University at Buffalo School of Law

Digital Commons @ University at Buffalo School of Law

Journal Articles

Faculty Scholarship

2002

Punishment Theory: Moral or Political?

Guyora Binder

University at Buffalo School of Law

Follow this and additional works at: https://digitalcommons.law.buffalo.edu/journal_articles

Part of the [Criminal Law Commons](#), and the [Political Science Commons](#)

Recommended Citation

Guyora Binder, *Punishment Theory: Moral or Political?*, 5 Buff. Crim. L. Rev. 321 (2002).

Available at: https://digitalcommons.law.buffalo.edu/journal_articles/293

Published as Guyora Binder, Punishment Theory: Moral or Political?, 5 Buf. Crim. L. Rev. 321 (2002). © 2002 by the Regents of the University of California. Copying and permissions notice: Authorization to copy this content beyond fair use (as specified

This Article is brought to you for free and open access by the Faculty Scholarship at Digital Commons @ University at Buffalo School of Law. It has been accepted for inclusion in Journal Articles by an authorized administrator of Digital Commons @ University at Buffalo School of Law. For more information, please contact lawscholar@buffalo.edu.

Punishment Theory: Moral or Political?

Guyora Bindert†

Is the justification of punishment a moral question? Much contemporary writing on punishment, whether by philosophers or legal scholars, treats it as such. Theories of punishment are taken to be moral theories, and the problem of justifying punishment is presented as a key battle-ground in the war between utilitarian and deontological ethics. The question of how and when the state should punish is reduced to the question of how and when particular persons should punish other persons. This question in turn is treated as just a special case of the more general question whether persons are morally obliged to govern their actions by the aim of maximizing human welfare or by rules of fair treatment.

But surely this is an odd way to think about punishment. Punishment is not a behavior, but an institution. It is part of a system that involves conduct norms, an authoritative procedure for generating these norms, an authoritative procedure for decisions to impose sanctions, and some measure of practical power over persons or resources. To punish someone is not just to harm them, nor even just to harm them because of something they have done. It is to stake a claim to a certain kind of institutional authority, even when the institution is only the family. To punish someone is to assert a right and accept an obligation to punish anyone similarly circumstanced and behaved, even if that other person be only a sibling. Punishment is never the isolated act of an individual: to punish is to act as an officer or agent participating in a system for enforcing an authoritatively promulgated norm.

Because punishment is part of a system of institutional authority, it is not amenable to a simple moral

† University at Buffalo Distinguished Professor of Law. This article benefited from discussions with Nick Smith, Markus Dubber, Fred Konefsky, and participants at a Buffalo law faculty workshop.

analysis. The legitimacy of punishment is bound up with the legitimacy of the norm it enforces and of the institutions promulgating the norm, imposing the punishment, and inflicting it. Thus, if we conceptualize the utilitarian and retributivist penologies merely as moral theories, both are implausible.

I. PUNISHMENT AND MORALITY

The implausibility of utilitarian morality as a guide to penal policy is famously illustrated by the textbook problem of framing the innocent. E.F. Carritt reasoned from utilitarian premises that

if some kind of very cruel crime becomes common, and none of the criminals can be caught, it might be highly expedient, as an example, to hang an innocent man, if a charge against him could be so framed that he were universally thought guilty; indeed this would only fail to be an ideal instance of utilitarian 'punishment' because the victim himself would not have been so likely as a real felon to commit such a crime in the future; in all other respects it would be perfectly deterrent and therefore felicitic.¹

H.J. McCloskey posed a similar problem:²

Suppose that a sheriff were faced with the choice of either framing a Negro for a rape that had aroused hostility to the Negroes (a particular Negro generally believed to be guilty but whom the sheriff knows not to be guilty)—and thus preventing serious anti-Negro riots which would probably lead to some loss of life and increased hatred of each other by whites and Negroes—or of hunting for the guilty person and thereby allowing the anti-Negro riots to occur, while doing the best he can to combat them. In such a case the sheriff, if he were an extreme utilitarian, would appear to be

1. E.F. Carritt, *Ethical and Political Thinking* 65 (1947).

2. H.J. McCloskey, *A Non-Utilitarian Approach to Punishment*, 9 *Inquiry* 249, 253 (1965).

committed to framing the Negro.³

McCloskey also provided a variant of the example, in which the utilitarian framer is a stranger passing through town rather than a public official.⁴ In all of these examples, utilitarianism is thought to impose a moral obligation on individuals always to choose the act that will maximize utility. In other words, by "utilitarianism" these authors mean "act-utilitarian ethics."⁵ The examples are contrived to compel individuals who are committed to such an act-utilitarian morality to act unfairly and dishonestly. These counter-intuitive results suggest there is something wrong with utilitarian premises.

Accordingly, these examples are widely thought to illustrate the implausibility of the utilitarian theory of punishment.⁶ Yet these examples proceed from a common

3. H.J. McCloskey, *An Examination of Restricted Utilitarianism*, 66 *Phil. Rev.* 468-69 (1957).

4. McCloskey, *supra* note 2, at 256.

5. For the distinction between act-utilitarian and rule-utilitarian ethics, see Richard B. Brandt, *Ethical Theory* 380 (1959).

6. G.E.M. Anscombe, *Modern Moral Philosophy*, in *Ethics* 206-07 (Judith J. Thomson and Gerald Dworkin eds., 1968); Philip Bean, *Punishment* 34 (1981); Carritt, *supra* note 1, at 65; Alan Donagan, *Is There a Credible Form of Utilitarianism?*, in *Contemporary Utilitarianism* 187-202 (Michael D. Bayles, ed., 1968); Charles Fried, *Right and Wrong* 59-64 (1978); Alan Goldman, *Beyond Deterrence Theory: Comments on van Den Haag's "Punishment as a Device for Controlling the Crime Rate,"* 33 *Rutgers L. Rev.* 721, 722 (1981); Lenn Goodman, *On Justice* 50 (1991); D.H. Hodgson, *Consequences of Utilitarianism* (1967); D.K. Lewis, *Utilitarianism and Truthfulness*, 50 *Australasian J. Phil.* 17 (1972); Steven Lukes, *Five Fables About Human Rights*, 40 *Dissent* 427 (1993); J.D. Mabbott, *Punishment*, 48 *Mind* 152 (1939); McCloskey, *supra* note 2; McCloskey, *supra* note 3; Richard McCormick, *Ambiguity in Moral Choice*, in *Doing Evil to Achieve Good* 7, 33 (Richard McCormick & Paul Ramsey eds., 1978); Edmund Pincoffs, *The Rationale of Legal Punishment* 33-35 (1966); Jeffrey Reiman & Ernest van Den Haag, *On the Common Saying that it is Better that Ten Guilty Persons Escape Than that One Innocent Suffer: Pro and Con*, in *Crime, Culpability and Remedy* 234-35 (Ellen Frankel et al. eds., 1990); George Schedler, *Behavior Modification and "Punishment" of the Innocent: Towards a Justification of the Institution of Legal Punishment* 11-20, 73-99 (1977); C.L. Ten, *Crime, Guilt and Punishment*, 13-14 (1987); Ernest van Den Haag, *Punishing Criminals: Concerning a Very Old and Painful Question*, 27-29 (1975); Ernest van Den Haag, *Punishment as a Device for Controlling the Crime Rate*, 33 *Rutgers L. Rev.* 706 (1981); Donald Beschle, *What's Guilt (Or Deterrence) Got To Do With It?: The Death Penalty, Ritual, and Mimetic Violence*, 38 *Wm. & Mary L. Rev.* 487 n.119

(1997); Michael Louis Corrado, Punishment and the Wild Beast of Prey: The Problem of Preventive Detention, 86 J. Crim. L. 778, 781 (1996); David De Gregorio, *People v. Marrero* and Mistake of Law, 54 Brooklyn L. Rev. 229, 279 n.220 (1988); Anthony Dillof, Punishing Bias: An Examination of the Theoretical Foundations of Bias Crime Statutes, 91 Nw. U. L. Rev. 1015, 1026 (1997) ("utilitarianism would recommend such counter-intuitive practices as... convicting those known to be innocent if sufficient deterrence was achieved."); James Duane, What Message Are We Sending to Criminal Jurors When We Ask Them to "Send a Message" With Their Verdict?, 22 Am. J. Crim. L. 565, 603 n.121 (1995) (asserting that punishing the innocent is a "perennial problem for the utilitarian model of punishment"); David Fried, Rationalizing Criminal Forfeiture, 79 J. Crim. L. 328, 386 n.269 (1988) ("From a utilitarian point of view, punishment of innocent persons is perfectly justified for its deterrent effect, at least if the suffering prevented by the deterrence of future crime outweighs the suffering inflicted by present punishment."); Steven Gey, Justice Scalia's Death Penalty, 20 Fla. St. U. L. Rev. 67 (1992); Kent Greenawalt, Punishment, in 4 Encyclopedia of Crime and Justice 1336, 1337-38 (1983) ("utilitarianism admits the possibility of justified punishment of the innocent"); Kent Greenawalt, "Prescriptive Equality": Two Steps Forward, 110 Harv. L. Rev. 1265, 1288 n.62 (1997) (asserting that utilitarians need to qualify their position by imposing deontological restraints if they are to avoid condoning the punishment of the innocent); Helene Greenwald, Capital Punishment for Minors: An Eighth Amendment Analysis, 74 J. Crim. L. 1471, 1508 n.200 (1983) (under "utilitarian rationale, a man whom the authorities knew to be innocent could be punished, if members of the community believed him guilty and threatened to seek personal revenge unless he were punished."); John Lawrence Hill, Law and the Concept of the Core Self: Toward a Reconciliation of Naturalism and Humanism, 80 Marq. L. Rev. 289, 314 n.94 (1997); Leo Katz, Blackmail and Other Forms of Arm-Twisting, 141 U. Pa. L. Rev. 1567, 1586 (1993) ("utilitarian account of punishment... would permit the punishment of mere innocents for the sake of some utilitarian goal."); Karen Lutjen, Culpability and Sentencing Under Mandatory Minimus and the Federal Sentencing Guidelines: The Punishment No Longer Fits the Criminal, 10 Notre Dame J.L. Ethics & Pub. Pol'y 389, 392 (1996); Leo Martinez, Federal Tax Amnesty: Crime and Punishment Revisited, 10 Va. Tax Rev. 535, 573 n.160 (1991); Toni Massaro, Shame, Culture, and American Criminal Law, 89 Mich. L. Rev. 1880, 1887 n.80 (1991); Christopher Peters, Foolish Consistency: On Equality, Integrity, and Justice in *Stare Decisis*, 105 Yale L.J. 2031, 2042 (1996) ("Bentham (although perhaps not Mill) would have considered the punishment of the innocent person to be a morally *erighti* act if ultimately it resulted in more total happiness than would have existed if the person had not been punished. Under Bentham's form of persons."); Tamara Piety, Scorched Earth: How the Expansion of the Civil Forfeiture Act Has Laid Waste to Due Process, 45 U. Miami L. Rev. 911, 973 (1991); Margaret Jane Radin, Market-Inalienability, 100 Harv. L. Rev. 1849, 1878 n.105 (1987) (claiming utilitarians need "fancy footwork" to avoid endorsing punishment of innocent people); Nancy Rhoden, The Judge In the Delivery Room: The Emergence of Court-Ordered Cesiarians, 74 Cal. L. Rev. 1951, 1995 (1986); David Richards, The Moral Criticism of Law 232-33 (1977); Ronald Rychlak, Society's Moral Right to Punish: A Further Exploration of the Denunciation Theory of Punishment, 65

misunderstanding of utilitarian penology.⁷ They depend on formulating the question of whether and whom to punish as a moral dilemma confronting an individual. If the question is reformulated as whether to establish an institution empowered to frame innocents whenever its agents determine that doing so will advance social welfare, utilitarianism will offer a very different answer. Such an institution would be exceedingly dangerous to the liberty and security of every citizen. Thus, Rawls:

