempt for the justice system in the reverse case: where it criminalized actions that the community thought were morally acceptable. The “experiment” with the prohibition of the consumption of alcohol demonstrated what can be the end result of a catastrophic mismatch between elements of the legal codes and the moral intuitions of large segments of the population directly affected by the

80. PAUL H. ROBINSON & JOHN M. DARLEY, *JUSTICE, LIABILITY, AND BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW* 42–50 (1995) [hereinafter ROBINSON & DARLEY, *JUSTICE, LIABILITY & BLAME*].

81. The contrast ratings were provided by control respondents who read similar stories but with endings that they perceived as just. For example, control respondents learned that the watcher was prosecuted as an accessory to the crime and received a one-year prison sentence. See Janice Nadler, *Flouting the Law*, 83 TEX. L. REV. 1399, 1417 (2005).

82. Erich J. Greene, *Effects of Disagreements Between Legal Codes and Lay Intuitions on Respect for the Law* iv (June 2003) (unpublished Ph.D. dissertation, Princeton University) (on file with Mudd Library, Princeton University).

83. *Id.* at v.

84. See Nadler, *supra* note 81, at 1399.

codes.⁸⁵ People continued to wish to drink, and the institutions that came into being to allow this were, of necessity, illegal. The inadvertent and unforeseen consequences of Prohibition included the rapid development of thriving criminal activity focusing on smuggling and bootlegging and the consequential clogging of the courts with alcohol-associated prosecutions. Crime increased and became “organized,” and crime mobs eventually engaged in criminal activities beyond those associated with alcohol consumption. Contrary to Prohibition’s goal of eliminating corrupting influences in society, it increased those influences because it brought many persons into contempt for the moral correctness of one law, which generalized to the institutions enforcing this law and, eventually, to the legal code in general.

It has been suggested that the U.S. Congress has continued in similar “overcriminalization” practices. John Coffee asserts that “the dominant development in substantive federal criminal law over the last decade has been the disappearance of any clearly definable line between civil and criminal law.”⁸⁶ He also concludes, with a note of anger, that “this blurring of the border between tort and crime predictably will result in injustice, and ultimately will weaken the efficacy of the criminal law as an instrument of social control.”⁸⁷ This is so because “the factor that most distinguishes the criminal law is its operation as a system of moral education and socialization. The criminal law is obeyed not simply because there is a legal threat underlying it, but because the public perceives its norms to be legitimate and deserving of compliance.”⁸⁸

Coffee cites other scholars who have made similar points:

[O]ur leading criminal law scholars—among them Henry Hart, Sanford Kadish, and Herbert Packer—have periodically warned of the danger of “overcriminalization”: namely, excessive reliance on the criminal sanction, particularly with respect to behavior that is not inherently morally culpable. . . .

. . . [A]ll three agreed that a basic “method” distinguished the criminal law. . . . [T]he principle elements of this method [included] . . . a close

85. See generally Paul L. Murphy, *Societal Morality and Individual Freedom*, in *LAW, ALCOHOL, AND ORDER: PERSPECTIVES ON NATIONAL PROHIBITION* 67 (David E. Kyvig ed., 1985) (discussing Prohibition as it relates to law and morality).

86. John C. Coffee, Jr., *Does “Unlawful” Mean “Criminal”?: Reflections on the Disappearing Tort/Crime Distinction in American Law*, 71 B.U. L. REV. 193, 193 (1991).

87. *Id.*

88. *Id.* at 193–94. See also TYLER, *supra* note 77 (using survey research to find that the public complies with the criminal law based not on its deterrents, but on its moral legitimacy).

linkage between the criminal law and behavior deemed morally culpable by the general community.⁸⁹

This applies not only to cases where civil law and criminal law have become blurred, but also to many “moralistic” criminal prohibitions. In an Op-Ed essay in the *New York Times*, Charles Murray comments on a recent bill that has been passed and signed into law that “will try to impede online gambling by prohibiting American banks from transferring money to gambling sites.”⁹⁰ He first comments that this will cost those who passed the law a good many votes from online gamblers who think that what they do is morally allowable. He then continues:

If a free society is to work, the vast majority of citizens must reflexively obey the law not because they fear punishment, but because they accept that the rule of law makes society possible. That reflexive law-abidingness is reinforced when the laws are limited to core objectives that enjoy consensus support

Thus society is weakened every time a law is passed that large numbers of reasonable, responsible citizens think is stupid.⁹¹

Resistance to particular laws occurs because they do not correspond with the prototype perception that the community holds of a moral offense of the sort worthy of treatment within the criminal justice system. The criminal law can most effectively maximize its moral credibility and thereby minimize resistance and subversion by adopting criminal rules that track shared community intuitions of justice. The danger of failing to harmonize criminal codes with intuitions of justice is that the code may lose credibility on a wide array of prohibitions if too many are perceived to be against notions of what is just.

4. Shaping Societal Norms: The Persuasive Power of the Law

In a society as diverse as ours, sustaining moral norms necessitates mechanisms that are able to transcend cultural differences. Criminal law is perhaps unique in its ability to inform, shape, and reinforce social and moral norms on a society-wide level. When criminal law harnesses societal intuitions, it can boost compliance with its dictates through the moral credibility it gains. This is the case because of the law’s interactions with social networks. Prohibitions shared by society and transmitted through social networks and individual’s own internalized conceptions of these

89. Coffee, *supra* note 86, at 197–98.

90. Charles Murray, Op-Ed., *The G.O.P.’s Bad Bet*, N.Y. TIMES, Oct. 19, 2006, at A27.

91. *Id.*

moral precepts help bring the individual's conduct into compliance with societal norms. When criminal law functions in concert with these shared societal intuitions, it taps into this powerful amplification to achieve greater utility. When the criminal justice system is viewed as just, it is most effectively situated to help shape and enforce societal norms.

In many discussions of the criminal justice system, it is noted that a characteristic that sets it apart from other societal institutions is that it has a monopoly on the legitimate use of force against citizens. We grant the truth of this, but note that the criminal justice system is not limited to just compelling behavior. The criminal justice system, as an institution involving legislatures, legal philosophers, various criminological experts, and enforcement agencies, also has the power to persuade citizens about the moral appropriateness of its enacted laws. Social scientists know a good deal about persuasion, both in terms of how it succeeds and how it fails, and applying their theories to the issues we are addressing can be quite illuminating.

Consider the criminal justice system as a "persuasive source." A large body of research suggests that a source that is seen as legitimate in its authority, expert in its knowledge, and trustworthy in its motives, is highly persuasive.⁹² If the criminal justice system, and more generally, the government's criminal liability and punishment rules, can be persuasive guides for conduct to citizens, knowledge of the law can play a powerful role in debates between groups of citizens about what the governing social norms should be.

5. Gaining Compliance in Borderline Cases

The criminal law may also be able to gain compliance when it earns a reputation for punishing what society believes to be blameworthy and punishing in proportion to the perceived blameworthiness of the offender. In a borderline case, where the acceptability of a particular act is unclear or undetermined, community members are more likely to give deference to the commands of the criminal justice system if the system is morally authoritative. This is of special concern for a complex society like ours, where some interactions may appear harmless, but when replicated on a large scale may have negative consequences warranting criminalization. If the criminal justice system lacks moral credibility, code drafters' attempts

92. See RICHARD E. PETTY & JOHN T. CACIOPPO, *ATTITUDES AND PERSUASION: CLASSIC AND CONTEMPORARY APPROACHES* 62–69 (1996) [hereinafter PETTY & CACIOPPO, *ATTITUDES & PERSUASION*].

to control such activities may be frustrated if citizens do not conform to the dictates of the law in instances where the rationale for the criminal liability imposed cannot be immediately intuited.

Social science research on persuasion supports this conclusion and illuminates the mechanisms of thought that bring it about. Lately, persuasion researchers have distinguished between two types: one more analytic and direct, the other more peripheral and heuristic.⁹³ The latter is more relevant to our discussion here. This path often involves giving one cues that lead that person to believe that the action in question should be regarded as wrong or right, criminalized or decriminalized, without presenting the direct arguments why this should be so.⁹⁴ It often relies heavily on the person's assessment of the source of the communication. Specifically, if the message comes from a credible and trustworthy source, people are likely to accept the rightness of the communication.⁹⁵ This is the immediate acceptance and obedience that Murray referred to as the "reflexive loyalty to the rule of law."⁹⁶ It is what the state relies on people giving when, for instance, the state passes laws against insider trading of stocks based on news that is not yet known by the public. Without knowing quite why insider trading is morally wrong, most of us accept the conclusion that it is wrong, because the relevant authorities have thought about it, and assert that it is wrong. It is what we all count on when, as we travel down a highway that is new to us, we slow down when we see a sign saying "caution, blind curve," and count on others to slow down as well. It is, in one sense, blind obedience, but it is extremely socially useful, and it functions based on attitudes of trust toward the source, in turn based on past experiences of the source providing messages that turned out to be credible.

That citizens in general regard the law as a credible guide to how they ought to behave is shown by the responses of the random sample of Chicago citizens tested in Tyler's study.⁹⁷ Eighty-two percent of the sample strongly agreed or agreed with the statement "[d]isobeying the law

93. See Tilmann Betsch, Henning Plessner & Elke Schallies, *The Value-Account Model of Attitude Formation*, in CONTEMPORARY PERSPECTIVES ON THE PSYCHOLOGY OF ATTITUDES 251, 251–52 (Geoffrey Haddock & Gregory R. Maio eds., 2004). See also Richard E. Petty & John T. Cacioppo, *The Elaboration Likelihood Model of Persuasion*, in 19 ADVANCES IN EXPERIMENTAL SOCIAL PSYCHOLOGY 123, 125 (Leonard Berkowitz ed., 1986) [hereinafter Petty & Cacioppo, *Elaboration Likelihood Model*] (describing the two types of persuasion as central and peripheral).

94. See Petty & Cacioppo, *Elaboration Likelihood Model*, *supra* note 93, at 125.

95. See ELIOT R. SMITH & DIANE M. MACKIE, SOCIAL PSYCHOLOGY 261–62 (2d ed. 2000).

96. Murray, *supra* note 90.

97. TYLER, *supra* note 77, at 46.

is seldom justified,” and 85 percent strongly agreed or agreed with the statement that “[p]eople should obey the law even if it goes against what they think is right.”⁹⁸ This survey also revealed that the people Tyler tested assigned a high degree of moral authority to legal codes,⁹⁹ and it is this moral credibility that we suggest gives the legal codes what social science calls their “informational influence”: their power to bring about law-abiding behavior even when the citizens are unsure in their own minds that the behavior is correct.

6. Conclusion

The ability of the criminal justice system to harness the power of stigmatization, to avoid subversion and vigilantism, to gain compliance in borderline cases, and to have a role in shaping societal norms is directly related to its ability to gain moral credibility from those to whom it applies. The moral credibility of the law is enhanced when the distribution of punishment it prescribes accords with the community’s own shared intuitions of justice. When the law is perceived as “doing justice,” assigning liability in proportion to the moral blameworthiness of the punished offender, it becomes more effective at controlling crime. In contrast, when criminal liability deviates from intuitions of justice, particularly when such deviations are dramatic, the loss of moral credibility undermines the ability of the criminal law to effectively perform a crime control function.¹⁰⁰

98. *Id.*

99. *Id.*

100. For more recent developments since the publication of *The Utility of Desert*, *supra* note 59, see Jeremy A. Blumenthal, *Who Decides? Privileging Public Sentiment About Justice and the Substantive Law*, 72 UMKC L. REV. 1, 9 (2003) (“[A]lthough there is of course a deterrent influence to the threat of punishment, data show that where citizens perceive a lawmaking authority to be unjust or to have promulgated unjust laws, their attitudes toward that authority or those laws will likely be less compliant than where perceived legitimacy is high.”); Dan M. Kahan, *Social Influence, Social Meaning, and Deterrence*, 83 VA. L. REV. 349, 351 (1997) (“The decisions of individuals to commit crimes are influenced by their perception of others’ beliefs and intentions; the law shapes information about what those beliefs and intentions are. It follows that a community that wants to deter crime should concern itself not just with the effect of particular policies on the price of crime but with the statements that those policies make (and enable others to make) about the public’s attitudes toward criminal behavior.”); Nadler, *supra* note 81 (supporting the view that the perceived injustice of one law or legal outcome increases people’s willingness to disobey unrelated laws); Greene, *supra* note 82, at iv (“Previous research has found that Americans are more likely to obey the law when they view it as a legitimate moral authority. . . . Participants rated themselves significantly less likely to cooperate with police and less likely to use the law to guide their behavior after reading an intuition-violating case.”).

