

## Joel Feinberg on Crime and Punishment: Exploring the Relationship Between *The Moral Limits of the Criminal Law* and *The Expressive Function of Punishment*

Bernard E. Harcourt\*

When I was originally approached to participate in this Symposium on the work and legacy of Joel Feinberg, I immediately began thinking about the influence of his essay *The Expressive Function of Punishment*<sup>1</sup> on contemporary criminal law theory in the United States. That essay has contributed significantly to a growing body of scholarship associated with the resurgence of interest in expressive theories of law.<sup>2</sup> In the criminal law area, the expressivist movement traces directly and foremost to Feinberg's essay. As Carol Steiker observes, "Joel Feinberg

---

\* Associate Professor of Law, University of Arizona; Visiting Professor of Law, Harvard University. Special thanks to Stuart Green for organizing this Symposium, and to Tom Christiano, Joel Feinberg, Toni Massaro, Carol Steiker, Bill Stuntz, and Lloyd Weinreb for discussion of this essay. This paper was originally presented at the panel on *The Moral Limits of the Criminal Law* at the AALS Annual Meeting in San Francisco on January 5, 2001. I thank the other participants on the panel—Jules Coleman, George Fletcher, and Stuart Green—for their reactions.

1. Joel Feinberg, *The Expressive Function of Punishment*, in *Doing and Deserving* 95-118 (1970) [hereinafter *The Expressive Function*].

2. For general accounts of expressive theories of law, see Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. Pa. L. Rev. 1503 (2000); Richard H. McAdams, *A Focal Point Theory of Expressive Law*, 86 Va. L. Rev. 1649 (2000); Richard H. Pildes, *Why Rights Are Not Trumps: Social Meanings, Expressive Harms, and Constitutionalism*, 28 J. Legal Stud. 725 (1998); Dan M. Kahan, *What Do Alternative Sanctions Mean?* 63 U. Chi. L. Rev. 591 (1996); Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. Pa. L. Rev. 2021 (1996); Elizabeth Anderson, *Value in Ethics and Economics* 17-43 (1993). For instances of expressive theories in other domains of contemporary law, see Deborah Hellman, *The Expressive Dimension of Equal Protection*, 85 Minn. L. Rev. 1 (2000) (equal protection context); Lawrence Lessig, *The Regulation of Social Meaning*, 62 U. Chi. L. Rev. 943 (1995) (First Amendment context); Richard H. Pildes & Richard G. Niemi, *Expressive Harms, "Bizarre Districts," and Voting Rights: Evaluating Election-District Appearances After Shaw v. Reno*, 92 Mich. L. Rev. 483 (1993) (voting rights context).

can be credited with inaugurating the “expressivist” turn in punishment theory with his influential essay, *The Expressive Function of Punishment*.<sup>3</sup> Matthew Adler, who offers a skeptical overview of expressive theories, similarly traces the movement back to Joel Feinberg:

The work of Professors Pildes, Kahan, and Sunstein . . . has given renewed salience and currency to expressive theories of law. But it bears emphasis that their scholarship is simply the most recent contribution to a much older and larger body of scholarly writing about the symbolic cast of legal decisions. For example, students of the criminal law have long debated the expressive dimension of punishment. The famous legal philosopher Joel Feinberg, in a 1965 article entitled *The Expressive Function of Punishment*, rejected the then standard definition of punishment as “the infliction of hard treatment by an authority on a person for his prior failing in some respect,” and asserted by contrast that punishment was essentially expressive—that it necessarily had a “*symbolic significance* largely missing from other kinds of penalties.” . . . Feinberg’s article touched off a still-flourishing debate within criminal law scholarship, prompting rebuttals by (among others) C.L. Ten, Michael Moore, and Michael Davis, and defenses by (among others) Robert Nozick, Jean Hampton, Igor Primoratz, Anthony Duff, and, now, Professor Kahan.<sup>4</sup>

I soon learned, though, that the intended focus of this Symposium is on Feinberg’s four-volume treatise *The Moral Limits of the Criminal Law* and that the

---

3. Carol Steiker, Foreword: Punishment and Procedure: Punishment Theory and the Criminal-Civil Procedural Divide, 85 Geo. L.J. 775, 803 (1997); see also Dan M. Kahan, The Secret Ambition of Deterrence, 113 Harv. L. Rev. 413, 419-20 n.16 (1999) (tracing expressive theories of punishment to Joel Feinberg’s work) [hereinafter Kahan, Deterrence]; Dan M. Kahan, What Do Alternative Sanctions Mean? 63 U. Chi. L. Rev. 591, 595 (1996) (same) [hereinafter Kahan, Alternative Sanctions]; Dan M. Kahan, Is Gun Control Illiberal? (Jan. 26-27, 2001) (unpublished paper presented at the Conference on Guns, Crime, and Punishment, University of Arizona) (on file with author) (same) [hereinafter Kahan, Gun Control].

4. Matthew D. Adler, Expressive Theories of Law: A Skeptical Overview, 148 U. Pa. L. Rev. 1363, 1369-70 (2000).

contributors are primarily moral and legal philosophers interested in the implications of *The Moral Limits*. The Symposium is intended to throw light on a remarkable body of work that has not received the attention that it may deserve in criminal law circles in the United States. The companion panel at the annual meeting of the Association of American Law Schools in January 2001 was motivated by a similar problematic. "Despite the great influence [*The Moral Limits of the Criminal Law*] has achieved among philosophers," the conference program announced, "*Moral Limits* has yet to attain the prominence that it deserves among legal academics and lawyers."<sup>5</sup>

After a pause and slight double take, I realized that this reflects a puzzling acoustic separation between *The Moral Limits* and Feinberg's writings on the expressive function of punishment—an acoustic separation in Feinberg's work, but also in the debates that his work has spawned. *The Moral Limits* itself does not develop the expressive claim, and, more generally, does not address theories of punishment in a systematic way. Although Feinberg made a few passing references to his earlier essay on the expressive function of punishment in *The Moral Limits*,<sup>6</sup> and although Feinberg on a few occasions addressed other theories of punishment in *The Moral Limits*,<sup>7</sup> he did not elaborate, nor fully integrate his expressive theory of punishment in his moral writings. At the same time, contemporary expressivist theorists in the criminal law area tend not to engage Feinberg's treatise on

---

5. Association of American Law Schools, Pursuing Equal Justice: Law Schools and the Provision of Legal Services, 2001 Annual Meeting 59 (2001).

