

A TRUCE IN THE DISTRIBUTIVE PRINCIPLE WARS?

EMPIRICAL DESERT, MORAL CREDIBILITY, AND THE INTERNALIZATION OF SOCIETAL NORMS

Paul H. Robinson¹

For a century before any of us were born, criminal law theory has been torn between two apparently irreconcilable distributive principles. Utilitarian principles like deterrence and incapacitation of the dangerous saw preventing crime as their primary goal. So-called “retributivist” principles, in contrast, set doing justice as their goal. In this view, doing justice – imposing liability and punishment according to each offender’s blameworthiness, his desert – was a value in itself that required no further justification.

There seemed little possibility of reconciliation between the two competing principles. In the utilitarians’ view, the imposition of punishment can be justified only by the practical benefit that it provides: avoiding future crime. In the retributivists’ view, doing justice for past wrongs is a value in itself that requires no further justification. The competing approaches simply give value to different things. They use different currencies: fighting future crime versus doing justice for past wrongs.

What I want to argue here is that the two are in fact reconcilable, in a fashion. I’m not suggesting that we declare a winner in the distributive principle wars but something more like a truce. Specifically, good utilitarians ought to support a distributive principle based upon desert, because the empirical evidence suggests that doing justice for past wrongdoing is probably the most effective and efficient means of controlling future crime.²

But “doing justice” here is not justice in the deontological desert sense of what moral philosophers think is deserved but rather justice in the “empirical desert” sense, as it is called, that reflects the community’s shared judgments of justice. The empirical research shows that people’s assessment of an offender’s overall blameworthiness takes into account a wide variety of factors, including the seriousness of the offense, the culpable state of mind of the offender (intentional, reckless, or negligent), and the offender’s mental, emotional, and physical capacities at the time the offense.³ Specifically, it is argued that by tracking the community’s principles of justice the system can build “moral credibility” with the community, which allows it to harness the powerful forces of social influence, community support, and internalized norms.

Recent empirical research confirms that a criminal justice system perceived by the community as conflicting with their principles of justice is one that provokes resistance and subversion, while a criminal justice system that earns a reputation for reliably doing justice – by tracking shared community judgments of justice – is one whose moral credibility inspires greater deference, assistance, and acquiescence, and is more likely to have citizens internalize its norms of what constitutes truly condemnable conduct.

I need to fill out this sketch of my main argument and to give some examples of empirical support. The most efficient way of doing that may be to address five questions that I think

¹ Colin S. Diver Professor of Law, University of Pennsylvania.

² See generally Paul H. Robinson, *Intuitions of Justice and the Utility of Desert* (Oxford University Press, 2013), Part II, 96–238.

³ See generally *id.*, Part III, 239–413.

people can reasonably ask about using empirical desert – the community’s shared principles of justice – as our primary distributive principle for criminal liability and punishment:

1. Is there any such thing as the community’s views of justice?
2. Aren’t the community’s views of justice brutish and draconian?
3. Why should a crime-control utilitarian care what the community thinks is just?

Why would an empirical desert distributive principle reduce crime?

4. Even if empirical desert can have some crime reduction effect, wouldn’t distributive principles of general deterrence or incapacitation of the dangerous have even more?

5. Should the criminal law ever deviate from the community’s shared principles of justice? If not, aren’t we stuck with the status quo, which social reformers might have good reasons to want to change?

1. IS THERE ANY SUCH THING AS THE COMMUNITY’S VIEWS OF JUSTICE?

Perhaps justice is such a complex judgment that everybody has their own personal view? Perhaps there is no community view, and thus no ability to construct a criminal law that reflects the community’s views of justice? The empirical evidence, however, suggests otherwise.⁴

Especially for issues that one might call the “core of wrongdoing” – physical injury to others, taking property without consent, and deceit in exchanges – there is in fact high agreement across all demographics. Consider one study that had subjects rank order 24 scenarios according to overall blameworthiness, deserved punishment. The results showed a Kendall’s W of 0.95 for in-person subjects and 0.88 for Internet subjects — an astounding result.

