The Role of Moral Philosophers in the Competition Between Deontological and Empirical Desert

Paul H. Robinson
THE ROLE OF MORAL PHILOSOPHERS IN THE COMPETITION BETWEEN DEONTOLOGICAL AND EMPIRICAL DESERT

PAUL H. ROBINSON*

INTRODUCTION

Desert has become increasingly attractive as a principle by which to distribute criminal liability and punishment. A number of modern sentencing guidelines have adopted it as their distributive principle.¹ Most recently, a committee of the American Law Institute proposed revising the Model Penal Code “purposes” section to adopt desert as the dominant distributive principle.²

---

¹ Colin S. Diver Distinguished Professor of Law, University of Pennsylvania. The author wishes to acknowledge the valuable research assistance of Desiree Liverseidge and the useful comments of Leo Katz, Stephen Morse, and participants of the Conference on Law and Morality, Institute of Bill of Rights Law, William & Mary School of Law, especially Claire Finkelstein and Michael Moore.

² According to the Model Penal Code,

The general purposes of the provisions governing sentencing and corrections, to be discharged by the many official actors within the sentencing and corrections system, are:

(a) in decisions affecting the sentencing and correction of individual...
But these reforms, and the current debates, are unclear as to whether the conception of desert under consideration is a deontological or an empirical one. The two can be quite different. A deontological conception of desert, based on reasoning from principles of right and good and aimed toward giving us a transcendent notion of justice, would distribute criminal liability and punishment differently than would an empirical conception of desert, based upon empirical research into the shared intuitions of justice of the community that is to be governed by the code or practice being formulated. For example, moral philosophers disagree about the significance of resulting harm, and each side of the debate has plausible arguments to make. In contrast, all offenders:

(i) to render punishment within a range of severity proportionate to the gravity of offenses, the harms done to crime victims, and the blameworthiness of offenders;

(ii) when possible with realistic prospect of success, to serve goals of offender rehabilitation, general deterrence, incapacitation of dangerous offenders, and restoration of crime victims and communities, provided that these goals are pursued within the boundaries of sentence severity permitted in subsection (a)(i); and

(iii) to render sentences no more severe than necessary to achieve the applicable purposes from subsections (a)(i) and (ii)....

MODEL PENAL CODE § 1.02(2) (Preliminary Draft No. 3, 2004) [hereinafter MODEL PENAL CODE].

3. Although deontological desert traditionally has carried the desert banner in academic circles, today's law- and policymakers often give people's-sense-of-justice explanations for desert-based legislation. In other words, legislators make empirical claims, not philosophical arguments. See, e.g., MODEL PENAL CODE, supra note 2, § 1.02(2) cmt. c (suggesting that because desert can be difficult to quantify at times, sentencing commissions should solicit a diverse range of community perspectives, and that doing so gives the commissions a "unique credibility").

4. But, as the reader will see in the text following note 19, only some moral philosophers—moral realists—conceive of desert as having such transcendent nature.

5. Those that have argued that resulting harm should matter include Leo Katz, Why the Successful Assassin Is More Wicked than the Unsuccessful One, 88 CAL. L. REV. 791, 806 (2000) (arguing by hypothetical that principled moral analysis suggests harm should be considered when assessing blameworthiness); Ken Levy, The Solution to the Problem of Outcome Luck: Why Harm Is Just As Punishable As the Wrongful Action that Causes It, 24 LAW & PHIL. 263, 303 (2008); and Michael S. Moore, The Independent Moral Significance of Wrongdoing, 5 J. CONTEMP. LEGAL ISSUES 237, 267-71 (1994) (posing that because we resent successful wrongdoers more than we do those who unsuccessfully attempt harm, 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, which suggests that "results matter" in the moral arena). Those that have argued that resulting
available data suggest a nearly universal and deeply held view among the community that resulting harm does matter, that it increases an offender's deserved punishment.\(^6\) This practical difference is only one of a host of issues on which a moral philosopher's conclusion might vary from the empirical data on lay persons' shared intuitions of justice.\(^7\)

Should we prefer one conception of desert over the other for use as a distributive principle for criminal liability and punishment? Each of the two competing conceptions of desert offer distinct advantages and disadvantages.

