1-1-2011

Mercy, Crime Control & Moral Credibility

Paul H. Robinson
University of Pennsylvania Carey Law School

Merciful Judgments and Contemporary Society

LEGAL PROBLEMS, LEGAL POSSIBILITIES

Edited by

Austin Sarat

Amherst College

CAMBRIDGE
UNIVERSITY PRESS

2

Mercy, Crime Control, and Moral Credibility

Paul H. Robinson

Should mercy have a role in the operation of the criminal justice system? That is the question this chapter will examine. As a preliminary matter, however, consider what *mercy* does and does not mean.

I. What Does *Mercy* Mean?

The standard definition of *mercy* – "kindness in excess of what may be expected or demanded by fairness"[1] – makes clear that much of what sometimes passes for "mercy" in the criminal law context is no such thing. When a very light sentence is given to a young first-time offender, it is not an exercise of mercy. If a criminal justice system purports to be in the business of doing justice,[2] then normal principles for assessing blameworthiness would take account of the youthfulness and inexperience of the offender. On a conclusion of significantly reduced blameworthiness, moral

[1] *Webster's New World Dictionary of the American Language* 921 (1962).

[2] Of course, there are many who would argue that the criminal justice system should have goals other than doing justice. The most common alternative distributive principle for criminal liability and punishment is one that would promote crime reduction without regard to whether it does justice. For discussion of these alternative principles, see Paul H. Robinson, *Distributive Principles of Criminal Law* chs. 3–6 (Oxford, 2008) (hereinafter Robinson, *Distributive Principles*).

The author thanks Eli S. Rubin, University of Pennsylvania Law School, class of 2011, for his very useful research assistance.

desert would demand a significantly reduced sentence. Such is not an exercise of mercy but simply careful application of principles of desert. Many writers have confused true and careful justice with mercy,[3] but others have pointed out the confusion.[4] It is an important distinction to get right in part because an offender has a right to justice, but classically mercy is a matter of grace.

Desert is a rich and nuanced concept that takes account of a wide variety of factors: not just the extent of the harm or wrongdoing but also the culpability level of the offender (purpose, knowing, reckless, negligent, or faultless as to the wrongdoing), as well as possible justifications for the offender's conduct (self-defense, law enforcement authority, or lesser evils) and possible excuses (mental illness, involuntary intoxication, duress, or in the previous example, immaturity). Any of these factors may have the effect of providing a mitigation in punishment, or even a complete defense. If a writer misconceives the demands of desert to be something more narrow, wooden, or objective,[5] then the writer is likely to label as "mercy" cases that are in truth cases of

[3] For examples of writers who confuse mercy with deserved mitigations, see, e.g., Claudia Card, "On Mercy," *Philosophical Review* 81 (1972): 184 ("Mercy ought to be shown to an offender when it is evident that otherwise *he would be made to suffer unusually more* on the whole, owing to his peculiar misfortunes, *than he deserves* in view of his basic character") (emphasis added); Martha C. Nussbaum, "Equity and Mercy," *Philosophy and Public Affairs* 22 (1993): 85, 90, 103 (Nussbaum sees mercy as an important part of overcoming the "we versus them" conflict inherent in retributivist analysis and as playing an inherent part in undertaking the "narrative approach" necessary in recognizing a wrongdoer's true guilt).

[4] See, e.g., Alwynne Smart, "Mercy," *Philosophy* 43 (1968): 345. Smart describes two meanings of mercy. Of the first, she says: "We exercise mercy to avoid an unduly harsh penalty which an insufficiently flexible legal system would impose upon the offender. In other words we exercise 'mercy' to avoid an injustice." She then concludes: "There is some impropriety in calling cases like these cases of genuine mercy." Id. at 358–59.

[5] This is not uncommon, especially among those who oppose assessing punishment according to desert. To construct an obviously unattractive conception of desert is to provide a straw man that is easier to discredit. See the discussion of the "vengeful" conception of desert at Paul H. Robinson, "Competing Conceptions of Modern Desert," *Cambridge Law Journal* 67 (2008): 147–48 (hereinafter Robinson, "Competing Conceptions").

mitigated desert. (In this way, the conception of mercy can serve as an interesting test of a writer's sophistication in understanding moral blameworthiness.)

Misconceiving desert as something wooden, narrow, or objective can be forgiven, however, because much of current American criminal law sometimes follows liability rules of this sort; the rules do not track a robust conception of desert. If a person assumes that American criminal justice uses desert as its distributive principle, then the conception of desert on display is clearly something less than the robust version described here. However, the current criminal law deviations from true desert are not the result of the drafters' lack of appreciation for the broader demands of desert but rather a rejection of desert as the guiding distributive principle. Desert and its demands are well understood, even by laypersons. (Indeed, the evidence is that laypersons share a quite sophisticated and nuanced understanding of desert.[6]) The current system's tendency to shortchange desert arises rather from its attempts to advance a utilitarian crime-control agenda at the expense of desert.[7]

To return to the original point, mercy does not include mitigations in punishment that are already demanded by a true conception of desert. Exercising mercy is different from avoiding injustice. It is forgoing punishment that is in fact deserved, even under a robust conception of desert that is fully sensitive to deserved mitigations. What can be said in favor of giving undeserved mitigations in punishment?

II. The Virtue of Mercy in the Exercise of Personal Judgment and Action

Although mercy may not be something deserved by the recipient, it can be a clear virtue in those who give it. The forgiving victim

[6] See, e.g., Paul H. Robinson and John M. Darley, *Justice, Liability, and Blame: Community Views and the Criminal Law* (Westview, 1995) (hereinafter Robinson and Darley, *Justice, Liability, and Blame*).