B. CRITICISMS OF EMPIRICAL DESERT AS A DISTRIBUTIVE PRINCIPLE

A variety of criticisms have been offered against our suggestion that the most effective crime control approach might be found in distributing criminal liability and punishment according to principles that mirror the community's shared intuitions of justice—a distributive principle we have called “empirical desert,” to distinguish it from the conception of desert offered by moral philosophers, “deontological desert.”¹⁰¹ A brief discussion of the most relevant criticisms may be useful here. We have noted previously, in Part I, two of these objections: that intuitions of justice are too vague and too much the subject of disagreement.¹⁰² Here we give a fuller response to these claims.

1. Notions of Desert as Hopelessly Vague

A common objection to empirical desert as a distributive principle is what is said to be its vagueness, as noted above in Part I.B. One author demonstrates this objection as follows:

[E]veryone may agree that five years in prison is unjustly harsh desert for shoplifting, or that a five dollar fine is unjustly lenient desert for rape, but beyond such clear cases our intuitions seem to fail us. Is two years, five years, or ten years the proper sanction for a rape? . . . Our sense of just deserts here seems to desert us.¹⁰³

Another author explains:

Insofar as we seek a morally sensitive scale in which to weigh subjective guilt, to classify the individual criminal on the long continuum from unblemished virtue to unmitigated evil . . . [t]he criminal law is unfitted for such issues. It faces an adequacy of difficulties without addressing such ethical nuances. It is necessarily generalized rather than related to the moral quality of the specific act. . . .

. . . .
 . . . Questions of guilt will thus be weighed on the imprecise scales of the criminal law which can allow for only a few subjective qualifications to the objective gravity of the crime.¹⁰⁴

101. See generally Robinson, *Competing Conceptions*, *supra* note 21 (discussing three conceptions of deserved punishment: vengeful desert, deontological desert, and empirical desert).

102. See *infra* Part I.B.

103. Leo Katz, *Criminal Law*, in A COMPANION TO PHILOSOPHY OF LAW AND LEGAL THEORY 80, 80–81 (Dennis Patterson ed., 1996).

104. NORVAL MORRIS, THE FUTURE OF IMPRISONMENT 74 (1974). Other writers have expressed similar views in different terms. See, e.g., PETER H. ROSSI & RICHARD A. BERK, JUST PUNISHMENTS: FEDERAL GUIDELINES AND PUBLIC VIEWS COMPARED 2–3 (1997) (“Norms are necessarily stated in

Some authors, such as Norval Morris,¹⁰⁵ may be willing to concede that desert is not a hopelessly vague concept, but that it has some meaning. They would make a related but slightly different criticism—desert cannot specify a particular amount of punishment that should be imposed; it can only identify a range of punishment that should not be imposed because it would be seriously disproportionate.¹⁰⁶

As noted in Part I.B, such complaints are based in part on a failure to distinguish two distinct judgments: setting the endpoint of the punishment continuum and, once that endpoint has been set, ranking all cases along that punishment continuum. Every society must decide what punishment it will allow for its most egregious case, be it the death penalty, life imprisonment,

general terms, in the sense of applying to classes of circumstances, defined fairly generally. Accordingly, there are many opportunities for disagreement to arise among members of a society over whether a norm is applicable in a specific circumstance. . . . [W]e can expect more agreement among Americans about which crimes deserve harsher punishments but less agreement about the specific sentences to be imposed.”); Hugo Adam Bedau, *Retribution and the Theory of Punishment*, 75 J. PHIL. 601, 613 (1978) (“Although there may be little difficulty in making uniform judgments of ordinal culpability . . . or of ordinal harmfulness . . . , there is no unique non-arbitrary way to combine these judgments into one judgment of ordinal seriousness. . . . Even if that problem is solved, and even if a plausible penalty scale can be constructed in terms of ordinal severity . . . there is no unique non-arbitrary way to identify the severity of the appropriate punishment, given only the seriousness of the crime . . .”).

105. Morris explains his position as follows:

Desert is, of course, not precisely quantifiable. There is uncertainty as to the judge’s role in its assessment, argument as to the extent to which he ought to reflect legislative and popular views of the gravity of the crime if they differ from his own. And further, views of the proper maximums of retributive punishments differ dramatically between countries, between cultures and subcultural groups, and in all countries over time. Nevertheless, the concept of desert remains an essential link between crime and punishment. Punishment in excess of what is seen by that society at the time as a deserved punishment is tyranny.

MORRIS, *supra* note 104, at 75–76.

106. Indeed, this is the underlying assumption of the American Law Institute’s recent proposal for a revised Model Penal Code Section 1.02(2)(a) that sets out the purposes of the sentencing provisions and the principles governing their interpretation and application:

Subsection (2)(a) embraces Morris’s observation that moral intuitions about proportionate penalties in specific cases are almost always rough and approximate—and that most people experience them as such. Even when a decisionmaker is acquainted with the circumstances of a particular crime, and has a rich understanding of the offender, it is seldom possible, outside of extreme cases, for the decisionmaker to say that the deserved penalty is *precisely* *x*. In Morris’s phrase, the “moral calipers” possessed by human beings are not sufficiently fine-tuned to reach exact judgments of condign punishments. Instead, most people’s moral sensibilities, concerning most crimes, will orient them toward a range of permissible sanctions that are “not undeserved.” Outside the perimeters of the range, some punishments will appear clearly excessive to do justice, and some will appear clearly too lenient—but there will nearly always be a substantial gray area between the two extremes.

Subsection 1.02(2)(a)(i) codifies the conception of latitude in morally permissible sentences when it speaks of a “range of severity” of proportionate punishments.

MODEL PENAL CODE § 1.02(2) cmt. b (Tentative Draft No. 1, 2007) (emphasis in original). The proposed subsection also signals its reliance upon this notion of desert as setting only outer boundaries of disproportionality, when subsection (2)(a)(ii) refers to “the boundaries of proportionality in subsection (a)(i).” *Id.* § 1.02(2)(a)(ii).

or fifteen years. Once that endpoint is set, the challenge for the adjudication system is to determine who should be punished and how much punishment each should be given. That process of distributing punishment requires only an ordinal ranking of offenders according to their blameworthiness. The result of that process is a specific amount of punishment, but that amount of punishment is not the product of some magical connection between that offense and that amount of punishment. Rather, it is the specific amount of punishment needed to set the offender's offense at its appropriate ordinal rank according to blameworthiness relative to all other offenses.¹⁰⁷ If the punishment continuum endpoint were changed, the appropriate punishment for each offender would change accordingly.¹⁰⁸ Thus, people's disagreements about the general issue of where the punishment continuum endpoint should be set masks their agreement about the relative blameworthiness of different cases.

This means that, once a society sets the endpoint of its punishment continuum, the ordinal ranking of cases along that continuum will produce quite specific punishments.¹⁰⁹ Those who complain about desert's vagueness seem to assume, incorrectly, that empirical desert needs to provide a universal, absolute amount of punishment deserved for a given offense.¹¹⁰ The content of empirical desert, though, is to ensure that offenders of different blameworthiness are each given different amounts of punishment that reflect those differences.¹¹¹ Ordinal ranking does not require a specific amount of punishment in an absolute sense. It requires

107. Robinson & Kurzban, *Concordance*, *supra* note 2, at 1855.

108. For a general discussion of these matters, see Robinson, *Competing Conceptions*, *supra* note 21.

109. *Id.*

110. For example, consider the following passage, in which the authors believe they are revealing a critical flaw in the theory of desert:

Retributivism cannot tell us what is the right punishment for murder, whether it should be 20 per cent higher or twice as high as that for burglary. The eighteenth-century judge who sentences the burglar to torture followed by death, the judge from Alabama who sentences him to ten years, and the judge from Amsterdam who sentences him to victim compensation all pronounce that they are giving the offender what he deserves. There is no retributivist answer as to which judge is right. On the retributivist's view, so long as they are all handing down sentences for burglary that are proportionately more than those for less-serious crimes and proportionately less than those for more-serious crimes, they could all be right.

JOHN BRAITHWAITE & PHILIP PETTIT, *NOT JUST DESERTS: A REPUBLICAN THEORY OF CRIMINAL JUSTICE* 178 (1990).

111. See JOHN KLEINIG, *PUNISHMENT AND DESERT* 114 (1973) ("What follows from [attempts to exactly define desert] is not the absurdity of specific desert claims but only the absurdity of expecting them to function like statements of empirical quantity."). See also ANDREW VON HIRSCH, *PAST OR FUTURE CRIMES: DESERVEDNESS AND DANGEROUSNESS IN THE SENTENCING OF CRIMINALS* (1985) (comparing the theory that sentencing should be based on desert and proportionality to the theory that sentencing should be based on the likelihood of reoffending).

imposition of only that specific amount of punishment that will put the offender at the appropriate ordinal rank *given the punishment continuum endpoint in that society*. That is, the uncertainty about the deserved punishment amount that Morris and others observe arises not because of any vagueness in the ordinal ranking of offenses according to offender blameworthiness, but rather because of differences in the punishment continuum endpoint that a society might adopt. It is the proper ordinal rank, though, that is the focus of empirical desert, not the punishment continuum endpoint. Once that endpoint is set, vagueness in deserved punishment amount disappears.

But this does not settle the vagueness complaint. Some writers argue that even ordinal ranking is something that can be done only in the vaguest terms, and that establishing specific rankings is impossible.¹¹² The claim is that the blameworthiness ranking of offenses is beyond the ability of people's intuitions of justice—that those intuitions are simply too vague to do more than to roughly distinguish between “serious” cases and “not serious” cases and cannot provide the nuance needed to do more.

This claim, however, is not consistent with the results of a wide variety of empirical studies.¹¹³ Subjects' judgments are nuanced, so that small changes in facts produce large and predictable changes in punishment.¹¹⁴ Alexis Durham summarizes the surveys as follows: “Virtually without exception, citizens seem able to assign highly specific sentences for highly specific events.”¹¹⁵ The conclusion suggested by the

112. An example of such an argument is as follows:

Perhaps, at best, retributivism can determine the roughly appropriate punishment by comparatively ranking offenses in such a way as murder warrants greater punishment than rape, which warrants greater punishment than armed robbery, and so on. But it cannot determine whether rape warrants twenty, thirty, forty years imprisonment. Though retributivism cannot set cardinal or absolute levels of punishment, its advocates insist that they can set ordinal, or relative, levels of punishment (for example, murder warrants greater punishment than larceny). But retributivism cannot even satisfactorily determine degrees of punishment ordinally. For example, even if we assume that, all other things being equal, murder warrants greater punishment than armed robbery, does negligent homicide warrant greater punishment than intentional rape or intentional armed robbery? . . . Retributivism has no answer to the issue of whether greater wrongdoing done with lower culpability (for example, negligence or recklessness) warrants more or less punishment than comparatively minor wrongdoing with a greater level of culpability (such as intention or purpose). Thus, retributivism can determine neither the ordinal nor the cardinal ranking of crimes and their concomitant degrees of punishment.