6. See, e.g., Joel Feinberg, Harm to Others 24, 248 n.20 (noting in passing that the punishment of imprisonment expresses censure and condemnation) [hereinafter Harm to Others]; Joel Feinberg, Harmless Wrongdoing at 149, 354 n.42 (discussing the relevance of moral blameworthiness to the expressive function of punishment) [hereinafter Harmless Wrongdoing]; id. at 295-303, 368 nn.21, 35 (discussing the relationship between the moral education theory of punishment and the expressive theory of punishment).

7. See, e.g., Harmless Wrongdoing, supra note 6, at 159-65 (discussing the retributive theory of punishment); id. at 300-05 (discussing the moral education theory of punishment).

moral limits.<sup>8</sup> Similarly, most of the scholarship surrounding *The Moral Limits* does not concern itself with Feinberg's expressive theory of punishment.<sup>9</sup>

This is odd. One would expect, after all, that a four-volume treatise on the moral bounds of the criminal sanction would include, at its heart, discussion of theories of punishment. Several commentators have expressed similar astonishment at this apparent acoustic separation. "It is intriguing that in the first three volumes of Feinberg's work there is no mention of the retributive function of punishment, and only a brief discussion of it in Volume IV," Jean Hampton remarks. "Given that generations of legal theorists have defended this conception of criminal punishment, why didn't Feinberg emphasize it in his discussion of the role and purposes of criminal legislation?"<sup>10</sup> Similarly, in his contribution to this Symposium, Hugo Bedau writes, "[o]ne might well have expected a more focussed discussion of punishment in any thorough account of 'the moral limits of the criminal law,' since some of the most conspicuous of those limits are to be found in the principles governing a system of fair and reasonable punishment."<sup>11</sup> Yet, *The Moral Limits* simply does not do that. It does not elaborate Feinberg's central insight about the expressive function of punishment, nor does it integrate punishment theory into the moral analysis.

This raises an interesting set of questions concerning the classic debate over the relationship between the moral limits of the criminal law and theories of punishment—the relationship between what H.L.A. Hart referred to as

---

8. See, e.g., Kahan, Deterrence, *supra* note 3, at 419-20.

9. For a good sample, see the following papers in this Symposium: R.A. Duff, Harms and Wrongs, 5 *Buff. Crim L. Rev.* 13 (2001); Hamish Stewart, Harms, Wrongs, and Set-Backs in Feinberg's Moral Limits of Criminal Law, 5 *Buff. Crim L. Rev.* 47 (2001).

10. Jean Hampton, Liberalism, Retribution and Criminality, in *In Harm's Way: Essays in honor of Joel Feinberg* 54 (Jules L. Coleman & Allen Buchanan eds., 1994).

11. Hugo Adam Bedau, Feinberg's Liberal Theory of Punishment, 5 *Buff. Crim L. Rev.* 103 (2001).

“primary laws setting standards for behavior and secondary laws specifying what officials must or may do when they are broken.”<sup>12</sup> What *is* the relationship in Joel Feinberg’s work between his moral-legal theory and his expressive theory of punishment? Specifically, what is the connection between Feinberg’s treatise on *The Moral Limits of the Criminal Law* and his essay *The Expressive Function of Punishment*? More generally, what *should* be the relationship between analyzing the moral limits of the criminal law and exploring the contours of punishment?

### I. FRAMING THE QUESTION

There are four broad categories of possible answers to the more general question. The first possible response is that there is no necessary relationship at all. Under this first approach, the argument would be that the moral limits of the criminal law do not help us define the proper type or amount of punishment, or even the circumstances that call for actual punishment—and vice versa. The issue of what conduct to criminalize does not help us resolve when, how and in what amount to punish. Criminalization is simply a different issue than enforcement and punishment: Defining the moral contours of the criminal law is one problem, figuring out whether to enforce the criminal law and how much to punish is another matter.

Jean Hampton’s discussion in her essay *The Moral Education Theory of Punishment* is a good illustration of this first category of possible responses. Hampton argues that punishment should be intended to teach the wrongdoer that her action is morally wrong and should not be repeated. In essence, the moral education theory of punishment “maintains that punishment is justified as a way to prevent wrongdoing insofar as it can teach both wrongdoers and the public at large the moral reasons for *choosing* not to perform an offense.”<sup>13</sup> Hampton defends the

---

12. See, e.g., H.L.A. Hart, Prolegomenon to the Principles of Punishment, in *Punishment and Responsibility* 7 (1968).

13. Jean Hampton, *The Moral Education Theory of Punishment*, 13 *Phil. &*

theory as a full and complete justification of punishment and discusses its implications as a punishment formula.

In her discussion, Hampton defends against the charge that her moral education theory would support legal paternalist criminal laws, namely criminal laws that restrict what a person could do to herself as opposed to others. Hampton's defense draws a sharp distinction between the definition of criminal law and punishment theory. The moral education theory of punishment, Hampton contends, "can give no answer" to the questions "what ought to be made law?" or "what is the appropriate area for legislation?" It can give no answer "for while the theory maintains that punishment of a certain sort should follow the transgression of a law, it is no part of the theory to say *what* ethical reasons warrant the imposition of a law."<sup>14</sup> In fact, Hampton emphasizes, this is one of its virtues. "Indeed, one of the advantages of the theory is that one can adopt it no matter what position one occupies on the political spectrum."<sup>15</sup> From Hampton's perspective, one set of normative values may define punishment, another the criminal law.

H.L.A. Hart was, of course, a strong proponent of distinguishing between the two questions,<sup>16</sup> and, in *Law, Liberty, and Morality*, he too articulated one version of this first category of responses. His discussion occurs in a passage responding to one of James Fitzjames Stephen's numerous arguments for legal moralism. In *Liberty, Equality, Fraternity*, Stephen had argued that one proof

---

Pub. Aff., Summer 1984, at 208, 213.

14. *Id.* at 219.