---

\(^6\) See, e.g., PAUL H. ROBINSON & JOHN M. DARLEY, JUSTICE, LIABILITY, AND BLAME: COMMUNITY VIEWS AND THE CRIMINAL LAW 14-28, 181-97 (1995) (reporting empirical studies); John H. Mansfield, Hart and Honnoré, Causation in the Law—A Comment, 17 VAND. L. REV. 487, 494-95 (1964) (concluding that "[t]he notion that there should be a difference in punishment [between unsuccessful attempts and completed crimes] is deeply rooted in popular conscience, and to ignore it is to risk [jury] nullification").

\(^7\) See generally ROBINSON & DARLEY, supra note 6 (comparing community views on a wide variety of criminal law issues to existing legal rules and discussing the points of disagreement). Finding wide and persistent disagreement among moral philosophers on many if not most significant issues is common. See supra note 5. Indeed, it is likely that on any issue over which law- or policymakers themselves disagree—prompting them to look to other disciplines for guidance—philosophers almost certainly will disagree among themselves. Other disciplines may have disagreements, but because most have some objective test by which a writer ultimately may be proven right or wrong, over time some coalescence tends to emerge around an accepted view. A proposed theory ends up either explaining more of the available data, and is accepted; or does not, and is rejected. But without such a clear test mechanism, moral philosophy lacks a path to coalescence. Because philosophers will disagree on nearly any significant issue, an outsider often has difficulty gaining something useful, in part because, to make an informed judgment as to which view ought to be given deference, the outsider must herself know something about moral philosophy. In other words, the outsider must become a bit of an insider. Ultimately, the moral philosophy literature is not terribly accessible, and thus its informed use is commonly costly.
I. DEONTOLOGICAL DESERT

Deontological desert can offer a unique and critically important value to the criminal justice law- or policymaker: it can provide a foundation for desert that transcends any particular case, community, or culture. That is, it can give us a means by which we can tell the truth of what is deserved, insulated from the vicissitudes of human irrationality and emotions. This deontological conception of desert gives us the ability to determine when our shared intuitions of justice may be wrong. Even though a liability or punishment rule may be popular, it nonetheless may be unfair or unjust, and the deontological conception of desert lets us spot these justice errors in people's intuitions.

The standard complaint against relying upon such a deontological conception of desert in distributing criminal liability and punishment is that it leads to disutility. Those consequentialists who seek to minimize future crime, for example, will be quick to point out that deontological desert as a distributive principle will allow future crimes to occur that could have been avoided under a utilitarian distributive principle. Traditionally, that fact has meant a utilitarian preference for distributing liability to optimize deterrence, rehabilitation, incapacitation, or some combination of them.

II. EMPIRICAL DESERT

Reliance upon an empirical conception of desert in the distribution of criminal liability and punishment prompts its own set of complaints. One primary objection is that people's intuitions of justice are too vague and suffer too much disagreement to be effectively operationalized. But empirical studies show this common wisdom to be false. In fact, people's shared intuitions of justice are quite nuanced: small changes in facts produce large and predictable


9. See ROBINSON & DARLEY, supra note 6, at 5-7.

10. See, e.g., MODEL PENAL CODE § 1.02(1) (1962) (listing the general purposes of the provisions).
changes in the assessment of blameworthiness. Further, an astounding level of agreement exists across cultures and demographics on the relative degree of blameworthiness. Although people and cultures disagree about the general level of punishment severity a criminal justice system should adopt, once the endpoint of a society’s punishment continuum is set, there is significant agreement on the ordinal ranking of cases along that continuum, at least for the core wrongs of physical aggression, unconsented-to takings, and deception or dishonesty in exchanges. Thus, empirical desert does not produce an indeterminate range of punishment, as some have suggested, but rather a specific amount. It is not that a particular violation necessarily deserves a specific amount of punishment in some absolute sense; rather, each violation, once placed on a fixed continuum of punishment, deserves a particular amount of punishment because that amount is required to give that violation its proper ordinal ranking among the range of possible violations. One can easily imagine how the erroneous common wisdom about disagreement developed: disagreement over general punishment severity—the continuum endpoint—masked the agreement on the ordinal ranking of violations.

As noted, the broad consensus on ordinal ranking exists primarily for the core wrongs: injury to others, the taking of property, and deceit or dishonesty in dealings. As the harm or wrong moves away from this core, disagreements appear across cultures and demographics, depending primarily upon the perceived strength of the analogy between the new conduct and the core wrongs. Ultimately, operationalizing empirical desert is quite feasible, more so than deontological desert, because of the higher level of agreement on the former than the latter.