Russell L. Christopher, *Deterring Retributivism: The Injustice of “Just” Punishment*, 96 NW. U. L. REV. 843, 893 (2002) (footnotes omitted).

113. For a general discussion of these studies, see Robinson & Kurzban, *Concordance*, *supra* note 2, at 1832–80.

114. *See id.* at 1832–46.

115. Alexis M. Durham III, *Public Opinion Regarding Sentences for Crime: Does It Exist?*, 21 J. CRIM. JUSTICE 1, 2 (1993).

empirical evidence is that people take account of a wide variety of factors and often give them quite different effect in different situations. That is, people's intuitions of justice are not vague or simplistic, as claimed, but rather quite sophisticated and complex.

2. Hopeless Disagreement as to Notions of Desert

The other common objection to using empirical desert as a distributive principle for criminal liability and punishment is the concern noted in Part I.B that even if individuals may have a clear notion of what desert demands, there is simply no agreement among people. Again, this objection simply does not match the empirical reality. The studies discussed above show widely shared intuitions that serious wrongdoing should be punished,¹¹⁶ and widely shared intuitions about the relative blameworthiness of offenders in different scenarios.¹¹⁷

One recent study may serve to illustrate the striking extent of the agreement on intuitions of justice. Participants ranked twenty-four crime scenario descriptions according to the amount of punishment the offender deserves.¹¹⁸ The researchers found that, despite the task being quite complex and demanding, requiring participants to compare the deserved punishment for each scenario to the deserved punishment for each of the other twenty-three scenarios, participants had little difficulty ranking the scenarios and in fact displayed a significant level of agreement in their ordinal rankings.

In this study, the amount of agreement in participants' rankings was measured using Kendall's coefficient of concordance ("Kendall's W"), a statistical measure of concordance in which 1.0 indicates perfect agreement and 0.0 indicates no agreement. The Kendall's W in this study was 0.95 (with $p < 0.001$),¹¹⁹ an astounding result. One might see Kendall's W coefficients this high when testing vision sensitivity, for example, by asking subjects to determine the relative brightness of images made up of different groupings of black and white spots.¹²⁰ Subjective or complex comparisons lead to lower Kendall's W values. For example, asking travel magazine readers to rank the safety of eight different travel destinations

116. See Robinson & Kurzban, *Concordance*, *supra* note 2, at 1848–54.

117. See *id.* at 1854–65.

118. See *id.* at 1866–74.

119. *Id.* at 1872.

120. Charles M. M. de Weert & Noud A. W. H. van Kruysbergen, *Assimilation: Central and Peripheral Effects*, 26 PERCEPTION 1217, 1219–21 (1997).

results in a Kendall's W of 0.52.¹²¹ When asking economists to rank the top twenty economics journals according to quality, one gets a very low Kendall's W of 0.095.¹²²

This study, as well as the other studies summarized in Part I.B, indicates the high level of agreement on intuitions of justice. The level of agreement is strongest for those core wrongs with which criminal law primarily concerns itself—physical aggression, taking property, and deception in exchanges—and becomes less pronounced as the nature of the offense moves farther from this core of wrongdoing.¹²³ The data overwhelmingly refutes the common perception that there is no clear agreement as to intuitions of justice.

This is not to say that there do not exist disagreements among people's intuitions of justice. Although people obviously do disagree about many things relating to crime and punishment, some instances of apparent disagreement are simply misleading. For example, if a test scenario is written ambiguously, different participants will have different perceptions of the same scenario, causing the extent of agreement to be underestimated.¹²⁴ So too will varying life experiences cause people to view differently cases with social or political implications. What one makes of the police testimony in the O.J. Simpson case or the Rodney King case, for example, may depend on one's opinion of police officers. Differing opinions on the liability or punishment deserved may result from people drawing different conclusions from the police testimony.

This leaves us with the above conclusion that there are stable, significant agreements as to the deserved degree of punishment, and that these agreements exist across cultures. One may wonder how this agreement among people about intuitions of justice, sometimes at astonishingly high levels even though it involves terribly subjective and complex matters, could have been missed for so many years. One may

121. Baruch Fischhoff et al., *Travel Risks in a Time of Terror: Judgments and Choices*, 24 RISK ANALYSIS 1301, 1302–03 (2004).

122. Kostas Axaroglou & Vasilis Theoharakis, *Diversity in Economics: An Analysis of Journal Quality Perceptions*, 1 J. EUR. ECON. ASS'N 1402, 1421 (2003) ("These results unveil significant diversity in the journal quality perceptions among groups of economists despite the fact that our sample focused on AEA members. To test the robustness of this claim, using Kendall's W we examined the correlation in journal quality perceptions between any two randomly selected economists in our sample. We found Kendall's W for the top ten journals in our rankings to be 0.396, which demonstrates a relatively low level of agreement among economists. Once we extended this exercise to the top 20 journals in our rankings, Kendall's W dropped to only 0.095.").

123. Robinson & Kurzban, *Concordance*, *supra* note 2, at 1883–92.

124. See ROBINSON & DARLEY, *JUSTICE, LIABILITY & BLAME*, *supra* note 80, at 219–23.

wonder how the common wisdom got it wrong. The discussion above may help explain. The failure to appreciate the high level of ordinal ranking agreement makes sense if people fail to distinguish between absolute severity and ordinal ranking. And it is easy to see how people's disagreements about the proper endpoint for the punishment continuum naturally obscure the existing agreement on the ranking of offenses along that continuum. That is, we are all well aware from news accounts and our own personal discussions that we frequently disagree with others when it comes to deserved punishment for cases in the news. The point here, however, is that those disagreements are often disagreements not about the relative blameworthiness of an offender as against other offenders, but rather disagreements about the facts of the case or about the general severity of punishment that the criminal justice system should use.

Because some disagreement does exist within the community, it follows that some deviation from some people's views is inevitable. In these situations, the logic of the utility of desert discussed above suggests that the law should adopt whatever rule will undermine its moral credibility the least. That commonly will mean adopting the majority rule over the minority rule, although lawmakers may need to take account of the strength of feeling in each of the competing views.

One might speculate about alternative explanations for the surprising evidence of shared intuitions of justice, as discussed previously.¹²⁵ Whether the agreement among people's intuitions of justice is the product of an evolutionarily developed mechanism or the product of social learning, or some combination of the two, that source is beyond the normal influence of culture or demographic. If it were not so insulated, one would see differences in intuitions of justice among different demographics and cultures. This insulation means that there may be serious limits on whether and how social engineers can manipulate intuitions of justice, at least those about core wrongdoing upon which there is broad agreement.

3. Utility of Desert as Based Upon the Assumption that People Intuitively Assess Punishment According to Desert When in Fact They Look to Deterrence or Incapacitation

Another criticism leveled against the utility of a desert-based model is that people's intuitions of justice are not distinct from conceptions of deterrence or incapacitation. That is, people making intuitive judgments of justice make them according to utilitarian models and not according to a

125. See *supra* text accompanying note 33.

level of punishment based upon blameworthiness. In several studies, however, we and others have examined the issue of what criteria people rely on when they make intuitive judgments of justice and found that it is desert, not deterrence or incapacitation, that drive people's intuitive assignments of punishment.¹²⁶

In the first set of studies, we explored the intuitive use of just deserts and incapacitation as models for judgments about justice. Participants were given ten short descriptions of criminal cases, which were generated by combining five levels of moral seriousness (theft of a CD, theft of a valuable painting, assault, homicide, and assassination), with two levels of perpetrator's criminal history (low likelihood of recidivism and high likelihood of recidivism).¹²⁷ Participants read each scenario separately and after reading were prompted to rate the case on two scales: first a seven-point scale of the severity of the punishment that should be imposed, ranging from "not at all" to "extremely severe"; and second a thirteen-point scale of criminal liability, ranging from "no liability" to "death penalty."¹²⁸ After submitting their liability judgments on these scales, participants were asked to reconsider the scenarios and assign punishments from a just deserts perspective and from an incapacitation perspective, respectively. The just deserts standard was described as assigning "the just punishment that the criminal *deserves* for the wrong that he did," while for the incapacitation standard, participants were instructed to assign "a sentence long enough to protect society from further harms by this person."¹²⁹

The results showed that participants were more responsive to case blameworthiness than to case recidivation. More importantly, participants' punishment assignments based on just deserts were closely aligned with their original intuitive decisions, while the punishments they assigned using the incapacitation model were not.¹³⁰ "What this suggests is that the default perspective of sentencing is indistinguishable from the just deserts perspective, but that both are significantly different from the incapacitation

126. See generally Kevin M. Carlsmith, John M. Darley & Paul H. Robinson, *Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment*, 83 J. PERSONALITY & SOC. PSYCHOL. 284 (2002) [hereinafter Carlsmith, Darley & Robinson, *Deterrence and Just Deserts*] (finding that individual sentencing decisions are motivated by just desert factors instead of by deterrence factors); John M. Darley, Kevin M. Carlsmith & Paul H. Robinson, *Incapacitation and Just Deserts as Motives for Punishment*, 24 L. & HUM. BEHAV. 659 (2000) [hereinafter Darley, Carlsmith & Robinson, *Incapacitation and Just Deserts*] (finding that just deserts was the primary sentencing motive in two experiments looking at desert motives and incapacitation motives).

127. Darley, Carlsmith & Robinson, *Incapacitation and Just Deserts*, *supra* note 126, at 662–63.

128. *Id.* at 663.

129. *Id.* at 681 (emphasis added).

130. See *id.* at 666–67.

perspective.”¹³¹

In another set of studies we explored participants’ reliance upon a just deserts model as against one of deterrence. In three studies, participants were given criminal scenarios that varied on a two by two format: the scenarios included different variations of two deterrence factors (detection rate and publicity) and two desert factors (severity of offense and extenuating circumstances).¹³² Participants were asked to rate the scenarios based on the same severity and sentencing scales described above, along with specific questions including: whether they endorsed or rejected general statements about the two theories; the moral acceptability of the crime committed; and how much resources should be expanded to either catch the perpetrator or prevent future occurrences.¹³³

The results of these studies reinforce the notion that people’s intuitive default for assigning criminal liability is consistent with a just deserts model. While participants explicitly endorsed deterrence justifications for punishment, in fact they meted out sentences “from a strictly deservingness-based stance.”¹³⁴ All three studies indicated that people assign an actor punishment in relation to the actor’s deservingness, rather than out of concern for deterrent effect on future offenses.¹³⁵ While participants support the concept and goals of deterrence, that support does not translate into differential assignments of punishment.

Finally, a third set of studies examined whether punishment intuitions were based on just deserts considerations by researching the types of information people find relevant when sentencing criminals.¹³⁶ In one of the studies, a participant is initially told only, for instance, that an offender has embezzled a certain amount from his employer. It is the participant’s task to determine the appropriate sentence by selecting and reviewing, one at a time, pieces of information about the crime that the participant considers most useful to know for punishing the criminal. The researcher then examines the order in which the subject acquires the information, consisting of information relevant to just deserts issues, incapacitation, or deterrence, inferring that the most important information is acquired

131. *Id.* at 667.

132. *See* Carlsmith, Darley & Robinson, *Deterrence and Just Deserts*, *supra* note 126, at 288.

133. *Id.* at 289, 291, 293.

134. *Id.* at 295.