15. *Id.*

16. See, e.g., Hart, *supra* note 12, at 6-8; *id.* at 6 (arguing that some "curious" theories of punishment "gain their only plausibility from ignoring" this distinction); *id.* at 8:

[T]he immediate aim of criminal legislation cannot be any of the things which are usually mentioned as justifying punishment: for until it is settled what conduct is to be legally denounced and discouraged we have not settled from what we are to *deter* people, or who are to be considered *criminals* from whom we are to exact *retribution*, or on whom we are to wreak *vengeance*, or whom we are to *reform*. . . .

## 2001] FEINBERG ON CRIME &amp; PUNISHMENT 151

that the criminal law is in fact based on legal moralist principles is that punishment is often tied to moral wickedness. The fact that judges traditionally focus on moral culpability in sentencing proves, Stephen argued, that the criminal law is actually aimed at moral wrongdoing—that the criminal law represents, in Stephen's words, the "persecution of the grosser forms of vice."<sup>17</sup>

In his lectures, Hart responds to this argument by drawing a clear distinction between the moral limits of criminal law and theories of punishment: "Surely this argument is a *non sequitur* generated by Stephen's failure to see that the questions 'What sort of conduct may justifiably be punished?' and 'How severely should we punish different offenses?' are distinct and independent questions."<sup>18</sup> Hart argues that it is not at all inconsistent to justify the criminal law on the basis of the harm principle but to evaluate moral culpability for purposes of measuring punishment. He explains:

It is in general true that we cannot infer from principles applied in deciding the severity of punishment what the aims of the system of punishment are or what sorts of conduct may justifiably be punished. For some of these principles, e.g., the exclusion of torture or cruel punishments, may represent other values with which we may wish to compromise, and our compromise with them may restrict the extent to which we pursue the main values which justify punishment. So if in the course of punishing only harmful activities we think it right . . . to mark moral differences between different offenders, this does not show that we must also think it right to punish activities which are not harmful.<sup>19</sup>

In essence, Hart argues, the two inquiries are different and may call for reliance on different moral principles.

A second category of possible answers is that both the moral limits of the criminal law and punishment theory

---

17. James Fitzjames Stephen, *Liberty, Equality, Fraternity* 162 (1873).

18. H.L.A. Hart, *Law, Liberty, and Morality* 36 (1963).

19. *Id.* at 37-38.

should be guided by some higher principle, whether it be justice, fairness, efficiency, or something else. They are both derivative of a moral, social, economic, or political theory that is overarching and controlling. Since they derive from the same theoretical framework, they are closely interrelated. They may be identical. They may be parallel. In either event, they are essentially fungible: If you know how one works, you can easily figure out how the other works.

Richard Posner's economic model of crime is a good illustration of this second category of responses. According to Posner, both the substantive doctrines of the criminal law (including the definition of crimes) and modes of punishment (fines, imprisonment, etc.) can be interpreted to—and in fact do—promote efficiency. “The major function of criminal law in a capitalist society,” Posner suggests, “is to prevent people from bypassing the system of voluntary, compensated exchange—the ‘market,’ explicit or implicit—in situations where, because transaction costs are low, the market is a more efficient method of allocating resources than forced exchange.”<sup>20</sup> Criminal law can be interpreted as a vehicle to deter market bypassing where other tools, such as tort law, do not work. Similarly, punishments, such as the death penalty, imprisonment, and fines, make economic sense insofar as they too can be interpreted as promoting efficiency. So, for instance, Posner suggests that it is “efficient to use different sanctions depending on an offender's wealth.”<sup>21</sup> Imprisonment is designed primarily for the non-affluent, who would not be deterred by tort law (since they are judgment proof). The affluent, in contrast, may be deterred by traditional tort remedies, and accordingly it makes economic sense to use these methods. Capital punishment, under Posner's view, makes sense as a

---

20. Richard A. Posner, *An Economic Theory of the Criminal Law*, 85 *Colum. L. Rev.* 1193, 1195 (1985) [hereinafter Posner, *Economic Theory*]; see also Richard A. Posner, *Economic Analysis of Law* ch. 7 (2d ed. 1977) [hereinafter Posner, *Economic Analysis*]; Robert D. Cooter & Thomas S. Ullen, *Law and Economics* 387-407 (2d ed. 1997).

21. Posner, *Economic Theory*, *supra* note 20, at 1205.

matter of marginal deterrence: It serves to deter, for example, prison murders. These aggravated murders would not be deterred if they were punished in the same way as other severe crimes that result in incarceration. In sum, most of the definitions of crimes, substantive criminal law doctrines, and methods of punishment “can be explained as if the objective . . . were to promote economic efficiency.”<sup>22</sup>

Although Posner explicitly denies that efficiency should be the *only* social value that courts and legislators promote through the criminal law, there is a normative dimension to his argument. Economic efficiency can be deployed as one among other justifications for limiting the criminal law and for measuring punishment. In this sense, Posner’s theory is an interesting mixture of descriptive and normative claims. It seeks both to explain the economic basis of the criminal law and punishment, and to suggest how the criminal law and punishment can be designed to promote the value of efficiency.

Variations of this second category of possible responses come in both purely normative and purely descriptive versions. Among the latter, an interesting variant is the theory of Georg Rusche and Otto Kirchheimer in *Punishment and Social Structure*—particularly interesting in its relation to Posner’s theory. Rusche and Kirschheimer’s thesis, stripped of much nuance, is that the combination of population demographics and labor markets significantly influences the way that social institutions criminalize and punish certain classes in society. The definition of crimes (for example, vagrancy, begging, or being a vagabond) and the modes of punishment (for example, corporal punishment, galley slavery, transportation, houses of correction, prison labor, or solitary confinement) are both heavily shaped by modes of production and labor supply. In fact, the history of penal legislation and punitive practices—as well as the contemporary criminal justice system—on their view,

---

22. *Id.* at 1195.

reflects the interplay of economic forces, or, more specifically, the class struggle.

Rusche and Kirchheimer's approach, in their own words, breaks the bond between crime and punishment in order to view the criminal sanction as a social phenomenon fundamentally shaped by economic forces. They wrote:

Every system of production tends to discover punishments which correspond to its productive relationships. It is thus necessary to investigate the origin and fate of penal systems, the use or avoidance of specific punishments, and the intensity of penal practices as they are determined by social forces, above all by economic and then fiscal forces.<sup>23</sup>

These economic and fiscal forces shape both the definition of crime and modes of punishment implemented during different historical periods. The net effect is that both inquiries—into the definition of crime and theories of punishment—are answered by one theoretical intervention.