[7] For a discussion of the instances and reasons for current criminal law's deviation from desert, see Paul H. Robinson and Michael T. Cahill, *Law without Justice: Why Criminal Law Doesn't Give People What They Deserve* (Oxford, 2006).

shows herself to be a virtuous person. Indeed, it is because the victim has every right to demand just punishment of her victimizer that her forgoing that right has such emotional power. That exercise of forgiveness tells us something about the forgiver that endears her to us. We would hope that we ourselves in a similar situation could emulate such forgiveness against our aggressors.

But note that the virtue of mercy is about the giver.[8] It is a virtuous person that forgives her aggressor. It is not necessarily virtuous for one to forgive another person's aggressor. What virtuous sacrifice has this forgiver made? What right has this nonvictim to forgive? If one has not been victimized, where is the virtue in forgiving? Thus, we may ask, where is the virtue in a sentencing judge's forgiving an aggressor's harm of others?

Beyond abstract virtue, one might also observe that the exercise of mercy, compassion, and forgiveness are of practical value. One spouse's compassion in response to errors by the other can help build an enriching and enduring relationship. The same can be said for compassion between friends. Indeed, a world in which strangers show compassion to one another would be a better world for all. But, of course, this is true because there is no societal cost in such exercise of compassion. There might be a personal cost to the victim of the transgression who forgives, but that is a cost that the victim has chosen to suffer, which she ought to be free to do. As discussed in sections 3 and 4, however, where the exercise of compassion frustrates a criminal offender receiving the punishment he truly deserves, it can have serious societal costs, in both frustrating justice and promoting avoidable crime.

[8] This is consistent with St. Anselm's view:

> Justice to sinners obviously requires that God punish them; but God's justice *to himself* requires that he exercise his supreme goodness in sparing the wicked. "Thus," Anselm says to God, "in saving us whom you might justly destroy . . . you are just, not because you give us our due, but because you do what is fitting for you who are supremely good."

Thomas Williams, "Introduction," in *Anselm's* Proslogion (p. X) (Hackett, 2001) (emphasis in original).

One may conclude that mercy has no proper place outside of the context of interpersonal relationships. An expression of mercy by an injured to her injurer is a quite different matter from the institutionalization of the failure to give truly deserved punishment by a criminal justice system. The practice is suspect not only because it is not the victim who is showing mercy, and not only because such institutionalization creates societal costs that do not occur in the context of private exercises of mercy, but also because, as section 7 discusses, by codifying mercy, one creates an expectation, if not a right, which may well undermine the virtue of giving mercy – it is a virtue to give it precisely because it is not required.

III. The Problem of Mercy in a Criminal Justice System Designed to Achieve Effective Crime Control

As noted, even if one can imagine enormous benefits to the exercise of mercy by individuals in their interpersonal dealings, it does not follow that mercy would be desirable as an institutionalized component of the criminal justice system. Certainly, the institutionalization of mercy would tend to undermine the effectiveness of traditional coercive crime-control mechanisms. Consider each of those standard mechanisms in turn.

A system built on the distributive principle[9] of efficient deterrence would set punishment such that the threat it signals will be the lowest that is sufficient to effectively deter others (or, in the case of special deterrence, to deter the offender) from future criminality. To impose less punishment than this deterrence principle calls for, in the exercise of mercy, would produce a threat of punishment that was less than that needed to effectively deter, and therefore would produce otherwise preventable crimes.[10]

[9] The phrase "distributive principle" is meant to refer to the guiding criteria by which the criminal justice system determines who is to be punished how much.

[10] Of course, any deterrence analysis makes sense only if it is likely that the prerequisites for effective deterrence exist. There are good reasons to suggest

Similarly, a system built on the principle of efficient rehabilitation or incapacitation of the dangerous would set punishment at the lowest sufficient to effectively rehabilitate or incapacitate. Again, an exercise of mercy to give the dangerous offender less punishment than is needed means releasing the offender while he is still dangerous or still in need of rehabilitation, thereby inviting otherwise preventable crime.

Thus, as a general matter, the exercise of mercy tends to undermine effective crime control, at least when it is promoted through the traditional coercive mechanisms of deterrence, rehabilitation, or incapacitation. One can imagine a mercy mitigation as less damaging in these situations only where the crime-control principles worked badly – that is, where they provide only a rough approximation of the efficient sentence and where one might allow a mercy mitigation limited to a gray area in which it is unclear whether that amount of punishment really is needed for effective crime control – that level of deterrent threat may or may not be necessary, or the offender may or may not need that length of incapacitation. But, of course, if the system is sufficiently bad as to be unclear as to the sentence that is really required, then it presumably also is sufficiently bad as to not know whether a mercy mitigation really could be allowed without undermining effective crime control. Such weakness in the system simply means that, at best, some mitigations for mercy create only risks of creating greater crime rather than ensuring it.

Although the foregoing analysis might suggest that some sort of mercy mitigation in some cases might (or might not) be tolerable, as not clearly undermining effective crime control, the analysis offers no positive justification for why such a risk of undermining crime control ought to be taken. At best it suggests that that the negative effect of a mercy mitigation might be limited. Unless one can identify some positive effect, it is hard to see why the system should take even a risk of undermining effective crime control. Section 2 discusses personal and societal benefits that may flow from the exercise of mercy in personal dealings by

that they commonly do not. See Robinson, *Distributive Principles*, at chs. 3 and 4.

individuals. What positive benefits flow from the institutionalization of mercy, in particular its institutionalization within the criminal justice system? What we can say here is that certainly there is no reason to think that the institutionalization of mercy will promote effective crime control, at least not through the traditional coercive mechanisms of deterrence, rehabilitation, or incapacitation of the dangerous.