135. *Id.*

136. *See generally* Kevin M. Carlsmith, *The Roles of Retribution and Utility in Determining Punishment*, 42 J. EXPERIMENTAL SOC. PSYCHOL. 437 (2006) (discussing people’s motives in sentencing based on the findings of three studies).

first.¹³⁷ Results showed that participants initially requested information related to desert considerations—only later did some seek information relevant to incapacitation, and rarely did any seek information relevant to deterrence.¹³⁸

Participants were asked their level of confidence in the punishment judgment each time they acquired information. Increments in the participants' confidence in their proposed sentence for the crime were most influenced by the just deserts information.¹³⁹ This effect could have occurred because the just deserts information was acquired first, when the participants' uncertainty would have been maximal. In a follow-up study, however, this confound was removed, and it continued to be the case that the desert information was the most effective in increasing the confidence of the respondents in their sentencing assignments.¹⁴⁰

These studies, taken together, provide support for the conclusion that intuitive judgments about punishment are based primarily on just deserts considerations.¹⁴¹

4. Empirical Desert as Inevitably Draconian

Some resistance to giving empirical desert a role in setting criminal liability rules derives from the following line of reasoning: (1) I dislike many of the modern crime control reforms—such as three strikes statutes, lowering the age of prosecution as an adult, high penalties for drug offenses, and the trend toward expanding the scope of criminalization to what had previously been regulatory offenses. (2) These reforms are the product of recent legislative action that reflects the community's views. (3) Giving explicit deference to empirical desert will increase the trend toward such draconian measures.

Such reasoning, however, both misconceives the nature of empirical desert and mistakenly assumes that modern American crime politics track people's intuitions of justice. In fact, empirical desert has little, if anything, to do with people's views of the cases and issues that make a political debate. And modern American crime politics have little, if anything, to do with people's intuitions of justice. On the contrary, the current political

137. *See id.* at 444.

138. *See id.* at 446.

139. *See id.*

140. *See id.* at 447.

141. Further research that demonstrates that citizens' judgments on crime sentences are strongly affected by their perceptions of what is justly deserved for the crime can be found in NORMAN J. FINKEL, *COMMONSENSE JUSTICE: JURORS' NOTIONS OF THE LAW* (1995).

trends noted directly conflict with shared intuitions of justice.

Empirical desert embodies those principles of justice that people intuitively rely upon in making judgments about relative blameworthiness. The most reliable method for determining these intuitions is by giving people a series of cases, determining how they treat the cases differently, then deducing from the factual differences among the cases the factors that shape people's blameworthiness judgments.¹⁴² People's political views or their reaction to cases in the headlines commonly have little to do with their intuitions of justice. As any social psychology researcher will confirm, what people claim their views are when asked about a principle in the abstract does not necessarily reflect their true intuition. Too often, a political position, or a case in the headlines, brings with it political baggage that colors people's views of the position or the case. Whether one supports or opposes the death penalty, for example, is an issue that people see as placing them in a larger complex of social and political dimensions, which involve much more than the person's intuitions of justice.

Perhaps more importantly, the engine driving American crime politics is not people's intuitions of justice. On the contrary, it is antidesert crime control theories—most notably deterrence and incapacitation—that have had the greatest influence in recent criminal justice reforms. Three strikes statutes, the lowering of the age of prosecution as an adult, high penalties for drug offenses, the expansion of the scope of criminalization to what had previously been regulatory offenses, and other recent popular reforms are typically justified by their purported ability to reduce crime. The increased criminalization and penalties are said to produce a stronger deterrent,¹⁴³ and the longer prison terms are said to more effectively incapacitate persons seen as likely to recidivate.¹⁴⁴ In fact, even with the incomplete studies we now have,¹⁴⁵ we have grounds to suspect that these reforms conflict with people's shared intuitions of justice.¹⁴⁶

142. For a more detailed description of the methodology of scenario research, see ROBINSON & DARLEY, *JUSTICE, LIABILITY & BLAME*, *supra* note 80, at 7–11, 217–28.

143. See 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, 963–69 (2003) (citing authorities that have used deterrence arguments in support of modern crime control legislation).

144. See Paul H. Robinson, *Punishing Dangerousness: Cloaking Preventive Detention as Criminal Justice*, 114 *HARV. L. REV.* 1429, 1429–31 (2001) [hereinafter Robinson, *Punishing Dangerousness*] (citing authorities that have used incapacitation arguments in support of modern crime control legislation).

145. The authors have under way a series of studies designed to more specifically test people's intuitions of justice against the modern crime control reforms noted in the text.

146. See, e.g., ROBINSON & DARLEY, *JUSTICE, LIABILITY & BLAME*, *supra* note 80, at 139–47, 189–97 (suggesting people's intuitions would neither support a trend toward reducing the age of

The utility of desert arguments above suggest that these kinds of crime-control reforms are shortsighted and that effective long-term control of crime lies not in trying to maximize deterrence or incapacitation in ways that conflict with people's intuitions of justice, but rather in doing justice, and thereby building the criminal law's moral credibility in order to harness the powerful forces of internalized norms and social influence.

5. Empirical Desert as Potentially Immoral

While empirical desert may not be hopelessly vague and the subject of eviscerating disagreement, it is subject to a quite valid and important criticism as a distributive principle: it is potentially unjust. While a community may share a view that certain conduct is moral or certain punishment is just, such views do not make it so. Witness the cases of slave holders in the pre-Civil War South. Empirical desert can only tell us what people think is just. It cannot tell us what is actually just in a transcendent sense. The latter is a task for moral philosophers, to determine "deontological desert."

This is, of course, an objection to all utilitarian principles for the distribution of liability and punishment. They set as their guiding purpose not the doing of justice but rather maximizing utility—most importantly, reducing future crime. Among utilitarian principles, empirical desert may be the least objectionable on these grounds, for it may best approximate a distribution according to true moral blameworthiness. Nonetheless, it is conceptually and practically distinct from deontological desert,¹⁴⁷ and can fairly be criticized as deviating from deontological desert.

What empirical desert produces is not justice, but only liability and punishment consistent with the community's views about what constitutes justice. The community's intuitions of justice could be wrong, even if there is a high degree of agreement about them. To protect against this error and to be able to identify when people's shared intuitions of justice are unjust, a system must look beyond the empirical desert social scientists can discover to the deontological desert of moral philosophy that will provide a

prosecution as an adult nor three strikes statutes).

147. As Michael Moore observes:

The familiar argument that in the design of punishment institutions we ought to capitulate to the felt resentments of most citizens has been answered many times before. If the popular judgement [sic] . . . is wrong, then we ought to educate citizens to the right view, not reinforce them in their error. Whatever one's meta-ethics, it is obviously false to equate the actual just deserts of offenders to what most people feel is their just deserts

MICHAEL MOORE, *PLACING BLAME: A GENERAL THEORY OF THE CRIMINAL LAW* 207–08 (1997) (footnote omitted).

transcendent check on the justness of its liability rules. It is only deontological desert that can give us the truth of what is deserved, insulated from the vicissitudes of human irrationality and emotions.

Unfortunately, an examination of the modern methodology of moral philosophers suggests that they fail to appreciate the importance of the difference between deontological and empirical desert. They commonly rely heavily upon shared intuitions of justice in their analyses, and thereby bias their conclusions in favor of principles of justice that match people's shared intuitions of justice. That reduces the extent to which moral philosophy can be relied upon to provide the transcendent check that empirical desert needs.¹⁴⁸

The current methodology of moral philosophers relies upon intuitions of justice in a variety of ways. Philosophers today test variations in a series of hypotheticals according to philosophers' own intuitions about the proper resolution of each as a basis for building moral principles, as in John Rawls' "reflective equilibrium."¹⁴⁹ The differences in their judgments about the intuitively proper resolution of different hypotheticals are used as data points from which philosophers derive a moral principle, which can in turn be refined by testing that moral principle against the philosophers' intuitions in new sets of hypotheticals.¹⁵⁰

But as one of us has stated previously,

The methodological reliance . . . on intuitions of justice creates a bias in favor of moral principles consistent with intuitions. Thus, moral principles with principled, reasoned support might nonetheless fail to gain currency among philosophers, or might be discarded, simply because philosophers as a group think their results inconsistent with intuitions—a practical veto by philosophers' shared intuitions.¹⁵¹

The failure of modern philosophy to provide a reliable guard against a

148. For a full discussion, see Paul H. Robinson, *The Role of Moral Philosophers in the Competition Between Deontological and Empirical Desert*, 48 WM. & MARY L. REV. 1831 (2007) [hereinafter Robinson, *Moral Philosophers*].

149. See JOHN RAWLS, A THEORY OF JUSTICE 48 (1971) (explaining that the best sense of justice is one which matches a person's judgments in reflective equilibrium—a state reached after consideration of various conceptions of justice).

150. Robinson, *Moral Philosophers*, *supra* note 148, at 1839. See also RAWLS, *supra* note 149, at 48. For an example of moral philosophers using intuitive analysis of case hypotheticals as a standard method, see Leo Katz, *Incommensurable Choices and the Problem of Moral Ignorance*, 146 U. PA. L. REV. 1465, 1482–83 (1998) (using a hypothetical, derived from the application of necessity defense to situations where the actor has culpably created the justifying situation, to argue that at times persons who are unavoidably morally ignorant can be blamed for making the wrong decision).

151. See Robinson, *Moral Philosophers*, *supra* note 148, at 1840.

community's shared but immoral intuitions of justice can make empirical desert both more and less attractive as a distributive principle. On the one hand, it means that moral philosophy is unavailable to provide the transcendent check on intuitions of justice that is needed. That makes empirical desert less attractive as a distributive principle. On the other hand, it means that deontological desert provides little or no advantage over empirical desert as a distributive principle. Thus, if empirical desert has any real competition, it is among other utilitarian theories. And among the utilitarian theories, empirical desert will be the one most likely to approximate justice. At the same time, for all the reasons summarized in Part III.A regarding the utility of desert, it also will be the one most likely to reduce crime in the long run.¹⁵²

To conclude, some criticisms of empirical desert as a distributive principle for liability and punishment are unfounded, but not all. Arguments that intuitions of justice are hopelessly vague or the subject of eviscerating disagreement either fail to appreciate the rank-order method of desert judgments or simply make assumptions that are inconsistent with the empirical evidence. Similarly inconsistent with available empirical evidence is the claim that people's intuitions about justice are based on notions of deterrence and incapacitation. It is fair, however, to complain that following empirical desert does not insure that true justice, in a transcendent sense, will be done.

C. DEVIATIONS FROM EMPIRICAL DESERT

While we think that there is utility in tracking shared intuitions of justice, it is not our view that deviations from empirical desert should never be tolerated. The good utilitarian seeking to reduce crime must remain open to the possibility that a deviation doctrine, while incurring a cost in marginally reducing the criminal law's moral credibility, nonetheless may generate a crime control benefit that exceeds that cost. Further, it is simply not the case that crime control is the only or even the most important

152. One last kind of objection to our utility of desert argument should be mentioned. Many critics of empirical desert assume that existing criminal legislation reflects the community's intuitions of justice. Because they do not like much of that legislation, they oppose a distributive principle based upon the criterion that they believe generates these laws. But this misconceives either the nature of empirical desert or the operation of legislatures in enacting criminal law legislation. As to the first, the utility of desert comes from tracking laypersons' intuitions of justice, as revealed under controlled studies by social scientists, not by tracking the decisions of politicians. As to the second, empirical studies make it quite clear that current criminal law legislation deviates regularly from the community's intuitions of justice in many important ways. See ROBINSON & DARLEY: JUSTICE, LIABILITY & BLAME, *supra* note 80, at 204–28.

societal interest. Privacy, fairness, limitations on governmental power, proper allocation of governmental authority, and a variety of other interests may be judged sufficiently important to outweigh the crime control interest furthered by empirical desert. Nor is it the case that reformers must remain content with the community's existing intuitions of justice. Part IV.D takes on the issue of how these intuitions might effectively be changed. Adopting criminal liability and punishment rules that deviate from existing intuitions of justice may be a necessary part of any such social engineering program.