A third possible set of answers would be that punishment theory is in some sense primary and should dictate or help define the moral limits of the criminal law. Under this third approach, what matters foremost is choosing the right punishment theory; the moral limits will follow. Hugo Bedau gives us a taste of this approach in his contribution to this Symposium when he suggests that “the most conspicuous of those [moral] limits [to the criminal law] are to be found in the principles governing a system of fair and reasonable punishment.”<sup>24</sup>

Jules Coleman, in his presentation at the AALS in January 2001, offered one variant of this third approach. Coleman criticized Feinberg for failing to first develop a theory of punishment in his analysis of the moral limits of the criminal law. Coleman argued that Feinberg had approached the very problem of the moral limits from the wrong vantage point—namely, by assuming a presumption

---

23. Georg Rusche & Otto Kirchheimer, *Punishment and Social Structure* 5 (1939).

24. Bedau, *supra* note 11, at 106.

in favor of liberty and by emphasizing the centrality of the criminal law as the mechanism through which the state imposes its coercive powers. A better approach, Coleman suggested, would have been to “argue back from theories of the justification of punishment.”<sup>25</sup> Coleman explained:

If we treat the criminal law as paradigmatic of the state’s coercive power, then a useful way of thinking about the limits of what can be criminalized—in a morally permissible way—begins not by asking the question in political philosophy about the constraints on liberty that are justifiable, but from the other end, namely, from the point of view of the justification of punishment.<sup>26</sup>

Coleman’s criticism of Feinberg parallels closely his argument in the area of tort law. In his forthcoming book, *The Practice of Principle: In Defense of a Pragmatist Approach to Legal Theory*, Coleman offers an account of tort law as an institution of corrective justice. His discussion focuses on the second-order duty of repair.<sup>27</sup> In contrast, he argues, “the relevant first-order duties are not themselves duties of corrective justice.”<sup>28</sup> These are the product of lengthy historical and institutional development and cannot be tied to one coherent theory. “While corrective justice presupposes some account of what the relevant first order duties are, it does not pretend to provide an account of them.”<sup>29</sup> Nevertheless, Coleman emphasizes, “corrective justice is not compatible with just any set of first-order duties.”<sup>30</sup> Working back from the second-order duty of repair, it is possible to evaluate and judge first-order duties.

The same argument applies by analogy to criminal

---

25. Personal Correspondence from Jules Coleman, Professor of Law and Philosophy, Yale University (Nov. 27, 2000) (on file with author).

26. *Id.* (emphasis added).

27. Jules Coleman, *The Practice of Principle: In Defense of a Pragmatist Approach to Legal Theory* (manuscript at 57, on file with author) (forthcoming).

28. *Id.*

29. *Id.*

30. *Id.*

law. Feinberg's treatise, *The Moral Limits*, is about first-order duties, Coleman argues. It is about defining criminal conduct. It is not an account of second-order duties of repair. In his words, "[it] is not, nor does it purport to be, an account of the grounds or justification for punishing those who fail to live by the criminal law's demands."<sup>31</sup> As in tort law, the criminal law's first-order duties are the product of lengthy historical and institutional development, and are not amenable to one coherent theoretical construct. Instead of approaching first-order duties head on, Coleman argues by analogy, it is better to work backwards from the justification of punishment. The fact is, theories of punishment impose constraints on what can be punished. "Retributivism is not compatible, for example," Coleman writes in *The Practice of Principle*, "with the abandonment of the various mental elements of a crime in favor of general strict liability."<sup>32</sup> The bottom line, according to Coleman, is that Feinberg would have been better advised to seek the moral limits of the criminal law in a theory of punishment.

A fourth category of possible responses is that the moral limits of the criminal law are in some sense primary and should dictate the substance of punishment theory. The idea here is that the moral definition of crimes is more fundamental and should guide our analysis of punishment. Under this fourth approach, the measure and method of punishment is derivative of the moral and legal theory that places limits on the criminal sanction.

#### *Joel Feinberg's Position*

In this essay, I will argue that Joel Feinberg's work is one version of this fourth category of potential answers. Although there is an acoustic separation between his *Moral Limits* and *Expressive Function* writings, it is not the case that there is no relationship at all between the two works.

---

31. Id. at 59 n.14.

32. Id. at 60.

The link is actually stronger than one would expect—and stronger than is generally recognized. Although Feinberg wrote the punishment essay first, Feinberg's expressive theory of punishment is collateral to, and derivative of, his writings on the moral limits of the criminal law. In *The Expressive Function* essay, Feinberg develops a theory of punishment that is based on his primary allegiance to the Millian harm principle.

Feinberg's theory of punishment involves both a descriptive and a normative claim. The descriptive is the now-popular claim that punishment necessarily has an expressive dimension, specifically an expressive element of moral condemnation.<sup>33</sup> The normative is the less-familiar<sup>34</sup> claim that the expressive dimension of moral condemnation should be calibrated to (1) the "amount of harm" generally caused by the criminal event and (2) the "degree to which people are disposed to commit it."<sup>35</sup> The amount of opprobrium that we should attach to a punitive practice, Feinberg argues, should be proportional to the amount of harm caused by the offender and whether others are likely to commit the offense. In other words, Feinberg's theory of punishment relates back, primarily, to the Millian harm principle.

Now, Feinberg's preferred method of analysis involves consideration, back and forth, between practice and principle—a method similar to legal analysis.<sup>36</sup> As a result,

---

33. This is the element that has been appropriated and developed most by Dan Kahan in the criminal law. See, e.g., Kahan, *Alternative Sanctions*, supra note 3; Kahan, *Deterrence*, supra note 3; Kahan, *Gun Control*, supra note 3.

34. Many commentators seem to dismiss or ignore this normative dimension of Feinberg's expressive theory. Igor Primoratz, for example, in his important essay on *Punishment as Language*, suggests that Feinberg's account is purely descriptive and, as a result, turns to other expressive theorists, including Durkheim, Hampton, and Stephen, to develop the normative aspects of the expressive account. See Igor Primoratz, *Punishment as Language*, 64 *Phil.* 187, 187-88 (1989).