IV. The Problem of Mercy in a Criminal Justice System Designed to Impose (Deontologically) Deserved Punishment

It may seem no surprise to conclude, as the previous section does, that the institutional exercise of mercy is likely to be inconsistent with effective crime control. That is the common wisdom surrounding mercy and crime control. In contrast, perhaps a mercy mitigation would be more at home if the distributive criteria for criminal liability and punishment were not effective crime control but desert? At least mercy and desert seem to share a nonutilitarian focus. Unfortunately, as is apparent from the discussion that follows, it may be just as clear that the institutionalized exercise of mercy can as seriously undermine the goal of imposing deserved punishment.

As noted in section 1, an offender's deserved punishment – proportioned to the offender's moral blameworthiness – is classically a function of the nature and extent of the wrongdoing, the accompanying culpable state of mind, justifying circumstances at the time of that wrongdoing, and a robust assessment of the offender's capacity to have avoided the wrongdoing. Giving an offender less punishment than is deserved according to a full assessment of these criteria is to fail in the goal of imposing deserved punishment – to fail to do justice. In other words, essentially by definition, giving mercy is inconsistent with giving deserved punishment, doing justice.[11]

[11] For a philosophical account of the natural tension between mercy and justice, see Jeffrie G. Murphy, "Mercy and Legal Justice," *Social Philosophy and Policy* 4 (1986): 1, reprinted in Jeffrie G. Murphy and Jean Hampton,

Could it be, however, in a way similar to the gray-area analysis previously, that in the context of crime control, one might allow a mercy mitigation within the bounds of some natural flexibility inherent in desert? There was in fact a school of thought of a decade or so ago that argued that the demands of desert are sufficiently vague that they set only outer bounds rather than a specific sentence, which leaves a good deal of flexibility within which one can take account of other goals and purposes. Those writers sought to use what they saw as the flexibility of desert to justify reliance on coercive crime-control principles, such as deterrence, rehabilitation, and incapacitation of the dangerous,[12] but one might make the same argument in favor of allowing mercy mitigations.[13]

Unfortunately (for this use of the mercy mitigation but fortunately for criminal law theory), this line of argument has, in my own view, been fully discredited. The notion that desert is an inherently vague distributive principle that sets only outer limits, not demands for specific amounts of punishment, is based on a failure to appreciate the specific demands of desert. The

Forgiveness and Mercy (Cambridge, 1988). Murphy concludes that if mercy is to stand as an autonomous virtue, and is to allow a wrongdoer to receive less than justice demands, it is not a virtue at all but a vice. He argues that "all the talk of mercy is not worthless" but rather appropriate in the "Private Law context," as opposed to the "Criminal Law context." He takes this distinction from P. Twambley, "Mercy and Forgiveness," *Analysis* 36 (1976): 84.

[12] Norval Morris, *The Future of Imprisonment* 75–76 (Chicago, 1974); American Law Institute (ALI), *Model Penal Code: Sentencing, Report* (2003): 4 (reporting that the ALI endorses "A new vision of sentencing purposes, borrowing from Norval Morris's theory of limiting retributivism, that organizes retributive and utilitarian goals and makes them applicable to decisionmakers throughout the sentencing system."); Paul H. Robinson, "The A.L.I.'s Proposed Distributive Principle of 'Limiting Retributivism': Does It Mean In Practice Anything Other Than Pure Desert?" *Buffalo Criminal Law Review* 7 (2003): 3–5, 10–12; Robinson, "Competing Conceptions," at 160–64.

[13] This is an aspect of the approach of Nathan Brett, "Mercy and Criminal Justice: A Plea for Mercy," *Canadian Journal of Law and Jurisprudence* 5 (1992): 94 (arguing that punishment need not be proportionate to blameworthiness and that desert sets only an upper boundary for punishment).

confusion arises in part from the failure to distinguish two distinct judgments: setting the high-end point of the punishment continuum and, once that high-end point has been set, ordinally ranking cases of different blameworthiness along that continuum. Every society must decide what punishment it will allow for its most egregious case, be it the death penalty or life imprisonment or fifteen years. Once that high-end point is set, the demand of desert becomes specific: it demands that more blameworthy offenders receive more punishment than less blameworthy offenders. That is, desert requires an ordinal ranking of offenders' punishment according to their relative blameworthiness.[14] The result is a specific amount of punishment for a particular offense and offender. That amount of punishment is not the product of some magical connection between that violator's offense and the corresponding amount of punishment deserved. Rather, it is the specific amount of punishment needed to set the offender's violation at its appropriate ordinal rank according to blameworthiness relative to all other offenses. If the high-end point of the punishment continuum were changed, for example, the deserved punishment for each offender would change accordingly.[15]

As it happens, empirical studies about people's intuitions of justice match this notion of desert as a specific ordinally ranking amount of punishment, not just a vague outer boundary of punishment. The studies show that people set punishment according to blameworthiness (not deterrence or dangerousness)[16] and have quite nuanced and sophisticated judgments about the

[14] Andrew von Hirsch, *Past or Future Crimes: Deservedness and Dangerousness in the Sentencing of Criminals* 39–46 (Manchester, 1985); Robinson, "Competing Conceptions," at 150–51.

