Harms, Wrongs, and Set-Backs in Feinberg's *Moral Limits of the Criminal Law*

Hamish Stewart*

1. INTRODUCTION

As a rough first approximation, one might divide the moral and political justifications for criminalizing conduct into two types: deterring harm-doing and punishing wrongdoing.[1] While there are many ways in which a system of penal justice might accommodate both ideas,[2] there will always be some tension between the two, in that deterring harm-doing may sometimes seem to demand the punishment of those who have done no wrong, while punishing wrong-doing may be, from the point of view of deterrence, too punitive or not punitive enough. This tension is closely related to a long-standing debate about the status of rights: Are our rights defined to serve purposes external to the system of rights itself and therefore vulnerable to redefinition as social and political goals change, or do some of our core rights trump other considerations and therefore indefeasible by

* Associate Professor of Law, University of Toronto, 78 Queen's Park, Toronto, ON M5S 2C5, Canada. B.A., 1983, University of Toronto; Ph.D., 1989, Harvard University; LL.B., 1992, University of Toronto. I am very grateful to David Dyzenhaus for his comments on a draft of this paper, to Stuart Green for his patience, and to the Wright Foundation for Legal Scholarship for financial support.

1. See Herbert Packer, The Limits of the Criminal Sanction 35-61 (1968); Ted Honderich, Punishment: The Supposed Justifications (1969); C.L. Ten, Crime, Guilt, and Punishment: A Philosophical Introduction 7-65 (1987).

2. See H.L.A. Hart, Punishment and Responsibility: Essays in the Philosophies of the Law 1-27 (1965); Jean Hampton, Liberalism, Retribution and Criminality, in In Harm's Way: Essays in Honor of Joel Feinberg 159 (Jules L. Coleman & Allen Buchanan eds., 1994); Alan Brudner, The Unity of the Common Law: Studies in Hegelian Jurisprudence 211-60 (1995); John Gardner & Stephen Shute, The Wrongness of Rape, in Oxford Essays in Jurisprudence, No. 4 (Jeremy Horder ed., 2000).

consequentialist considerations? For if rights are non-instrumental trumps over other considerations, then the extent to which our rights are protected by the criminal law, both as potential victims of crime and as potential criminal defendants, would, it would seem, have to be defined without reference to the possibility of deterrence; while if our rights are instrumental to something else, then the possibility of deterring harm would be highly relevant to their definition.

Each side of this contrast has an unimpeachable liberal pedigree beginning not later than the 18th century, with Bentham and Kant respectively.[3] In his magisterial and richly argued *Moral Limits of the Criminal Law*,[4] Joel Feinberg makes a sustained case for the proposition that liberals should confine the justifications for criminalizing conduct to those that can fit within two principles, which he calls "the harm principle" and "the offense principle." Feinberg's project therefore sounds as though it belongs in the harm-deterring category. But, as his readers are well aware, Feinberg is steeped in the rights culture of the United States, and far from treating rights as mere instruments, Feinberg is on occasion inclined to treat them as trumps. It is not my purpose to assess the overall success or failure of Feinberg's project; rather, I would like to explore Feinberg's understanding of "harm" and relate that understanding to the role that rights play in his arguments for or against criminalizing conduct.[5]

Feinberg is sometimes read as defining "harm" as a

---

3. Jeremy Bentham, An Introduction to the Principles of Morals and Legislation 143-86 (J.H. Burns & H.L.A. Hart eds., Oxford U. Press 1996) (1789); Immanuel Kant, The Metaphysics of Morals 331-35 (Mary Gregor trans., Cambridge U. Press 1996) (1797).

4. Joel Feinberg, The Moral Limits of the Criminal Law, Volume One: Harm to Others (1984) [hereinafter Harm to Others]; Volume Two: Offense to Others (1985) [hereinafter Offense to Others]; Volume Three: Harm to Self (1986) [hereinafter Harm to Self]; Volume Four: Harmless Wrongdoing (1988) [hereinafter Harmless Wrongdoing].

5. Like Feinberg, in this article I will be strictly concerned with the sort of conduct that ought to attract criminal sanction and not with questions of criminal law doctrine, such as whether fault requirements should be subjective or objective, or whether there should be liability for attempts.

set-back to an individual's serious welfare interests,[6] but while harm in this sense motivates much of Feinberg's analysis, he clearly defines "harm" as a *wrongful* set-back to interests.[7] This definition may seem to point to an interesting way of combining the deterrence of harm and the punishment of wrong-doing, but I want to suggest that it is not the most felicitous for the purpose of thinking about the reasons for criminalizing conduct or indeed for Feinberg's own purposes. Again and again, and usually implicitly rather than directly, Feinberg either abandons the wrongfulness component of his definition of harm, or assumes rather than justifies its operation. It would be more fruitful for Feinberg's project to identify clearly the separate roles that the concepts of wrong and set-back to interests might play in arguments for or against criminalization.