35. Feinberg, *The Expressive Function*, supra note 1, at 118.

36. Feinberg described this method of analysis as follows:

The best way to defend one's selection of principles is to show to which positions they commit one on such issues as censorship of literature, 'moral offenses,' and compulsory social security programs. General principles arise in the course of deliberations over particular problems, especially in

his style of argument is somewhat eclectic.<sup>37</sup> There are passages in Feinberg's writings that might suggest a different category of response. In fact, there is even a passage in *The Moral Limits* where Feinberg argues back from retributive theory to refute legal moralism—the precise rhetorical move that Jules Coleman advocates.<sup>38</sup> But the more consistent position that Feinberg takes in his work is that the moral limits of the criminal law are primary and guide his discussion of the expressive function of punishment. In his writings, Feinberg respects the distinction between first-order definitions and second-order remedies,<sup>39</sup> and, most consistently, argues that resolution of first-order duties should guide the analysis of punishment. In effect, Feinberg takes the fourth approach in my typology of possible responses.

I will flesh out this limited claim in this Symposium essay. I should emphasize, though, that my purpose here is not to discuss all of Feinberg's writings on moral limits or on punishment. Hugo Bedau, in his contribution to this

---

the efforts to defend one's judgments by showing that they are consistent with what has gone before. If a principle commits one to an antecedently unacceptable judgment, then one has to modify or supplement the principle in a way that does the least damage to the harmony of one's particular and general opinions taken as a group. On the other hand, when a solid, well-entrenched principle entails a change in a particular judgment, the overriding claims of consistency may require that the judgment be adjusted. This sort of dialectic is similar to the reasonings that are prevalent in law courts.

Joel Feinberg, *Moral Enforcement and the Harm Principle*, in *Social Philosophy* (1973), reprinted in *Ethics and Public Policy* 287 (Tom L. Beauchamp ed., 1975).

37. See generally Hampton, *supra* note 10, at 159-61 (discussing the methodology Feinberg uses to define and argue for principles).

38. See *Harmless Wrongdoing*, *supra* note 6, 164-65 (arguing that the retributive punishment theory is not consistent with the Stephen-Devlin argument for the legitimacy of victimless crimes).

39. In *The Moral Limits*, Feinberg is acutely aware that his focus is on the first-order question of the moral limits of the criminal law, and not on the second-order question of remedies. At one point in *The Moral Limits*, Feinberg draws on Jean Hampton's analogy between punishments and electrical fences (punishments, like electrical fences, are intended to demarcate the boundary of acceptable conduct). Feinberg then emphasizes: "The question of this book . . . is not (or not only) why Hampton's fence should be electrified, but why it should be located where it is." *Harmless Wrongdoing*, *supra* note 6, at 301.

Symposium, collects Feinberg's writings on punishment and offers a more complete account of Feinberg's philosophy of punishment—"such as there is one,"<sup>40</sup> as Bedau emphasizes—including comments on the death penalty, and other topics. My contribution here will not attempt to cover this ground again, but will focus instead on the narrow claim concerning the link between *The Moral Limits* and Feinberg's earlier essay, *The Expressive Function of Punishment*.

## II. FROM THE HARM PRINCIPLE TO PUNISHMENT

My claim, then, is that for Feinberg, the moral limits of the criminal law are foundational and guide his writings on punishment. The place to begin to explore this relationship is Feinberg's definition of punishment.

### *The Descriptive Claim*

Feinberg argues that the definition of legal punishment should take account not only of the pain or hard treatment that is meted out, but also of the moral condemnation and reprobative symbolism of that treatment. Hard treatment without reprobative symbolism is not necessarily legal punishment, nor is reprobation without hard treatment. Although reprobation is often itself painful, Feinberg emphasizes, "we can conceive of ritualistic condemnation unaccompanied by any *further* hard treatment, and of inflictions and deprivations which, because of different symbolic conventions, have no reprobative force."<sup>41</sup> The symbolic meaning and the particular conduct are two facets of punishment, and it is crucial to have both if we are to have legal punishment. "In short, punishment expresses blame, and it is through this expression that we recognize certain actions as punishment."<sup>42</sup>

---

40. Bedau, *supra* note 11, at 105.

41. *The Expressive Function*, *supra* note 1, at 98.

42. Steiker, *supra* note 3, at 803.

The basic insight of Feinberg's descriptive claim can be illustrated with a simple example. Suppose that someone gives another person a Heimlich maneuver. If the recipient is choking on her dinner, it is likely that the Heimlich maneuver will be interpreted as an act of good samaritanism and will be rewarded. The expressive dimension of that act is compassion, assistance, and support. If the recipient is a total stranger walking in the street, it is likely that the Heimlich maneuver will be interpreted as an assault and battery, a crime. The expressive dimension here is hostility and aggression. However, if the recipient is convicted of a crime and the person administering the Heimlich is doing so under court order, it is likely that the maneuver will be interpreted as punishment. The expressive function of that act is moral condemnation, Feinberg would argue. Feinberg's point is that the act of administering the Heimlich—the human behavior—is identical in all three cases. What differs is the moral condemnation expressed in the third instance. Only when the maneuver is associated with moral condemnation does it fall within the scope of legal punishment.

Another way to say this is that the *meaning* of the maneuver may largely determine whether or not it is considered to be punishment. "Even floggings and imposed fastings do not constitute punishments," Feinberg explains "where social conventions are such that they do not express public censure ..., and as therapeutic treatments simply, rather than punishments, they are easier to take."<sup>43</sup> This, I take, is now very familiar, due in large part to the expressivist turn in punishment theory and the numerous contemporary writings drawing on Feinberg's insight concerning the expressive function of punishment.<sup>44</sup>

### *The Normative Claim*

What is less familiar, but more important here, is that

---

43. The Expressive Function, *supra* note 1 at 114.

44. See *infra* note 2.

## 2001] FEINBERG ON CRIME &amp; PUNISHMENT 161

Feinberg builds a normative claim on the basis of the expressive function of punishment. His claim is that the amount of public censure that we attach to any particular physical punishment—floggings, fastings, incarceration—should be proportional to the harm caused by the offender and the degree to which others are likely to commit the offense.