[15] For a general discussion, see Robinson, "Competing Conceptions"; Robinson, *Distributive Principles*, at ch. 7; Paul H. Robinson, "Empirical Desert," in *Criminal Law Conversations*, 29, 31–32, ed. Paul H. Robinson, Stephen P. Garvey, and Kimberly Kessler Ferzan (Oxford, 2009) (hereinafter Robinson, "Empirical Desert").

[16] Kevin M. Carlsmith, John M. Darley, and Paul H. Robinson, "Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment," *Journal of Personality and Social Psychology* 83 (2002): 292; John M. Darley, Kevin M. Carlsmith, and Paul H. Robinson, "Incapacitation and Just Deserts as Motives for Punishment," *Law and Human Behavior* 24 (2000): 676.

relative blameworthiness of different offenders based on quite subtle differences. Even the average layperson can make an enormous number of distinctions among offenders in similar situations.[17] To fit the high number of distinguishable cases on the limited continuum of punishment, once a continuum high-end point has been set, each offender must receive a specific amount of punishment to properly distinguish that offender from other offenders of distinguishable degrees of blameworthiness.

The point here is that mitigations of punishment based on mercy – factors unrelated to the offender's blameworthiness – will set the offender in the wrong rank order as compared to other offenders. The exercise of mercy as clearly undermines desert as it undermines effective crime control.

V. The Interesting Case of Mercy in a System Based on Empirical Desert, Designed to Enhance the Criminal Law's Moral Credibility

What has been said in the previous section applies to the classic notion of moral desert – moral blameworthiness as reasoned out by the principles of moral philosophy. However, the analysis is interestingly different when one considers mercy from the point of view of what has been called empirical desert, by which is meant determining blameworthiness according to the shared intuitions of justice of the community. Before working through this analysis, let me lay a foundation by highlighting the difference between deontological and empirical desert and by explaining why one might want to rely on the latter, rather than the former, as the distributive principle for criminal liability and punishment.

A. Empirical Desert and Its Rationales

The difference between deontological desert, derived from the reasoned analysis of moral philosophy, and empirical desert, derived from the shared intuitions of justice of the community

[17] See, e.g., Robinson and Darley, *Justice, Liability, and Blame*.

to be bound by the law, can produce important differences in the distribution of liability and punishment. For example, moral philosophers disagree about the significance of resulting harm, and each side of the debate has plausible arguments to make.[18] In contrast, all available data suggest a nearly universal and deeply held view among laypersons that resulting harm does matter. The absence of a resulting harm or evil reduces the actor's blameworthiness; the presence increases it.[19] Thus, the two alternative distributive principles disagree about whether resulting harm should even be an element of an offense definition, whether it should affect an offense's grade, and whether completed offenses should be punished more than unsuccessful or interrupted attempts.[20] This is only one of a host of issues on which moral philosophy's

[18] Those arguing that resulting harm should matter include: Leo Katz, "Why the Successful Assassin Is More Wicked Than the Unsuccessful One," *California Law Review* 88 (2000): 806 (arguing by hypothetical that principled moral analysis suggests that harm should be considered when assessing blameworthiness); Ken Levy, "The Solution to the Problem of Outcome Luck," *Law and Philosophy* 24 (2005): 263; Michael S. Moore, "The Independent Moral Significance of Wrongdoing," *Journal of Contemporary Legal Issues* 5 (1994): 267–71 (positing that our own experiences – we feel more guilty about our own completed misdeeds than we do about attempts, and we are dissatisfied with reasonable moral choices that produce undesirable consequences – suggest that "results matter" in the moral arena). However, there is significant disagreement in this arena. See, e.g., Joel Feinberg, "Equal Punishment for Failed Attempts: Some Bad but Instructive Arguments Against It," *Arizona Law Review* 37 (1995): 119; Sanford H. Kadish, "The Criminal Law and the Luck of the Draw," *Journal of Criminal Law and Criminology* 84 (1994): 686 ("punishing attempts and completed crimes differently makes no sense insofar as the goal of the criminal law is to identify and deal with dangerous offenders who threaten the public"); Stephen J. Morse, "The Moral Metaphysics of Causation and Results," *California Law Review* 88 (2000): 879.

[19] See, e.g., Robinson and Darley, *Justice, Liability, and Blame*, at 14–28, 181–96 (reporting empirical studies).

[20] For a discussion of the issue of the significance of resulting harm in the construction of modern criminal law, see Paul H. Robinson, "The Role of Harm and Evil in Criminal Law: A Study in Legislative Deception?" *Journal of Contemporary Legal Issues* 5 (1994): 304–21; Paul H. Robinson and John M. Darley, "Objectivist vs. Subjectivist Views of Criminality: A Study in the

analytic conclusions are likely to vary from the empirical data on laypersons' intuitions of justice.[21]

Perhaps even more important than such differences in blameworthiness judgments are the differences between the underlying theories that drive these two conceptions of desert and thereby shape their application. The special value of the deontological conception of desert is its ability to produce, at least theoretically, transcendent principles of justice – true justice in an ideal sense. Unfortunately, there is no mechanism by which this goal can be achieved as a practical matter – to produce more just criminal codes, for example – because philosophers famously disagree among themselves about much, if not most, issues relating to desert. How can code drafters determine which of these competing philosophical conclusions is the "right" conclusion, which represents true justice?