II. HARMS AND WRONGS

"What sorts of conduct may the state rightly make criminal?"[8] Feinberg approaches this question by considering the application of several "liberty-limiting principles" to various types of conduct that might plausibly be prohibited. A principle is "liberty-limiting" if it points to "a given type of consideration [that] is always a morally relevant reason in support of penal legislation even if other reasons may in the circumstances outweigh it."[9] Feinberg defines and discusses several liberty-limiting principles, and argues that the liberal position on criminalization is

6. See, e.g., Brudner, supra note 2, at 211-12; David Dyzenhaus, Liberalism, Autonomy, and Neutrality, 42 U. Toronto L.J. 354, 355 (1992). This reading is quite understandable, given the way Feinberg relates his project to Mill's. Compare Harm to Others, supra note 6, at 11-14, with John Stuart Mill, On Liberty, in Utilitarianism, Liberty, Representative Government 63, 72-75 (H.B. Acton ed., 1972); but see John Kleinig, Criminally Harming Others, 5 Crim. Just. Ethics, Winter/Spring 1986, at 3, 4 (noting the ways in which Feinberg's harm principle and Feinberg's project differ from Mill's).

7. Harm to Others, supra note 4, at 36.

8. Id. at 3.

9. Id. at 9.

that "[t]he harm and offense principles, duly clarified and qualified, between them exhaust the class of good reasons for criminal prohibitions."[10] These principles are defined as follows:

> *1. The Harm Principle*: It is always a good reason in support of penal legislation that it would probably be effective in preventing (eliminating, reducing) harm to persons other than the actor (the one prohibited from acting) *and* there is probably no other means that is equally effective at no greater cost to other values.
>
> 2. *The Offense Principle*: It is always a good reason in support of a proposed criminal prohibition that it is probably necessary to prevent serious offense to persons other than the actor and would probably be an effective means to that end if enacted.[11]

Thus, principles that rest on arguments of paternalism (prohibiting conduct to prevent harm to the actor himself or herself) or of moralism (prohibiting immortal conduct that is neither harmful nor offensive) are rejected.

In this paper, I am concerned with Feinberg's articulation of the idea of "harm" and how this articulation relates to his arguments for and against criminalizing conduct. I am not directly concerned with the offense principle; nothing in my analysis of his use of the harm principle turns on accepting or rejecting his definition of offense, or his argument that the offense principle may also justify criminalization. Nor am I directly concerned with Feinberg's specific conclusions on particular issues, partly because Feinberg treats the harm and offense principles not as decisive arguments but as considerations that have weight in an argument for or against criminalization,[12] partly because my agreement or disagreement would, in some cases, depend on the resolution of empirical

10. Id. at 26.
11. Id. (emphasis in original).
12. Id. at 9-10.

questions,[13] and partly because his specific conclusions are frequently sensible, liberal, and humane.

Feinberg considers three possible senses of the word "harm": harm as damage, harm as set-back to interests, and harm as wronging. He concludes that for the purposes of the harm principle, "harm" should be thought of as combining the last two senses:

> [T]he harm principle as a guide to the moral limits of the criminal law does not licence liability for acts that tend to cause only nonharmful wrongs.... no plausibly interpreted harm principle could support the prohibition of actions that cause harms without violating rights, for example setbacks to interests incurred in legitimate competitions, or harms to the risk of which the "victim" freely consented. The sense of "harm" as that term is used in the harm principle must represent the overlap of senses two and three: only setbacks of interests that are wrongs, and wrongs that are setbacks to interests, are to count as harms in the appropriate sense.[14]

By defining the harm principle in this way, Feinberg has set himself a task that is very difficult in at least two respects. First, he has apparently ruled out the possibility of defining rights just to protect persons from certain setbacks to interests. He cannot say, for instance, "X ought to be considered wrong because it is a setback to interests," or "X ought to be considered a set-back to interests because it is wrongful." Yet the thought that rights might be defined this way has been central to harm-oriented accounts of criminal liability; indeed, Feinberg himself says at one point, "The whole purpose of the criminal prohibition [for certain core criminal offences] is to discourage the particular antisocial behavior that is forbidden, and that behavior can be characterized quite independently of the legal statute that forbids it."[15]

Second, since "harm" is a combination of set-back to

13. Id. at 17.
14. Id. at 36.
15. Id. at 20.

interests and wrongness, Feinberg requires an independent account of each. He cannot say "X is harmful just because it is a set-back to interests," or "X is harmful just because it is wrong,"[16] because each of these statements involves the reduction of one to the other,[17] and it is supposed to be their combination that produces the distinctive form of harm that is the appropriate target of criminal sanction.