Feinberg develops this argument in the concluding paragraph of *The Expressive Function*. Feinberg is discussing whether there is any merit to a particular type of retributivist theory of the form that “the wicked should suffer pain in exact proportion to their turpitude.” Feinberg rejects this “pain-fitting-wickedness version of the retributive theory,”<sup>45</sup> but he nevertheless does suggest that this version of retributivism rests on some element of common sense. He writes:

[J]ustice *does* require that in some (other) sense “the punishment fit the crime.” What justice demands is that the *condemnatory aspect* of the punishment suit the crime, that the crime be of a kind that is truly worthy of reprobation. Further, the degree of disapproval expressed by the punishment should “fit” the crime only in the unproblematic sense that the more serious crimes should receive stronger disapproval than the less serious ones, **the seriousness of the crime being determined by the amount of harm it generally causes and the degree to which people are disposed to commit it.**<sup>46</sup>

Feinberg’s emphasis, of course, is on the symbolic function of the punishment, not on the actual physical pain or treatment associated with the punishment. “[I]t is social disapproval and its appropriate expression that should fit the crime,” Feinberg emphasizes, “and not hard treatment (pain) as such.”<sup>47</sup> His argument is that the measure of symbolic disapproval should correspond to the amount of harm caused and to the general predisposition to commit

---

45. *The Expressive Function*, supra note 1, at 118.

46. *Id.* (italics in original; bold added).

47. *Id.*

the crime.

The first of these two elements maps neatly onto the Millian harm principle that Feinberg refines in his first volume, *Harm to Others*. It corresponds well to Feinberg's argument that it is generally a good reason to criminalize conduct if the conduct causes harm—or, more technically for Feinberg, causes a wrongful set back of interests.<sup>48</sup> Feinberg's treatment of the harm principle in *Harm to Others* is, of course, far more nuanced than his reference to "the amount of harm" in *The Expressive Function*. In fact, in *Harm to Others*, Feinberg makes clear that harm alone is not determinative and that "non-wrongful harms"—such as certain harms where the victim consents—should not form the basis of penal legislation. Nevertheless, it is interesting to see that harm forms the core of the degree of moral opprobrium that Feinberg believes we should attach to punishment. The heart of his theory of punishment draws on his later work on the Millian harm principle.

The second of the two elements—"the degree to which people are disposed to commit" the crime—is less easy to interpret. Feinberg does not really develop the idea of predispositions to commit crime in his work—a sociological and criminological concept that is most frequently addressed from an empirical perspective. Moreover, Feinberg is addressing here dispositions at both an individual and at a general level. He is not solely concerned with the actual person accused, but also with people in general. Feinberg is treating *categories* of crime, not just individual crimes.<sup>49</sup> As he explains in *The Moral Limits*:

The punishment expresses condemnation of classes of crimes too, not only of particular criminals for committing

---

48. See, e.g., Joel Feinberg, *Offense to Others* x (1985).

49. This is true of the harm caused by the conduct as well. Feinberg is not suggesting that the moral opprobrium should be calibrated solely to the harm caused by *this particular crime*, but by the harm generally caused by this type of crime. So, for instance, if a terrorist act fortuitously results in no deaths, the opprobrium should nevertheless be great because it is the type of crime that generally results in a lot of harm.

those crimes; the act as well as the actor is condemned. Even if the actor's motives were entirely good so that he is not blameworthy at all, the condemnation is to impress on him the community's moral judgment that the *act* he intentionally performed from such innocent motives was nevertheless wrong.<sup>50</sup>

Along this line of thought, it could be that the second element refers to general predispositions to commit crime that are reflected in the different degrees of felony or misdemeanor statutes.

Under another reading, this second element could track the offense principle—the idea that it is generally a good justification to criminalize conduct if the conduct causes serious offense to others. If people are very much disposed to commit the crime, it is unlikely that the conduct is going to be viewed as *seriously* offensive. Seriously offensive conduct generally, I would argue, is behavior that people are not highly predisposed to commit. The advantage of this interpretation is that Feinberg's theory of punishment would then track closely the two principles that he is most attached to—the harm and offense principles. It would be possible to argue, then, that Feinberg ties his theory of punishment more closely to his treatise *The Moral Limits*. The difficulty with this interpretation is the conjunction “and.” In order to make this interpretation work, Feinberg would have had to use “or” instead. In Feinberg's work, harms and offenses are exclusive, in the sense that actionable offenses are defined as not causing harm. If offenses cause wrongful harm, then they can be dealt with under the harm principle. The offense principle is only truly operative in the case of wrongful but harmless serious offenses. As a result, if the second element truly mapped on to the offense principle, Feinberg would have had to use “or” as the conjunction in the sentence quoted (and bolded) above.

To be sure, the meaning of the second element is not as clear as that of the first. However, what is clear is that, in

---

50. Harmless Wrongdoing, *supra* note 6, at 149.

Feinberg's writings, the harm principle (at the very least) serves as the basis of his theory of punishment: harm is the measure of the expressive element of moral condemnation. The concept of harm which is at the heart of the moral limits of the criminal law gives substance to Feinberg's expressive theory of punishment.

Feinberg reaffirms this point in *The Moral Limits*, in his discussion of legal moralism. Feinberg is addressing Stephen's argument, discussed earlier, that sentencing practices prove that the criminal law aims at immoral conduct. After rehearsing Hart's response, Feinberg offers his own "much simpler and direct reply" to Stephen.<sup>51</sup> The harm principle, as a liberty-limiting principle, Feinberg argues, is in fact a moral principle. It serves not merely to minimize harm, but to minimize *morally wrongful* harm. As a result, moral theory enters at the stage of defining criminal law as well as at the stage of punishment (which should be based in part on moral blameworthiness). With regard to punishment, though, the measure of moral condemnation must be *that* harm—the wrongful harm that forms the basis of the criminal law—and not other harms. As Feinberg explains:

When [the consistent liberal] approves gradations in punishment based on different degrees of *blameworthiness* (as opposed to responsibility) he must not permit the types of blameworthiness which he excludes at the legislative level to sneak in the back door at the sentencing level. In both cases the moral blameworthiness that *is* relevant is the harm-threatening, right-violating kind, dispositions to feel or act in ways condemned by grievance morality. And in both cases also, moral blameworthiness based on the principles of nongrievance morality must equally be excluded.<sup>52</sup>

For Feinberg, then, the concept of wrongful harm serves as a moral limit of the criminal law, and that very

---

51. *Id.* at 151.