Try to imagine, then, an institution (which we may call "telishment") which is such that the officials set up by it have authority to arrange a trial for the condemnation of an innocent man whenever they are of the opinion that doing so would be in the interests of society. The discretion of officials is limited, however, by the rule that they may not condemn an innocent man to undergo such an ordeal unless there is, at the time, a wave of offenses similar to that with which they charge him and telish him for. . . . Once one realizes that one is involved in setting up an *institution*, one sees that the hazards are very great. For example, what check is there on the officials? How is one to tell whether or not their actions are authorized? How is one to limit the risks involved in allowing such systematic deception? How is one to avoid giving anything short of complete discretion to the authorities to telish anyone they like? In addition to these considerations, it is obvious that people will come to have a very different attitude towards their penal system when telishment is adjoined to it. They will be uncertain as to whether a convicted man has been punished or telished. They will wonder whether or not they should feel sorry for

Tul. L. Rev. 299, 324 n.93 ("If a judge were actually to follow a utilitarian theory of punishment, the judge might actually be required to punish a defendant who was widely believed to be guilty, even if the judge knew the defendant to be innocent."); Michael Seidman, *Soldiers, Martyrs, and Criminals: Utilitarian Theory and the Problem of Crime Control*, 94 Yale L.J. 315, 320 n.11 (1984); Kenneth Simons, *Rethinking Mental States*, 72 B.U. L. Rev. 463, 504 n.141 (1992). R.J. Spijt, *The Relevance of Culpability to the Punishment and Prevention of Crime*, 19 Akron L.J. 197, 199 (1985).

7. This point is developed at length in Guyora Binder & Nicholas J. Smith, *Framed: Utilitarianism and Punishment of the Innocent*, 32 Rutgers L.J. 115 (2000).

him. They will wonder whether the same fate won't at any time fall on them. If one pictures how such an institution would actually work, and the enormous risks involved in it, it seems clear that it would serve no useful purpose. A utilitarian justification for this institution is most unlikely.⁸

Robert Goodin adds:

Once it becomes public knowledge that, as a matter of policy, we are willing to hang innocent people to assuage a baying mob . . . then everyone starts worrying: Who will be next? The anxieties associated with such thoughts systematically occurring across the whole population will more than suffice to cancel the utility advantages of . . . throwing one prisoner to the mob on any given occasion.⁹

Utilitarianism, which presumes that individuals

8. John Rawls, *Two Concepts of Rules*, 44 *Phil. Rev.* 3, 11-12 (1955).

9. Goodin explains further:

My point is . . . that public officials cannot systematically violate people's rights, as a matter of policy, and expect that policy to continue yielding the same utility payoffs time and again. Take the case of punishing criminal offenders, for example. The criminal sanction deters crime only in so far as it is imposed on the guilty and only the guilty. Introducing any probability that the innocent will be punished along with the guilty narrows the expected utility gap between criminal and noncriminal conduct, and increases the temptation for everyone to commit a crime. Thus, if we were as a matter of policy to punish people whether or not they were guilty, just according to some utilitarian calculation of public convenience on a case-by-case basis, then the utilitarian advantages of punishing the occasional innocent person would quickly diminish, and probably soon vanish altogether.

The reason utilitarian policy makers are precluded from violating the rights of the innocent, as a matter of policy, is that policies soon become public knowledge. If nothing else, they are easily inferred from past practices. Once news of such a policy gets out, people revise their expectations in the light of it—in the case of criminal punishment, their expectations of being punished even if not guilty. There are major utilitarian payoffs to be had from sustaining certain sorts of expectations and from avoiding others. Settled policies of one sort or another are characteristically required to produce socially optimal effects in both directions. That is one reason for utilitarian policymakers to abide by settled policies, even when greater utility gains might be achieved in any given instance by deviating from them.

Robert E. Goodin, *Utilitarianism as a Public Philosophy* 70-71 (1995).

maximize their own self-interests, predicts that officials with the means to deceive the public will do so for their own private interests rather than for the public interest. Thus a utilitarian analysis of institutions disfavors a policy of framing the innocent. Carritt's and McCloskey's examples therefore do not prove that utilitarian penology leads to unacceptable consequences. But they do show that we should not apply an act-utilitarian morality to questions of punishment. Utilitarian penology must rest on some other ground.

We can make a similar point about retributivist penology by considering the question of vigilante justice. If retributive punishment rests on the obligations of individuals to punish all and only those who deserve it, then it would seem to condone some instances of lynching. Suppose we are in a society that accepts capital punishment as an appropriate punishment for the most aggravated murders. Now, suppose a group of persons have witnessed an atrocious crime, perhaps a multiple murder. They have clearly identified the murderer and satisfied themselves that he was conscious, sane, and fully culpable for his act. They have disarmed and securely captured the killer, who now poses no further threat. Moreover, let us assume the society we live in has a well-administered impartial legal system that can be counted upon to duly condemn and execute the murderer. Thus there is no practical reason why the captors should not wait until the constituted authorities arrive, and then hand the offender over for legal proceedings. But why, from the standpoint of deontological morality, must they? Suppose instead they reason that the murderer deserves to be killed, and that shifting responsibility for this unpleasant task onto others would be cowardly. Suppose on this reasoning, they simply kill him. In so doing, they will have carried out the hypothesized moral duty to punish the guilty.

Nevertheless, I think most of us would say that this kind of gratuitous vigilante justice would be very wrong, an act of murder. But why? Is it morally wrong? If so, how

can it be morally right for state officials to execute the criminal based on the same considerations? I think the answer is that vigilante justice is not morally wrong, and that legally authorized punishment of the guilty is not morally right. The wrong of vigilante justice is a political wrong and the right to punish conferred by law is a political right. The offender in this case deserves punishment, and the persons punishing him know he deserves punishment. But they have not publicly established that he deserves punishment. They have not proven his guilt before any one who might seek to question it, nor have they publicly established reasons why it is necessary to punish—here to execute—one who is guilty of this crime. Nor have they insured that anyone stands ready to review their actions and to remedy their mistakes or abuses. The problem with our hypothetical lynching is not with the consequence that the murderer gets punished, but with the institution doing the punishing. The defect is not of morality, but of legitimacy.

The preceding examples reveal that neither utilitarian nor deontological morality adequately accounts for our intuitions about punishment, because these intuitions seem to involve political as well as moral ideas. Thus we may hypothesize scenarios in which a private person somehow has absolutely reliable knowledge that a particular punishment would be both utility-maximizing and deserved, and still we may refuse to recognize such punishment as politically legitimate. And, as McCloskey's sheriff example suggests, the same might be true of punishment by a public official. Judge Herbert Stern found himself in such a situation in 1979, when he was sent to occupied West Berlin to preside over a special court convened to try some Polish hijackers of German ethnicity, who had forced a Polish plane to land in West Berlin by threatening the crew with a toy gun. West German authorities were reluctant to prosecute these refugees from communism, but the American executive was unwilling to condone air piracy, no matter how benign the motives or gentle the means. As a court of occupation, Judge Stern

was obliged to apply local law.¹⁰ But he was reluctant to do so, given his position as a temporary, ad hoc court, whose judgments would be executed by an occupation army rather than a constitutionally organized, popularly representative, and judicially supervised executive. At the sentencing of one of the hijackers, Tiede, Judge Stern addressed U.S. government lawyers saying,

Under these circumstances, who will be here to protect Tiede if I give him to you for four years? Viewing the constitution as nonexistent, considering yourselves not restrained in any way, who will stand between you and him? What Judge? What independent magistrate do you have here? What independent magistrate will you permit here? . . . I sentence the defendant to time served. You . . . are a free man right now.¹¹

Stern's position was neither that the hijackers did not deserve punishment, nor that policymakers could not reasonably conclude that social welfare demanded their punishment. The difficulty was that there was no way to ensure that they would be punished according to law. Indeed, there were no policymakers prepared to take responsibility for the decision that utility or desert demanded the punishment of hijackers—not the Germans, who were unwilling to prosecute, and not the Americans who seemed unwilling to afford the defendants the procedural protections of American law.¹²

One of Stern's controversial rulings was to extend to the accused hijackers the American right of trial by jury, which German law did not provide. By recruiting a panel of Berliners to sit in judgment of Tiede, Stern arguably

10. Art. 64, Convention Relative to the Protection of Civilian Persons in Time of War, Concluded at Geneva, 12 August 1949, 75 U.N.T.S. 287.

11. Herbert Stern, *Judgment in Berlin* 370 (1984) (quoted in Robert Cover, *Violence and the Word*, 95 *Yale L.J.* 1620 (1986)).

12. This is a potentially serious defect in international criminal tribunals. See generally Scott T. Johnson, *On the Road to Disaster: The Rights of the Accused and the International Criminal Tribunal for the Former Yugoslavia*, 10 *Int'l Legal Persp.* 111 (1998).

rendered his court more politically responsive and responsible. He ensured that the application of German law would take some account of the sympathy for refugees from communism that was widespread in Germany. The German government had been reluctant to flout such sympathy by prosecuting Tiede, and Stern ensured that American prosecutors would also have to face representatives of the German public. Stern intuited that the German proscription of hijacking had, if not a different meaning, then at least a different force, when applied by Germans. Stern's action implied that legitimate punishment requires not only the procedural check of an independent court, but also the political check provided by a democratic public.

The notion that punishment should be not only morally deserved but also politically legitimate is reflected in the American legal system's tolerance for jury nullification.¹³ Of course one function of jury nullification is to provide an additional check against undeserved punishment. Juries can resist proscriptions that infringe liberty and punishments that are disproportionate, either in general, or as applied to the particular case before them. But jury nullification can also serve the quite different function of checking procedural unfairness in the formulation or administration of the laws.¹⁴ One example of such nullification is provided by Paul Butler's controversial suggestion that black jurors should refuse to convict black men of drug crimes, even when they believe them guilty beyond a reasonable doubt. Butler does not argue that such offenders do not deserve some punishment, if guilty. Instead he proposes jury nullification as a political protest against two features of the criminal justice system. Butler's first target of protest is the persistence of unfair discrimination against African-Americans in decisions about whom to investigate, arrest, and prosecute,

13. See generally Darryl K. Brown, *Jury Nullification Within the Rule of Law*, 81 Minn. L. Rev., 1149 (1997).

14. See Gerald F. Uelman, *Perspective On Justice: Why Some Juries Judge the System*, L.A. Times, Jan. 24, 1996, § B.

what behavior to criminalize, and how severely to punish it. Butler's second target is our society's increasing reliance on a social control strategy that focuses on punitive sanctions rather than education and economic opportunity. The combination of these two features produces devastatingly high rates of incarceration among young African-American males.¹⁵ Butler argues that African-Americans should not simply disagree with these policies, they should regard them as illegitimate. They should regard them as policies that, like slavery, segregation, and disfranchisement, deny the full civic status of African-Americans and so absolve them of any political obligation to respect, or to participate in enforcing, their country's laws.¹⁶ This kind of challenge to the legitimacy of criminal laws and procedures is invited, even institutionalized, by the practice of tolerating jury nullification. This particular device for enhancing political legitimacy is, of course, a double-edged sword. Sometimes nullification by jurors from one group erodes the legitimacy of the criminal justice system among members of another group, as in O.J. Simpson's murder trial, or the first trial of Rodney King's assailants.