But it is also true that some of the doctrines of deviation are not in fact supportable under any of these justifications, or that the interests they promote can as easily and as effectively be promoted by means that do not require criminal liability and punishment rules that deviate from empirical desert.

Consider the wide range of rules and practices in the current criminal justice system that regularly and intentionally deviate from shared intuitions of justice. Some of these deviations from community perceptions of desert are not a rejection of justice as the goal. They reflect a keen interest in promoting justice, but seek to take account of the complexity of the world in which we live in reaching that end. Other deviation rules, though, openly sacrifice doing justice in order to promote other interests.¹⁵³

1. Promoting Justice in a Complex World

Some rules that deviate from perceived desert are adopted out of fear that a more desert-based rule would be subject to easy manipulation and abuse, which would ultimately produce less justice, not more. For example, the law typically rejects a defense for a reasonable mistake of law. Some states reject an insanity defense or limit its availability to a set of cases more narrow than community views would support. It also is common for states to ignore the individual characteristics of a defendant in making liability judgments such as those involving provocation or negligence—including limitations on capacities that would have made it difficult, if not impossible, for the person to have avoided the violation.¹⁵⁴

Or, a rule that deviates from desert may be adopted because a more desert-based rule would encounter evidentiary problems that would reduce the reliability of the liability and punishment judgment. Statutes of limitation are adopted to avoid the dangers of stale evidence. Strict liability

153. For a general discussion of these matters, see PAUL H. ROBINSON & MICHAEL T. CAHILL, *LAW WITHOUT JUSTICE: WHY CRIMINAL LAW DOESN'T GIVE PEOPLE WHAT THEY DESERVE* (2006).

154. *See id.* at 27–51.

is imposed in cases where culpability may be difficult to prove, but is likely to exist. And coerced confessions and uncounseled lineups are excluded to avoid false recriminations.¹⁵⁵

Or, finally, a deviation rule may be adopted because, while it fails to do justice in one case, it allows justice to be done in many other cases or in another more important case. Most notably, plea bargains and witness immunity are granted, even for quite serious offenses, if the cooperation thereby obtained will allow the successful prosecution of even more serious offenses by others.¹⁵⁶

While each of the deviation rules might normally produce a cost in the system's moral credibility, there at least exists the possibility that the cost can be reduced through a public education campaign that explains why the apparent deviation does greater justice overall, assuming that the arguments in support of such claims do indeed have empirical support, which is not always the case.¹⁵⁷ What may ultimately save the system is that it is at least *trying* to do justice, as best it can in a complex world. Good intention counts a good deal in setting reputation.¹⁵⁸

2. Sacrificing Justice to Promote Other Interests

Some other rules that deviate from perceived desert are justified on the grounds that they advance interests other than doing justice. The legality principle bars conviction for offensive conduct that was not specifically described in a previously existing prohibition, even if most people, including the offender, believed that the conduct was prohibited. The flipside to this commitment to living by the rules is a common tendency to avoid acquittals if a prohibition has been clearly violated. This tendency is at its strongest where the strength and clarity of the prohibition is at its weakest. That is, courts may be reluctant to allow an acquittal based upon lack of culpability or the presence of excusing conditions if such acquittal may serve to undercut the clarity of a prohibition that is in need of reinforcement.¹⁵⁹

Other doctrines deviate from desert in order to advance a traditional utilitarian crime control program, such as distributing criminal liability and punishment to optimize deterrence, rehabilitation, or the incapacitation of

155. *See id.* at 52–71.

156. *See id.* at 72–86.

157. *See infra* Part III.D.

158. *See* Robinson & Darley, *The Utility of Desert*, *supra* note 59, at 495–96.

159. *See* ROBINSON & CAHILL, *supra* note 153, at 89–116.

the dangerous, even where such a distribution conflicts with shared intuitions of justice.¹⁶⁰

Other deviation doctrines are designed to control police and prosecutors. For example, the exclusionary rule may bar the use of clearly reliable evidence in order to discourage police from engaging in unauthorized searches or seizures, even if it lets a clearly guilty offender go free.¹⁶¹ Speedy trial rules designed to discourage prosecutorial delay can have a similar effect.¹⁶² The bar against “double jeopardy” operates to limit abuse through multiple prosecutions, even if it means that clearly guilty offenders will escape the punishment they deserve.¹⁶³ Similarly, the entrapment defense is designed to discourage overzealous police conduct, even if the offender is a career criminal looking for an opportunity to commit the offense.¹⁶⁴

Still other deviation rules are justified on grounds unrelated to criminal justice or the operation of the criminal justice system. The unique condemnatory power of criminal conviction is used to boost the prohibition of what would otherwise be minor regulatory violations.¹⁶⁵ Similarly, the prohibition of violations by nonhuman corporate entities is given greater effect by enlisting the moral force of the criminal law.¹⁶⁶ Finally, defenses like diplomatic and official immunity are recognized in order to promote interests unrelated to criminal justice. The diplomat deserving of punishment goes free, but his immunity is thought to promote the establishment of effective diplomatic relations between nations.¹⁶⁷

While each of these deviation rules may have some justification, there is also reason to believe that each such deviation incurs a cost to the criminal justice system’s ability to reduce crime in the long term, a cost that has not traditionally been fully appreciated. Each may deserve a reevaluation to determine whether the benefits derived from the rule do indeed outweigh the costs to the law’s moral credibility and its resulting detrimental effect on its crime control effectiveness, an inquiry sketched below.

160. *See id.* at 117–36.

161. *See id.* at 149–55.

162. *See id.* at 155–57.

163. *See id.* at 166–69.

164. *See id.* at 180–83.

165. *See id.* at 187–91.

166. *See id.* at 192–95.

167. *See id.* at 195–201.

D. STRATEGIES FOR AVOIDING DEVIATIONS FROM SHARED INTUITIONS

If, as Part III.A suggests, there are costs incurred by deviating from the community's shared intuitions of justice, then the doctrines of deviation described above ought to be examined to determine whether their benefits exceed their costs and, if so, whether they can be modified or should be abandoned.

1. Criminal Justice Reforms

Some of the deviation doctrines may no longer be valid on their own terms. For example, regarding statutes of limitation, there may have been a time when the trial process was inadequate to exclude evidence that was unreliable because it was too old. But with modern trial procedures that give defense counsel every opportunity to point out the weakness in the prosecution's evidence, there seems a less obvious need for statutes of limitations today. For a long time, there was little need to abandon the doctrine because the proof beyond a reasonable doubt standard and the fact that the strength of a prosecutor's case tends to degrade over time meant that prosecutors had little interest in pursuing old cases. But scientific advances, such as fingerprinting and DNA analysis, increasingly make it possible for prosecutors to present overwhelmingly reliable evidence in old cases. It may be no surprise then, that jurisdictions today are repealing or lengthening their statutes of limitation.¹⁶⁸

For many other doctrines of deviation, however, an awareness of their potential costs simply suggests that they should be narrowed in some way or that their objectives be achieved in a way that does not deviate from community notions of desert.

For example, rather than rejecting a mistake excuse for a reasonable mistake of law or permitting the use of strict liability, the system can avoid the potential for abuse and manipulation by simply shifting the burden of proof to the defendant.¹⁶⁹ Rather than rejecting legitimate excuse defenses out of fear that an acquittal will muddle the clarity of a prohibition, the system can create a verdict system that distinguishes between an acquittal that condones the defendant's conduct and an acquittal that condemns that conduct but excuses the actor.¹⁷⁰ Rather than imposing punishment that conflicts with community views of justice in order to rehabilitate,

168. *See id.* at 53–63 (describing a case that prompted Illinois to dramatically extend its statute of limitation for sexual assault).

169. *See id.* at 205–09.

170. *See id.* at 210.

incapacitate, or deter, the system can impose the amount of punishment that is deserved but select a punishment method for doing so that best furthers these other crime control goals. A wide range of punishment methods are available other than incarceration, many of which can more effectively serve these alternative purposes.¹⁷¹

2. Employing Civil Rather than Criminal Processes

While criminal justice reforms are possible, it is likely that the most effective and most important reforms will not be those of the criminal justice system, but rather reforms that shift the responsibility for advancing a societal interest away from the criminal justice system to a civil remedies system. For example, given that the penalties imposed by the criminal justice system for regulatory violations typically are only fines, there is little gained by using the criminal justice system for such cases and much to lose. By using criminal conviction in cases obviously devoid of conduct that merits moral condemnation, such convictions only serve to dilute the condemnatory effect of criminal liability. In other words, the criminalization of regulatory violations undermines that very characteristic of criminal liability that such regulatory criminalizations seek to harness.¹⁷² An analogous point can be made about the use of criminal liability for nonhuman legal entities such as corporations. Because legal fictions cannot make moral decisions, there is at least some question as to whether morally condemning a nonhuman entity undermines the condemnatory value of criminal conviction.¹⁷³ But whether this is true or not will require empirical research into community reactions to criminal convictions of corporations. The more successful the personification of corporations is in the layperson's mind, the less damaging the effect on the condemnation inspired by criminal conviction.

An even more important reform may be to use civil rather than criminal methods to control overzealous police and thereby avoid the costly deviations that arise from excluding reliable evidence because of police error or an entrapment defense given to a career criminal looking for a crime opportunity. Administrative sanction of offending officers by citizen review boards or quick, easy, and substantial compensation of citizens whose rights are violated, or both, will avoid the deliberate deviations from desert that undermine the criminal justice system's moral credibility (and at the same time may provide a more effective means of controlling police

171. For a general discussion of what a system like this might look like, see *id.* at 212–17.

172. See *id.* at 221.

173. See *id.* at 192–94, 220.

than our current system of letting blameworthy offenders go unpunished).¹⁷⁴

Another important reform is to avoid the practice of using the criminal justice system as a mechanism of preventive detention. The practice is problematic because, by engaging in preventive detention of dangerous offenders while pretending it is punishment for past crimes, we end up not only with a criminal justice system that does not do justice, but also with a preventive detention system that is both unfair and ineffective.¹⁷⁵ This occurs because rather than looking directly and openly at the dangerousness of an offender, as an open preventive detention system would do, we have the criminal justice system look instead at his prior criminal record, for example, as a substitute for his dangerousness. But it is a poor substitute—one that produces both false positives, which wastefully and unfairly imprison nondangerous persons, and false negatives, which let dangerous offenders go free. And by using the criminal justice system for preventive detention, we pervert the procedures that ought to be in place when restrictions on liberty are based upon merely a prediction of future dangerousness. For example, in a criminal justice system designed to punish for a past wrong, there is little reason for the use of nonpunitive conditions, a principle of minimum restraint, periodic review of present dangerousness, or a preference for treatment of all detainees, yet these characteristics would logically govern in a preventive detention system. If we are indeed interested in preventively detaining dangerous offenders, it would be better for both detainees and for the community to do it openly in a civil preventive detention system.¹⁷⁶

The larger point here is that, once it is understood that deviations from empirical desert have crime control costs, the system ought not permit such deviations unless it is clear that they provide some greater benefit in advancing crime control or some other core interest.

IV. IMPLICATIONS FOR CHANGING INTUITIONS OF JUSTICE

A third implication of the universal and intuitive nature of core judgments about justice is what it predicts for attempts to change people's intuitions, as is often a feature of social engineering programs aimed at changing people's behavior. That nature suggests that one ought not be

174. See *id.* at 222.

175. See *id.* at 225. See also Robinson, *Punishing Dangerousness*, *supra* note 144 (discussing the trend in the criminal justice system toward the detention of dangerous criminals).

176. See ROBINSON & CAHILL, *supra* note 153, at 226–27.

overly optimistic that arguments or education necessarily will produce the change in judgments about justice that one seeks. While some kinds of changes are possible, it may be that they can be brought about most effectively not by trying to subvert people's intuitions, but rather by harnessing them. That is, rather than trying to "talk people out of" their existing intuitions, a more effective approach may be to fight fire with fire—to use stronger, core intuitions to help modify other, weaker intuitions.