These two difficulties would be avoided if we understood Feinberg as defining a right as a way of protecting particularly crucial welfare interests. This interpretation might be supported by some of Feinberg's analysis; for example, his discussion of the way in which death is a set-back to interests[18] might be understood as supporting a right against being killed. But, there is little evidence for this interpretation in Feinberg's overt discussion of rights: More than once, he distinguishes the question of whether conduct sets back interests from the question of whether conduct violates rights.[19]

Feinberg does, of course, provide a very detailed and careful discussion of the idea of a setback to interests,[20] which he quite rightly relies upon in his analysis of the various types of conduct that might appropriately be criminalized. But, his discussion of the idea of wrongfulness, though insightful in many respects, is left incomplete, with the result that much of his analysis relies on an implicit and rather intuitive specification of what a person's rights are.

16. A statement of this sort has appealed to some criminal law theorists, see, e.g., R.A. Duff, Subjectivism, Objectivism and Criminal Attempts, in Harm and Culpability 19, 37 (A.P. Simester & A.T.H. Smith eds., 1996). I mean neither to agree nor to disagree with the proposition that to wrong someone is to harm him or her but to point out how difficult it is to express in Feinberg's framework.

17. See Harm to Others, supra note 4, at 215.

18. Id. at 92-93.

19. See id. at 34-35 (discussing trespass to property); id. at 99 (discussing the distinction between harming and wronging an unborn child); id. at 155 (distinguishing between harm and wrong in the context of consensual activity); id. at xxix (discussing the meaning of the phrase "harmless wrongdoing" and noting that there is no such thing as "harmless wrongdoing" if "harm" is understood as wrongful adverse effect on another's interest).

20. Id. at 31-104.

Feinberg defines a wrong as an act, performed with an appropriate type of fault, that violates another's right without excuse or justification.[21] I am going to accept, arguendo, Feinberg's theory of what constitutes an act,[22] and what constitutes fault,[23] in order to focus on what he says about rights. Feinberg notes that "we must take care not to identify wrongs, simpliciter, with violated *legal* rights"[24] because that would lead his project into a vicious circle in which "we cannot know which harms can properly be prevented by legal coercion until we know which harms can properly be prevented by legal coercion."[25] A natural law theory would certainly provide an account of right and wrong not dependent on the positive law of any jurisdiction,[26] but Feinberg is hesitant to go down this seductive path, partly because of hesitations about its coherence,[27] and partly because it seems to demand that he develop "a rather complete moral system"[28] before proceeding. Instead, Feinberg suggests that we rule out "[c]ertain kinds of morally disreputable interests" such as "interests in causing pain and suffering for their own sakes"[29] and take seriously the thought that "*any interest at all* (apart from the sick and wicked ones) is the basis of a

21. Harm to Others, supra note 4, at 105-06. Note that this formulation suggests that an excused violation of another's right is not a wrong. It is probably more common to think of an excused violation of another's right as an unpunishable wrong. See, e.g., George P. Fletcher, Rethinking Criminal Law 759-817 (1978). Feinberg minimizes the importance of this distinction. See Harm to Others, supra at 108-09.

22. See Harm to Others, supra note 4, at 118-86.

23. Id. at 108.

24. Id. at 110 (emphasis in original).

25. Id. at 111.

26. For a contemporary version of the natural law project in criminal law, see Brudner, supra note 2, at 211-60.

27. See Harm to Others, supra note 4, at 111. It is no accident that, working as he does on the harm side of the distinction I began with, Feinberg quotes Bentham's famous excoriation of natural rights theory: "*Natural rights* is simple nonsense: natural and imprescriptible rights, rhetorical nonsense—nonsense upon stilts." Id. at 254-55 (quoting Jeremy Bentham, Anarchical Fallacies, in Nonsense Upon Stilts 53 (Jeremy Waldron ed., 1987).

28. Harm to Others, supra note 4, at 111.

29. Id.

valid claim against others for their respect and non-interference."[30] The conclusion is that a wrong can be defined as "*any* indefensible invasion of another's interest (excepting of course the sick and wicked ones)."[31]

This definition of a wrong, as Feinberg recognizes, puts a lot of weight on the question of moral indefensibility.[32] As a result, some of his argument is structured around a notion of what is fair, or at least defensible, conduct in advancing one's interests. For example, one is quite entitled to set back another's interests if one does so in a fair competition. When the Baltimore Ravens beat the New York Giants in the 2001 Super bowl, certainly the Giants' interests had been set back, but they had not been wronged. "[D]efeat in a competition is a harm only when the form of competition itself is somehow morally illegitimate or else the defeat was inflicted in an unfair way—by fraud, treachery, or force."[33] Similarly, one is morally entitled to set back another's interest if the other has truly consented to the conduct that created the set-back (or the risk of the set-back).[34] In both cases, an independent though ill-defined norm of fairness makes the difference between a harmful and a non-harmful set-back to interests.