Consequentialists might offer a second kind of objection to empirical desert as a distributive principle, similar to the disutility complaint they make against deontological desert: such a desert distribution of criminal liability and punishment allows future

12. Id. pt. II.
14. See supra note 12 and accompanying text.
crime that could be avoided with a distribution that optimizes deterrence, rehabilitation, incapacitation, or other traditional utilitarian crime control mechanisms.\footnote{See Gruber, \textit{supra} note 8, at 454-68 (discussing consequentialist theories of punishment, including deterrence, rehabilitation, and incapacitation).} But, as I have suggested elsewhere, strong arguments suggest greater utility in a distribution based on shared intuitions of justice than in a distribution based upon optimizing deterrence, rehabilitation, or incapacitation.\footnote{See Paul H. Robinson & John M. Darley, \textit{The Utility of Desert}, 91 \textit{Nw. U. L. Rev.} 453, 456 (1997).}

To briefly summarize those “utility of desert” arguments: deviating from a community’s intuitions of justice inspires resistance and subversion among participants—juries, judges, prosecutors, and offenders—when effective criminal justice depends upon acquiescence and cooperation.\footnote{\textit{Id.}} Relatedly, some of the system’s power to control conduct derives from its potential to stigmatize violators—with some persons this is a more powerful, yet essentially cost-free, control mechanism compared to imprisonment.\footnote{\textit{ROBINSON \\& DARLEY, supra note 6, at 202.}} Yet the system’s ability to stigmatize depends upon it having moral credibility with the community; for a violation to trigger stigmatization, the law must have earned a reputation for accurately assessing what violations do and do not deserve moral condemnation.\footnote{\textit{Id.} at 201-02.}

Liability and punishment rules that deviate from a community’s shared intuitions of justice undercut this reputation.\footnote{\textit{Id.} at 202.}

Perhaps the greatest utility of desert comes through a more subtle but potentially more influential form.

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

\footnote{\textit{Id.}}
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 society-wide 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.

The 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 ... perceive[s] 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 characteristic of ours, an apparently 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 are disposed to believe that the law is an accurate guide to appropriate prudential and moral behavior.

The extent of the criminal law's effectiveness in [all] these respects—in [avoiding resistance and subversion of an unjust system, in bringing the power of stigmatization to bear,] in facilitating[,] communicating[, and maintaining] societal consensus on what is and is not condemnable, and in gaining compliance in borderline cases through deference to its moral authority ... is to a great extent dependent on the degree of moral credibility that the criminal law has [gained] in the minds of the citizens governed by it. Thus, ... the criminal law's moral credibility is essential to effective crime control, and 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 [their shared
intuitions of justice]. Conversely, the system's moral credibility, and therefore its crime control effectiveness, is undermined by a distribution of liability that deviates from community perceptions of just desert.22

The important point here is that distribution according to the moral philosophy conception of desert is not only unnecessary for these utilitarian crime control benefits, but indeed ineffective in gaining them. The beneficial consequences of a desert distribution, described above, flow not from following a deontological desert distribution, but only from following an empirical desert distribution—one that tracks the community's shared intuitions of justice. It is the community's perception that justice is being done that pays dividends, not the system's actual success as measured by a deontological conception of desert.23

On the other hand, empirical desert can be criticized on the ground that it is not a reliable source for determining what is truly deserved, as deontological desert can claim. In other words, generally tracking a community's shared intuitions of justice may well build some credibility within that population that the criminal justice system is doing justice, but it does not follow that in fact justice is being done. Shared intuitions might simply be wrong; they tell us only what lay persons believe is just. Only deontological desert can reliably tell us what is just.24

III. THE ROLE OF MORAL PHILOSOPHERS IN THE COMPETITION BETWEEN DEONTOLOGICAL AND EMPIRICAL DESERT

What role do moral philosophers play in this competition between deontological and empirical desert? We nonphilosophers might well assume that they stand on the side of deontological desert, reasoning out justice from principles of right and good, facing on the other side the social psychology researchers mapping people's shared intuitions of justice to determine empirical desert. In fact, the situation is somewhat more complex than this, in ways that reflect

22. Id. at 587-88.
23. See id. at 7.
24. Id. at 6.
both well and badly on the usefulness of the current moral philosophy project.