A similar conditioning of criminal justice on political legitimacy is implied by the principle of legislativity,¹⁷ found in many legal systems. The California Penal Code, for example, prescribes that "No act or omission . . . is criminal or punishable, except as prescribed or authorized by this code, or by some of the statutes which it specifies as continuing in force . . . or by some ordinance, municipal, county or township regulation."¹⁸ This requirement

15. John Kaplan, Robert Weisberg, & Guyora Binder, *Criminal Law: Cases and Materials* 66 (2000).

16. Paul Butler, *Racially Based Jury Nullification: Black Power in the Criminal Justice System*, 105 *Yale L.J.* 677 (1995); see also Jeffrey Rosen, *One Angry Woman*, *New Yorker*, Feb. 24/Mar. 3, 1997, at 54; Jeffrey Rosen, *The Bloods and the Crits*, *New Yorker*, Dec. 9, 1997, at 27.

17. I use my colleague Markus Dubber's term "legislativity" in preference to the more common, but less precise "legality," which includes the additional requisites of prospectivity and specificity.

18. Cal. Penal Code § 6 (West 2001).

ensures that conduct can only be criminalized by an elected, representative body. When combined with constitutional prohibitions on legislative bills of attainder, it ensures that the body criminalizing conduct will not also have the judicial power to impose or the executive power to inflict punishment. Our system further separates the power to punish by relying on an accusatorial procedure in which the power to investigate and charge vests in elected officers of the executive power rather than in the judiciary. This separation of powers ensures that punishment can only follow a number of independent authorizing decisions, each an occasion for judging the legitimacy of decisions farther upstream.

Yet another aspect of the American system which suggests that moral obligations to punish are subordinate to concerns about political legitimacy is our practice of suppressing reliable,¹⁹ but illegally obtained evidence. A consequence of this practice is to acquit some defendants who are clearly guilty, deserving of punishment, and a threat to public safety. One purpose of suppressing illegally obtained, but nevertheless reliable, evidence is to deter illegal methods of investigation. Yet other sanctions—fining or dismissing rogue investigators—could accomplish this purpose at less cost in social protection and deserved punishment. The practice of ignoring illegally obtained reliable evidence in effect holds the court responsible for the conduct of the police. It precludes the judge from saying “I can conscientiously punish this guilty offender, for I have done no wrong.” This chain of responsibility binds together the institutionally separate processes of investigation and punishment imposition into a single system. That system cannot legitimately punish even the morally deserving offender unless it has respected his civil rights at every stage.

19. Of course one reason to forbid some investigative methods is their general unreliability: torture is apt to produce false confession. But when torture reveals the location of physical evidence or incriminating knowledge of crime scene details unavailable to the public, we are faced with a dilemma between condoning torture and suppressing reliable evidence.

In sum, it makes more sense to think of punishment as a political institution than as a type of private act; and it makes more sense to think of the justification of punishment as a problem of political theory than as a problem of ethics. But does this require junking our received theories of punishment and starting fresh? Must we desist from considering utility and desert when we formulate the criminal law? Not at all. I will argue that both utilitarianism and retributivism were originally political theories concerned with the legitimacy of punishment as an institution. Properly understood, Bentham's utilitarianism was not a moral theory at all, but a political theory, about democratically legitimate lawmaking. Unlike Bentham, Kant was centrally concerned with moral philosophy. But as a consequence, he had much less to say about punishment, which he classified as a problem of justice rather than morality. While Kant required that just actions had to conform to moral standards, he required more: he also held that the justice of particular actions depends on their being authorized by stable, effective, and, to the extent possible, democratic legal institutions. Thus Kant's retributivism was also a theory of legitimate lawmaking with a democratic component. So it seems that both of our received theories of punishment have a good deal to say about the requisites of legitimate punishment and provide useful starting points for developing a theory of punishment as a political institution.

II. UTILITY AND LEGITIMACY

Let us first examine the political dimension of utilitarian penology. The originators of utilitarian penology were Cesare Beccaria and Jeremy Bentham. But both derived their conception of utilitarianism chiefly from Baron Helvetius. Helvetius was quite clear that his utilitarianism had nothing to say about the moral duties of individuals. Human behavior was driven by the rational pursuit of pleasure and avoidance of pain, and so could only

be influenced by incentives rather than preaching. Helvetius therefore concluded that the general happiness was best pursued by establishing efficacious institutions and generally applicable laws.

If morality hitherto has little contributed to the happiness of mankind it is not owing to any want of perspicuity . . . in the moralists; but . . . it must be owned, that they have not often enough considered the different vices of nations as necessarily resulting from the different form of their government; yet it is only by considering morality in this point of light, that it can be of any real use to men.²⁰

For Helvetius, bad behavior resulted from bad government. The proper subject of the moral philosopher was therefore government, rather than ethics.

I say, that all men tend only toward their happiness; that it is a tendency from which they cannot be diverted; that the attempt would be fruitless, and even the success dangerous; consequently, it is only by incorporating personal and general interest, that they can be rendered virtuous. This being granted, morality is evidently no more than a frivolous science, unless blended with policy and legislation: whence I conclude that, if philosophers would be of use to the world, they should survey objects from the same point of view as the legislator. . . . The moralist is to indicate the laws, of which the legislator insures the execution, by stamping them with the seal of his authority.²¹

In their efforts to reform the criminal law, Beccaria and Bentham adopted the perspective recommended by Helvetius. They approached the subject from the standpoint of the legislator, and they viewed criminal law as an institution of government rather than a set of moral norms. Beccaria's famous reformist tract, *On Crimes and Punishments* began with a discussion of political theory,

20. Claude A. Helvetius, *De L'Esprit, Or Essays on the Mind and its Several Faculties* 120 (Burt Franklin Research & Source Works Ser. 1970) (1761).

21. *Id.* at 124-25.

inspired by Rousseau's *Social Contract*.²² Beccaria argued that government was legitimate insofar as rationally consented to, and rational persons would consent only to so much public coercion and injury as served their common interests.²³ While law should serve the common welfare, "the greatest happiness of the greatest number," lawmakers inevitably served their own interests and profited by the public's ignorance and misplaced trust.²⁴ Hence legitimate law could only arise from the will of an enlightened general public, and the function of the philosopher was to enlighten the public as to its own interest.²⁵ Beccaria used utility as the touchstone of policy analysis, because it was a principle on which individuals of differing views could rationally agree;²⁶ hence it could provide the basis for a social contract.²⁷

Beccaria argued that public coercion and injury served the common interest only to the extent that they prevented greater private coercion and injury.²⁸ It followed that deterrence of crime was the only legitimate basis for punishment, and then only where noncoercive measures would not suffice.²⁹ And the most important noncoercive crime control device was the "public tranquility" achieved by establishing legitimate government.³⁰ Like Helvetius, Beccaria assumed that citizens were by nature ruled by passion and self-interest,³¹ and that all social achievement proceeded from using legislation to harness or enable these energies.³²

Beccaria insisted that both laws and their enforcement

22. David Young, Introduction to Cesare Beccaria, *On Crimes and Punishments* xii (David Young trans., Hackett Pub. Co. 1986) (1764).

23. Beccaria *supra* note 22, at 5, 7, 8.

24. *Id.* at 5, 38.

25. *Id.* at 5.

26. *Id.* at 2.

27. *Id.* at 9.

28. *Id.* at 8.

29. *Id.* at 16, 79.

30. *Id.* at 22, 43.

31. *Id.* at 14, 75.

32. *Id.* at 41, 16.

be public³³ and regular.³⁴ He reasoned that certain punishment deterred more effectively than severe punishment, for two reasons. First, certain punishment did not allow the offender the hope of escaping punishment. Second, both severity and discretion undermined deterrence and security by delegitimizing the law. If laws were too severe, citizens would refuse to cooperate with the investigation, prosecution, and punishment of crimes. Hence, the more severe punishments became, the less certain, and so the less deterrent.³⁵ Moreover, if the law could be bent, citizens would seek advantage by turning their energies to intrigue rather than productive accomplishment.³⁶ Citizens would lose respect for law and perhaps oppose it by force.³⁷ Rulers would criminalize dissent, causing unnecessary unhappiness and squelching enlightenment, art, and science. Thus Beccaria was profoundly wary of discretion in the administration of justice.³⁸ Beccaria accordingly insisted that justice was public and so private forgiveness of crimes should play no role in the administration of criminal law. In a well-conceived regime of punishment, characterized by mildness and regularity, pardons would also be unnecessary.³⁹

Thus Beccaria concluded that optimal deterrence of crime depended on avoiding over-criminalization, on certain but mild punishment, and above all on regularity of its administration. But Beccaria's concerns transcended optimal deterrence. His larger concern was with rational, utility-maximizing governance. Here, governmental rationality depended on popular enlightenment,⁴⁰ popular participation in lawmaking, and public scrutiny of the administration of the laws.⁴¹ In sum, Beccaria's theory is a

33. *Id.* at 28, 81.

34. *Id.* at 81.

35. *Id.* at 46.

36. *Id.* at 75.

37. *Id.* at 51-52.

38. *Id.* at 10, 53.

39. *Id.* at 80.

40. *Id.* at 76.

41. *Id.* at 5.

theory about the legitimate use of public coercive power, rather than the private moral obligations of officials. "[O]nly the law may decree punishments for crimes, and this authority can rest only with the legislator, who represents all of society united by a social contract."⁴² Officials are expected to openly serve utility as an enlightened public defines it, by rigorously adhering to rules.⁴³ If they do so, an enlightened public will accept and abide by the laws,⁴⁴ the great aim of public security will be achieved,⁴⁵ the productive energies of society will be freed, and public happiness will flourish.⁴⁶

Utility plays two primary roles in this argument. First, utility is a principle of psychology: thus elites, potential offenders, and the wider public are presumed to be bent on maximizing their self-interest.⁴⁷ Elites are presumed either to know their own self-interest or to profit unwittingly from received arrangements. The general public depends on philosophical enlightenment, for knowledge of the public interest, however. Thus, the second role for the idea of utility is evaluative and forensic. In arguing to the general public for law reforms, the philosopher should advert to the effect of the proposed reforms on the general welfare. If a proposed reform serves the public welfare, it may be rationally consented to. If rationally consented to, it is legitimate. Thus the normative significance of utility rests on a contractarian theory of the legitimacy of laws, not on a general theory of value, or a theory of moral obligation. There is no room for officials to secretly pursue utility since this contributes nothing to the legitimacy of government and risks destroying it. "The right to inflict punishment does not belong to an individual, but to all citizens, or the

42. *Id.* at 9.

43. *Id.* at 54.

44. *Id.* at 22.

45. *Id.* at 12.

46. *Id.* at 74-75.