On the other hand, Feinberg's discussion of "aggregative" and "accumulative" harms does not have this structure. In this discussion, conduct can be defined as "harmful" for the purposes of penal law even if it is not independently wrongful or unfair. Harms are "aggregative" where an activity as a whole is harmful, but many specific instances of it are not harmful and may even be beneficial.[35] Drinking alcohol is a central example: "Probably more harm is done generally by the practice of imbibing alcoholic beverages than would occur if no one

30. Id. at 112 (emphasis in original).
31. Id.; see also id. at 114, 215.
32. Id. at 107-09.
33. Id. at 220.
34. Id. at 115-17.
35. Id. at 193.

ever drank,"[36] but "if the use of liquor is simply banned across the board, millions of citizens will be deprived of their wholly innocent and harmless pleasures."[37] The solution is not a blanket prohibition, but a licensing scheme that attempts to target the dangerous conduct while permitting the harmless conduct. Under any such scheme, some conduct that does no independent wrong is going to be criminalized—*e.g.,* a person drives while intoxicated but does not cause any damage or injury—while other conduct that does set back interests is not going to be criminalized—*e.g.,* a person becomes drunk and behaves obnoxiously but does not drive. So the proscribed conduct does not necessarily violate another's rights, while the scope of the right to engage in the conduct is defined by policy considerations. In other words, the justifications for the scheme are in terms of setbacks to interests; we end up defining the right to engage in the activity for instrumental reasons. So, in the case of aggregative harms, "harm" must mean simply a setback to interests because the "wrong" is defined by, not independently of, the regulatory scheme.

An activity causes "accumulative" harms where it is harmful if a lot of people engage in it, but not harmful if only a few do.[38] Familiar problems of collective action, such as excessive use of a commonly held resource or free-riding on a collectively beneficial scheme, are plausible candidates for such activities. Accumulative harms, like aggregative harms, cannot be defined in strict accordance with the definition of harm Feinberg began with. Discussing the question of regulating environmental pollution, Feinberg says:

> For *A* to harm the public interest is for *A*'s wrongful conduct to cause a setback to that interest. In the context of industrial polluting, "wrongful" must mean unlawful as judged by a regulative agency applying rules for allocating permits in accordance with specified requirements of

36. Id.
37. Id. at 194.
38. Id. at 228.

> fairness and efficiency. There is nothing inherently wrongful or right-violating in the activity of driving an automobile, generating electricity, or refining copper.[39]

Here, again, the right to engage in the activity in question is defined not by an independent standard of wrongfulness, but in terms of set-backs to interests.[40]

Thus, we might ask: Why does Feinberg talk about rights at all? If a wrong is going to be just a certain type of unjustified invasion of an interest, and if the circumstances that make it permissible to invade an interest are themselves defined in terms of set-backs to interests, then why not simply drop the rights talk altogether, define harm as set-back to interests, and focus on the effectiveness of law in containing those set-backs? The two examples that follow are intended to enable further thought about these questions. In the first, rights play a powerful, trumping, role over set-backs to interests. In the second, Feinberg comes close to equating "harm" with "setback to interest" simpliciter, rather than with "wrongful setback to interest." These examples are meant to suggest that Feinberg does require an independent account of rights to complement his account of set-backs to interest, in order to clearly show what is given up when one trumps the other.

III. HARMS AS SET-BACKS: TWO EXAMPLES

A. *The Harm of Pornography*

Liberals have traditionally taken a stringent view of the circumstances under which speech can be legally restrained, let alone criminalized, and Feinberg is no exception. But given the way in which Feinberg has previously defined "harm," it is not clear that this stringency is justified.

39. Id. at 230.
40. See also Andrew Kernohan, Liberalism, Equality, and Cultural Oppression 71-79 (1998).

Feinberg explores two possible grounds for criminal restrictions on pornography.[41] First, Feinberg considers criminalizing pornography on the ground that it is "obscene" in his sense of that term—that is, within the scope of the offense principle. "But obscenity (extreme offensiveness) is only a necessary condition, not a sufficient condition, for rightful prohibition. In addition, the offending conduct must not be reasonably avoidable, and the risk of offense must not have been voluntarily assumed by the beholders."[42] This approach leads naturally to time, place, and manner restrictions on the display of pornography, not to content-based criminal sanctions.[43]

Second, Feinberg considers the feminist argument that "pornography degrades, abuses, and defames women, and contributes to a general climate of attitudes toward women that makes violent sex crimes more frequent."[44] He is willing to contemplate criminalizing pornography that involves "wanton and painful violence against helpless victims,"[45] but not on the defamation-based ground that pornography induces "false beliefs" about women's subservience to men. In Feinberg's view, that rationale cannot be adequately confined to violent pornography:

> Are fat men defamed by Shakespeare's picture of Falstaff? Are Jews defamed by the characterization of Shylock? Could any writer today even hope to write a novel partly about a fawning corrupted black, under group defamation laws, without risking censorship or worse? The chilling

41. In criminalizing any form of expression, there are important legal and political questions about whether criminal sanctions should apply only to the production and distribution, or also to the mere possession, of the offensive expressive material in question. See, e.g., The Queen v. Sharpe, 194 D.L.R. 4th 1 (Can. 2001), rev'g 175 D.L.R. 4th 1 (B.C. 1999) (Can.) [Dominion Law Reports] (upholding a statute criminalizing simple possession of child pornography). For the purposes of my discussion of Feinberg, I set those questions aside to focus on the structure of Feinberg's argument as applied to expression. If on Feinberg's approach expression may in principle be criminalized, those questions would become relevant to determining the precise scope of the criminal sanction.