The empirical conception of desert, in contrast, is not "true" justice but only the community's conception of justice. It may be the best approximation of deontological desert that it is practical to hope for in the real world, but it is not true justice in a transcendent sense. It is considered an attractive distributive principle for criminal liability and punishment, however, not only because it might serve as a workable approximation of true justice but also because, by distributing criminal liability and punishment in this way, the criminal law enhances its moral credibility with the community, which, in turn, can help the law harness the powerful forces of social influence and internalized norms. Indeed, there is reason to believe that such normative crime control can be more powerful, more effective, than the traditional mechanisms of coercive crime control of deterrence, rehabilitation, and incapacitation of the dangerous. This is because of in part the strength of the normative forces that arise from criminal

Role of Social Science in Criminal Law Theory," *Oxford Journal of Legal Studies* 18 (1998): 442–44.

[21] For community views on a variety of criminal law issues that may conflict with moral philosophers' views, see generally Paul H. Robinson, "The Role of Moral Philosophers in the Competition between Philosophical and Empirical Desert," *William and Mary Law Review* 48 (2007): 1831.

law's moral credibility and in part the practical limitations and weaknesses of traditional coercive crime-control mechanisms.[22] Let me try to summarize the practical attraction of empirical desert.[23]

I have argued elsewhere that empirical desert is an attractive distributive principle because by building the moral credibility of the system it can promote cooperation and acquiescence with it, harness the powerful social influences of stigmatization and condemnation, and increase the criminal law's ability to shape societal and internalized norms. (Others have argued that empirical desert is an attractive distributive principle because it promotes democratic ideals or because it is the best approximation of deontological desert that is reasonable to expect in the real world.[24])

Some of the system's power to control conduct derives from its potential to stigmatize violators – with some potential offenders this is a more powerful, yet essentially cost-free control mechanism as compared to imprisonment. Yet the system's ability to stigmatize depends on it having moral credibility with the community. That is, for a conviction to trigger community stigmatization, the law must have earned a reputation for following the community's view on what does and does not deserve moral condemnation. Liability and punishment rules that deviate from a community's shared intuitions of justice undermine this reputation.

[22] For a full discussion, see Robinson, *Distributive Principles*, at ch. 10.

[23] For a fuller account, see Paul H. Robinson and John M. Darley, "Intuitions of Justice: Implications for Criminal Law and Justice Policy," *Southern California Law Review* 81 (2007): 1–67; Robinson, *Distributive Principles*, at chs. 8 and 12; Robinson, "Empirical Desert," at 29–39; Paul H. Robinson, Geoff Goodwin, and Michael Reisig, "The Disutility of Injustice," *New York University Law Review* 85 (2010) (hereinafter, Robinson, Goodwin, and Reisig, "Disutility").

[24] See, e.g., Andrew E. Taslitz, "Empirical Desert: The Yin and Yang of Criminal Justice," in *Criminal Law Conversations* 56, ed. Paul H. Robinson, Steve P. Garvey, and Kimberly Kessler Ferzan (Oxford, 2009); Adil Ahmad Haque, "Legitimacy as Strategy," id. at 57.

The effective operation of the criminal justice system depends on the cooperation, or at least the acquiescence, of those involved in it – offenders, judges, jurors, witnesses, prosecutors, police, and others. To the extent that people view the system as unjust – as in conflict with their intuitions about justice – that acquiescence and cooperation is likely to fade and be replaced with subversion and resistance. Vigilantism may be the most dramatic reaction to a perceived failure of justice, but a host of other less dramatic (but more common) forms of resistance and subversion have shown themselves. Jurors may disregard their jury instructions. Police officers, prosecutors, and judges may make up their own rules. Witnesses may lose an incentive to offer their information or testimony. And offenders may be inspired to fight the adjudication and correctional processes rather than participating and acquiescing in them.

Criminal law also can have effect in gaining compliance with its commands through another mechanism: if it earns a reputation as a reliable statement of what the community perceives as condemnable, people are more likely to defer to its commands as morally authoritative and as appropriate to follow in those borderline cases in which the propriety of certain conduct is unsettled or ambiguous in the mind of the actor. The importance of this role should not be underestimated; in a society with the complex interdependencies that characterize ours, a seemingly harmless action can have destructive consequences. When the action is criminalized by the legal system, one would want the citizen to respect the law in such an instance, even though he or she does not immediately intuit why that action is banned. Such deference will be facilitated if citizens believe that the law is an accurate guide to appropriate prudential and moral behavior.

Perhaps the greatest utility of empirical desert comes through a more subtle but potentially more influential mechanism. The real power to gain compliance with society's rules of prescribed conduct lies not in the threat of official criminal sanction, but in the influence of the intertwined forces of social and individual moral control. The networks of interpersonal relationships in which people find themselves, the social norms and prohibitions

shared among those relationships and transmitted through those social networks, and the internalized representations of those norms and moral precepts control people's conduct. The law is not irrelevant to these social and personal forces. Criminal law, in particular, plays a central role in creating and maintaining the social consensus necessary for sustaining moral norms. In fact, in a society as diverse as ours, the criminal law may be the only societywide mechanism that transcends cultural and ethnic differences. Thus, the criminal law's most important real-world effect may be its ability to assist in the building, shaping, and maintaining of these norms and moral principles. It can contribute to and harness the compliance-producing power of interpersonal relationships and personal morality, but it will be effective in doing so only if it has sufficient credibility.