A. PROPOSALS TO CHANGE INTUITIONS OF WRONGDOING AND BLAMEWORTHINESS

A standard form of political and social activity, and social norms theory debate, is the wish to change people's views and ultimately their conduct to that which the reformer sees as preferable. With the advent of recognition of an expressive function of criminal law and punishment, it is also common to have criminal law play a central role in this social engineering process. The typical plan is, with the help of criminal law manipulation, to get people to view conduct disapproved by the social engineers as condemnable or more condemnable (or, alternatively, to dissolve the condemnation with which people currently view the conduct sought to be promoted). In other words, the reform programs are in the business of trying to alter people's intuitions of justice.

The list of recent reform programs is substantial. Reformers concerned about date rape have attempted to eradicate the notion that women often pretend to withhold consent to intercourse to appear more alluring or simply to avoid appearing "promiscuous," rather than as a genuine indication of not wanting to engage in sexual activity.¹⁷⁷ Antismoking activists have sought to stigmatize smoking in the presence of nonsmokers who do not consent to inhaling secondhand smoke.¹⁷⁸ The campaign against domestic violence sought to raise the general level of societal condemnation of the abuser.¹⁷⁹ A coalition of governmental and nongovernmental organizations hoped to alter people's views on drug use in order to reduce the conduct.¹⁸⁰ The nonprofit group Mothers Against

177. See STEPHEN J. SCHULHOFER, UNWANTED SEX: THE CULTURE OF INTIMIDATION AND THE FAILURE OF LAW 20–23 (1998); Dan M. Kahan, *Gentle Nudges vs. Hard Shoves: Solving the Sticky Norms Problem*, 67 U. CHI. L. REV. 607, 607–08 (2000) [hereinafter Kahan, *Gentle Nudges*].

178. See *infra* Part IV.D.1.

179. For a review of antidomestic violence norm-changing techniques, see Kahan, *Gentle Nudges*, *supra* note 177, at 628–31.

180. See Freevibe.com, Drug Facts, http://www.freevibe.com/Drug_Facts/lowdown.asp (last

Drunk Driving (“MADD”) sought to change public perceptions about the harms inherent in drunk driving.¹⁸¹ Reformers pressed for changes in public views on the acceptability of sexual harassment in the workplace.¹⁸² Finally, proponents of same-sex marriage have sought to build public acceptance of both same-sex intercourse and the legal recognition of same-sex unions.¹⁸³

B. DIFFICULTIES IN CHANGING INTUITIONS

Because of the universal and intuitional nature of core judgments about justice as discussed in Part I, we suggest that these judgments cannot easily be changed. We do not deny, though, that intuitions sometimes can be overridden by the reasoning system. One of the functions of the reasoning system is to monitor the quality of the mental products of the intuitive system. But apparently, this monitoring is quite intermittent, perhaps because of the feelings of certainty produced by the automaticity of the intuitive system. Reviewing many studies, Daniel Kahneman and Shane Frederick conclude that the reasoning system’s “monitoring [of the products of the intuitive system] is normally lax, and allows many intuitive judgments to be expressed, including some that are erroneous.”¹⁸⁴ Because it is likely that we only infrequently, if ever, conclude that our intuitions of justice are in error, our reasoning system does not habitually scan our intuitive judgments of justice.

What is it that might trigger the reasoning system into a long-lasting

visited Nov. 1, 2007) (“Users will likely experience dry mouth, rapid heartbeat, some loss of coordination and poor sense of balance, and slower reaction times . . .”).

181. See MADD, <http://www.madd.org> (last visited Nov. 1, 2007) (“MADD recognizes its fundamental responsibility to give a voice to victims who have been affected by the violent crime of drunk driving—both to acknowledge the reality of their losses and to bring that reality to a society numbed by statistics.”).

182. See Kahan, *Gentle Nudges*, *supra* note 177, at 634–40.

183. See Human Rights Campaign, Questions About Same-Sex Marriage, <http://www.hrc.org/issues/5517.htm> (last visited Nov. 1, 2007) (“Married couples have the automatic right to visit each other in the hospital and make medical decisions. Same-sex couples can be denied the right to visit a sick or injured loved one in the hospital.”).

184. Daniel Kahneman, *Maps of Bounded Rationality: Psychology for Behavioral Economics*, 93 AM. ECON. REV. 1449, 1450 (2003) (citing Daniel Kahneman & Shane Frederick, *Representativeness Revisited: Attribute Substitution in Intuitive Judgment*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 49 (Thomas Gilovich, Dale Griffin & Daniel Kahneman eds., 2002)). It is also the case, however, that the products of intuitive thinking can be quite accurate. Malcolm Gladwell’s thesis in his book, *Blink*, is that intuitive judgments are often startlingly accurate, sometimes more accurate than reasoned judgments. See MALCOLM GLADWELL, *BLINK: THE POWER OF THINKING WITHOUT THINKING* 14 (2005). The problem, however, is knowing which of our many intuitive judgments are accurate.

override of our moral intuitions? Certainly not frequent discoveries of disagreements with other citizens. As we demonstrate above,¹⁸⁵ there is a remarkable degree of consensus among citizens about the relative severity of punishments due wrongdoers for various transgressions. If one became convinced that, for instance, a Benthamite version of deterrence theory was the appropriate guide to sentencing either as a matter of principle or as a conclusion of the best way to reduce the rate of criminal acts, one might then conclude that one's judgments of the appropriate punishment for any crime should be guided by deterrence calculations.¹⁸⁶ But this strikes us as both a rare event, perhaps limited primarily to academics or policy wonks, and a point of view that would be difficult to sustain. As discussed, the two contending explanations for a citizen's intuitions of justice are cultural socialization and evolutionary readiness, and both produce intuitions that are quite deep-seated.

Evidence suggests that it takes a dramatic, concerted effort to fundamentally alter a person's intuitive notions of justice. Such changes in core judgments have been notably observed in cases of coercive indoctrination, often referred to as "brainwashing." While there remains scholarly dispute over the precise definition and nature of brainwashing, the mechanisms to produce effective changes in an individual's views have been persuasively demonstrated, for example, in their effective application by the Chinese military on American soldiers captured during the Korean War. Effective coercive indoctrination manifested itself in the acceptance of the Chinese Communist ideology by POWs—some of whom, at least temporarily, had their personal ideologies seriously altered.¹⁸⁷

Coercive indoctrination commonly can be achieved through three stages: (1) isolation; (2) destruction of the preexisting self; and (3) construction of a new, indoctrinated self.¹⁸⁸ Isolation is used to exert complete control over the subject's access to information, physical condition, and environment. Destruction of self is then achieved through the psychological and physical degradation of the subject. The subject is

185. See *supra* Part I.B.

186. Jeremy Bentham, who developed the classic deterrence rationale for punishment, believed that the amount of punishment given should be determined by how much is needed to deter that crime in the future. Robinson & Darley, *The Utility of Desert*, *supra* note 59, at 454–55.

187. See generally EDGAR H. SCHEIN WITH INGE SCHNEIER & CURTIS H. BARKER, COERCIVE PERSUASION: A SOCIO-PSYCHOLOGICAL ANALYSIS OF THE "BRAINWASHING" OF AMERICAN CIVILIAN PRISONERS BY THE CHINESE COMMUNISTS (1961) (discussing the findings of a study regarding coercive persuasion in the context of the Chinese Communists' "brainwashing" of the American POWs).

188. Cf. *id.* at 119–39 (describing the three phases as "unfreezing, changing, and refreezing") (emphasis in original).

destabilized through a combination of physiological stressors (starvation, insufficient sleep, poor sanitation, physical abuse) and psychological pressures (ritual confession, peer pressure, renunciation of values). Finally, the subject's identity is reconstructed into a new self, adopting as the subject's own the beliefs and values of the indoctrinator.

But even such highly controlled and intrusive coercive indoctrination procedures have success only within limits. As applied by the Chinese, one of the more successful and systematic employments of the technique, brainwashing was found to work only on a small number of soldiers. Out of 4428 Korean War POW survivors, only 192 were deemed chargeable of a serious misconduct offense for siding with the enemy—only about 4 percent.¹⁸⁹ Thus, brainwashing would hardly seem effective if the goal is to change the views of large populations. Further, even on the persons who succumb to brainwashing, the effects often fail to hold without continued reinforcement, leading to a reversion in beliefs. After the war, without constant reinforcement, most indoctrinated POWs lost their devotion to the communist cause and, when repatriated to the United States, abandoned their attitudes concerning the superiority of communism.¹⁹⁰

Thus, even using a method as extreme as coercive indoctrination, one can fail to permanently separate a person from his preexisting beliefs and values. To effectively change such views would require intensive and intrusive coercion to bring about the initial change and a continuous reinforcement of the newly adopted views. As a practical matter this appears impossible to replicate broadly, and as a moral and political matter the coercion involved would not be tolerated in a modern liberal democracy.

What might be seen as a weaker form of coercive indoctrination is the well-documented "Stockholm Syndrome."¹⁹¹ In this, victims of hostage situations come to positively associate themselves with their captors, often adopting a sympathetic view of the captor's ideals or even, as in the heavily publicized case of Patty Hearst, joining the captor.¹⁹² In many respects,

189. Donald L. Manes, Jr., *Barbed Wire Command: The Legal Nature of the Command Responsibilities of the Senior Prisoner in a Prisoner of War Camp*, 10 MIL. L. REV. 1, 2 (1960).

190. See *Where Are Korean War Defectors Now?*, GRAYBEARDS, July–Aug. 2002, at 10, 10–11, available at http://www.kwva.org/graybeards/gb_02/gb_0208_final.pdf.

191. See generally DEE L. R. GRAHAM WITH EDNA I. RAWLINGS & ROBERTA K. RIGSBY, *LOVING TO SURVIVE: SEXUAL TERROR, MEN'S VIOLENCE, AND WOMEN'S LIVES* 1–29, 267–71 (1994) (describing the Stockholm incident and classic Stockholm Syndrome); Thomas Strentz, *The Stockholm Syndrome: Law Enforcement Policy and Hostage Behavior*, in *VICTIMS OF TERRORISM* 149 (Frank M. Ochberg & David A. Soskis eds., 1982) (same).

192. See ROBINSON, *CASE STUDIES*, *supra* note 79, at 725–32.

Stockholm Syndrome is induced and functions similarly to the more extreme version of coercive indoctrination described above. The hostage, faced with constant anxiety and/or physical abuse (or threat thereof) begins to self-deconstruct and seek the protection of the captor, who controls the situation. In the process, the captive adopts the perspective of the captor, creating both emotional and ideological bonds.¹⁹³

Thus in Stockholm Syndrome, we see that the degradation and reconstruction processes are vital to changing basic intuitions and beliefs. Isolation and anxiety alone, while capable of producing elevated levels of personal sympathy, are likely not sufficient to effect dramatic changes in a person's belief structure. This demonstrates the resilience of intuitive views. While captives may identify on a personal level with captors from isolation alone (Stockholm Syndrome can develop in a relatively brief amount of time so long as there is sufficient contact between the captor and the captive), it takes extreme interventions in order to relegate a captive to the condition necessary to be able to fully internalize a new set of values.

This again points to the difficulty in effecting societal-wide changes in intuitive notions. While it may be possible to find isolated nations where leaders manipulate the anxiety of their citizens to induce feelings of loyalty as Stockholm Syndrome writ large, even this weak form of coercive indoctrination would be incompatible with the functioning of a modern liberal democracy and, in any case, is likely to produce little other than affection for the person of the leader.

