HOW YOU CAN BE BOTH A LIBERAL AND A RETRIBUTIVIST: COMMENTS ON *LEGAL MORALISM AND LIBERALISM* BY JEFFRIE MURPHY

Jean Hampton*

One of the questions Jeffrie Murphy puts to us in his paper *Legal Moralism and Liberalism* is whether a liberal society that endorses *the harm principle*, i.e.: "that legal coercion is justified only to prevent one person from harming another [nonconsenting] person," can also endorse *the retributive principle*, i.e.: "that punishment of the guilty is justified insofar as the guilty deserve this response."

Murphy answers that question with a "no." However, his paper does not actually *show* any inconsistency between the harm principle and retribution. Instead it (merely) gives a series of reasons why it seems that such an inconsistency exists. In these comments I will argue that Murphy cannot show that there is an inconsistency between the harm principle and retribution because, in fact, there is no such inconsistency, and I will establish this using some passages from the writings of Joel Feinberg who, I think, understands this point. However I shall go on to argue that Murphy's paper does raise the worry that one particular *motivation* for endorsing the harm principle is inconsistent with retribution. As we shall see, however, it is possible to motivate both the harm principle and retribution such that they are not only consistent, but also mutually supportive. Indeed, although Feinberg dislikes what might be called "foundational theorizing" on moral matters, he says enough in his four volume work *The Moral Limits of the Criminal Law* to allow us to identify his form of liberalism as one that can consistently accommodate both retribution and the harm principle.

I. WHY THE TWO PRINCIPLES ARE NOT INCONSISTENT

The best place to look for an account of why retribution and the harm principle aren't, as they are stated above, inconsistent is Feinberg's fourth volume of *The Moral Limits of the Criminal Law*,[1] in which he notes that not only retribution but even the more "morally extreme" theory of punishment that I have called the "moral education" theory is consistent with the harm principle. Consider that when they have objected to legal paternalism, liberals have rejected only a view that allows the state to restrict what an individual can

* A.B. Wellesley, 1976; M.A. and Ph.D. Harvard, 1980; Professor of Philosophy, University of Arizona.
    1. JOEL FEINBERG, HARMLESS WRONGDOING, 4 THE MORAL LIMITS OF THE CRIMINAL LAW 300-305 (1988).

do to himself; they have not objected to the view that when the state justifiably punishes a person after he has broken a law whose content is only concerned with harm or offense to others that it should have a moral purpose.[2] Moral education theorists and retributivists answer a logically different question from the one that concerns legal paternalists and anti-paternalists: Whereas legal paternalists and anti-paternalists answer the question, "What ought to be made law?", the moral education theorist and retributivist answer the question, "What moral attitude should the state take toward someone who has violated its criminal code?" Someone can answer the first question by insisting that only (what Feinberg calls) "grievance morality" is the appropriate subject for criminal legislation, where grievance morality is defined by the harm principle. But grievance morality is still *morality*, and hence one can also insist that when the state punishes a person who has violated grievance morality, it can consider that person morally blameworthy and deserving of a moral response implicit in the punishment, where that moral response can be either retributive in nature, or morally educative in nature, or both.[3]

Feinberg himself endorses a kind of (expressivist) retributivism, and therefore makes the following (remarkably) strong endorsement of a moral role for the state in punishing and responding to criminals:

> The liberal...can and must concede that the criminal *process* in its very conception is inherently moral (as opposed to nonmoral)—a great moral machine, stamping stigmata on its products, painfully "rubbing in" moral judgements on the persons who had entered at one end as "suspects" and emerged from the other end as condemned prisoners. The question the liberal raises about this moral machine is: "which actions should cause their doers to be fed into it?", and his answer is: "only those actions that violate the rights of others." There is no doubt in his mind that the law may "enforce morality". The question is *"which* morality (or which sector of morality) may it properly enforce?", and he restricts the criminal law to the enforcement of "grievance morality".[4]

So Feinberg, who is arguably the strongest endorser of the harm principle in our time, nonetheless calls the process of the criminal law "a great moral machine." The Feinberg liberal does not disagree with the legal paternalist about *whether* to "enforce morality"; he disagrees with the legal paternalist about *which* morality to enforce.

To put the point as succinctly as possible, there is no inconsistency between the harm principle and the principle of retribution (or, for that matter, moral education) because the harm principle is a theory of the content of legitimate law, and the retributive principle is a theory of the justification of punishment for (legitimate) law breakers.

---

2. *See* Jean Hampton, *The Moral Education Theory of Punishment*, 13 PHIL. & PUB. AFFAIRS 208-38 (1984).