One can’t normally get this level of agreement except in observational studies, as with asking subjects to judge the relative brightness of dot clusters. Where subjects are asked for something beyond the purely observational, the analytic task requested of them must be almost intuitional. For example, one can get a similarly high Kendall’s W by asking subjects to rank order the standard pain images used for medical diagnosis with nonverbal patients. See Figure 1.

FIGURE 1 – FACES GRAPHIC – TOP OF PAGE 1

The empirical studies also suggest that even uneducated people have very sophisticated and nuanced judgments about justice. Small changes in facts produce predictable changes in blameworthiness judgments.⁵

People’s judgments don’t tie a particular punishment to a particular level of blameworthiness. However, because ordinary people can distinguish so many cases along the blameworthiness continuum and because the punishment continuum contains a finite number of points (meaningful differences require larger units as the length of imprisonment gets longer), people’s judgments about the relative blameworthiness of a particular case against all other cases end up at a particular point on the punishment continuum, not because there is some magical connection between that amount of punishment and that blameworthiness but rather because it is that single point that puts the case in its proper ordinal rank with all other cases. If the endpoint of the punishment continuum changes, so too will the punishment location of each case on continuum.

⁴ See *id.*, 18–34.

⁵ See *id.*, Part III for a wide range of such studies.

I don't think the endpoint of the punishment continuum is something on which people's judgments are fixed. We see significant endpoint differences among different societies, which confirms how malleable the endpoint judgment is. Judgments of relative blameworthiness, in contrast, especially concerning the core wrongdoing, are not so malleable. This is confirmed by the fact that we find the same rank ordering of crime scenarios across demographics and cultures.⁶

This high level of agreement on relative blameworthiness within the core of wrongdoing is not a surprise when one considers that people's judgments of justice are in some significant part a feature of human evolutionary development.⁷ And this is consistent with evidence suggesting that many justice judgments are in large part intuitional, rather than the product of conscious reasoning.⁸ Danny Kahneman has famously distinguished intuitional judgments as having attributes somewhere between pure perception and reasoned judgment, as seen in Figure 2.

FIGURE 2 – DANNY KAHNEMAN GRAPHIC – BOTTOM OF PAGE 1

As one moves out from the core of wrongdoing, disagreements among people do appear. Downloading music from the Internet without a license can be seen as analogous to traditional theft but is not itself a physical taking without consent. Thus, while there may be strong agreement on issues relating to the core of physical taking, there will be disagreement on the downloading issue depending upon the extent to which a person has accepted the analogy between unlicensed downloading and physical taking.

Whenever the intuitional justice judgment is supplemented by some reasoned gloss, disagreements will appear. But the point is that, contrary to the common wisdom of a decade ago, justice judgments are not something about which everybody disagrees about everything. There is a strong core of agreement. And to the extent that there is disagreement, we have the methodology to reliably determine the center of the bell curve on any criminal liability or punishment issue.

That is, it is indeed possible to construct a criminal law that best approximates the community's view.

2. AREN'T THE COMMUNITY'S VIEWS OF JUSTICE BRUTISH AND DRACONIAN?

Just because we can reliably determine the community view, it doesn't follow that we would want to follow that view in setting criminal liability and punishment rules. Certainly many American academics are horrified at such a prospect because they see a series of policies in current criminal law that they find to be highly objectionable – policies that do injustice, not justice, as they see it.

If you look at Figure 3, the fourth column lists a variety of common American criminal law doctrines that progressive academics frequently criticize as unjust. Each row of the table represents one of 12 real-world cases that illustrate the operation of one of these crime-control doctrines. (I refer to these as "crime-control doctrines" because they are justified by and have been adopted upon crime-control grounds rather than upon a claim that they produce carefully modulated deserved punishment.)

⁶ See *id.*, 18-34.

⁷ See *id.*, 35-62.

⁸ See *id.*, 5-17.