47. *Id.* at 14, 75.

sovereign."⁴⁸ Accordingly, officials charged with the administration must be closely observed, to prevent their betraying the public interest in favor of their own.⁴⁹

It should be clear that Beccaria views punishment as a governing institution, to be evaluated from the standpoint of political legitimacy rather than morality. Because of his institutional focus, Beccaria would be bound to frame the problem of punishing the innocent in the same way that Rawls did: as a question about the utility of an institution like telishment. Because of his ultimate concern with political legitimacy, Beccaria would be bound to ask whether citizens would rationally agree to establish an institution so insulated from popular scrutiny and control. For Beccaria, utility is not an objective standard of value capable of being deployed by anyone. Utility means the welfare of a particular population, as it appears to them in the context of public debate. It is a standard of value to be deployed by a democratic public at large. Hence, when officials employ utility as a standard, they may do so only by the express delegation of this public decisionmaker, and only within the limits set by that public. The utilitarian standard of value, as it is developed by Beccaria, presupposes a democratic decisionmaker and process of decision.

The same was true of Bentham's utilitarianism, although it was not built on a foundation of contractarian political theory. Indeed, as I have elsewhere demonstrated, Bentham did not build his policy science on any philosophical foundation, and certainly not an ethical one.⁵⁰ Readers have been misled about the structure of Bentham's utilitarianism by his statement in *The Introduction to the Principles of Morals and Legislation*, that "[t]he principle of utility is the foundation of the present work. . . ."⁵¹ Most

48. *Id.* at 55.

49. *Id.* at 78.

50. See generally Binder & Smith, *supra* note 7.

51. Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* 11-12 (J.H. Burns & H.L.A. Hart eds., Cambridge U. Press 1996) (1823) [hereinafter Bentham, *Introduction*].

readers assume that this means that the greatest happiness of the greatest number is Bentham's definition of value and so provides the foundation for any normative inquiry. But Bentham was not interested in philosophical questions about the ultimate nature of the good. He was interested in the question how to rationally settle political disagreement and develop legitimate law. He regarded the utility principle as a useful premise for political debate about legislation because he thought everyone could agree that other things being equal, it would be better to increase rather than decrease the general welfare.

The utility principle is offered as the common ground for political debate about legislation, rather than as a foundation for ethics. Consider the book's opening paragraph:

Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do. . . . They govern us in all we do, in all we say, in all we think. Every effort we can make to throw off our subjection will serve but to demonstrate and confirm it. In words a man may pretend to abjure their empire: but in reality he will remain subject to it all the while. The *principle of utility* recognizes this subjection, and assumes it for the foundation of that system, the object of which is to rear the fabric of felicity by the hands of reason and law. Systems which attempt to question it, deal in sounds instead of sense, in caprice instead of reason, in darkness instead of light.⁵²

Bentham here reasons that individuals are self-interested and can only be induced to be good by laws offering rational incentives. Arguments supporting such laws either rationally explain how such laws will gratify the audience or they are empty cant, merely indicating that such laws will gratify the speaker. When equated with utility, "the words *ought*, and *right* and *wrong*, and others

52. Id. at 11.

of that stamp, have a meaning: when otherwise, they have none."⁵³ This is not a claim about the nature of the good, but about the nature of political discourse among self-interested actors. As H.L.A Hart says of this passage,

it is . . . intended to convey an idea which is central to his whole argument, namely that when so used [the words ought and right and wrong] raise a rationally settleable issue because only then do they invoke an external standard which reasonable men would accept for the determination of right and wrong.⁵⁴

It is in this limited sense that the principle of utility provides the foundation of Bentham's argument. Bentham's argument is about good legislation. The audience for such argument is necessarily a public audience, and such an audience can only be rationally persuaded by arguments about the consequences of laws for the public welfare. And evaluation of acts according to the public welfare is the meaning of the utility principle when the actor one is trying to persuade is the public. When the actor one is trying to persuade is an individual, the principle of utility has the quite different meaning of self-interest. Here is how Bentham draws the line between the subjects of ethics and legislation:

Private ethics teaches how each man may dispose himself to pursue the course most conducive to his own happiness, by means of such motives as offer themselves: the art of legislation (which may be considered as one branch of the science of jurisprudence) teaches how a multitude of men, composing a community, may be disposed to pursue that course which upon the whole is most conducive to the happiness of the whole community, by means of motives to be applied by the legislator.⁵⁵

53. *Id.* at 13.

54. H.L.A. Hart, *Bentham's Principle of Utility*, in Bentham, *Introduction*, *supra* note 51, at xc.

55. *Id.* at 293.

In making sense of Bentham's principle of public utility, it is important to remember its rhetorical function, as the focus of a deliberative process of public discussion and collective decision. Because the public utility principle is merely a device for guiding the desired discursive process, it is not an end in itself. Therefore, it cannot be defined in ways that would undermine the conditions of reasoned public discourse and rational collective decisionmaking. In utilitarian political thought, process values take precedence over substantive values.

Thus, Bentham defines public happiness in such a way that it inherently requires certain institutional conditions. For Bentham, the key feature that differentiates human from animal experience is the ability to anticipate both pleasure and pain.⁵⁶ This trait is important to the utilitarian picture of human beings as rational calculators of future consequences. But when humans anticipate future pleasure and pain they are not simply dispassionate calculators. The expectation of pleasure is itself a pleasure, and the expectation of pain is itself painful. Thus, when anticipated, pain can be greatly magnified by the apprehension that precedes it. Similarly, contentment can be greatly magnified by a sense of security in its continuation. Accordingly, the greatest portion of the happiness humans seek is security.⁵⁷ This concern for security of expectations makes utilitarianism interested not only in beneficial consequences, but in the stability of the institutional means by which we might pursue these consequences.

Accordingly, Bentham was less interested in advocating particular policies as utility-maximizing than he was in advocating a policy *process* that would guarantee

56. See Jeremy Bentham, *Principles of the Civil Code*, in 1 *The Works of Jeremy Bentham* 308 (John Bowring ed., 1962).

57. Frederick Rosen, *Jeremy Bentham and Representative Democracy: A Study of the Constitutional Code 211-20* (1990) [hereinafter Rosen, *Bentham and Representative Democracy*]; Frederick Rosen, *Introduction*, in Bentham, *Introduction*, *supra* note 51, at xxxvi [hereinafter Rosen, *Introduction*]; Binder & Smith, *supra* note 7, at 174-75.

to the public that policy would systematically serve public utility. This process required five interrelated institutional conditions that, together would serve as "securities against misrule." (1) institutionalized investigation of social conditions; (2) a transparent language of policy analysis and legislation; (3) separation of legislative and administrative functions; (4) publicity of government decision-making; and (5) democratic election of legislators.

The first requisite of a utilitarian policy process is information about social conditions and their effect on the welfare of the society. Such information might be gathered by public officials, pursuant to legislation, by legislative commissions and hearings, or by scholars, journalists, and publicists.⁵⁸ But if legislation is to be judged by its expected contribution to the public welfare, it must be framed in light of information about society. Once implemented, legislation must be reevaluated based on its actual effects. Bentham's concern with the gathering of information is evident in his famous "Panopticon" proposal, that prisons and other custodial institutions be designed as laboratories for the observation and study of their inmates.⁵⁹ Bentham imagined policy as a kind of controlled experiment, in which both the interventions and the observations of the investigators would be carefully recorded and catalogued for future study. One difficulty with the common view that utilitarianism obliges individuals to act so as to maximize the public welfare is that it leaves unexplained how individuals are expected to know what actions will best serve the public welfare. Yet individuals are likely to have limited information and their perceptions of society are likely to be distorted by their own interests.⁶⁰ By contrast, utilitarianism envisions an

58. Jeremy Bentham, *Constitutional Code*, in 9 *The Works of Jeremy Bentham* 41-46 (John Bowring, ed. 1962) [hereinafter *Bentham, Constitutional Code*]; T.P. Peardon, *Bentham's Ideal Republic*, in 3 *Jeremy Bentham: Critical Assessments* 621, 641 (Bikhu Parekh ed., 1993); Rosen, *Bentham and Representative Democracy*, *supra* note 57, at 27 n.50.

59. Jeremy Bentham, *Panopticon; Or, the Inspection-House*, in 4 *The Works of Jeremy Bentham* 37-172 (John Bowring ed., 1962).

60. See generally Russell Hardin, *Morality Within the Limits of Reason*

institutionally organized collective process of gathering information and subjecting it to expert analysis and public deliberation.⁶¹ This process culminates in the promulgation of generally applicable conduct rules, rather in individual decisions about how to act.

Thus, a second requisite of utilitarian policymaking is the rational and impartial analysis of the information gathered. This required a certain kind of language of policy analysis.⁶² Such a language should be as dispassionate and value-free as possible. While Bentham bet on the widespread normative appeal of collective happiness, he assumed that values were ultimately subjective and that there was no point in debating them. A useful language of policy analysis should therefore avoid evaluative terminology, should clearly identify publicly observable facts and should always attribute value judgments to particular persons. The concept of utility itself was an effort to replace vague, hortatory evaluative claims with descriptive claims about the hedonic states of particular persons.⁶³ Bentham complained that much moral philosophy was obfuscatory cant, that persuaded through emotional appeal rather than reasoned argument. He also argued that most reformers had an incentive to disguise the true failures of prevailing policy. Given human nature, prevailing policy was likely to benefit only those in a position to make law. Reformers were likely to be persons interested in replacing current lawmakers so as to benefit from prevailing or similar policy. Their interest lay in deceiving the public into thinking that permitting them to make law would benefit the public.⁶⁴ Bentham devoted considerable energy to developing and defining new terms to use in utilitarian policy analysis, so as to permit a clear comparison of the costs and benefits of

(1988).

61. See generally L.J. Hume, *Bentham and Bureaucracy* (1981).

62. Binder & Smith, *supra* note 7, at 176-84.

63. *Id.* 184; Bentham, *Introduction*, *supra* note 51, at 26-28.

64. Bentham, *Constitutional Code*, *supra* note 58, at 44.

alternative policies.⁶⁵ An important purpose of developing a uniform and clear language of policy analysis was to permit the public to assess the impact of policies on its own interests. This in turn would enable the public to make rational choices of policies and of elected representatives. In other words, Bentham's project of reforming the language of political debate presupposed a role for the democratic public in assessing and influencing policy. Utilitarianism was in this sense a project of public enlightenment, aimed at educating the public politically, so as to qualify it for its new role of political participation in the emergent democratic state.

If the pursuit of the public welfare required a clear language of policy analysis, it also required a clear language of legislation.⁶⁶ It was important that legislation be clear to at least four different audiences. The meaning of legislation needs to be clear to legislators, scholars, and publicists, so they can study and analyze its likely effects. The meaning of legislation must be clear to members of the public so that they can evaluate it, express their opinion of it, and vote for or against legislators who propose or enact it. The effect and import of legislation also must be clear to those subject to it so that they can understand what conduct it demands of them and what positive and negative incentives it provides. Finally, it must be clear to officials charged with implementing it, so they have no occasion to exercise legislative discretion. If the meaning of the law is clear to the public, it should also be possible for the public to monitor official compliance with law, and so to identify illicit exercises of political discretion by officials. In order to render the law clear and accessible to all these audiences, it was necessary that it be codified, so that gaps, redundancies and conflicts could be identified and eliminated. It was also necessary that all important terms be defined and used consistently within and across codes. Bentham pursued this program by criticizing the legal

65. Binder & Smith, *supra* note 7, at 183-84.

66. *Id.* at 182-86.

terminology of the common law as obfuscatory and by developing and defining terms of his own for use in his model codes.