3. *See* FEINBERG, *supra* note 1, at 302. He even wonders whether the moral education view presupposes a denial of legal paternalism, given that the theory wants to make the lesson that harm to others is wrong the *point* of the punishment, and it would not make sense for the point of the punishment to be that one should refrain from harms to others for one's own sake. Although rights-violations are bad for the agents who perform them, that kind of harm doesn't make such actions morally wrong—what makes them wrong is the fact that they harm other people. Hence this is what a retributive punishment would be expressing. *Id.*

4. Joel Feinberg, *Some Unswept Debris from the Hart-Devlin Debate*, 72 SYNTHESES 260 (1987).

So why did Murphy think that there was any inconsistency between these two principles?

II. NEUTRALITY

I believe it is not really the harm principle itself, but what might be thought of as its philosophical underpinnings, that motivate Murphy's worries that there is some kind of inconsistency between that principle and retribution. So our next question is: does the philosophical defense of the harm principle contain something in it that would also, as it happens, rule out retribution? Since Feinberg is concerned with middle-level theorizing, and not with foundational moral justifications for the principles he endorses, even if he is right that, as stated the two principles aren't inconsistent, might he nonetheless have missed something about the proper philosophical foundation for the harm principle that (unfortunately) rules out the legitimacy of retributive punishment?

Murphy's remarks suggest that the reason why a society might decide to endorse the harm principle as the principle governing the proper construction of criminal legislation is that it accepts the view that a liberal society must, as far as possible, be neutral between competing moral and religious conceptions, in its coercive practices. Call this understanding of the proper role of the liberal state "neutrality liberalism."

Now there are different versions of neutrality liberalism, and different slogans associated with them. One slogan, used often by Ronald Dworkin, is that the liberal state should be neutral toward different conceptions of the good. However, Dworkin also assumes, and sometimes argues quite passionately, that the liberal state should *not* be neutral toward different conceptions of the right. Dworkin's commitment to justice as the foundation of the liberal state motivates much of his theory of judicial reasoning and decision-making, and explains his strong objections to the communitarian-style political theorizing of people such as Michael Walzer. Indeed, in *A Matter of Principle* Dworkin objects to Walzer's *Spheres of Justice*[5] because it is insufficiently committed, in Dworkin's view, to the idea of an objective and universal concept of justice as the basis for any well-functioning legal system:

> [I]t is part of our common political life, if anything is, that justice is our critic not our mirror, that any decision about the distribution of any good—wealth, welfare, honors, education, recognition, office—may be reopened, no matter how firm the traditions that are then challenged, that we may always ask of some settled institutional scheme whether it is fair. Walzer's relativism is faithless to the single most important social practice we have: the practice of worrying about what justice really is.[6]

But anyone who bases a liberal state on the pursuit of "what justice really is" has plenty of room to accommodate retribution. Such a Dworkinian liberal is ready to be neutral about all sorts of things, but not justice, and if justice includes *retributive* justice, then a just liberal state that is neutral between competing goods is nonetheless required, by virtue of its commitment to being just, to punish retributively.

---

5. MICHAEL WALZER, SPHERES OF JUSTICE (1983).
6. RONALD DWORKIN, A MATTER OF PRINCIPLE 219 (1985).

So we need a more radical form of neutrality liberalism to endanger retribution, and something like the theory we need is suggested in certain writings of John Rawls[7] (but not, as I'll explain, explicitly put forward in his book *Political Liberalism*[8]). On this view, the essence of a liberal state is its acceptance of the principle of toleration: in particular, divergent lifestyles, and different religious, moral and metaphysical commitments are to be tolerated in a liberal state, and this diversity protected, so that the state must reject the idea of operating from any particular metaphysical (and potentially sectarian) moral or religious doctrine, and should instead operate from and seek to develop an "overlapping consensus" among the state's disparate groups on principles of justice, which would be the basis for legislation and for the adjudication of conflicts. Such a consensus in a pluralistic society will serve as the intellectual cement binding the community together and will be supported by each group in the community for different reasons in virtue of their different moral, religious or metaphysical beliefs and practices. This consensus is, at the very least, a *modus vivendi* for a pluralist community, but it is, at least arguably more than that, since its support from each group in the community derives from the fact that each thinks (albeit for different moral or religious reasons) that the consensus is *correct*, as well as useful. So on this view, the state which operates from an overlapping consensus and which strives in all its dealings with the many different groups composing it to respect their different views and lifestyles, is properly called "liberal" precisely insofar as that state eschews commitment to any particular moral or religious code. Henceforth, I'll refer to this sort of view as "radical neutrality liberalism."