FIGURE 3 – CRIME-CONTROL CASES – PAGE 2

But what the research reveals is that these common liability and punishment rules clearly do not reflect community views. Just the opposite; they dramatically conflict with them.⁹ They may well be consistent with crime-control strategies of general deterrence or incapacitation of the dangerous but they have the effect of disconnecting criminal punishment from community notions of justice. See Figure 4, which shows the community judgments on the relative blameworthiness and appropriate punishment for each of the cases contained in the previous table.

FIGURE 4 – RANKING OF CRIME CONTROL AND MILESTONE CASES – PAGE 3

In the first column of this table, the twelve crime-control cases with the draconian sentences are shown indented and in italics. The twelve cases in bold are what might be called “milestone” cases – crime scenarios ranging from the most minor to the most serious that taken together provide a continuum of blameworthiness against which the test crime-control cases can be compared. In the study, subjects were asked to rank order all twenty-four of the cases, the twelve milestone cases and the twelve crime-control cases. The result is the order of cases that you see in the table.

The important point here is that the crime-control cases that have the draconian penalties in law, are in fact perceived by the subjects as being dramatically less serious and blameworthy than the law treats them. For example, Case F (*Rummel*: air-conditioner fraud), involving a minor fraud by an offender who had been convicted for a series of previous such minor frauds, was seen by the subjects as somewhat more serious than stealing a microwave from a house and somewhat less serious than a minor assault at a record store, offenses for which the subjects gave a sentence of 2.3 years and 3.9 years respectively, as compared to the life sentence that Rummel actually got.

Note that the crime-control cases here are not cases in which some renegade prosecutor or rogue judge but rather are cases where the crime-control doctrine is being lawfully applied as designed. The *Rummel* case went to the U.S. Supreme Court where the conviction and sentence were affirmed.

FIGURE 5 – CONFLICT BETWEEN CRIME-CONTROL DOCTRINES and COMMUNITY VIEWS – PAGE 4

Figure 5 visually displays the dramatic nature of the law-community conflict revealed on the previous table. Take, for example, Case F (the *Rummel* case) on the right hand margin. The solid line to the center indicates where on the punishment continuum the subjects place this case, close to the three-years mark. The dashed sloping line indicates the punishment that was actually imposed, life imprisonment.

The important point here is to see on the right-hand side the dramatic difference between the solid lines and the corresponding dashed lines for the same case. The enormity of the law-community conflict is emphasized by the fact that the punishment scale in this graphic is exponential, not linear. Each one of the large dots, 1 through 8, represents typically a doubling of punishment – the standard structure of criminal code offense grade categories in the United States. Thus, if the difference between the solid line and the dashed line for any case were only the difference between 4 and 5 on the punishment scale, that small difference on the scale means that the person is getting twice the punishment that the subjects thought was deserved. In fact, community-law differences are all dramatically more than that.

⁹ See *id.*, 110–140.

How could such a conflict occur in a democracy? It is not the draconian justice judgments of ordinary people that are producing these doctrines but rather politicians' reliance on coercive crime-control theories like general deterrence and incapacitation of the dangerous – crime-control theories developed and pressed in the past by academics.¹⁰ Having criminal liability and punishment rules track community views could be an effective way of short-circuiting those injustice-producing doctrines.

3. WHY SHOULD A CRIME-CONTROL UTILITARIAN CARE WHAT THE COMMUNITY THINKS IS JUST? WHY WOULD AN EMPIRICAL DESERT DISTRIBUTIVE PRINCIPLE REDUCE CRIME?

The past 50 years in the United States have seen reformers sufficiently concerned about crime-control that they have been happy to sacrifice justice in order to attempt to avoid future crime. This has brought greater reliance upon the utilitarian crime-control distributive principles, primarily those of general deterrence and incapacitation of the dangerous, despite their conflict with deserved punishment.

But the empirical research suggests that the use of criminal liability and punishment rules that conflict with the community's principles of justice may be self-defeating.