Bentham opposed discretion in the application of law because of another requisite for a properly utilitarian policy process: the separation of legislative from judicial and executive functions.⁶⁷ Bentham sought to isolate the exercise of policymaking authority in one body so as to render it visible and responsive to the public. This required that law be enacted by the legislature only: there could be no customary or common law.⁶⁸ Laws needed to be precise and detailed so as to minimize opportunities for judicial construction or administrative discretion.⁶⁹ Bentham proposed that official application of the law be monitored and that misapplication of law be punished. Thus, for Bentham, separation of powers was a means to legislative supremacy. Legislative supremacy, in turn, insured that law would be made in the form of general rules.

Bentham's insistence on a legislative monopoly on policymaking fits with two other "securities" against misrule: publicity and democracy. The reason to locate policymaking authority in one place was to permit public monitoring of policymaking. The reason to locate it in an elected, representative legislature was to insure that policy would respond to popular will.

Publicity was essential to this system.⁷⁰ Bentham

67. *Id.* at 184-89.

68. John R. Dinwiddy, *Bentham* 64 (1989); Jeremy Bentham, *Rationale of Judicial Evidence*, in 7 *Works of Jeremy Bentham* 1, 122 (John Bowring ed., 1962).

69. Jeremy Bentham, *Constitutional Code*, *supra* note 58, at 383-417; Jeremy Bentham, *Securities Against Misrule and Other Writings for Tripoli and Greece* 42, 252 (Philip Schofield ed., 1990) [hereinafter *Bentham, Securities Against Misrule*]; G.J. Postema, *The Principle of Utility and the Law of Procedure: Bentham's Theory of Adjudication*, in 3 *Jeremy Bentham: Critical Assessments*, *supra* note 58, at 308, 316; L.J. Hume, *Jeremy Bentham and the 19th Century Revolution in Government*, in *id.* at 829, 825; Frederick Rosen, *Jeremy Bentham and Constitutional Democracy: A Study of the Constitutional Code* 65, 86 (1983).

70. Binder & Smith, *supra* note 7, at 200-08; Bentham, *Securities Against Misrule*, *supra* note 69, at 25-28; Dinwiddy, *supra* note 68, at 89.

insisted that every government action be duly recorded, along with the reasons for which it was taken and the evidence relied on. All government proceedings, including every stage of the criminal justice process, should be open to the public and reported in the press.⁷¹ Bentham reasoned that the combination of institutionalized public observation of government proceedings and a free press would create an informed public. He referred to such an informed public as "the Public Opinion Tribunal," and considered it a crucial feature of a legitimate constitutional order.⁷² The effectiveness of public inspection of government depended, however, on clearly delineated government functions and the use by government officials of a clear and common language for reporting and justifying government action. Bentham reasoned that government officials would surely pursue their own interests rather than the public interest, unless they were checked by the threat of removal by the public and unless that threat was rendered effective by public monitoring.

But effective public monitoring of government action had another advantage as well. Not only would it force officials to pursue the public welfare, as the public understood it; it would also reassure the public that officials were pursuing the public's understanding of its welfare and that they were constrained to do so. In other words, an open, transparent process of policymaking was crucial to providing the public with the sense of security that Bentham regarded as the greater part of happiness.

The threat of removal of legislators and other public officials from office required democratic election of legislators, to whom all government officials were, in turn, subordinate.⁷³ Bentham was initially suspicious of democracy, because of his disapproval of the natural rights language by which it was justified in the rhetoric of the

71. Bentham, *Securities Against Misrule*, *supra* note 69, at 11, 50.

72. *Id.* at 28, 57-61; Bentham, *Constitutional Code*, *supra* note 58, at 41-46; Frederick Rosen, *Jeremy Bentham and Representative Democracy* *supra* note 57, at 27.

73. Binder & Smith, *supra* note 7, at 189-200.

American revolution. But his views on democracy warmed as he experienced frustration in winning the support of British leaders for his proposed codes. He also became very impressed with America's constitutional arrangements.⁷⁴ Although in many ways an elitist, who thought the legitimacy of policy depended on its rationality rather than its popularity, Bentham became convinced that elites could not be trusted to serve the public welfare unless they were tethered to popular will. And while the public might be deceived about its interests in the short run, public judgment would improve with political experience. In an open process of discussion, a roughly correct understanding of the public interest would eventually prevail.⁷⁵ But even if the policy developed through a democratic political process was not the most efficacious possible, utilitarianism would nevertheless demand such a process. Again, only democratic checks on government action could reassure the public that policy was not being hijacked by private interest. Only such checks could secure the public against concerted attacks on the public welfare by tyrants.

Bentham is commonly understood to have simply proposed an ethical criterion, an evaluative principle to be applied by every actor in making every decision. In fact, he did nothing of the sort. What he instead proposed was a complex institutional process for making and implementing policy. Part of that institutional process was a uniform discourse of policy analysis. The utility principle was central to that policy discourse, but it was not prior to or more fundamental than the institutions necessary to render such a policy discourse a useful instrument of the

74. Bentham, *Constitutional Code*, *supra* note 58, at 63; Bentham, *Plan of Parliamentary Reform*, in 3 *The Works of Jeremy Bentham* 433, 447, 494 (John Bowring ed., 1962); Bentham, *The King Against Edmonds and Others*, in 2 *The Works of Jeremy Bentham* 239, 246 (John Bowring ed., 1962); H.L.A. Hart, *Essays on Bentham*, 101 (1982); Frederick Rosen, *Introduction*, *supra* note 57, at xlv; Elie Halevy, *The Growth of Philosophic Radicalism* 254-56 (Mary Morris trans., 1966); M.H. James, *Bentham's Democratic Theory at the Time of the French Revolution*, 10 *The Bentham Newsletter* 5-16 (1986).

75. Binder & Smith, *supra* note 7, at 199-200; Bentham, 1 *Constitutional Code* 36 (F. Rosen & J.H. Burns eds., 1983).

public welfare. Utilitarianism does not instruct individuals to maximize utility. Maximizing utility is exclusively the job of a democratically controlled, publicly monitored legislature, in its role of enacting clear, stable, legal rules of general application. What utilitarianism demands of the individual is merely that she help realize utility by supporting utilitarian institutions, by making available information about local social conditions and her own preferences, by monitoring government, and by informing herself about public issues before casting a vote.

This revised understanding of utilitarianism as a democratically controlled policy process does not by itself imply very definite prescriptions with regard to criminal justice. But it does imply that, contrary to prevailing perceptions, utilitarian criminal justice will be highly formalistic, consistent, transparent, and procedurally scrupulous. Accordingly, it rules out some common misunderstandings of the implications of utilitarianism. Thus, it disposes of the oft-repeated objections of McCloskey and Carritt.⁷⁶ Properly understood as an institutional process rather than an ethical doctrine, utilitarianism neither compels nor authorizes individuals to punish the innocent. A practice of punishing the innocent could only enhance deterrence if it involved misleading the public into believing the persons punished were guilty. For the reasons that Goodin and Rawls develop, public awareness that officials were pursuing such a policy of framing the innocent would destroy the security that it is the aim of utilitarianism to achieve. And for such a policy to be developed and maintained without public knowledge would require institutional conditions completely inconsistent with such security. Bentham's key securities against misrule—legislative supremacy, publicity, and democratic accountability—would preclude any policy involving the systematic deception of the public. Still, one might argue that the goal of maximizing the public welfare might, from a moral point of view, justify an

76. Binder & Smith, *supra* note 7, at 188, 208-09, 212.

individual official or witness in framing an innocent person. The rejoinder is that such a moral view does not follow from utilitarianism, which was a theory about how to design institutions so as to secure the procedural conditions for identifying and collectively pursuing the public welfare, by means of law.⁷⁷

The received understanding and critique of utilitarian penology presumes that "utility," or the public welfare, is just there, available to be inspected and maximized by any well-meaning moral agent. Generations of political and economic theorists have raised questions about the notion that group preferences can be treated as independent of the institutions by which they are measured and defined. And some have even doubted that individual preferences can be viewed as independent of the institutions that measure and aggregate them into collective preferences.⁷⁸ Thus, the conception of the public welfare that critics ascribe to utilitarianism is not one that many contemporary political or economic theorists hold. What these critics fail to realize is that Bentham himself had a more complicated conception of the public welfare. While he probably thought of "private utility" as fairly straightforward fact about subjective states, he thought of the "public utility" that was to guide legislation as an institutional construct. Public utility was the product of an institutionally organized and regulated discursive process of social research and policy analysis. The utilitarian imperative is not for individuals to maximize the aggregate happiness of society as they see it. The utilitarian imperative is for politics to realize the institutional conditions for public utility.

77. *Id.* at 210-11.

78. Jane J. Mansbridge, *Beyond Adversary Democracy* (1980); Kenneth Arrow, *Social Choice and Individual Values* (1963); Richard H. Pildes & Elizabeth Anderson, *Slingshot Arrows at Democracy: Social Choice, Value Pluralism, and Democratic Politics*, 90 *Colum. L. Rev.* 2121 (1990); Stephen Holmes, *Passions and Constraint: On the Theory of Liberal Democracy* (1995).

III. RETRIBUTIVE JUSTICE AND LEGITIMACY

Contemporary retributivism traces its lineage to Kant,⁷⁹ who discussed punishment briefly in the course of laying out his theory of justice in the first half of *The Metaphysics of Morals*.⁸⁰ For Kant, just actions are a subset of moral actions, so that legitimate punishment must satisfy the dictates of both morality and justice. Thus, clarifying Kant's rather hazy statements on punishment requires explicating his conceptions of morality and justice.

For Kant, moral action is action undertaken because of a belief that it is required by a principle that can be coherently seen as binding on everyone.⁸¹ Examples of principles that Kant would have considered contradictory when universalized are the following:

Take whatever you want to have.

Tell others whatever you wish them to believe.

For Kant, these principles are self-contradictory when universalized, because they are self-defeating. Thus, if everyone takes whatever she wants, no one will be able to keep and have what she takes.⁸² And if everyone routinely says what she wishes others to believe rather than what she actually believes, what people say is true will have no probative value. It will simply be taken as evidence of what the speaker wishes the hearer to believe, rather than as evidence of what is true. Thus it would become impossible to induce others to believe something is true

79. Herbert Morris, *Persons and Punishment*, 52 *The Monist* 475 (1968); see also e.g., Jeffrie G. Murphy, *Three Mistakes About Retributivism*, 31 *Analysis* 166, 166-69 (1971); Andrew Von Hirsch, *Doing Justice* 47-48 (1976).

80. Immanuel Kant, *The Metaphysics of Morals* (Mary Gregor trans., Cambridge U. Press 1991) (1785).

81. Immanuel Kant, *Groundwork of the Metaphysics of Morals* 15, 31 (Mary Gregor trans., Cambridge U. Press 1998) (1797) [hereinafter Kant, *Groundwork*].