Is radical neutrality liberalism inconsistent with retributive punishment? It seems so, because such a neutral society would appear to be required to eschew *any* kind of moral role for itself, because such a role means adopting a moral theory which defines and justifies such a role. On this view, the liberal state should strive to operate in a way that excludes the appeal to moral ideals, so that it can avoid improperly siding with only some of the many different moral and religious communities in that society. To be properly neutral, and thus tolerant of the variety of moral and religious beliefs of its members, this conception of liberalism requires that the state not explicitly take on any role driven by a particular—and inevitably contestable—moral doctrine, nor any particular—and inevitably contestable—metaethical position on the possibility of an objective moral theory which would be used either to develop the notion of retributive desert or to condemn criminal behavior as (objectively) wrong.

So for this kind of liberal, the idea that the state ought to inflict retributive punishment, understood as I have defined it, is a real problem. It is not merely that some members of the society may have moral views that oppose punishment so conceived. More fundamentally, if the state understands itself as the institution which inflicts deserved suffering upon wrongdoers (where it was either obliged or required to do so) it would be invoking moral notions and undertaking a moral role as it punished. As it acted upon that role, it would be committed to the idea that those it punished had not only acted harmfully but

---

7. This view is suggested in John Rawls, *Justice as Fairness: Political not Metaphysical*, 14 PHIL. & PUB. AFFAIRS 223-251 (1985), and John Rawls, *The Domain of the Political and Overlapping Consensus*, THE IDEA OF DEMOCRACY 245-269 (David Copp, Jean Hampton and John E. Roemer eds., 1993).
8. JOHN RAWLS, POLITICAL LIBERALISM (1993).

also *wrongfully*, such that they were morally blameworthy for their actions, and it would be committed to the idea of moral desert, an idea which would guide the way in which it designed and implemented punishment sentences. So not only would its punishment activities be concerned to effect social order; more controversially the state would be aiming to secure a moral result, namely, the infliction of suffering upon all and only those who deserve it in view of their wrongful actions. And when the state becomes such a "servant" of morality, it would seem to become a servant of a particular theory of what morality involves, not only because it would categorically reject those conceptions of the good which dictate conduct it regards as wrong and deserving of a retributive response, but also because it would be committed both to the idea of retributive desert and the idea that there is a right moral answer to questions about acceptable conduct. So if the state accepted that it should punish retributively, it would not only be partisan in the way it thought about retribution, but more fundamentally, it would be partisan in thinking that there were objectively right moral answers about conduct, about human value, and about the appropriate response to moral injury, and thus it could not claim to be neutral in its dealings with all its citizenry, some of whom would likely disagree with one or all of these ideas.

So what theory should animate the punishment of the radically neutral liberal state? Some people might believe that it fits best with a deterrence theory of punishment, insofar as that theory allows the radically neutral state to legally prohibit and punish anti-social behavior that people in the society don't like (no matter what their moral or religious views)—where their reasons for not liking this behavior may have nothing to do with morality or religion. Hence on this view, punishment by the liberal state aims to stop and deter such behavior, and nothing more. So understood, there is no contestable moral role involved in the punishment, and no inadvertent endorsement of moral ideas that are rejected by some members of the pluralist society governed by this state. Hence it is a conception cleansed of partisanship and, it seems, perfectly suited to the task of justifying the punishment activities of a liberal state.

However, isn't deterrence itself a partisan theory of punishment? I proposed above that a deterrence theory may seem appropriately neutral, and hence a suitable justification for punishment in such a state, but consider that even a deterrence justification is controversial insofar as some people believe that a purely deterrent response to a law-breaker is morally disrespectful to that person, treating him more like a dog to be conditioned than a human being who acts on reasons.[9] For this reason, and for the reason that the deterrence theory could justify punishing the innocent and, more generally, punishing people in ways that are not proportionate to their offense when doing so has desirable deterrent effects in the society, many people in our society (including me) would reject the idea that punishment must only be justified and constructed on deterrence grounds (even if most of us would accept that the deterrent effects of any punishment system are relevant to its justification and design). So it would seem that the liberal state cannot be neutral if it only relies on the deterrence theory in its justification of punishment.

---

9. This idea was first suggested in HEGEL'S PHILOSOPHY OF RIGHT 246 (T.M. Knox trans., 1st ed., 1976).

But if it cannot be neutral punishing people for deterrence, retributive, or moral education reasons, then how can the radically neutral liberal state punish at all? It cannot say "there is an overlapping consensus on the fact that the guilty should be punished, so that entitles us to do so", because *how* it does so is a function of the punishment theory it accepts, and any punishment theory in our times is going to be supported by only some of the many groups in our pluralist society.