Setting aside the accumulating evidence that general deterrence may work in principle but not in practice,¹¹ recent research suggests that crime-control effectiveness depends in some significant part upon the criminal law's moral credibility with the community. A criminal justice system with a good reputation for reliably doing justice and avoiding injustice is one that will inspire cooperation, support, deference, and the internalization of its norms. In contrast, a criminal justice system that earns a reputation for deviating from the community's principles of justice – deviating, that is, from "empirical desert" – is a system that will inspire resistance and subversion and will lose the ability to harness the powerful forces of social influence and internalized norms.¹²

It is easy to see this principle at work on a large scale anecdotally. The discredited Soviet criminal justice system gained compliance only through an overwhelming and ever-present police state. The worse its reputation for reliably doing justice, the less the deference it earned from its citizens.

But the more recent research suggests that the relationship between moral credibility and community deference and compliance is much more widespread and nuanced. Even small incremental losses in moral credibility can produce corresponding incremental losses in deference and compliance.

In Figure 6 are the results of a study, using a within-subjects design, in which subjects were asked a number of questions relating to various ways in which moral credibility is thought to affect deference, compliance, and the internalization of the law's norms. Will citizens assist police by reporting crimes? Will they assist in the investigation and prosecution of crimes? Do people take the imposition of criminal liability and punishment as a reliable sign that the defendant has done something truly condemnable? Do people take the extent of the liability

¹⁰ See Paul H. Robinson, *Distributive Principles of Criminal Law: Who Should Be Punished How Much?* (Oxford University Press, 2008) chs. 3-6.

¹¹ *Id.* at 21-98. For a discussion on the difficulties with using incapacitation of the dangerous as a distributive principle see Robinson, *Distributive Principles*, 109-134.

¹² See Robinson, *Intuitions of Justice*, 141-207.

imposed as a reliable indication of the seriousness of the offense and the blameworthiness of the offender?

FIGURE 6 – REDUCTION IN DEFERENCE FROM DISILLUSIONMENT (within-subjects design) – TOP AND BOTTOM OF PAGE 5

With a baseline established, the subjects were then disillusioned by exposing them to accounts of the system’s failures of justice and doing of injustice. Finally, later retesting showed that the measures of deference, compliance, and internalization of norms had all decreased.

A follow-up study used a between-subjects design, giving different levels of disillusionment to three different groups and then testing their levels of deference, compliance, and internalization. Figure 7 reports the results, which confirm the conclusions of the earlier within-subjects design. The greater the disillusionment, the greater the loss in deference, compliance, and internalization.

FIGURE 7 – REDUCTION IN DEFERENCE (between-subjects design) – TOP OF PAGE 6 (Figure 8 gives the results of a study using large datasets, which came to a similar conclusion using regression analysis.)

FIGURE 8 – REGRESSION ANALYSIS – BOTTOM OF PAGE 6

The results in the two experimental studies are particularly striking because subjects came to the study with pre-existing views on the system’s reputation for being just. The experimenters, within the context of the study, could only nudge those pre-existing views slightly. Yet, even that incremental disillusionment produced corresponding reductions in deference and compliance.

This is important because it means that no matter what the current state of a criminal justice system’s moral credibility with the community, any incremental reduction in credibility can produce an incremental reduction in deference – and any increase can produce an increase in deference.

4. EVEN IF EMPIRICAL DESERT CAN HAVE SOME CRIME-REDUCTION EFFECT, WOULDN’T DISTRIBUTIVE PRINCIPLES OF GENERAL DETERRENCE OR INCAPACITATION OF THE DANGEROUS HAVE EVEN MORE?

Part of the attraction of empirical desert as a distributive principle comes from the weaknesses of general deterrence, incapacitation the dangerous, and other alternative principles. I have written a good deal on the subject,¹³ but let me quickly sketch the nature of my criticisms.

General deterrence can be an effective crime-control mechanism in principle, but rarely in practice. Having a criminal justice system that imposes punishment on wrongdoers certainly has a general deterrent effect. Less clear, however, is the effectiveness of general deterrence as the distributive principle for criminal liability and punishment – that is, *setting liability and punishment rules* so as to maximize efficient general deterrent effect.