52. *Id.* at 154.

same concept guides the measurement of the condemnatory expressive element of punishment. It is what gives substance to his normative claim concerning the expressive function of punishment.

*A Few Collateral Points on Punishment*

It is important to note, in this context, that Feinberg draws a sharp distinction between the purposes and function of punishment. He admits that he does not know how the expressive function always relates to purposes:

The relation of the expressive function of punishment to its various central purposes is not always easy to trace. Symbolic public condemnation added to deprivation may help or hinder deterrence, reform, and rehabilitation—the evidence is not clear. On the other hand, there are other functions of punishment, often lost sight of in the preoccupation with deterrence and reform, that presuppose the expressive function and would be difficult or impossible without it.<sup>53</sup>

Moreover, when he does discuss the expressive function in the *Moral Limits*, Feinberg makes clear that the expressive function should not solely determine the type and degree of punishment. There are, he argues, “other social functions of punishment, notably deterrence, that have a bearing on the decision” about the degree of punishment.<sup>54</sup>

Finally, it is also important to mention that Feinberg does express a preference—at least, a personal preference, he calls it a “fantasy”—for avoiding the physical aspects of punishment.<sup>55</sup> As Toni Massaro correctly points out in her essay, *The Meaning of Shame*, Feinberg “is so convinced of the usefulness of official condemnation that he would

---

53. The Expressive Function, supra note 1, at 101. Feinberg then discusses these other expressive functions, including disavowal, nonacquiescence, vindication and absolution.

54. Harmless Wrongdoing, supra note 6, at 149.

55. The Expressive Function, supra note 1, at 116.

prefer punishment that preserved this aspect *solely*, and dispensed altogether with the physical media of incarceration and corporal mistreatment, if this were feasible.”<sup>56</sup> Feinberg does not develop this personal preference into a full-blown argument, but it does shed light on the direction that his punishment theory would take.

In sum, for Feinberg, there is an expressive dimension to punishment that is separate and distinct from the purposes of punishment, as well as from the physical media of incarceration and corporal mistreatment. By drawing these distinctions, Feinberg develops a theory about the justification and calibration of punishment that is independent of the traditional debates over the purposes of punishment, and also that leaves open the important question of methods of punishment.

### III. AN ASSESSMENT OF FEINBERG’S ARGUMENT

This reading of Feinberg emphasizes a missing, but strong link between his discussion of moral limits and his discussion of punishment. In contrast to the three other categories of responses, Feinberg gives pride of place to the moral limits of the criminal law. His theory of punishment is collateral to, and derivative of, his primary allegiance to the Millian harm principle, which is at the very core of his theory of first-order duties.

This interpretation of Feinberg, in effect, reads *The Moral Limits* into the earlier essay *The Expressive Function of Punishment*. This reading, of course, is not seamless. It overemphasizes, perhaps, a link that was never properly or fully developed by Feinberg. It forces a connection that Feinberg himself did not consider important enough to emphasize. And it is terribly incomplete. What is missing is any indication as to what type of punishment we should actually administer.

---

56. Toni Massaro, *The Meanings of Shame: Implications for Legal Reform*, 3 *Psychol., Pub. Pol’y & L.* 645, 702 (1997).

Although we have some vague idea of how we should calibrate the opprobrium in relation to harm, we still have no idea what type of pain and suffering to inflict. We are left, then, with a skeletal theory of punishment.

The next question, of course, is whether Feinberg is right—in at least two respects. Specifically, is Feinberg right that we should calibrate moral condemnation to harm (and propensity)? More generally, does Feinberg take the right approach in suggesting that theories of punishment should be derivative of moral-legal theories concerning the limits of the criminal law?

With regard to Feinberg's specific argument, I have some reservations.<sup>57</sup> On the one hand, I agree with Feinberg that harm must play a central role in the determination of how we administer formal social control and, more specifically, the criminal sanction. In a previous article, *The Collapse of the Harm Principle*,<sup>58</sup> I argue that the rise of the Millian harm principle in the debates over the legal enforcement of morality has triggered a proliferation of harm arguments in contemporary legal and political discourse. On all sides of the political spectrum, advocates have turned to harm arguments as their main justification for or against legal prohibition and regulation of conduct. This rhetorical shift is evident in a wide range of political debates, from prostitution and pornography, to drinking and illicit drug use, to homosexual and heterosexual promiscuity, to loitering and adultery. One important consequence of this proliferation of harm arguments is that the harm principle itself has lost its critical edge. Claims of non-trivial harm have become so pervasive in political debate that the harm principle no longer serves the function of a *critical principle*. It no

---

57. Of course, I am not alone in this respect. Hugo Bedau, in his contribution to this Symposium, similarly sounds a note of deep skepticism. Bedau writes, "I must confess I fail to see how we are supposed to scale degrees of condemnation any better than we can scale degrees of deserved punishment straightaway." Bedau, *supra* note 11, at 122-23.

58. Bernard E. Harcourt, *The Collapse of the Harm Principle*, 90 *J. Crim. L. & Criminology* 109 (1999).

longer really excludes much conduct from the ambit of the criminal law.

The collapse of the harm principle, I argue, is beneficial to contemporary political theory—and, especially, in this context, to punishment discourse. It forces us to assess, compare, and weigh harms in a manner that had previously been obfuscated by the harm principle. It forces us to deal with complex and competing claims of harm—claims of harm presented, generally, on both sides of the debates. In this respect, I endorse Feinberg's focus on harm as a measure of certain elements of punishment, and agree that the criminal sanction must be related to these complex assessments of harm. I would emphasize, though, that the focus on harm is likely to produce highly contentious debates. Putting aside the more conventional cases of unjustifiable homicide or mass killings, the "harm" associated with the more difficult cases (drug use, sodomy, prostitution, etc.) is more a term of *contestation* than a term of *consensus*. The term 'harm' is likely to trigger social disagreement and cleavage, rather than unity. It is by no means an easy path.