The extent of the criminal law's effectiveness in all these respects – in bringing the power of stigmatization to bear; in avoiding resistance and subversion to a system perceived as unjust; in gaining compliance in borderline cases through deference to its moral authority; and in facilitating, communicating, and maintaining societal consensus on what is and is not condemnable – is to a great extent dependent on the degree to which the criminal law has gained moral credibility in the minds of the citizens governed by it. Thus, the criminal law's moral credibility is essential to effective crime control, and it is enhanced if the distribution of criminal liability is perceived as "doing justice" – that is, if it assigns liability and punishment in ways that the community perceives as consistent with its shared intuitions of justice. Conversely, the system's moral credibility, and therefore its crime-control effectiveness, is undermined by a distribution of liability that conflicts with community perceptions of just desert.

Confirming the findings of previous studies,[25] the most recent set of studies show that many modern crime-control doctrines seriously conflict with the community's shared intuitions of justice, that this conflict does indeed undermine the criminal law's moral credibility, and that this loss does indeed have practical

[25] See Robinson, Goodwin, and Reisig, "Disutility," at pt. V.F.

consequences that undermine the criminal justice system's crime-fighting effectiveness.[26]

B. Exercising Mercy to Enhance the Criminal Law's Moral Credibility

If a criminal justice system were to adopt as its goal enhancing its moral credibility with the community it governs, would mercy appropriately have a role in the system? At first blush, one might suspect that the analysis would run parallel to what has been said here in the context of crime control and deontological desert. That is, to the extent that reducing punishment in the exercise of mercy causes the resulting punishment to conflict with the community's shared intuitions of justice, then that exercise of mercy would undermine the criminal law's moral credibility with the community, and thereby undermine its crime-control effectiveness.

What makes things quite interesting here, however, is that empirical studies of laypersons' intuitions suggest that some circumstances or factors that commonly prompt inclinations toward the exercise of mercy commonly have the support of laypersons in assessing punishment. They might not consider the factor as reducing the offender's blameworthiness, but they nonetheless may feel that his or her punishment ought to be reduced.

For example, a recent study showed that a majority of laypersons supported a mitigation of punishment when, after the offense was complete and before being arrested by police, the offender showed true remorse, acknowledged guilt, and offered a sincere apology.[27] On average they more than halved the punishment

[26] See Robinson, Goodwin, and Reisig, "Disutility," at pts. V and VI.

[27] Paul H. Robinson, Matthew Majarian, Thomas Gaeta, and Daniel M. Bartels, "Extralegal Punishment Factors: A Study of Forgiveness, Hardship, Good-Deeds, Apology, Remorse, and Other Such Discretionary Factors in Assessing Criminal Punishment," pt. 2 (in progress, draft available) (hereinafter Robinson et al., "Extralegal"). The percentage of subjects giving a mitigation varied according to the type and seriousness of the offense, varying from a high of 74 percent (in a case involving theft with trespass) to a low of 38 percent (in a case involving intentional killing and abduction), for an average of 57 percent across five kinds of cases. Id.

that they would otherwise have given.[28] (A truly remorseful public acknowledgment of guilt and apology but only after arrest also supported a mitigation by a majority of subjects, but not in all cases.[29]) In another example, a majority of laypersons supported a mitigation when the punishment of the offender would render a hardship on the offender's family, at least in some cases.[30] Those giving the mitigation more than halved the punishment they otherwise would have given.[31] A final example of a mitigation supported by a majority of subjects was when the offender had already paid substantial civil compensation to the victim, again in some cases only.[32] Again, the subjects giving

[28] Of those who gave the mitigation, the average mitigation across all five kinds of cases was −1.124, where −1 is a halving of punishment and +1 is a doubling of punishment from what they would have given without the conditions prompting the mitigation. When all subjects are taken into account – with a subject not giving a mitigation calculated at 0 – the average mitigations varied from a high of −1.08 in a case involving governmental corruption to a low of −.316 in a case involving an intentional killing and abduction, with the average mitigation for all subjects across all five scenarios as −0.67. Id.

[29] The percentage of subjects giving a mitigation was 50 percent in a case involving theft with trespass and dropped to a high of 45 percent in a case involving personal injury to a low of 20 percent in a case involving an intentional killing and abduction. Id.

[30] The percentage of subjects giving a mitigation was 50 percent for a case involving personal injury and 49 percent for a case involving theft with trespass. Id.

[31] In the two scenarios noted in note 30, the average mitigation for those who gave one was −1.35 and −1.21, respectively, with an average mitigation across all five scenarios of −1.03. When all subjects are taken into account – with a subject not giving a mitigation calculated as 0 – the average for the two scenarios mentioned was −0.66 and −0.63, respectively, with the average mitigation for all subjects across all five scenarios −0.51. Id.

[32] The percentage of subjects giving a mitigation was 60 percent for a case involving theft with trespass and 51 percent for a case involving personal injury. Id.

mitigations more than halved the punishment they otherwise would have given.[33]

Some mitigations may not have majority support among a broad national sample but may have support among more local groups. For example, a majority of the Chicago subject pool supported a mitigation in which the offender's punishment caused significant hardship for the offender's family even in the case of governmental corruption, whereas the full national sample would not.[34]

None of the mitigations can be said to be simply an exercise in properly assessing an offender's moral blameworthiness for the offense. In the first example, the offense is complete (thus, we know the offender's culpability and capacities at the time of the offense), and the harm or evil of the offense cannot be undone. In the second example, the operative factor has nothing to do with the offense or the offender but concerns the effect of the offender's punishment. And even then, it concerns the effect of his punishment on others, not on the offender him- or herself. In the last example, the fact that the offender has previously paid civil compensation may affect what restitution a criminal court might otherwise have required but arguably does not address the matter of punishment for wrongdoing. Even if the harm of an offense were undone – for example, the property stolen is recovered – such does not undercut the blameworthiness of the offender for the offense. In this instance, it is not even that the offender has voluntarily compensated the victim but rather has

[33] In the two scenarios noted in note 30, the average mitigation for those who gave one was −1.36 and −1.12, respectively, with an average mitigation across all five scenarios of −1.03. When all subjects are taken into account – with a subject not giving a mitigation calculated as 0 – the average for the two scenarios mentioned was −0.82 and −0.56, respectively, with the average mitigation for all subjects across all five scenarios −0.46. Id.