For a rule formulation to enhance general deterrence, it must meet at least three prerequisites. First, the intended audience must know of the rule. Second, the intended audience must be rational calculators who can and will behave in a way that promotes their

¹³ See Robinson, *Distributive Principles*, 21–98, 141–207.

self-interest in light of the rule. And, third, their cost-benefit analysis under the rule must suggest that the costs of the contemplated violation outweigh its benefits.

Unfortunately, rarely do these prerequisites exist in the real world. First, the empirical research suggests that the target audience rarely knows the law. Even when they think they know, they typically have it wrong.¹⁴ Academics and politicians spend a good deal of time agonizing over the adoption and formulation of utilitarian crime-control doctrines, such as a felony-murder rule, the three-strikes rule, the use of strict liability, and the other crime-control doctrines listed in the fourth column of Figure 3. But if one asks people on the street, or even offenders in particular, whether their jurisdiction adopts such a rule and, if so, which formulation of the rule it has adopted, people will not know or, if they think they know, will have the answer wrong. Instead, the research suggests that people generally assume the criminal law rules are as they think they should be: formulated to give deserved punishment based upon an offender's overall blameworthiness.

Second, even if the target audience did know the legal rules, available research suggests that the target audience is more often than not anything but rational calculators. Instead, their decisions are heavily influenced by mental or emotional disturbance; drug use or addiction; group influence, especially by gangs; impulsiveness; and an indifference or inattentiveness to consequences.

Finally, even if the target audience did know the legal rules and were rational calculators, a general deterrent effect is possible only if the rational calculations suggest that the costs of the wrongdoing outweigh the benefits. Yet, the capture and punishment rate for most offenses is so low – commonly less than 100 to 1 for offenses other than homicide – that the target audience frequently sees the benefits as outweighing the costs. More importantly, the result of the calculation depends not on the reality of the situation but rather on the potential offender's perception of it. Thus, when the empirical evidence suggests that many if not most potential offenders generally overestimate their ability to avoid detection and punishment, the general deterrence project can have limited effect even if it dramatically improves its punishment rates.

What makes a general deterrence distributive principle even less attractive is the fact that there is a general deterrent effect inherent in a desert distribution of punishment. The only way that a general deterrence distribution of punishment can provide more deterrent effect than that inherent in a desert distribution is by deviating from desert – in other words, by doing injustice.

Yet, it is just these instances of deviations from desert in which a general deterrence distribution has its worst performance. As noted previously, people assume the criminal law follows a desert distribution. Thus, it takes a special educational campaign to make the target audience aware of a rule based upon a general deterrence distribution that deviates from desert. The evidence suggests that such special education is extremely difficult, especially for the target audience of potential offenders.

Incapacitation of the dangerous is as problematic as a distributive principle of general deterrence, but for different reasons. Unlike general deterrence, which has real difficulty producing greater deterrence than that already inherent in a desert distribution, incapacitation does in fact work. Putting people in prison does prevent further victimization of the community.

But the problem with the incapacitation distributive principle is that behavioral scientists are at present relatively poor in reliably predicting future criminality in a specific individual. False

¹⁴ This is a particular problem in the United States where there are 51 American criminal codes.

positive rates are very high, which creates enormous costs and many intrusions on personal liberties with no crime-control benefit. The incapacitation distributive principle is particularly disadvantaged in the United States where constitutional limitations imposed by courts limit the open use of such preventive detention and require instead that it be cloaked in criminal justice terms. Thus, for example, instead of being able to openly evaluate an offender's predicted future dangerousness in setting a criminal sentence, liability and sentencing rules commonly use substitutes like prior criminal record, which have turned out to be even worse approximations of future dangerousness.

Finally, even if there were a situation where such coercive crime control principles as general deterrence or incapacitation could provide a crime-control benefit by deviating from desert, any such advantage is likely to be wiped out by the loss of crime-control effectiveness that comes when such deviations from desert undercut the criminal justice system's moral credibility with the community.

Good reputations, as social psychologists make clear, are hard to build and easy to destroy. A continuous stream of cases that deviate from deserved punishment in order to promote general deterrence or incapacitation is just the sort of thing that can seriously undermine the criminal law's moral credibility and thereby undermine its ability to harness the powerful forces of social influence and internalized norms.