42. Offense to Others, supra note 4, at 142.

43. Cf. id. at 37-44, 157-64.

44. Id. at 143.

45. Id. at 146.

effect on the practice of fiction-writing would amount to a near freeze.[46]

He therefore turns to the argument that "violent pornography . . . promotes rape and physical violence."[47] Feinberg is concerned about the possible scope of criminalization this argument would permit,[48] and about the strength of its empirical premises.[49] But, while he is willing to concede that "if there is a clear enough causal connection to rape, a statute that prohibits violent pornography could be a morally legitimate restriction of liberty,"[50] he defines the required causal connection so stringently that it could rarely, if ever, be demonstrated; the relationship between pornography and rape must be akin to incitement under the criminal law.[51]

An argument for criminalizing pornography (or any other form of expression) in Feinberg's terms would have to be structured as follows. Recall that "harm" is defined at the outset of Feinberg's analysis as a wrongful set-back to interests. Rape and other forms of sexual violence undoubtedly qualify as harms in this sense: "Rape is a harm and a severe one. Harm prevention is definitely a legitimate use of the criminal law."[52] Pornography should be criminalized because it sometimes encourages sexual crimes, even though the one who wrongfully sets back the victim's interests will not be the pornographer himself, but a consumer (and probably not all consumers, though that is an empirical question).[53] Society as a whole would be

46. Id. at 148; but cf. The Queen v. Sharpe, 194 D.L.R. 4th at paras 87-94 ("cognitive distortions" induced by child pornography among the grounds for holding that criminalization of possession of child pornography was a justified limit on the right to freedom of expression).

47. Offense to Others, supra note 6, at 149.

48. Id. at 149-50.

49. See id.

50. Id. at 154.

51. Id. at 156-57, 232-43.

52. Id. at 154.

53. Cf. The Queen v. Butler, [1992] 1 S.C.R. 452 (Can.) [Decisions of the Supreme Court of Canada] (accepting an argument of this type to justify the limit on freedom of expression created by the obscenity offenses in the Criminal Code, R.S.C., ch. C-46,

better off without pornography, although many individual instances of the use of pornography are harmless. In Feinberg's terms, pornography causes aggregative harms. Or, as Andrew Kernohan has argued: "pornography, as a cultural practice, reinforces men's false beliefs about sex and male dominance. It contributes to eroticising the submission of women, and indirectly to women's beliefs about their own self-worth. Yet the solitary consumption of pornography by one individual man may be totally harmless."[54] Using Feinberg's terms, Kernohan suggests that "a culture of pornography . . . [is] an accumulative harm, not an individual one."[55] And, as we have seen above, in Feinberg's account, when we are trying to control aggregative or accumulative harms, it is acceptable to define the rights of potential makers and users of pornography in terms of setbacks to the interests of others, rather than in terms of the wrongs they cause directly.

But, Feinberg would reject this argument. His general point against criminalizing expression on this type of ground is as follows:

> Even where there is strong statistical evidence that a certain percentage of communications of a given type will predictably lead the second party to harm third parties, so that in a sense the resultant harms are not "fortuitous," that is not sufficient warrant for prohibiting all communications of that kind.[56]

The difference between consuming pornography and consuming alcohol, then, must have something to do with the status of the activity itself. What is the value of free expression in pornography that enables it to trump the

§ 163 (1985) (Can.) [Revised Statutes of Canada]).

54. Kernohan, supra note 40, at 84-85.

55. Id. at 85. Kernohan makes this argument as part of an interpretation of liberalism that is quite different from Feinberg's. See id. at 88-109. Kernohan is expressly concerned with a problem that Feinberg sets aside at the outset of his project, namely that Mill's recognition that society can be as oppressive as the state. Compare Harm to Others, supra note 4, at 3, with Mill, supra note 6, at 68.