The best known case in which the legal authorities of the United States sought to produce change in thought and action patterns of the populace was the "noble experiment" of Prohibition, noted previously.¹⁹⁴ In 1920, the 18th amendment to the Constitution came into effect, followed

193. As in the coercive indoctrination of POWs, the victims are isolated with their captors, subjected to severe anxieties, and come to shift their previous views. It can, however, be termed a "weaker" form, as the captives exhibiting Stockholm Syndrome do not necessarily go through the extreme transformation of personal ideology exhibited in systematic coercive indoctrination. This may be due to the fact that often the captives are not degraded to the same extent as POWs in a hijacking or hostage situation, as was the case in the standoff that gave the syndrome its name. The Patty Hearst case may help illustrate this difference. As part of her ordeal, Hearst was held captive in a closet and physically and sexually assaulted. From this added component of physical degradation and subsequent self-reconstruction, Hearst came to adopt the entire political and social agenda of her captors. In the Stockholm incident, where the captives were merely isolated with their captors but not subjected to abuse, the captives became emotionally attached to their captors as individuals, but aside from being sympathetic to their captors' personal plights, they did not radically alter their beliefs about the world at large. See J. Michael Davis, *Brainwashing: Fact, Fiction and Criminal Defense*, 44 UMKC L. REV. 438, 439 (1976) (comparing Patty Hearst to a POW).

194. See *supra* Part III.A.3.

soon by the National Prohibition Act (the Volstead Act) and the sale and consumption of alcoholic drinks was banned. On our account, the consumption of alcohol was a rather poor candidate for criminalization in the minds of the citizens. Alcohol consumption does not fit the paradigm of an act that deliberately inflicts harm on others. Within many of the subcultures of citizens of the United States, alcohol consumption was perceived as a harmless activity, embedded in treasured subcultural practices. Realizing this, some of the advocates of Prohibition raised the specter of impoverished wives and children abused by drunken fathers, to make salient the prototype of the act of alcohol consumption inflicting harms on others. But this was not persuasive, or at least not persuasive enough, to maintain support for Prohibition.¹⁹⁵

C. WHAT INTUITIONS OF JUSTICE CAN BE CHANGED?

The picture is not entirely bleak for social engineers. Some intuitions of justice can be changed in some respects using certain methods. As has been our recurring theme, the means by which intuitions of justice can be changed is a function of their intuitive nature. Acknowledging and learning from that nature will in the end be more effective than ignoring or denying it.

The intuitive nature of justice judgments we describe in Part I suggests the intuitions that are most likely to be resistant to change and those that might be more changeable. Those wrongs closest to the “core”—that is, those prohibitions on which there is the greatest level of agreement within and across cultures, those matters upon which a society would need to have prohibitions in force in order to continue to exist—are likely to be the most difficult to change. That there is agreement across demographics and cultures suggests that these intuitions are the most immune from the influence of normal human experience that can vary so greatly across demographics and cultures.

Studies showing high levels of agreement suggest that the core intuitions include three kinds of violations: (1) physical aggression against others; (2) the taking of property; and (3) dishonesty or deception in exchanges.¹⁹⁶ Other kinds of studies support similar conclusions. The

195. See generally EDWARD BEHR, PROHIBITION: THIRTEEN YEARS THAT CHANGED AMERICA (1996) (describing the history of Prohibition).

196. Note, for example, that nearly all of the scenarios used in the recent Robinson and Kurzban study, which showed such a high level of agreement on scenario ranking, involved these core wrongs. Robinson & Kurzban, *Concordance*, *supra* note 2, at 1894–98.

intuitive core of wrongdoing seems to start with direct personal action such as hitting or stabbing another individual.¹⁹⁷ Joshua Greene's neuroscience studies support a core of such direct aggression or pushing another in a way that places him in the path of harm.¹⁹⁸

As the reader will recognize, there is a phrase common in older law books that captures this idea: offenses that are *mala in se*—actions that are wrong in and of themselves. Although this is a phrase that has gone out of common usage in modern criminal law theory,¹⁹⁹ it has some claim to describe ordinary citizens' responses when they are asked, for instance, why it is wrong to murder another. Their answer, often, is that it is just wrong. Moreover, it is an answer often given with some emotion, and brain imaging work suggests that emotional areas of the brain are involved in responding to these sorts of core wrongs.²⁰⁰

As this perspective suggests, the nearer acts are to doing direct harm to others, the more the difficulty there is in getting them perceived as not wrong. For example, the intuition that taking “candy from a baby”—more generally taking sustenance from a weaker person—is a wrong action is easily connected to a “do not do harm to others” prototype.

The other candidates mentioned above for prohibitions that we think are difficult to change, such as the one on breach of agreement or deception in exchanges, may have this status for one of two reasons. First, they may exist because of their analogical closeness to inflicting direct personal harm on another. As we suggested above, unjust takings of another's foodstuffs could have this character. More generally, crimes against property may be perceived as wrong because it is intuitively easy to make the connection between physically taking property and physically harming another as

197. As military trainers know, it is extremely difficult to get a military recruit to shoot to kill an enemy. Dave Grossman, who has extensively studied the problem, suggests that there is “a powerful, innate human resistance toward killing one's own species” DAVE GROSSMAN, ON KILLING: THE PSYCHOLOGICAL COST OF LEARNING TO KILL IN WAR AND SOCIETY xxix (1996). He does, however, go on to sketch the training systems that the military has developed that successfully overcome this resistance. But he reports that those who do kill pay a high psychological price, often of post-traumatic stress disorder. *See id.* at 87–93, 141–92, 250–61, 281–89.

198. *See* Joshua D. Greene et al., *The Neural Bases of Cognitive Conflict and Control in Moral Judgment*, 44 NEURON 389, 389 (2004). *See also* Joshua D. Greene et al., *An fMRI Investigation of Emotional Engagement in Moral Judgment*, 293 SCI. 2105, 2106–07 (2001) [hereinafter Greene et al., *fMRI Investigation*] (finding that more “personal” moral dilemmas are associated with stronger emotional responses).

199. *See, e.g.*, WAYNE R. LAFAYE, PRINCIPLES OF CRIMINAL LAW § 1.6(b) (2003) (reviewing difficulties in giving workable meaning to the distinction between *malum in se* and *malum prohibitum*).

200. *See generally* Greene et al., *fMRI Investigation*, *supra* note 198 (discussing two studies conducted of people's emotional engagement in response to “moral” and “nonmoral” dilemmas).

related forms of aggression generally. In the case of intuitions such as prohibitions against failure in keeping one's commitments, an independent underlying prototype may have been evolutionarily prepared for. At this moment, research on prototypes of wrongdoing is not far enough advanced to enable us to know how many prototypes there are of wrongdoing.

It is also the case that the development of a generalized desire to punish those who transgress these core rules is an extremely useful mechanism for a society to develop. Two sets of studies have examined the underpinnings of the urge to punish those who transgress these rules. Experimental gaming studies have demonstrated that participants have a propensity to punish defectors in prisoners' dilemma and other games that allow a person to make a "defecting" choice. These defecting choices are generally interpreted by those that suffer from them as a breach of contract, which generates considerable anger at the defector. The important finding, for our purposes, is that a player will expend resources to punish a defector. This is true even when steps are taken to insure that the player has no rational reason to inflict the punishment. One would rationally expend resources to punish a person who has defected if that would establish a reputation of being a person who does not tolerate double-crossers. Thus, those constructing the games make clear that the player will never play against that defector again, and make clear to the player that he will never encounter the defector outside the game context again. Thus, there is no self-interested reason for the player's costly infliction of punishment to establish a reputation of being willing to retaliate against defections. Still, the experiments demonstrate, the player will expend resources to punish the soon-to-disappear defector.²⁰¹

The second set of relevant studies is brain imaging studies. They illuminate what could motivate this punishment response. These studies have demonstrated that a person who has been defected against expresses a strong desire to punish when presented with an opportunity to inflict costly punishment of the defector and often takes the opportunity to administer that punishment. The magnitude of activation in the reward center of the brain predicts the magnitude of the punishment administered.²⁰²

201. For a description of a set of gaming studies, see Joseph Henrich et al., *Costly Punishment Across Human Societies*, 312 *SCI.* 1767, 1768–70 (2006). For a general discussion of these issues in the context of intuitions of justice, see Robinson, Kurzban & Jones, *supra* note 31.

202. See Dominique J.-F. de Quervain et al., *The Neural Basis of Altruistic Punishment*, 305 *SCI.* 1254, 1256–57 (2004).

D. HOW CAN INTUITIONS OF JUSTICE BE CHANGED?

If an intuition is of a sort that can be meaningfully altered, the previous discussions on the origin and changeability of intuitions hints at how change might most effectively be brought about. We note two potentially effective methods: (1) public education campaigns to manipulate the strength of analogies to core wrongs; and (2) the internalization of group beliefs. These are not mutually exclusive; a reform program might combine aspects of each approach.

1. Public Education to Manipulate the Analogy to Core Wrongs

First, it is possible through public education to educate people about negative effects of conduct that had not previously been fully appreciated and by analogizing the conduct sought to be condemned with conduct that already is seen as condemnable. This approach changes intuitions of what constitutes wrongdoing not by fighting the existing intuition but by harnessing it, by demonstrating that the conduct at issue really does have the condemnable character or effect that people's intuitions abhor.

Two examples illustrate how intuitions may be changed by such a public education campaign. For both drunk driving and cigarette smoking in banned areas, which have been either criminalized or given more severe penalties in recent years, criminalization was "successful" in that the community came to think of those actions as appropriately condemnable and criminalized.

For drunk driving, the path to criminalization involved changing the identity of the victim of drunk driving from the driver himself to the innocent bystander injured by the driver's actions. Groups such as MADD, formed of mothers who had children killed or injured by drunk drivers and were therefore tragically motivated to make drunk driving criminal, educated the public as to why it was that these actions fit the "moral wrong" prototype.²⁰³ Previously, the stereotype of a drunk driver was one of a foolish fellow, likely to crash into large objects, and if injuring anybody, injuring only himself. The MADD parents, holding up childhood or graduation pictures of their dead children, or photographs of the horrible results of car accidents caused by drunk drivers, provided a powerfully persuasive message that drunk driving was indeed conduct highly dangerous to others.

203. See generally MADD, *supra* note 181.

Cigarette smoking was a bit more complicated case, but similar in that it worked by building the strength of the analogy to condemnable physical aggression. Rather than be able to use the direct linkages to obvious victims that were effective in anti-drunk driving campaigns, the process of changing intuitive notions concerning cigarette smoking required incrementally expanding the understanding of who suffered from the smoking.

Initially, as evidence began to accumulate that smoking had remarkably harmful effects on cigarette smokers²⁰⁴ and that it was, for at least some people, highly addictive, cigarette sales were prohibited to minors. This was justified on the basis that minors are conceived of as a group unable to reason rationally about risks. This did not, however, warrant prohibiting the sale of cigarettes to adults or restricting where smokers could smoke. Smoking was seen as creating risks for the smoker, but conventional wisdom coded these as risks that adult persons could not be stopped from *choosing* to take, since it was a risk only to themselves.²⁰⁵

The next expansion of cigarette prohibitions followed the release of research showing that secondhand smoke—smoke inhaled by (innocent) bystanders who were in rooms in which smokers had filled with smoke—produced similar, although reduced, health risks to the nonsmokers. As a result, many jurisdictions turned to banning cigarette smoking in private spaces such as offices and meeting places as well as public spaces in which smokers and nonsmokers congregated. As the public became educated on the potential harm caused by secondhand smoke, people came to see that smokers inflicted real harm on other, nonconsenting people. Therefore, laws were passed to prohibit the infliction of these harms. Regulations, however, did not extend to certain public establishments such as restaurants and bars, where other options such as smoke-free sections or the ability of patrons to simply stay home were seen as appropriate reasons to continue to allow public smoking indoors.