First, consider the special usefulness moral philosophy provides. Unfortunately, social psychologists are rather unsophisticated about what drives people's intuitions of justice. Most of the studies that they have done without the involvement of moral philosophers or criminal law theorists are nearly useless, because the investigators are testing concepts that muddle together what moral philosophers know to be distinct and importantly different issues. The moral philosophy literature is the richest and most sophisticated source about lay intuitions of justice that exists today, and it is the starting point that I recommend to any social psychologist doing research in the area.

The reason for this superiority is clear: the current methodology of moral philosophers relies heavily upon intuitions of justice, both informally and formally, as in Rawls's "reflective equilibrium."\textsuperscript{25} A standard analytic form, if not the standard form, among moral philosophers today is to use hypotheticals and philosophers' own intuitions about the proper resolution of the hypothetical as a basis for building moral principles.\textsuperscript{26} Their judgments about the intuitively proper resolution of each of a series of hypotheticals are used as data points, as it were, from which philosophers derive a moral principle, which can then be tested and refined by comparing the moral principle results to philosophers' intuitions on other hypotheticals.\textsuperscript{27} The ultimate effect of this standard methodology is that philosophers have thought more carefully about intuitions of

\textsuperscript{25} See John Rawls, A Theory of Justice 42-43 (rev. ed. 1999) (explaining that the best sense of justice is one that matches a person's judgments in reflective equilibrium—a state reached after consideration of various conceptions of justice).

\textsuperscript{26} See generally Kadish, supra note 5 (employing numerous hypotheticals in an attempt to prove the harm doctrine unsupportable); Katz, supra note 5 (employing hypotheticals to counter Kadish's view of the harm doctrine); Paul H. Robinson, Some Doubts About Argument by Hypothetical, 88 Cal. L. Rev. 813 (2000) (critically analyzing Professor Katz's use of the "argument-by-hypothetical" method).

\textsuperscript{27} See also Leo Katz, Incommensurable Choices and the Problem of Moral Ignorance, 146 U. Pa. L. Rev. 1465, 1480, 1482-84 (1998) (providing an example of moral philosophers using intuitive analysis of case hypotheticals as a standard method by using a hypothetical, derived from the application of the necessity defense to situations where the actor has culpably created the justifying situation, to argue that at times persons can be blamed for making the wrong decision in a state "of unavoidable moral ignorance").
justice than any other group, and their literature reflects this sophistication.

But this methodology is problematic, for several reasons. First, presumably philosophers want to rely upon intuitions that accurately capture the shared intuitions of the community, not some idiosyncratic intuition that only philosophers share. The danger here is not only that philosophers as a group may be different from the rest of the community—some nonphilosophers would think this an obvious truth—but also that, even if philosophers are not idiosyncratic, their methods of testing their own intuitions violate many rules of reliable empirical testing. Presumably, in no situation would moral philosophers be happy to use inaccurate representations of intuitions of justice. But, as I have described elsewhere, the methods by which moral philosophers think they are learning intuitions of justice are simply bad research techniques, giving good reason to believe that they produce unreliable results in assessing intuitions of justice. If philosophers think intuitions of justice are useful to their enterprise, they ought to at least get them right. They ought to look to a more reliable source, or adopt more reliable methods of social psychology research, and not “wing it” on their own.

A second, more problematic feature of moral philosophy’s heavy reliance upon intuitions of justice is that it compromises philosophy’s ability to reliably spot community intuitions of justice that are wrong, in the sense of conflicting with a notion of justice that transcends shared intuitions. The methodological reliance of moral philosophy 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.

28. See Robinson, supra note 26, at 823 (“[I]n some cases, the results we get ... are probably not intuitive judgments of blameworthiness but more likely intellectualized answers generated by applying the professor’s resident collection of theoretical positions—[for example,] whether resulting harm ought to be judged significant.”).
29. Id. at 825.
30. Robinson & Darley, supra note 6, at 5-7.
But providing this transcendent check on intuitions is how philosophers are most useful to law- and policymakers. It is for this check—to assure that a shared intuition of justice does not violate a transcendent principle of justice—that philosophy is given deference. Yet moral philosophers, by their heavy reliance upon intuitions of justice, have become unreliable in performing just this task.

Many nonphilosophers may be shocked to hear that many, if not most, of today's moral philosophers no longer see themselves as being in the business of trying to provide this transcendent check. The moral relativists have given up the enterprise entirely; only the moral realists continue to see it as an explicit and attainable goal. Everyone in between sees themselves as providing some kind of useful guidance to law- and policymakers, but guidance of a sort that is different from the transcendent check on people's intuitions of justice that law- and policymakers need.