[34] Fifty-one percent of the Chicago subject pool supported the mitigation in this case, whereas only 38 percent of the full sample did, that is, the combined subject pools of the Chicago sample and the national Mechanical Turk sample. Id.

been compelled by civil lawsuit to pay.[35] Thus, it seems that all three factors are examples of laypersons' intuitions about proper punishment that call for a mitigation in punishment, even though a factor does not suggest reduced blameworthiness. Such giving of an undeserved mitigation in punishment can fairly be considered an exercise of mercy. And the goal of enhancing the criminal law's moral credibility may thus give a practical justification for an institutionalization of mercy in the criminal justice system.

To the extent that non–desert factors affect the community's judgment about the moral propriety of an offender's punishment, a goal of enhancing the criminal law's moral credibility with the community argues that those factors should be taken into account. To fail to do so would be to undermine the criminal law's moral credibility, and thereby to incrementally undermine its crime-control effectiveness through normative forces.[36]

VI. The Problem of Fairness and Consistency in Application

Assume for the sake of argument that the empirical data suggest that a criminal justice system designed to enhance the system's moral credibility should properly include a mercy mitigation, at least for those factors for which there is empirical evidence of community support. The crime-control benefit earned by the enhanced moral credibility provides a positive benefit that justifies the exercise of mercy in the criminal justice system.

[35] The best argument that one can make with regard to the final example is that the earlier civil compensation somehow reduces the harm of the offense and, therefore, reduces the blameworthiness of the offender in the same way that the occurrence of other results beyond the offender's control may nonetheless affect the assessment of blameworthiness.

[36] It would be somewhat awkward to describe this as following a distributive principle of empirical desert, given that even the laypersons supporting it would not necessarily describe the basis for the mitigation as one that reduces the offender's blameworthiness and thereby his desert. However, the motivation for recognizing such a mercy mitigation certainly shares the underlying justification that underlies a distributive principle of empirical desert, namely, to enhance the criminal law's moral credibility.

It seems likely, however, that a new sort of objection will emerge – a common objection made against the ad hoc exercise of mercy by sentencing judges under current practice. Even if one were to believe in the benefits of exercising mercy, our society has always put a high value on equality of treatment in exercise of governmental authority. That is the point, of course, of the constitution's equal protection clause. In the context of sentencing, this is the primary reason for the most substantial sentencing legislation of the past half century: the Sentencing Reform Act of 1984, which created the U.S. Sentencing Commission and mandated its promulgation of sentencing guidelines that would apply to all federal sentencing judges.[37]

The reasoning behind the concern for equality in sentencing is clear and compelling: an offender's punishment ought to depend on what she did and her culpability in doing it, not on the good or bad luck she has in the sentencing judge she is assigned. The ideal of modern criminal sentencing is to have all differences among sentences based on rational offense and offender factors, not on the personal preferences and philosophies of individual sentencing judges.

The exercise of mercy may particularly invite concern about inappropriate disparity because of its reputation as being something quite apart from articulated rules. Indeed, by its nature, mercy may be something that is necessarily unexpected and unpredictable. Many would suggest that an offender is never entitled to mercy. Mercy is perhaps a product of a special dynamic between the person giving and the person receiving the mercy.[38]

[37] U.S. Sentencing Commission, *Guidelines Manual* §1A1.3 (Nov. 2009) (reporting that "Congress sought reasonable uniformity in sentencing by narrowing the wide disparity in sentences imposed for similar criminal offenses committed by similar offenders" and that allowing for judicial discretion would risk "a return to the wide disparity that Congress established the Commission to reduce and would have been contrary to the Commission's mandate set forth in the Sentencing Reform Act of 1984," at 2–3).

[38] See *infra* section 2.

Mercy's reputation for being unpredictable and subject to complex and unarticulated criteria seems only confirmed by the recent empirical data. The evidence suggests a good deal of disagreement among the community regarding most factors that might prompt the exercise of mercy. Even the factors for which there is the most support – such as remorse or hardship for the offender's family, discussed here – draw a minority of people who would oppose their recognition. At the same time, there are a host of factors for which some significant minority would support their recognition while a majority would disapprove.[39]

Given this natural tendency toward disagreement among the community, and presumably among judges, our strong historical commitment to avoiding unjustified disparity in sentencing would seem to suggest that it is that much more important to provide some guidance in the exercise of a mercy mitigation. The guidance need not be tight and binding, but the system at least ought to identify those factors that should be taken into account and those that should not, and it ought to give some general sense of the conditions under which a factor should be taken into account and the general extent of mitigation that it normally should provide. To authorize a mercy mitigation without providing any guidance is to invite each individual sentencing judge to formulate his or her own set of policies on the issues – on which we know there is wide disagreement – and to subject offenders to punishment based on the luck of the draw.