56. Offense to Others, supra note 4, at 155.

harm done to third parties, whereas the value of consuming alcohol in large quantities does not trump the harm done to third parties? Feinberg is deeply committed to freedom of expression and has a lot of very compelling things to say on its application to particular issues,[57] but to a large degree the value of freedom of expression is *assumed* rather than *defended*. In the discussion of pornography, for instance, he says:

> Given that "communication" is a form of expression, and thus has an important social value, obviously it cannot be rightly made criminal simply on the ground that it may lead some others on their own to act harmfully. Even if works of pure pornography are *not* to be treated as "communication," "expression," or "speech". . . , but as mere symbolic aphrodisiacs or sex aids without further content . . . , they may yet have an intimate personal value to those who use them, and a social value derived from the importance we attach to the protection of the private erotic experience.[58]

Note that understanding a right as a way of protecting particularly important welfare interests would not help here, as it is implausible to suppose that the value of using pornography could outweigh the aggregative or accumulative harms it causes to women (assuming, of course, that empirical evidence would support the argument that pornography causes aggregative or accumulative harm). Particularly given Feinberg's insistence that pornography is not art but "simply consists of all those pictures, plays, books, and films whose *raison d'etre* [sic] is that they are erotically arousing,"[59] a deeper justification is required to explain why these values outweigh the harms caused to third parties.[60]

57. See, e.g., id. at 281-87.
58. Id. at 156-57 (emphasis in original).
59. Id. at 142.
60. In other works, Feinberg has used a metaphor of weighing to describe situations in which other values compete with freedom of expression. See Joel Feinberg, Freedom and Fulfillment 124-51 (1992). But even on this account, one of the main values of freedom of expression is "the public interest in the discovery and dissemination of all information that can have any bearing on public policy, and of all

That deeper justification might be sought in a political theory of freedom of expression such as Ronald Dworkin's. For Dworkin, freedom of expression is justified "not just in virtue of the consequences it has, but because it is an essential and "constitutive" feature of a just political society that government treat all its adult members, except those who are incompetent, as responsible moral agents."[61] On this constitutive justification, the government can never infringe upon a person's freedom of speech based on the content of the views expressed, because that would be to deny that the person was a responsible moral agent.[62] Therefore, even if pornography "'silences' women and so decreases their voice and role in democratic politics,"[63] "even if it could be shown, as a matter of causal connection, that pornography is in part responsible for the economic structure in which few women attain top jobs or equal pay for the same work,"[64] criminalization of pornographic expression could not be justified because it would be inconsistent with the liberal democratic commitment to treating everyone as a responsible moral agent.[65]

---

opinions about what public policy would be." Id. at 150. Pornography, in Feinberg's view, contributes little to this value and so its value on this ground could not have much weight as against the harms it causes to women (assuming again that the empirical evidence supports the existence of those harms).

61. Ronald Dworkin, Freedom's Law 200 (1996).

62. Hence, Dworkin's stunned reaction to The Queen. v. Butler, [1992] 1 S.C.R. 452 (Can.):

> The Canadian Court conceded that the effect of its ruling was to narrow . . . constitutional protection [for freedom of expression], but said that 'the proliferation of materials which seriously offend the values fundamental to our society is a substantial concern which justifies restricting the otherwise full exercise of freedom of expression.' That is an amazing statement. It is the central, defining, premise of freedom of speech that the offensiveness of ideas, or the challenge they offer to traditional ideas, cannot be a valid reason for censorship; once that premise is abandoned, it is difficult to see what free speech means.

Dworkin, supra note 61, at 206 (alteration in original).

63. Dworkin, supra note 61, at 205.

64. Id. at 219.

65. See David A.J. Richards, Liberalism, Free Speech, and Justice for Minorities, in In Harm's Way: Essays in Honor of Joel Feinberg 92 (Jules L. Coleman & Allen Buchanan eds., 1994) (offering another account of the importance of freedom of expression, and supporting Feinberg's position on free speech issues).

I do not mean to suggest that Feinberg must adopt Dworkin's understanding of freedom of expression, or even that Dworkin's approach necessarily leads to Dworkin's own conclusion about pornography.[66] I do mean to suggest that some such account is necessary. What Feinberg provides instead is an insistence on the value of individual choice,[67] an important value to be sure, but one that we regularly limit in other criminal contexts and that cannot by itself explain why liberals are reluctant to criminalize pornographic expression.

B. Fraud and Rape

Recall that one of the limits on the concept of harm is that a set-back to interest is not a harm if the one whose interest is set back has voluntarily consented to (the risk of) the set-back. Most of Volume Three of Feinberg's *magnum opus* is concerned with how to define "voluntariness" and "consent" for the purpose of this limit. Here, once again, the question of whether Feinberg can usefully maintain the definition of harm as wrongful set-back to interests must be asked. I will limit myself to Feinberg's analysis of one issue, the effect of fraud on a man's liability for rape.[68]

Traditionally, fraud would vitiate a woman's consent to sexual intercourse, and thus make the man guilty of rape, if the fraud went to the "nature and quality of the

66. Dworkin's statement of the conditions under which criminalizing pornography would be acceptable is quite similar to Feinberg's. See Dworkin, supra note 61, at 219. For what might be called a Dworkinian defense of a criminal statute prohibiting pornography and of the approach in Butler, see David Dyzenhaus, Pornography and Public Reason, 7 Can. J.L. & Juris. 261 (1994); see also Dyzenhaus, supra note 6, at 364-65.