On the other hand, I am not sure that Feinberg is right to focus on *moral condemnation* as the principal expressive element of punishment. It is not clear to me that the expressive dimension of punishment is exclusively, primarily, or even importantly, *moral opprobrium*. In other words, while I agree with Feinberg that there is an expressive dimension to punishment, I disagree that morality is in fact central to that function. Punishment usually also communicates, importantly, political, cultural, racial and ideological messages. The meaning of punishment is not so coherent or simple. Many contemporary policing and punitive practices, for instance, communicate a racial and political, rather than moral, message—a message about who is in control and about who gets controlled.<sup>59</sup> Feinberg's descriptive claim, in essence,

---

59. See, e.g., Dorothy E. Roberts, Foreword: Race, Vagueness, and the Social Meaning of Order-Maintenance Policing 89 *J. Crim. L. & Criminology* 775 (1999).

may truncate a significant portion of what is expressed in punishment.

Moreover, I doubt that the expressive dimension of punishment is easily amenable to engineering or manipulation. It is not clear to me that we *could* calibrate the moral message attached to punishment in the first place. The message expressed by punishments, especially the moral message, is likely to be shaped in large part by the perceived legitimacy of the criminal justice system. The moral condemnation that attaches to punishment for certain drug offenses, for example crack or non-violent drug use, is shaped largely by political and racial dimensions that have little to do with moral theory. In other words, it is not clear that we can transform Feinberg's descriptive claim into a normative one.

As for the broader question—whether Feinberg takes the proper approach to the relationship between the moral limits of the criminal law and punishment theory—I have even greater reservations. In my work, I take a different approach and draw on different intellectual traditions, principally for two reasons.

First, I am deeply skeptical that we could successfully derive rules about what to criminalize from moral principles—from relatively abstract theorizing about first-order duties. In my experience, moral principles in the criminal law most often run out on us before we have reached the end of our analysis. A poignant example of this—especially in the context of the earlier discussion of the Millian harm principle—is how a moral principle like “lessening human suffering” can be deployed both in support of *and* in opposition to capital punishment. John Stuart Mill, in Parliamentary debates in 1868, specifically argued in favor of capital punishment because it “effects its purposes at a less cost of human suffering than any other.”<sup>60</sup> Mill's argument, in essence, was that executing the offender in the most egregious cases is far less cruel a

---

60. John Stuart Mill, Speech in Favor of Capital Punishment 273 (1868), reprinted in *Philosophical Perspectives on Punishment*, (Gertrude Ezorsky ed., 1972).

punishment than imprisonment at hard labor for life. Naturally, the same moral principle has been commonly deployed in opposition to the death penalty. In this respect, I am more inclined to agree with Stanley Fish that neutral principles “don’t have the constraining power claimed for them” and are most often used “to disguise substance so that it appears to be the inevitable and nonengineered product of an impersonal logic.”<sup>61</sup>

Secondly, I would argue that the two inquiries—into the moral limits of the criminal law and punishment theory—are not only integrally related to each other, but are themselves just a fraction of the proper inquiry we should be having. The proper scope and application of the criminal sanction is itself just a small part of the larger question of social control. The criminal sanction is one tool in a larger tool box of ways of formally and informally controlling and shaping people. To be sure, it is unique in its devastating potential for the individual—as evidenced by executions, solitary confinement, deportation, etc. And, in this sense, it is important to be able to differentiate between the death penalty, incarceration, fines, alternative sanctions, and other forms of social control. There is a delicate relationship of difference and similarity between the criminal sanction and civil remedies, and it is no doubt important, for many reasons, to maintain the uniqueness of different criminal sanctions, especially to retain the distinctively condemnatory aspect of certain forms of criminal punishment. But, in thinking about the scope of the criminal sanction, it is important to conceptualize punishment within the larger framework of social control. To address the question “what should be the limits of the criminal sanction?” we need to explore the comparative advantage of the criminal penalty versus other legal mechanisms—such as tort law, or abuse and neglect law—as well as other social institutions—the school, the church, the family, the workplace—to accomplish the goals that we have in mind.

---

61. Stanley Fish, *The Trouble with Principle* 4 (1999).

## 2001] FEINBERG ON CRIME &amp; PUNISHMENT 171

In this light, the proper inquiry, I believe, is an exercise in utopian social theory and in the history of knowledge. Instead of focusing on moral principles, and instead of isolating moral principles from theories of punishment, we need to look at the distributional consequences of proposed criminal sanctions and at the type of society, social relations, and subject that we are shaping with our policies.<sup>62</sup> This type of inquiry draws more on social and political theory, than on moral theory. It involves a close historical analysis of subject creation. Along these lines, I find inspiration in the aphoristic writings of Friedrich Nietzsche, especially certain thoughts on punishment in the *Genealogy of Morals*:

As its power increases, a community ceases to take the individual's transgressions so seriously, because they can no longer be considered as dangerous and destructive to the whole as they were formerly: the malefactor is no longer "set beyond the pale of peace" and thrust out; universal anger may not be vented upon him as unrestrained as before. . . . It is not unthinkable that a society might attain such a *consciousness of power* that it could allow itself the noblest luxury possible to it—letting those who harm it go *unpunished*. "What are my parasites to me?" it might say. "May they live and prosper: I am strong enough for that!"<sup>63</sup>

## CONCLUSION

My purpose in this Symposium essay is not so much to develop my own views—which would take much longer—as it is to simply excavate the relationship between *The Moral Limits* and the earlier essay *The Expressive Function of Punishment*. My thesis is that there is, in fact, a stronger connection than is usually recognized between Feinberg's discussion of moral limits and his theory of punishment. It

---

62. I make this argument in my book, *Illusion of Order: The False Promise of Broken Windows Policing* (2001).

63. Friedrich Nietzsche, *On the Genealogy of Morals* 72 (Walter Kaufmann & R.J. Hollingdale trans., 1989).

172      *BUFFALO CRIMINAL LAW REVIEW* [Vol. 5:145

is a modest, but provocative connection. It provides, in essence, that the moral expression of condemnation in punishment should be calibrated primarily to the harm associated with the crime. In this way, Feinberg tied his expressive theory of punishment to his primary allegiance to the Millian harm principle. My hope is that this short contribution will assist future Feinberg scholars in integrating more seamlessly two important bodies of work.