82. Immanuel Kant, *Metaphysical Elements of Justice* 139 (2d ed. John Ladd trans., Hackett Pub. Co. 1999) (1796) [hereinafter Kant, *Metaphysical Elements*].

merely by saying it is.⁸³ Kant therefore concludes that stealing undermines the institution of property, lying undermines the institution of truth telling, and bad faith undermines the institution of promising. In all these cases, the principle of exploiting a cooperative institution to gain an unfair share of the benefits created by that institution is rejected as contradictory because, if universalized, it would destroy the institution and thereby preclude the very benefit sought.⁸⁴

Kant further refines his test of morality by treating morality itself as such a cooperative institution. A moral action must not only conform to a universalizable principle: it also be undertaken because of a universalizable principle.⁸⁵ This implies that a moral action may not be done because of inclination (the present desire so to act) or interests (the hope of bringing about later consequences desirable to one's self or another). If a moral action may not be taken because of inclination or self-interest, it follows that it may not be taken because of a coercive threat, which merely creates a motive of self-interest. So to act morally requires a capacity to resist inclination, to reason out one's moral obligations (using the universalization test), and to make an uncoerced choice to act according to the resulting moral judgment. It follows that to act morally is always to act on the basis of two principles: the universalizable principle identified by one's moral judgment, and the metaprinciple that one should freely act on the basis of one's moral judgment. Universalized, this metaprinciple becomes the rule that all persons should act according to their own moral judgments. But doing so requires that they be permitted the opportunity to develop their own moral views and to choose to act on them—or not—without coercion. It follows that no moral principle passes the universalization test unless it is compatible with the exercise of autonomy by all others. Morality therefore requires not only fair cooperation with

83. Kant, *Groundwork*, *supra* note 81, at 15, 32.

84. Ralph Walker, *Kant and the Moral Law* 36 (1999).

85. Kant, *Groundwork*, *supra* note 81, at 36.

others, but also equal respect for the autonomy of all others.⁸⁶ Respecting this autonomy requires granting to others the same freedom to develop into rational moral agents that engenders one's own moral agency. It also requires according others equal freedom to pursue the nonmoral ends that constitute self-interest, because one cannot choose to act on moral ends without the freedom to act on nonmoral ends.⁸⁷

Kant's principle of equal autonomy generates difficulties, however, in that if individuals are completely free to act as they choose, they will infringe one another's freedom.⁸⁸ Thus, the efforts of all to act autonomously are incompatible unless they all freely choose to act morally.⁸⁹ And they cannot be constrained to act morally, because action cannot be moral unless it is freely willed.

The inevitable conflict of freely willed actions explains the necessity of law, or collective coercion; the tension between law and the moral autonomy of those subject to it frames the problem of justice, or legitimate coercion. In order to protect the autonomy of its members, society must coerce its members to behave according to the dictates of morality, but thereby appears to deprive them of their autonomy and so of any opportunity to act morally. Kant's solution to this paradox is a social contract, modeled on Rousseau's, in which society's members freely subject themselves to law.⁹⁰ Having consented to the imposition of legal coercion, they can follow the dictates of the law and still act morally. They can also coerce one another through the medium of the state in a way that they cannot do as individuals, because the state represents the will of those

86. Kant, *Metaphysical Elements*, *supra* note 82, at 38; Allen Rosen, *Kant's Theory of Justice* 64-65 (1965).

87. Jeffrie G. Murphy, *Kant: The Philosophy of Right* 93, 95, 101 (1970).

88. Rosen, *supra* note 86, at 16. This problem is thoroughly ventilated in Joseph Singer, *The Legal Rights Debate in Analytical Jurisprudence from Bentham to Hohfeld*, 1982 *Wis. L. Rev.* 975.

89. The condition of universal freedom reconciled by free submission to the moral law is Kant's "Kingdom of Ends." See Walker, *supra* note 84, at 41; Barbara Herman, *The Practice of Moral Judgment* 85, 227 (1993).

90. Rosen, *supra* note 86, at 33.

coerced.⁹¹

Since an effective legal system is necessary to protect one's own and others' autonomy, the exercise of morally autonomous choice implies a moral obligation to subject one's self to law.⁹² Moreover, law's interference with autonomy is limited. Where morality regulates intentional states or internal acts of will, law governs only external conduct.⁹³ It does not require individuals to act out of moral motives, but simply to conform their conduct to universalizable and so moral principles of action. In essence, law forces individuals to comply with rather than exploit cooperative institutions.

Punishment is legal coercion, threatened and inflicted because of the morally culpable violation of a conduct norm. By coercion, Kant means a sanction that effectively deprives an actor of the advantage she hoped to gain by violating a cooperative norm of morality. In this way it not only deters others from violating the norm, it frustrates the criminal's selfish purpose in violating the norm. By frustrating the criminal's purposes, punishment represents, in the criminal's own experience, the contradiction that would arise from universalizing the criminal's principle of action. Here is how Kant explains the meaning of imprisoning a thief:

Inasmuch as someone steals, he makes the property of everyone else insecure, and hence he robs himself (in accordance with the law of retribution) of the security of any possible property. He has nothing and can also acquire nothing, but he still wants to live, and this is not possible unless others provide him with nourishment. But because the state will not support him gratis, he must let the state have his labor at any kind of work it may wish to use him for (convict labor), and so he becomes a slave, either for a

91. *Id.* at 62; Murphy, *supra* note 87, at 108.

92. Rosen, *supra* note 86, at 10.

93. Kant, *The Metaphysics of Morals*, *supra* note 80, at 56; Rosen, *supra* note 86, at 89-90.

certain period of time, or indefinitely, as the case may be.⁹⁴

Kant holds that a threatened sanction does not count as punishment unless it actually has a coercive effect in one of the two senses discussed above. That is, it must actually frustrate the immoral purposes of those on whom it is imposed, or its threat must deter some persons from committing the crime. Kant apparently did not consider whether the probability of punishment must be so great as to deter any rational person from committing the crime. But he did seem to think of punishment as inherently deterrent. Kant revealed this aspect of his conception of punishment in a discussion of the doctrine of necessity. Kant posed the hypothetical of two shipwrecked sailors adrift in the ocean, struggling over a plank large enough to support only one of them. He concluded that neither one is justified in pushing the other off, to his death, because such a principle of action, if universalized, would be self-defeating. But he also concluded that it is impossible to punish such a killer, because the threat of later punishment could not possibly deter the killing, since its execution would not wholly deprive the survivor of the advantage he would gain by killing his competitor.

This imagined right [of necessity] is supposed to authorize me to take the life of another person when my own life is in danger, even if he has done me no harm. It is quite obvious that this conception implies a self-contradiction within jurisprudence. . . . It is clear that this allegation is not to be understood objectively, according to what a law might prescribe, but merely subjectively, as the sentence might be pronounced in a court of law. There could be no penal law assigning the death penalty to a person who has been shipwrecked and finds himself struggling—both of them in equal danger of losing their lives—and in order to save his own life pushes the other person off the plank on which he had saved himself. For no threatened punishment from the law could be greater than losing his life in the first instance.

94. Kant, *Metaphysical Elements*, *supra* note 82, at 139.

Now a penal law applied to such a situation could never have the effect intended, for the threat of an evil that is still uncertain (being condemned to death by a judge) cannot outweigh the effect of an evil that is certain (being drowned). Hence, we must judge that, although an act of self-preservation through violence is not inculpable, it still is unpunishable. . . .⁹⁵

Kant's claim here is not that such punishment would be pointless or wasteful or cruel, but that it would not even be punishment at all. The threatened harm must exceed the expected benefits of the crime, or the actual harm must exceed the actual benefits of the crime, for either to count as coercive, in Kant's sense. Both imposing a conditional threat and forcibly depriving someone of a sought after benefit are ways of interfering with their autonomous choice. In its coercive effect we may analogize punishment to coercive offenses like robbery, or rape, which may violate consent by either force or the threat of force.

It may seem surprising that Kant's conception of punishment presumes its deterrent effect, since he is often thought to be indifferent to consequences. Certainly, he professed to consider consequences morally irrelevant. Thus, in the *Groundwork of the Metaphysics of Morals*, he inveighs that "the moral worth of an action does not lie in the effect expected from it and so too does not lie in any principle of action that needs to borrow its motive from this expected effect."⁹⁶ On the other hand, he acknowledged in other passages that morality aims at ends: for example, morality requires us to treat each human being as an end by according her as much freedom as is compatible with our own.⁹⁷ But even if consequences were irrelevant to the morality of an act, this would not entail that consequences were irrelevant to justice, which is a different standard of value. And justice, rather than morality, is the standard

95. Id. at 35-36.

96. Kant, *Groundwork*, supra note 81, at 14.

97. Walker, supra note 84, at 8-9. For a general argument that Kant's moral theory only makes sense as a partly consequentialist method of practical judgment, see Barbara Herman, *The Practice of Moral Judgment* 73-112 (1993).

Kant proposes for assessing punishment.

So what does Kant have to say about the consequences of punishment? Here his views are more complex than is commonly recognized, in that he rejects only certain kinds of consequences as irrelevant to justice. In a widely quoted and colorful passage, he objected to what he imagined to be the implications of utilitarianism for punishment:

Judicial punishment can never be used merely as a means to promote some other good for the criminal himself or for civil society, but instead it must in all cases be imposed on him only on the ground that he has committed a crime; for a human being can never be manipulated merely as a means to the purposes of someone else and can never be included among objects of the law of things. His innate Personality protects him against such treatment, even though he may indeed be condemned to forfeit his civil Personality. He must first be found deserving of punishment before any consideration is given to the utility of this punishment for himself or for his fellow citizens. The law concerning punishment is a categorical imperative, and woe to him who rummages around in the winding paths of a theory of happiness looking for some advantage to be gained by releasing the criminal from punishment or by reducing the amount of it. . . . If legal justice perishes, then it is no longer worthwhile for humans to remain alive on this earth. If this is so, what should one think of the proposal to permit a criminal who has been condemned to death to remain alive, if, after consenting to allow dangerous experiments to be made on him, he happily survives such experiments and if doctors thereby obtain new information that benefits the community? Any court of justice would turn down such a proposal with scorn . . . for justice ceases to be justice if it can be bought for a price.⁹⁸

A slightly later passage, however, qualifies this seeming rejection of consequentialism:

Anyone who is a murderer—that is, has committed a

98. Kant, *Metaphysical Elements*, *supra* note 82, at 138.

murder, commanded one, or taken part in one—must suffer death. This is what justice as the Idea of judicial authority wills in accordance with universal laws that are grounded a priori. The number of accomplices in such a deed might, however, be so large that the state would soon approach the condition of having no more subjects if it were to rid itself of these criminals, and this would lead to its dissolution and a return to the state of nature, which is much worse, because it would be a state of affairs without any external legal justice whatsoever. Since a sovereign will want to avoid such consequences and above all, will want to avoid adversely affecting the feelings of the people by the spectacle of such butchery, he must have it within his power in case of necessity to assume the role of judge and to pronounce a judgment that, instead of imposing the death penalty on the criminals, assigns some other punishment that will make the preservation of the mass of the people possible, such as, for example, deportation. Such a course of action would not come under a public law, but would be an executive decree, that is, an act based on the right of majesty, which, as an act of reprieve, can be exercised only in individual cases.⁹⁹

How can we reconcile these seemingly contradictory statements about consequentialism? We must distinguish among different types of consequentialism. The type of consequentialism that Kant regards as irrelevant to punishment is hedonism, or the “theory of happiness.” Generally speaking, happiness is not an ultimate, but only a contingent value, for Kant. What makes us happy depends on our interests and inclinations, but it is the nature of a moral agent to choose her ends, not to accept its own desires uncritically.¹⁰⁰ Thus Kant is more interested in securing the conditions for persons to freely choose their ends than he is in insuring their happiness. And to the extent that morality requires us to seek the happiness of others, it is in order to affirm their equal worth as choosers of their own ends, not because happiness is an intrinsic

99. *Id.* at 141.