67. Compare Brudner, supra note 2, at 212, with Dyzenhaus, supra note 6, at 355.

68. In Canada, as in some other jurisdictions, the crimes of "rape" and "indecent assault" have been replaced with the general, gender-neutral crime of "sexual assault." Since Feinberg's account is cast in terms of rape alone, I will maintain his terminology. For the purposes of this paper, nothing turns on whether the offense at issue is rape or sexual assault.

act" or to the identity of the man.[69] But fraud concerning the consequences of the act, *e.g.,* whether the parties would marry after the act, or whether the woman would be paid for the act, or even whether a sexually transmitted disease might be communicated, did not vitiate consent.[70] As Feinberg puts it, the traditional view was that fraud in the factum vitiated consent, while fraud in the inducement did not.[71] This seems puzzling. If the prosecution can establish that the woman would not have consented in either case, and if there was dishonesty in either case, why is there criminal liability in one and not the other?[72] Feinberg resolves the puzzle by placing three kinds of non-consensual intercourse on a spectrum of harmfulness. First, rape in its paradigmatic form, involving violence or threats of violence, "is a violent imposition of one person's will on another's, which is not just an alternative means to the same harm as others, but *an important part of the harm itself.*"[73] Second, "[s]exual participation produced by fraud in the factum is equally as involuntary as that produced by coercive violent rape," perhaps more so because "it is totally involuntary, not involving the will whatever."[74] Both of these forms of non-voluntary sexual intercourse are harmful enough to amount to rape. But the third type of non-consensual intercourse, that caused by fraud in the inducement, is different:

> The sorts of fraudulent inducements that would tend to reduce voluntariness most markedly, it is commonly supposed, create at the same time a presumption that the sexual relations that result are *not very harmful*, if harmful at all. It is implausible, for example, to think that a woman who consents because of a false promise of a cash gift has

69. See, e.g., The Queen. v. Maurantonio, [1968] 1 O.R. 145 (Can.) [Ontario Reports].

70. See The Queen v. Clarence, [1888] 22 Q.B.D. 23 (Eng.) [Queen's Bench Division].

71. Harm to Self, supra note 4, at 293.

72. Id. at 295.

73. Id. at 293.

74. Id. at 295.

> suffered a grievous harm consisting in "unwanted sexual intercourse as such." Indeed her consent to the terms that later proved fraudulent indicates that at most the sexual intercourse was an "evil" or a "cost" to her for which a certain modest amount of cash was considered adequate compensation, so it could not have been thought to be a very great evil in itself.[75]

What is the status, in Feinberg's approach, of a claim that a result is "not very harmful"? Recall that "harm" is defined at the outset as a wrongful set-back to interests. Does "not very harmful" then mean "not very wrongful" or "not a serious set-back to interests"? "Rightful" and "wrongful" have a binary quality that is difficult to quantify. Although some wrongs are in some sense worse than others, in most cases that will be because the associated set-back to interests is greater; murder is worse than assault and assault is worse than trespass to property, not because trespass is less wrongful than murder but because the consequence of murder is always a tremendous set-back to the victim's interests, while the consequence of a trespass may be a quite trivial set-back to interests.

So it seems that "not very harmful" here must mean "not a serious set-back to interests." One might well disagree with the argument that engaging in fraudulently induced sexual intercourse is an insufficiently serious set-back to interests to be the subject of the criminal law.[76] But

75. Id. at 300; see also Jeffrie G. Murphy, Some Ruminations on Women, Violence, and the Criminal Law, in In Harm's Way: Essays in Honor of Joel Feinberg, supra note 65, at 209 (offering a related, but not identical approach).

76. The Supreme Court of Canada recently wrestled with this question in the case of an accused man who allegedly failed to disclose his HIV-positive status to two sexual partners. The court held unanimously that if the facts alleged by the Crown were established, the accused would be guilty of aggravated assault, but divided on the reasons for that holding. The majority held that consent to sexual contact would be vitiated by fraud regarding "a significant risk of serious bodily harm." The Queen v. Cuerrier, [1998] 2 S.C.R. 371 at 137. Two concurring judges took a narrower approach, holding that the traditional test of "fraud going to the nature and quality of the act" should be extended to encompass non-disclosure of the risk of sexually transmitted diseases. Id. at pages 58-75. A third concurring judge took a broad approach, holding that fraud vitiating consent would be established by proof "that

once again, the point of this paper is to note that it is the set-back to interests, rather than the combination of wrong and set-back, that seems to be driving the analysis on this point.

IV. A SIMPLIFYING AND COMPLICATING SUGGESTION

Suppose Feinberg were to define "harm" as a "set-back to interests," perhaps with some quantitative dimension, while defining "wrong" as a "violation of right." These definitions would simplify Feinberg's task in one way but complicate it in two others. Notwithstanding the complications, I want to suggest that the overall effect would be to clarify the reasons that Feinberg wants to offer for the positions he takes up.