Fortunately, there is no reason to think that there would be serious difficulty in providing such general guidance, as to both the eligible factors and their general application. As a point of comparison, note that many factors included in current sentencing guidelines are complex and subjective in nature. Those qualities certainly limit the precision of the guidance that can be provided, but they do not preclude attempts at identifying eligible and ineligible factors and at articulating general application

[39] Robinson et al., "Extralegal," at pt. II.

criteria. Consider, for example, two provisions of the current federal sentencing guidelines:

> §5K2.12. Coercion and Duress
>
> If the defendant committed the offense because of serious coercion, blackmail or duress, under circumstances not amounting to a complete defense, the court may depart downward. The extent of the decrease ordinarily should depend on the reasonableness of the defendant's actions, on the proportionality of the defendant's actions to the seriousness of coercion, blackmail, or duress involved, and on the extent to which the conduct would have been less harmful under the circumstances as the defendant believed them to be. Ordinarily coercion will be sufficiently serious to warrant departure only when it involves a threat of physical injury, substantial damage to property or similar injury resulting from the unlawful action of a third party or from a natural emergency. Notwithstanding this policy statement, personal financial difficulties and economic pressures upon a trade or business do not warrant a downward departure.
>
> §5K2.13. Diminished Capacity
>
> A downward departure may be warranted if (1) the defendant committed the offense while suffering from a significantly reduced mental capacity; and (2) the significantly reduced mental capacity contributed substantially to the commission of the offense. Similarly, if a departure is warranted under this policy statement, the extent of the departure should reflect the extent to which the reduced mental capacity contributed to the commission of the offense.
>
> However, the court may not depart below the applicable guideline range if (1) the significantly reduced mental capacity was caused by the voluntary use of drugs or other intoxicants; (2) the facts and circumstances of the defendant's offense indicate a need to protect the public because the offense involved actual violence or a serious threat of violence; (3) the defendant's criminal history indicates a need to incarcerate the defendant to protect the public; or (4) the defendant has been convicted of [certain sexual offenses].[40]

The matters at issue – coercion and diminished capacity – are quite subjective and complex, and the guidance provided is necessarily general in nature. However, the existence of the provisions is likely to increase uniformity in application by different judges because the provisions do give answers to certain central questions. They explicitly acknowledge the propriety of mitigations in such cases, which some judges might otherwise have thought their use to be inappropriate. At the same time, they exclude the use of such factors in certain cases, where some judges might otherwise have thought their use was appropriate. Further, the provisions articulate some criteria for applying the factors, which provides more guidance, and hence more uniformity, than if such guidance had not been provided.

One might conclude, then, that a criminal justice system, with some careful effort, might be able to construct guidance mechanisms that would allow it to fairly institutionalize the exercise of mercy in those situations in which the community supports its exercise.

VII. Does Institutionalizing Mercy Undermine Its Proper Nature?

Does the need to articulate the criteria for a mercy mitigation – the codification of the exercise of mercy – itself create a problem for mercy? Perhaps. If one supports the criminal justice system's exercise of mercy for reasons parallel to those justifying its use in personal relationships, as discussed in section 2, then its institutionalization may be problematic. The giving of such a mercy mitigation may be highly dependent on the fact that the giver has no obligation to give it.[41] Thus, the institutionalization of mercy

[40] U.S. Sentencing Commission, *Guidelines Manual*, at §5K2.12 & §5K2.13.

[41] See, e.g., Murphy, "Mercy and Legal Justice," at 3 ("[Mercy] is never owed to anyone as a right or a matter of desert or justice. It always, therefore,

may tend to destroy the very characteristic of mercy that gives it value.

The larger point here is that the nonentitled character of mercy may be simply fundamentally incompatible with the use of mercy in a system of criminal justice like ours that demands equality of treatment among similarly situated offenders. Every offender is entitled to a sentence that is no more and no less than that which other similarly situated offenders receive. And the criteria that are used to determine whether offenders are "similarly situated" must have some rational connection to the criminal justice system's purposes and goals, be that doing justice or controlling crime or something else. It would be hard to justify different amounts of punishment for two offenders who were similar with regard to the system's purposes and goals and were different only with respect to some inarticulable or unpredictable dynamic tied to a nonentitled notion of mercy.

However, the same problem does not exist for the institutionalization of mercy when done as a means of enhancing the moral credibility of the criminal law, and thereby harnessing the normative forces of social influence and internalized norms for more effective crime control. The codification of mercy would increase the criminal law's ability to more accurately track the community's notions of appropriate punishment. The more specific the articulation of the criteria for mercy, the more reliable and predictable its application and the better the system's reputation in the community for getting it right. Thus, the codification of mercy – at least those principles of mercy supported by the community – only improves the system's effectiveness in gaining greater moral credibility and, thereby, more effective crime control.

This distinction also seems to apply to the larger point: when mercy is urged for its purely deontological value, it conflicts with the deontological value of doing justice, as discussed in section 4.

transcends the realm of strict moral obligation and is best viewed as a free gift – an act of grace, love or compassion that is beyond the claims of right, duty, and obligation").

To prefer mercy in this conflict would require a philosophical account of why the value of giving mercy is greater than the value of doing justice. It is not clear what that account would be. Interestingly, however, when one shifts to a utilitarian crime-control justification for mercy, there is a clear argument for why giving mercy should trump doing justice: because it better tracks the community's intuitions of proper punishment and therefore better enhances the criminal justice system's moral credibility, and thereby its crime-control effectiveness.