100. Kant, *Groundwork*, *supra* note 81, at 36.

good.¹⁰¹ So Kant does not reject consequentialism, but in evaluating consequences, he completely prioritizes the goods of justice and autonomy over the good of happiness.¹⁰²

In the context of punishment, it is impossible to avoid considerations of happiness since, as we have seen, a penal sanction must, by definition, threaten to make the defendant worse off as a consequence of committing the crime. But Kant wants to avoid calculating the unhappiness of the person punished as a *cost*, since he conceives of that unhappiness as the frustration of an immoral desire to violate the autonomy of others. If a society allows the criminal to enjoy the benefits of his crime, it becomes an accomplice. Perhaps Kant also wishes to avoid the notion that punishment is a sort of therapeutic benefit. This idea might tempt us to coercively impose therapeutic treatment on those who have not consented to it and who have not yet earned coercive treatment by actually violating the law.

While society may not trade justice to achieve the end of happiness, it may apparently trade justice for justice. Thus, it may mitigate deserved punishment in order to preserve the legal order on which justice depends. This is exactly the choice that many societies have recently faced in trying to transform themselves from dictatorships into democracies. Dictatorial regimes often maintain themselves by systematizing atrocity, a practice that may render a very large proportion of the population complicit in crimes. Unless the bulk of these accomplices are amnestied, it may be impossible to win the population's assent to a just regime of prospective laws.¹⁰³ Kant so prioritized the authority of the legal order that he objected categorically to the trial of overthrown tyrants, on the

101. Walker, *supra* note 84, at 12-13.

102. Murphy, *supra* note 87, at 106-07.

103. For general discussions of this problem, see Pablo De Greiff, *Trial and Punishment, Pardon and Oblivion: On Two Inadequate Policies for the Treatment of Former Human Rights Abusers*, 22 *Phil. & Soc. Criticism* 93 (1996); Bruce A. Ackerman, *The Future Liberal Revolution* (1992); Ruti Teitel, *Transitional Justice* (2001).

grounds that all of their unjust actions were lawful at the time committed.¹⁰⁴ Society faces a more prosaic dilemma between competing claims of justice when it tries to induce underlings in a criminal organization to cooperate in providing evidence against their leaders. A measure of justice may have to be sacrificed in order that any justice may be done. Kant seems to accept this kind of trade off in the quoted passage about pardoning accomplices.

So when it comes to the important value of autonomy, Kant is willing to be very practical. The exercise of autonomy realistically depends upon the rule of law, preferably the rule of just law. Because the rule of law is merely a means to autonomy, it is more important to ensure its future survival than to insist on its retrospective vindication. Yet, Kant insists, while current justice may have to be compromised to insure the survival of just institutions, the decision to do so lies outside of those institutions. So Kant was a consequentialist about law, but apparently believed that consequentialism had no place within law.

Kant's insistence on the regularity and consistency of law as such, flows from his conception of law as an institution that actually universalizes principles of action. This is one of the differences between morality and justice. An act is moral if it is undertaken because of a principle of action that *can* be universalized, even if it is the only such act ever performed. An actor of bad character can still perform a moral act, because an act can be moral in isolation from other acts. By contrast, an act cannot be just in isolation, because justice is a value that inheres in regularity and systematicity. As Jeffrie Murphy explains,

[i]t is a necessary truth about the institutional rules prescribing punishment, and not merely a moral observation about them, that they should be justly enforced—that like cases be treated alike. For only in this way can law attain its primary social function: the control of social behavior through rules. . . . These rules can perform

104. Kant, *Metaphysical Elements*, *supra* note 82, at 128.

their socials function only if coupled with a justly enforced system of authoritative punishment.¹⁰⁵

That death is a fit penalty for murder (as Kant believed) does not make the execution of, say, only black murderers just, if equally guilty whites are spared.¹⁰⁶ This is why Kant could not have regarded even deserved punishment as just, if meted out on an ad hoc basis by vigilantes.¹⁰⁷ This requirement of regularity in the imposition of punishment helps us to resolve the puzzle raised earlier about the extent to which Kant required that punishment actually deter. To count as *just*, punishment must be successfully inflicted in response to every crime, and must always deprive the criminal entirely of the benefit of the crime. This means that, by definition, a *just* regime of punishment always suffices to deter a rational actor. In this way, just punishment ideally deters criminal acts and effectually protects the equal autonomy of potential victims. Just coercion depends on stable, regular legal institutions, which render spheres of autonomous choice real, by reliably protecting them.¹⁰⁸

Thus far, we have seen that to qualify as just, punishment must be imposed by law because of the freely chosen violation of a morally obligatory norm of conduct. Such punishment must be imposed systematically and equally on all similarly situated actors. The punishment must be proportionate to the crime—harsh enough to deprive the criminal of the benefit of the crime, but no harsher than the evil inflicted by the crime.¹⁰⁹ There is, in

105. Murphy, *supra* note 87, at 117.

106. See *Furman v. Georgia*, 408 U.S. 238 (1972) (arbitrary and capricious capital sentencing scheme risks racial discrimination); *McCleskey v. Kemp*, 481 U.S. 279 (1987) (dismissing challenge to capital sentencing scheme based on statistical evidence of discrimination on the basis of race of the victim).

107. Murphy, *supra* note 87, at 117.

108. Rosen, *supra* note 86, at 89-91.

109. Kant, *Metaphysical Elements*, *supra* note 82, at 172.

The only time a criminal cannot complain that he is treated unjustly is when he draws the evil deed back onto himself, and when he suffers that which according to the spirit of the penal law—even if not to the letter thereof—is the same as what he has inflicted on others.

addition to these conditions, one more requisite of just punishment. In order for coercion to be just, the person suffering coercion must have consented to the law imposing it. This requirement of consent follows from the moral principle that every person should be treated equally, as an end—that is, as an autonomous person who cannot be coerced in her moral choices.¹¹⁰ However, such a requirement raises the difficulty that citizens do not personally consent to any laws. Kant's solution to this problem is to substitute the collective consent of majoritarian democracy for personal consent:

The legislative authority can be attributed only to the united Will of the people. Since all of justice is supposed to proceed from this authority, it can do absolutely no injustice to anyone. Now, when someone orders something against another, it is always possible that he thereby does another an injustice, but this is never possible with respect to what one decides for oneself (for *volenti non fit injuria*). Hence, only the united and consenting Will of all—that is, a general united Will of the people by which each decides the same for all and all decide the same for each—can legislate. The members of such a society (*societas civilis*), that is, of a state, who are united for the purpose of making laws are called *citizens* (*cives*). There are three juridical attributes inseparably bound up with the nature of a citizen as such: first, the lawful freedom to obey no law other than that to which he has given his consent; second, the civil equality of having among the people no superior over him except another person whom he has just as much of a moral capacity to bind juridically as the other to bind him; third, the attribute of civil self-sufficiency that requires that he owe his existence and support, not to the arbitrary will of another person in the society, but rather to his own rights and powers as a member of the commonwealth. . . .¹¹¹

Of course the collective consent of the citizenry seems imperfect in several ways. First, and most obviously, the

110. Rosen, *supra* note 86, at 14, 56-57, 62-63.

111. Kant, *Metaphysical Elements*, *supra* note 82, at 119, 120.

result of an election is preferred only by those in the majority. A majoritarian process of decision does not yield consensus. Second, in a large and complex society, citizens cannot be expected to inform themselves about and personally vote on every law. As a result, Kant proposed a representative democracy, in which voters would only exercise consent vicariously, through representatives, who cannot possibly follow the inevitably inconsistent wishes of all of their supporters on each and every issue. Third, potential voters are not always in a position to choose autonomously. Kant recognized the impingement of circumstance on consent by limiting the franchise to citizens who were materially independent.¹¹² Like many seventeenth and eighteenth century political thinkers, Kant assumed that granting the vote to servants, tenant farmers, and married women would only increase the influence of their powerful employers, landlords, and husbands.¹¹³ Of course the result was that some persons subject to the criminal laws would not—in Kant's view could not—consent to them.

These limitations on the practicability of consent to the criminal laws reveal that even for Kant, the legitimacy of punishment was a matter of degree. Punishment can be morally deserved, but it cannot be entirely just, because the institution inflicting it is ultimately irreconcilable with the moral autonomy of those on whom it is inflicted. In this sense, retributive justice is a practical end that can be achieved more or less well, and the degree of success depends not only on whether a particular sanction is deserved, but also on the extent to which the law imposing it is democratically legitimate.¹¹⁴

Given Kant's emphasis on individual autonomy, why does he accept majority consent as a substitute for individual consent?¹¹⁵ And why is he satisfied with an approximation of retributive justice? I think the answers

112. *Id.* at 120-21.

113. Rosen, *supra* note 86, at 14, 35-36.

114. *Id.* at 57.

115. *Id.* at 49.

to these important questions are to be found in his conception of moral imperatives as cooperative institutions. We have already seen that Kant views legal institutions like property as core examples of such cooperative institutions. For Kant, all our efforts to pursue our chosen ends depend upon a sphere of freedom of action protected by law. Without freedom from violence and secure possession of the tools and resources we require, such planning and striving would be futile. To pursue almost any practical end is therefore to take advantage of the security provided by law, and to imply assent to law.¹¹⁶ Thus, to refuse assent to the criminal laws is almost always to exploit their benefits and so to act according to a non-universalizable principle of action, thereby violating a moral imperative.¹¹⁷ As far as Kant was concerned, there was no practical alternative to a legal order: he agreed with Hobbes that banditry or civil war would surely involve greater inhibitions on autonomy than authoritarian rule.

Thus far, Kant has demonstrated a moral obligation to support law in general, but not a moral obligation to respect any particular laws or legal regime. Yet, Kant insisted that citizens have a moral obligation to obey and defend any legal order they find themselves subject to, whether or not it is just.¹¹⁸ The next premise in Kant's argument is that freedom from coercion is necessarily a political achievement, not available in a state of nature. In such a condition, individuals might find respite from coercion by hiding from their fellows, but the legal "right" to be free from coercion can only be available within a legal system.¹¹⁹ Thus, a complaint of unjust coercion address an organized political community, and appeals to a set of authoritative norms. Such a complaint therefore presupposes the authority of some actual legal system, and so assent to the authority of such a legal system is logically

116. Rosen, *supra* note 86, at 124-25, 128.

117. Murphy, *supra* note 87, at 135.

118. Kant, *Metaphysical Elements*, *supra* note 82, at 124-25; Rosen, *supra* note 86, at 19, 119.

119. Rosen, *supra* note 86, at 10.

necessary to avoid self-contradiction. Moreover, this legal system must be one recognized by the audience being addressed. One difference between law and morality is that law is systematic. An individual can subject herself to moral principles on her own, even if no one joins her. But because law is an institution, an individual can assent to law only by joining with others: the act of making law is inherently a collective, political act. Accordingly, the obligation to support law implies the obligation to cooperate with others in supporting some particular legal system.

So Kant was—like Bentham—an institutional positivist, who considered the concept of consent virtually meaningless outside of some kind of legally authorized procedure for expressing consent. On such reasoning, to express political consent or opposition to a particular law or legal regime an individual must be constituted with others, *by law*, as citizens of a polity.¹²⁰ No individual can make, alter, or overthrow law on her own. A law is not merely an opinion. Indeed, no group can legislate simply by expressing a preference, because laws can only be made through an authoritative procedure. Laws become part of a system of rules joined by what H.L.A. Hart called a rule of recognition.¹²¹ On this reasoning, the act of lawmaking always presupposes preexisting legal institutions. Accordingly, Kant would argue that to advocate or support any particular legal regime is to accept the idea that opinions and preferences only become legally authoritative by legal means. So one may advocate for legal reform, but in so doing one must accept the authority of the existing legal system.¹²²

This explains why majoritarian democracy satisfies the requirement of consent. When we vote, we join a game bound by rules. In casting our votes we do not merely express opinions, as do participants in an opinion poll. We carry out an official lawmaking function within a legal

120. *Id.* at 123, 181-82.