5. SHOULD THE CRIMINAL LAW EVER DEVIATE FROM THE COMMUNITY'S SHARED PRINCIPLES OF JUSTICE? IF NOT, AREN'T WE STUCK WITH THE STATUS QUO, WHICH SOCIAL REFORMERS MIGHT HAVE GOOD REASONS TO WANT TO CHANGE?

One reason to worry about having criminal law rely upon community views is that such a system may tend to impede social change – or, more accurately, impede using law to bring about social change. Relying upon community views presumably means relying upon people's *current* community views.

We know from history that existing community views are not always the best for society. Changing those views can sometimes bring a better world. And criminal law can sometimes be useful in helping to change views. Consider, for example, the recent decriminalization of same-sex intercourse and increased criminalization of domestic violence and date rape.

However, the problem is that if criminal law gets out in front of community views, the disparity between the two can potentially lead to undermining criminal law's moral credibility. This could be particularly tragic, not just for effective crime control but also for social reform, for the greater the moral credibility of the criminal law, the greater the law's power to help shift community views.

Does reliance upon an empirical desert distributive principle condemn society to live with existing views forever? Not necessarily. As the criminal law improves its moral credibility with the community – as it “earns moral credibility chips” with the community – it can selectively “spend” those chips by having criminal law lead rather than follow on selected issues of special importance to social reformers.¹⁵

In other words, reliance upon an empirical desert distributive principle might be an enormous help to social reformers because it creates a powerful mechanism for changing social norms—a mechanism that did not previously exist when moral credibility was low.

¹⁵ See Robinson, *Intuitions of Justice*, 70–95, 189–207.

However, social reformers will want to follow some particular strategies. Spending the criminal law's earned moral credibility chips to help change societal norms has to be done carefully. If community views do in fact shift as reformers want, then the conflict with the law disappears and there is no long-term credibility damage. On the other hand, as we saw in the American Prohibition movement of the 1920s, if the law gets too far out in front of community views and does not successfully shift community views, then the law will lose moral credibility and that will translate into reduced crime-control effectiveness. In fact, crime rates during Prohibition went up, and not just for alcohol-related offenses but rather for a wide range of offenses unrelated to alcohol, which is exactly what one would expect when the criminal law's credibility has been undermined by showing itself to be continually imposing punishment that conflicts with the community's justice judgments. People become habituated to lawbreaking. Perhaps worse, pushing too far ahead without successfully shifting views can undermine the law's reputation in such a way as to reduce law's usefulness to social reformers in the future.

CONCLUSION

Neither the crime-control utilitarians nor the desert retributivists can claim "victory" in the distributive principle wars. For the crime-control utilitarians, it makes good sense to adopt empirical desert as the dominant distributive principle even though it would substitute for the classic utilitarian crime-control theories, such as general deterrence and incapacitation the dangerous. The good utilitarian will follow the numbers, even though it may seem to take them dangerously close to a desert distribution.

On the other hand, neither can the retributivists claim victory because the desert distribution being relied upon is not that of the deontological desert produced by moral philosophers. Instead, it is the empirical desert distribution produced by social psychologists' study of ordinary people and is justified by its crime-control effectiveness rather than by a transcendent notion of just desert as a value in itself. The retributivists can take some comfort, however, in the fact that an empirical desert distribution is probably the best practical approximation of deontological desert that could be produced to guide the formulation of criminal law. Moral philosophers disagree among themselves on many (if not most) key issues and for criminal code drafters there is no reliable mechanism for determining which philosophical camp ought to be followed. In other words, there is no practical possibility of adopting deontological desert as a distributive principle, which is what the retributivists seek. They should be well satisfied, however, with empirical desert as its best practical approximation.¹⁶

Thus, while there is no victor in the distributive principle wars, empirical desert as a distributive principle represents the basis for a truce that both parties ought to feel comfortable with as the best practical means of achieving their goals.

¹⁶ See Robinson, *Distributive Principles*, 175–212, 247–260.