To shift the public's intuition on the appropriateness of smoking in restaurants and bars, public education was used to change the focus from

204. It is interesting to note that the cigarette companies continued to search for “experts” who would deny this evidence, which was a sound technique for resisting what would be the otherwise persuasive, unanimous message that smoking was harmful.

205. The current film, *Thank You For Smoking*, in fact reproduces the contention of the cigarette companies that smoking is an entirely voluntary act, an argument that exists in uneasy relationship with the fact that smoking is addictive, and that many who are trying to stop smoking are in the grip of an addiction that they wish they had never started. THANK YOU FOR SMOKING (Fox Searchlight Pictures 2005).

the effects on other patrons to the effects on the employees. Unlike other patrons, the waitresses, waiters, and bartenders who were exposed to secondhand smoke were portrayed as struggling young workers, often the sons and daughters of people like us, who had not voluntarily consented to contracting cancer; they were innocent victims working in needlessly dangerous areas, driven by the need for their wages. This resulted in a growing number of cities to ban smoking in these last bastions of cigarette smokers. The rate at which the immoralization of smoking has spread through our society is remarkable; fueled by the ability of antismoking advocates to demonstrate the harmful effects of smoking on discreet categories of nonsmokers.

The ability to replicate the results of these two public education campaigns in order to change intuitions on other actions relies on the ability to analogize a given action to a clearly condemnable harm suffered by another person or group of people. The current effort to change intuitions concerning “insider trading” is an example of both the possibilities of and the limitations on changing intuitions. Insider trading, the buying and selling of stocks or bonds based on information that a person has that is not yet known to the public, has been criminalized. However, despite high profile prosecutions for violations, public intuitions of the criminality of insider trading remain somewhat complex.²⁰⁶

As predicted by our argument, it is easier for the public to view the action of insider trading as intuitively criminal when it results in the selling of a stock shortly prior to the price of the stock falling. This occurs because a victim can be established in the subsequent purchasers who suffer personal or institutional economic harm. Hence, in the recent case of the collapse of Enron, a great deal of attention was paid to the fact that many employees had their retirement savings invested in the company.

In the opposite situation, buying stock on insider information prior to it gaining value, it is more difficult to identify a victim class and thus more difficult to change intuitions of the criminality of such action. In this case, insider trading is an action that is some distance away from what we have referred to as the core prohibition, in which one person harms another. The action harms all other, later, buyers of the stock, because the buying actions of those who bought on insider information slightly raised the price of the

206. See Joan MacLeod Heminway, *Save Martha Stewart? Observations About Equal Justice in U.S. Insider Trading Regulation*, 12 TEX. J. WOMEN & L. 247, 249–54 (examining the public indifference to insider trading and the Justice Department’s actions in high-profile cases); R. Foster Winans, *Let Everyone Use What Wall Street Knows*, N.Y. TIMES, Mar. 13, 2007, at A19 (arguing that insider trading laws are impossible to enforce and should be removed).

stock at which the later buyers bought in. These later buyers constitute a large and nonpersonified class of actors who are harmed in some fairly abstract ways. Given this, it could be predicted that criminalizing this version of insider trading would be met with some skepticism from the general public.

2. Internalization of Group Beliefs

Another mechanism for changing intuitions operates in a more indirect way. Once a substantial group comes to see the conduct differently, their views can be used to help change the views of others. This might be predictable given the evolutionary and social group utility of a person adopting the intuitions of justice of others in that person's group.²⁰⁷

There is considerable social psychological literature on conformity that demonstrates that an individual who is surrounded by a group of people who voice a uniform and unanimous opinion comes to express the same opinion.²⁰⁸ And there are specific conformity studies that demonstrate that people do conform on judgments of a moral sort, when they face uniform disagreement from others. When an action is tagged by the group that surrounds the person, as the sort of act that should not be done, it may become persistently and perhaps permanently perceived by our person as morally wrong.

During the childhood period in which the child is socialized into, among other things, the moral thinking of the culture, the child often accepts the norms presented in an initially uncritical way. It is often the task of the child to articulate the underlying reasons for the moral norm in question.²⁰⁹ For example, the child, cued by the expressions of disgust on the faces of adults, comes to understand that spreading baby food on the living room rug is a transgression.

Alternately, it is possible that the person's group-induced conclusion about the act will fade over time. In psychological terms, this is the question of whether the intuition produced by the group has been "internalized" and made part of the individual's own thinking, or is only the result of an identification with the group that is temporary in nature and

207. See Robinson, Kurzban & Jones, *supra* note 31.

208. See SOLOMON E. ASCH, *SOCIAL PSYCHOLOGY* 450–500 (Oxford Univ. Press 1987) (1952).

209. See generally Richard A. Shweder & Nancy C. Much, *Determinations of Meaning: Discourse and Moral Socialization*, in *MORAL DEVELOPMENT THROUGH SOCIAL INTERACTION* 197 (William M. Kurtines & Jacob L. Gewirtz eds., 1987) (discussing children's construction of moral norms).

the intuition will revert to some previously held content as the memories of the group opinion fade. A broad generalization from the research literature is that opinions or intuitions voiced while in the orbit of groups tend to revert to prior states.²¹⁰

But there may be reason to doubt that moral opinions initially expressed because of conformity pressures from others will always revert. Groups provide a particularly powerful source of persuasive cues for moral intuitions. If many others think that some novel, complex action is wrong, this is a good cue to a person that the action is wrong if reflected on. One then may be motivated to search for reasons why the others think the action is wrong, and the search is likely to be biased to find confirmation. If certain processes are initiated while the individual is under the influence of the group consensus, then the change in the intuition can become more far-reaching or “internalized” by the individual, to use the current vocabulary.

An example of this process comes from social scientists studying the developmental steps in the radicalization of those who protest against abortion. For a considerable group of Americans, the story, in brief, is as follows. Many who opposed abortion did so initially because they thought abortion was an act in opposition to their religious beliefs. They then began to involve themselves in protests, at which point they were surrounded by others who opposed abortion. The sharing of their antiabortion attitudes shifted the group norms against abortion in a more extreme direction.

Apparently many antiabortion sympathizers became much more extreme in their beliefs when they watched a film of a sonogram of an abortion of a fetus, looking childlike, and apparently struggling to avoid the abortionist’s needle.²¹¹ Witnessing this event with others who already thought that abortion was a wrong amplified the processes by which everyone’s belief became more extreme. It was from people with this set of experiences that many of the “rescuers” came. “Rescuers” were those who became willing to inflict violence on those who performed abortions and several did so.²¹²

As this suggests, it is often the case that the group will actively provide the reasons that an action should be considered morally wrong, generally by analogizing it to a paradigm of one person doing physical

210. See PETTY & CACIOPPO, *ATTITUDES & PERSUASION*, *supra* note 92, at 259.

211. *THE SILENT SCREAM* (1984) (depicting the abortion procedure via ultrasound where the fetus appears to react in pain).

212. See generally JAMES RISEN & JUDY L. THOMAS, *WRATH OF ANGELS: THE AMERICAN ABORTION WAR* (1998) (discussing the antiabortion movement).

harm to another. This predisposition to take account of others' expressed views on wrongdoing and blameworthiness can create a potential for change, in concert with the processes of recruiting beliefs that support that change. In our modern society, one's relevant group may be a social, political, ethnic, or religious group, whatever it is with which the person identifies.

These processes may work in the service of either immoralizing behavior previously thought allowable, or demoralizing behavior previously thought wrong. As an example of the latter, consider the rookie on the high school football team. He begins with the usual taboos about harming others. But he discovers that all the older team members, the people he looks up to and from whom he desperately wishes acceptance, express that certain aggressive rule-breaking behavior on the sports field is the right way to play. He may then begin to provide reasons, which will be justifications, for that view from his own mind. It would occur to him that enemy players are well padded by the protective gear they wear and will not really be hurt. He might be reminded of other times he has learned that the real rules that govern behavior are not always the ones in the rulebooks, and so on. In other words, the person is conducting a search for information about the moral rightness of an action, but the search is a biased one, seeking to find the reasons that the group might think it right.

Another example may make this clearer, and an unpleasant but realistic one comes to mind. In some so-called male "affinity groups," young men engage in partially organized, partially spontaneous, aggressive sexual assaults on women. Investigations often discover that the ideology of the group into which they socialize new members centers around what social scientists call "rape myth beliefs." What is contested by these beliefs is whether nonconsensual sexual intercourse, inflicted on a woman, is harm. An example of this is the view that girls often say no when they really mean yes.²¹³ One holding these beliefs considers sexual relations to be based on getting what one wants from the other, and that coercive force is an appropriate way to get it.

Thus the mechanism of conformity that is used to change intuitions can motivate individuals to internalize either enhanced or diminished views of the criminality of certain behaviors. The difficult challenge for this approach to change is not the initial changing of the view under group influence, but rather is causing the view to be internalized such that it will

213. See generally Martha R. Burt, *Cultural Myths and Supports for Rape*, 38 J. PERSONALITY & SOC. PSYCHOL. 217 (1980) (describing a study conducted regarding rape myths).

survive a diminishing influence of the group.

V. CONCLUSION

One of the great issues in the foundations of criminal law concerns whether and to what degree the moral intuitions of the community should find a place in the legal codes governing that community. In this Article, we have argued that these moral intuitions have a considerable claim to be incorporated in criminal codes. The justifications that we cite are somewhat unusual, at least in legal doctrinal circles. We suggest that, at least for the core wrongs, the *mala in se* principles of the criminal law and evolutionary and cultural pressures coalesce to cause community members to hold those principles in a way that gives them a privileged place in their hierarchy of moral judgments. Specifically, those judgments arise intuitively, which means that they are arrived at automatically, quickly, and without any perceived expenditure of complex thought. This gives them a “force” in the mind of the judgment maker that is quite powerful.

Across individuals in a culture, and often across individuals in different cultures, there is a remarkable degree of consensus in these judgments, particularly in the relative seriousness rankings of the degree of blameworthiness of various moral transgressions. This means that a society has available to it a possible principle for doing justice, which is to punish according to this societally shared sense of the moral blameworthiness of the offender. This is what we have referred to as “empirical desert,” to distinguish it from the “deontological desert” that retributivist moral philosophers urge.

As a way of allocating punishments, the empirical desert principle has various advantages of a utilitarian sort. This is somewhat counterintuitive since the principle is a version of a just deserts system for distributing justice, which is generally thought to be the one principle for the distribution of sanctions that is deontological rather than utilitarian. The advantages generally lie in sustaining the community’s belief in and commitment to the justice system, and avoiding the contempt-produced disengagement that follows perceived violations of doing justice. Empirical evidence begins to accumulate that a belief that a society’s system does do justice fosters society members’ voluntary obedience to its rules.

This reliance on community intuitions of justice would indeed be morally repugnant if lawgivers were to solely rely on those intuitions, fixed and unchangeable, to determine liability assignments. This is so because there are numerous occasions when specific societies have held shared

moral intuitions and designed criminal penalties around those intuitions that we now think are morally wrong—slavery was one such example we cited. But that static system is not the system we are advocating. Evidence is that moral intuitions can be changed, and we devoted considerable analysis explaining how that process works. We suggested that this process of moralization, or occasionally demoralization, is most likely to succeed when it works in connection with underlying images of core wrongs. We also pointed out that, by enlisting reasoning processes, it is possible for persons to override the products of their intuitive system, although this override response sometimes can be difficult to sustain over time if not internalized.

We examined these processes of changing moral intuitions because it is clear that such processes are used in practice and should be understood, and because such processes are important to allow a society to reform its criminal law doctrines. To do this intelligently requires reformers to make their case for the wrongness of certain actions and for why certain other actions, that intuitively seem wrong, are in fact not so. The task is to modify the intuitions of others in the society, or persuade them that this is one of those times that intuitions should be overridden—neither an easy task nor an impossible one.