The simplifying effect of this suggestion would be that Feinberg would be able to consider conduct and ask of it: "Should we make this conduct criminal because it is very harmful?" or "Should we not make this conduct criminal because it is rightful, or because a plausible system of rights would permit it?" While these questions would require independent accounts of "wrong" and "harm," they would for that very reason help to identify qualitatively different kinds of reasons for criminalization.

The two complicating effects are closely related to each other (and indeed to the simplifying effect). The first is that Feinberg would have to be clearer about the source and the role of rights in his analysis, indeed to reconsider his reluctance to engage with natural rights theorists. If, as Feinberg often indicates, the right to engage in a certain activity can trump the direct or indirect set-backs to interests that the activity produces, one is entitled to ask where the right gets this force. The answer would generally have to invoke the role of the right in structuring relationships between individuals or between the individual and the state, and would explain why no gains

the accused acted dishonestly in a manner designed to induce the complainant to submit to a specific activity, and that absent the dishonesty, the complainant would not have submitted to the particular activity. . . ." Id. at.para.16.

in welfare or other consequential measures of value could offset violations of right.[77]

The second complicating effect is that Feinberg would have to develop an account of which rights can be limited in the cause of preventing set-backs to interests, and under what circumstances. It might be said that Feinberg already does this; indeed that the entire work provides just such an account. There is a way in which this is true; much of Feinberg's work could be *transformed into* such an account, but my concern is that this aspect of Feinberg's account is often implicit or assumed rather than explained and argued.

I want to suggest that notwithstanding the complicating effects, the suggested redefinition of terms would be beneficial to Feinberg's project. By clearly separating questions about rights from questions about set-backs to interests, one can isolate the reasons for and against criminalization more clearly than Feinberg's harm principle does.[78] For example, what Feinberg calls

77. There are many different theories which might fall under the umbrella of natural right theories, and I do not mean to suggest that there is only one such theory or that Feinberg need adopt any one in particular. For three rather different versions of the kind of theory I have in mind here, see Ronald Dworkin, Taking Rights Seriously 81-130 (1977); John Rawls, Political Liberalism 173-211 (1993); Ernest Weinrib, The Idea of Private Law 56-114 (1995). Richards has previously urged Feinberg to go in a Rawlsian direction. See David A.J. Richards, The Moral Foundations of Decriminalization, 5 Crim. Just. Ethics, Winter/Spring 1986, at 11, 13 (arguing that Feinberg's harm principle can be grounded in Kantian constructivism rather than in utilitarianism). Feinberg regarded this suggestion as "tempting" but unnecessary to his project. Joel Feinberg, Harm to Others—A Rejoinder, 5 Crim. Just. Ethics, Winter/Spring 1986, at 16, 24. Feinberg's discussions of rights in his other works show that he does not regard rights as mere instruments for utility maximization or some other goal, but it is not clear how those discussions would interact with the problems discussed in The Moral Limits of the Criminal Law. See Feinberg, Freedom and Fulfillment, supra note 60, at 97-259; Joel Feinberg, Rights, Justice, and the Bounds of Liberty 130-58 (1980).

78. Andrew von Hirsch's restatement of Feinberg's approach does that; significantly, von Hirsch uses the word "harm" where Feinberg's definition would, I think, demand the phrase "setback to interests." See Andrew von Hirsch, Extending the Harm Principle: Remote Harms and Fair Imputation, in Harm and Culpability, supra note 16, at 259, 261. I take it that Hampton was also urging Feinberg to be more explicit about the normative elements of his project that were not strictly concerned with preventing set-backs to interests. See Hampton, supra note 2.

"aggregative" and "accumulative" harms have to be understood as set-backs to interests, and if we see them this way, it is easier to see what we are gaining and losing if we define rights for the very purpose of controlling them rather than defining rights as trumps. If this way of defining rights seems to grant too much to an instrumental understanding of the value of rights, and to grant too little to the importance of rights in our political and legal culture, then a different sort of account of rights is required.

V. CONCLUSION

The word "harm" in Feinberg's "harm principle" is sometimes interpreted as "set-back to interests." It is clear that this interpretation is mistaken; Feinberg's harm principle is centrally concerned with harm understood as wrongful set-back to interests. But, as Feinberg's analysis proceeds, the mistaken interpretation begins to become quite plausible as the "wrongfulness" element of the definition of harm tends to drop out (as in the discussion of aggregative and accumulative harms), or be taken for granted (as in the argument against criminalizing pornography). Thus, contrary to Feinberg's stated intention, the idea of right tends either to become simply instrumental to the prevention of set-backs to interests or to be insufficiently analyzed. I have tried to suggest that Feinberg's already rich discussion of the arguments for and against criminalizing various types of conduct could be enriched further by not allowing the idea of right to be submerged in this way, so that Feinberg could more explicitly engage with non-instrumental accounts of the value of rights.