121. H.L.A. Hart, *The Concept of Law* (1961).

122. Kant, *Metaphysical Elements*, *supra* note 82, at 182.

system. In so doing, we expect and intend that if our views are shared by a majority of those voting, they will prevail and be enacted into laws that will be enforced coercively. But this intention implies that if some other view is preferred by a majority, that view will be enacted into law. Accordingly, those who vote do actually assent to the results of the election and to the authority of the laws ultimately enacted. In other words, when we vote, we cooperate with others in achieving a public good, by constituting an authoritative lawmaking body with the power to coerce our fellows—and ourselves. We don't simply express a preference about what law should be made, we assume the power to make law.

What about those who choose not to vote? While voting clearly signifies assent to law, nonvoting does not necessarily imply repudiation of the rule of law: voluntary nonvoters are usually just delegating their lawmaking discretion to their peers. This makes perfect sense, given Kant's view that there is no rational alternative to assenting to law. So as long as the franchise and the conditions for its autonomous exercise were made available to all adults, Kant was willing to ascribe to those adults consent to the resulting laws. The important consideration, then, is not whether the entire adult population actually supported each law, but whether the entire adult population had the opportunity to share in lawmaking.¹²³ If they were thus enfranchised as members of the sovereign lawmaking body, the resulting law is self-imposed.

It may seem that Kant would need to make an exception for laws that, although supported by majority will, systematically oppress discrete minorities. Yet recall that consent is only one of Kant's criteria for a justly enforceable law. The law must be a moral prescription, requiring universal compliance with a cooperative institution from which all benefit, and the punishment must be proportionate. So a justly enforceable law must be fair as well as democratically legitimate, and we need only

123. Rosen, *supra* note 86, at 37.

inquire into the consensual basis of laws that meet this requirement of fairness. Justice does not leave democracy unrestricted, but requires that the democratic sovereign act only through generally applicable and fairly administered legal rules.¹²⁴ Kant viewed the resulting constitutionally limited democracy as a "republic," in which a democratically representative legislative power was separated from an independent judiciary and executive.¹²⁵

According to Kant, persons have a moral obligation to consent to law. In a tyrannical regime, however, the law can require persons to act immorally. Only a regime of moral laws can reconcile the citizen's obligations to obey law and act morally. But if such a regime is not also democratic, the citizens' obedience to morality is coerced rather than freely motivated by a good will. Only democracy grants citizens the freedom to choose and assent to moral laws. Democracy does not guarantee just laws, because democratic majorities are also free to pass immoral laws. But democracy is a necessary condition to legal justice, as Kant conceives it.

IV. COMPARING UTILITARIAN AND RETRIBUTIVIST MODELS OF POLITICAL LEGITIMACY

Most criminal law theory treats normative questions about the imposition of punishment as moral questions. Accordingly, criminal law theory has been organized as a debate between utilitarian and deontological ethics. But these moral conceptions of punishment have paradoxical implications. Thus, an act utilitarian conception of punishment seems to endorse framing the innocent, while a deontological conception of punishment seems to endorse vigilante justice. I have therefore argued that punishment should be seen as an institution rather than a behavior and should be evaluated politically rather than ethically. In keeping with this program, I have reinterpreted the two

124. *Id.* at 50.

125. Rosen, *supra* note 86, at 34.

main traditions of normative theorizing about punishment as theories about politically legitimate institutional action rather than as theories about morally correct individual action.

I have argued that we should not view the "utility principle"—which tests actions by their consequences for the public welfare—as a moral test for the choices of individuals. The founders of utilitarian penology, Beccaria and Bentham, did not use the utility principle in this way. They were not interested in preaching morality to individuals, whom they saw as inevitably and appropriately self-serving. Thus, they did not propose utility as a secret test to be used in the private forum of conscience. Instead they offered it as a public criterion, to guide political debate about laws in a new era of democratic revolution. In order to understand Bentham's conception of utility, it is necessary to grasp two essential points. First, he did not propose this criterion in isolation. He proposed it as part of a comprehensive new decision-making technology, involving bureaucratic investigation, democratic oversight and participation, legislative supremacy, and a clear and common language of policy analysis and legislation. Second, Bentham gave a particular content to the hedonistic element in his utility criterion. He did not mean by happiness simply satisfaction of desire or unrestricted choice. Instead, he meant freedom from fear and anxiety, and the secure foreknowledge that one had the means to meet one's future needs. This meant creating a powerful, but democratically controlled and rule-bound welfare state. Despite the fact that Bentham saw each individual as the best judge of her own happiness, he did not really leave people free to define and pursue their own happiness as individuals. Instead, he compelled them to participate in a collective, political process of defining and securing the public welfare. This public welfare was not simply there to be discovered, a matter of adding up private wants. Instead, it was a political construct, to be developed by following a certain institutional process. It inhered not just in the

consequences of the policies chosen, but in the means by which those policies would be chosen and implemented.

Utilitarian penology is premised on this model of utility as an institutionally defined and legally implemented conception of the public welfare. Accordingly, utilitarian punishment depends on a legislatively codified body of specific and prospective criminal proscriptions, and a regime of legally regulated and publicly monitored procedures for imposing and inflicting punishment. A utilitarian system of punishment would certainly neither mandate nor permit the framing of the innocent. It is conceivable that such a system could make use of other consequentialist policies criticized by retributivists, such as preventive detention or other decisions conditioning harsh treatment on judgments of future dangerousness. But these kinds of policies are far less attractive from the utilitarian standpoint than commonly supposed. Insofar as they rely on discretionary, speculative judgments by public officials, they are incompatible with the severe limits on administrative discretion required in the utilitarian policy process. In general, critics of utilitarianism have greatly exaggerated its totalitarian potential because they have assumed it is concerned only with the threats individuals might pose to the public's security. They have not realized the extent to which utilitarianism is concerned with controlling the possibly much greater threat that unfettered state officials might pose. In so doing, critics have forgotten utilitarianism's historical origins in an age of liberal democratic revolt against absolutist monarchy. Utilitarian penology began as a movement to reduce the harshness and arbitrariness of punishment. As such, it was part of a larger project of legitimating state force by bringing it under democratic and constitutional control.

Kantian retributivism was born in the same era. Like Bentham, Kant considered punishment as a legal institution rather than an individual behavior. Punishment was a form of legal coercion, an infringement of moral autonomy. As such, it was justifiable only as a collectively self-imposed means of securing to everyone a

limited, but fair scope for the exercise of autonomy. Like Bentham, Kant conceived the problem of punishment as a problem of political theory, rather than a moral problem. For Kant, justice was an evaluative criterion that presupposed a legal system. As a result, Kant's conception of justice combined considerations of morality and legitimacy. A law met the moral test if it used or threatened to use coercive force against persons committing immoral actions—actions that hypocritically exploited a cooperative institution to achieve an unfair advantage. But the fact that a law's proscriptions conformed to Kantian morality did not suffice to make them just. If the criminal laws did not effectively constrain the immoral behavior they forbade, they did not achieve justice. While good intentions may suffice to make actions moral, Kantian justice depends on results. But even moral and effectual laws are not just unless they are also accepted by those subject to them. Thus Kantian justice depends on his conception of freedom as submission to a self-imposed law. A law meets Kant's collective consent test fully if it is made by a representative legislature, freely chosen by a majority of those choosing to vote, among an electorate consisting of all adults subject to the law. An electoral choice counts as free only if it is based on adequate information and was not constrained by material dependence. Kant recognized that this kind of democratic legitimacy was a matter of degree, and held that citizens were obliged to obey laws and uphold legal systems that were partially, or even wholly illegitimate.

While Kant and Bentham both conditioned the legitimacy of state force on democratic lawmaking, they disagreed about the features of state force that needed legitimation. For Bentham, it was the state's infliction of pain or unhappiness. For Kant, it was the state's use of coercion, which infringed the individual's autonomy. For Bentham, the value of utility was the key to both the problem of legitimacy, and its solution. For Kant, autonomy was the master value that defined and solved the problem. But these two concepts have considerable

overlap. For Bentham, utility was not just the fulfillment of desire. It included the assurance of future satisfaction derived from having options secured. Since utility required the expectation of having choices, it included a kind of autonomy. On the other hand, Kantian autonomy turned out to include consequentialist and hedonistic elements that we associate with utility. Thus Kantian autonomy depends not just on having a choice, but on satisfaction in the choice. That is why a coercive threat, by depriving an actor of satisfaction in a chosen course of action, infringes her autonomy. Thus, autonomy depends not just on having options, but on the concrete possibility of fulfilling desires by choosing those options. To the extent that Benthamite utility requires the security provided by having certain choices, while Kantian autonomy includes a measure of satisfaction in the choices made, the two concepts are not as different as they might seem.

On a conventional understanding, the concepts of utility and autonomy confer legitimacy in quite different and even incompatible ways. Thus, it is commonly assumed that only the consequences of policy matter on a utility analysis and that only the origin of policy matters on an autonomy analysis. Thus, we tend to assume that a utilitarian analysis imposes no a priori constraints on government, making legitimacy entirely contingent on what citizens actually desire. Conversely, we generally assume that an autonomy analysis is entirely a matter of such a priori constraints, and that it ignores consequences.

In fact, however, we have seen that neither assumption is correct. Bentham's peculiar conception of utility as security, and his use of this conception as a public standard for settling political disagreements, in effect converted utility into a process value. A utilitarian policy had to arise as a result of a particular kind of political process and had to preserve the requisite conditions for that process. It follows that this utilitarian policy process imposes many a priori constraints on utilitarian policy. The result is that individuals are guaranteed an elaborate

regime of rights involving free speech,¹²⁶ political participation, and protections against arbitrary or excessive punishment. On the other hand, we have seen that Kant deployed his conception of autonomy not just as a procedural constraint on the making of policy, but also as a desired end. Thus, even if criminal law were to infringe unjustly the autonomy of those it threatened and punished, it would still be necessary to prevent the greater infringements of autonomy that would result from a state of anarchy. And a criminal law that failed to influence behavior would fail to fulfill its proper function of preventing these violations of autonomy. Accordingly, Kant can fairly be categorized as a rights-consequentialist, who judges legal systems by their success in securing to each person, her fair share of autonomy.

Thus Bentham and Kant both assess legal force generally, and criminal punishment in particular, as political institutions rather than moral acts. Both require that law secure to each citizen a sphere of freedom from both private and public interference and both require that legal force be democratically legitimate. Yet Bentham and Kant required these broadly similar conditions for different reasons. Bentham hoped that democratic deliberation would be guided by the public welfare, a complex good that includes security rights and process values, as well as collective wealth. Kant hoped democratic deliberation would be guided by a morality of cooperating in the achievement of public goods and an ideal of equal autonomy. It seems to me that there is likely to be a good deal of overlap in the lawmaking procedures endorsed by these two models of political legitimacy, and a good deal of common ground between these two substantive criteria of political value. Accordingly, there is reason to hope that debate about utility and autonomy in criminal lawmaking will become more productive once it is redefined as a political debate about institutions rather than a moral debate about the conduct of criminals and officials.

126. Mary Peter Mack, *Jeremy Bentham: An Odyssey of Ideas* 315 (1963).