The useful guidance they think they provide is, in a sense, to "rationalize" intuitions, as, for example, in translating a set of shared intuitions of justice on a set of cases into a general principle. But social psychologists do that when they interpret data from lay intuition studies to construct a principle that seems to explain how subjects are thinking about the test cases. Philosophers might argue that they also examine and resolve conflicts between competing intuitions, in part by taking account of the relative depth of our commitment to the intuitions in conflict. But, of course, that too is just an empirical question—to which intuition do people have greater allegiance when two conflict?—that social psychologists can more reliably investigate.

But one can imagine that moral philosophers might respond that they are doing something more here than just resolving conflicts between intuitions by testing the relative depth of commitment—the relative strength—of the conflicting intuitions. They might claim that their analysis here goes beyond the empirical to bring to bear some more fundamental, transcendent analysis, relying upon objective principles of right and good. And if they did this, they would have something useful to say to law- and policymakers. They would be providing that needed transcendent check on people's intuitions of justice. Unfortunately, most of today's moral philosophers do not do this, and do not claim to do this. Perhaps they do
not because they think it impossible to do, which is fine, but then they ought to accept their limited usefulness to law- and policymakers, which many of them may do.

A defense of the intuition-dependent methodology that many moral philosophers appear to make is found in a claim that intuitions of justice provide some validating effect in assessing true moral principles of justice: that many people share an intuition means that a moral principle consistent with that intuition is thereby made stronger. Social psychologists would find this an odd claim, for one has good reason to believe that a person’s intuitions of justice are simply behavioral phenomena. It is well documented that people hold strong intuitions of justice even though the reasons for their holding those intuitions are inaccessible to them. When asked to explain an intuition, many people will have nothing to offer, other than perhaps “It’s obvious.” Others, perhaps those who prize their self-image as a rational being, will offer an explanation, yet different people offer different explanations even though their intuitions are identical.

In other words, the research suggests that the source of intuitions of justice is not rational reasoning but rather the effects of evolutionary and social forces. And such a source of intuitions provides no reason to think that intuitions have any claim to validate a moral principle in any transcendent philosophical sense. What gave evolutionary advantage six million years ago on the savanna hardly justifies enshrining as a moral truth today. Monkeys and other primates have intuitions of a similar phenomenological sort, even intuitions about fairness. Are we to assume

31. See, e.g., Rawls, supra note 25, at 42-43; Judith Jarvis Thomson, The Realm of Rights 20 (1990) (“We certainly act as if we thought of many of our moral beliefs as necessary truths.”).


33. See supra note 32.

34. See, e.g., Frans De Waal, Chimpanzee Politics: Power and Sex Among Apes 38-39
that monkey "intuitions" about fairness contribute to the validation of moral principles for monkeys?

CONCLUSION

Do I think that moral philosophers should cease their reliance upon intuitions of justice as they construct what they offer as the deontological conception of desert? Not entirely, for I see great benefit from their work in the research to map shared intuitions of justice. But my guess is that moral philosophers themselves would want to contribute something more than what social psychologists already can do. If they are to provide a philosophical conception of desert that transcends our intuitions of justice, they must adopt a methodology that is more skeptical of reliance upon those intuitions.

(rev. ed. 1998) (reporting evidence that some nonhuman primates have capacities to think purposefully). In a recent experiment, brown capuchin monkeys (Cebus apella) refused to participate in an exchange if they observed another monkey receiving a better deal than they received. Some researchers suggest that this refusal implies not only understanding of exchange and unfairness, but a willingness to endure a cost in what can be interpreted as a kind of protest. Sarah F. Brosnan & Frans B.M. de Waal, Monkeys Reject Unequal Pay, 425 NATURE 297, 297-98 (2003). Indeed, Brosnan and de Waal mention cases of monkeys "[t]hrowing the token at the experimenter." Id. at 299. Similarly, some evidence from the field indicates that rhesus macaques (Macaca mulatta) are subject to harassment if they do not let others know when they have found food, an intriguing potential example of moralistic punishment. Marc D. Hauser & Peter Marler, Food-Associated Calls in Rhesus Macaques (Macaca mulatta): II. Costs and Benefits of Call Production and Suppression, 4 BEHAV. ECOLOGY 206, 211-12 (1993).