Does that mean a radically neutral state cannot punish its citizens because it cannot do so neutrally? I take that to be a disastrous implication for this kind of liberalism, for how can a viable theory of liberalism imply that we not punish those who break criminal laws? However, note that even a decision *not* to punish is not neutral between competing moral and political conceptions, insofar as it is a policy that opposes the moral and religious views of those who insist that punishment of the guilty is morally desirable or necessary.

Given these problems, we may begin to wonder whether (unlike theorists such as Dworkin, who eschew radical neutrality) those who support the idea that the liberal state must be radically neutral have a coherent conception of neutrality on which to base their policy recommendations. Can that conception be developed so as to insure that the view generates coherent policy recommendations? If so, will that development allow the neutrality liberal to accommodate retribution?

We can pursue these questions by considering John Rawls' recent book *Political Liberalism*, which attempts to define a conception of neutrality liberalism that is more sophisticated than the conception I have just sketched. While he advocates that the state take a neutral attitude as it governs its citizens, he requires this neutral attitude only between *reasonable* competing conceptions among the citizenry. Does this qualification enable him both to insure that radical neutrality liberalism is conceptually coherent and to find room for retribution after all?

To answer this question, we must explore what the qualification means.[10] Like other liberals, Rawls accepts that freedom and equality are the two ruling concepts in a liberal political regime. Nonetheless Rawls says he wants to affirm them merely as "political" values, and avoid endorsing them as part of some kind of comprehensive or partially comprehensive moral view. They are supposed to be "free-standing" values that are implicit in the political culture of the liberal society. Perhaps most importantly, Rawls says that a liberal society's commitment to these values should *not* be a function of the fact that this society believes it has in hand a moral argument establishing them as *true*. While these values are supported by what Rawls terms "public reason" in a liberal society, this reason is not understood as the discloser of universal moral truth, nor the vehicle through which citizens of a liberal society should seek to get access to the moral values that should animate it.[11] Instead, public reason is something common to all members of the society because it is constructed out of commonly accepted values. Public reason involves appeal to fully public

---

10.   I discuss this Rawlsian term at much more length than I am able to do here in Jean Hampton, *The Common Faith of Liberalism*, 75 PACIFIC PHIL. Q., (forthcoming, 1995).

11.   For a discussion of this conception of reason, see Richard Rorty, *The Priority of Democracy to Philosophy*, THE VIRGINIA STATUTE FOR RELIGIOUS FREEDOM 262, 268 (Merril D. Peterson & Robert C. Vaughan eds., 1988).

reasons to which all citizens appeal in their political arguments, and that are concerned with the good of the public.[12] The implication of Rawls's discussion is that such reasons are commonly accepted because there is an overlapping consensus on them, whereas traditional liberals assume that the reasons to which people appeal in political argumentation are commonly accepted because reasoning has shown them to be "right." A liberal society, says Rawls, needs "a conception of justice that may be shared by citizens as a basis of reasoned, informed, and willing political agreement."[13] Once constructed, such a conception "expresses their shared and public political reason."[14]

So on Rawls' view, liberals are not supposed to say to one another, "It is a moral fact that x; therefore we must have policy y." Instead, they are supposed to work together as reasonable people to develop an overlapping consensus, accepting that reasoning yields plural results, and looking for political values and policies that all of them, despite their different starting points and moral views, can accept.

Moreover, even though public reason is constructed out of the beliefs of the citizenry, this public reason is supposed to be vigorously non-partisan:

> [T]o attain such a shared reason, the conception of justice should be, as far as possible, independent of the opposing and conflicting philosophical and religious doctrines that citizens affirm. In formulating such a conception, political liberalism applies the principle of toleration to philosophy itself. The religious doctrines that in previous centuries were the professed basis of society have gradually given way to principles of constitutional government that all citizens, whatever their religious views, can endorse.[15]

Such non-partisanship makes a liberal state governed by this public reason tolerant and respectful of the competing moral views of all its citizenry.

What are the implications for Rawls's political liberalism on the issue of the neutral liberal state's punishment policies? Do we get a clear answer to the question "Can the state punish, and if it can, how must it do so?" It seems that we do, insofar as we can say that in a Rawlsian state that is properly neutral punishment policy is defined by the shared public reason in the community. And this seems to mean that if that shared public reason authorizes retributive punishment, then the state is acting properly (and neutrally) by inflicting it on its criminal law-breakers. Of course there seems nothing necessary nor inevitable about the affirmation of retribution, so the Rawlsian state would be precluded from punishing retributively if shared public reason did not endorse it. So even though there is nothing in Rawls's view that would require a liberal state to embrace retributive principles, it would seem possible for it to do so.

One might worry, however, that because no theory of punishment is endorsed by all reasonable conceptions now in existing liberal societies, that whatever counts as "shared public reason" in these societies now cannot include any particular theory of punishment. Perhaps Rawls would want to suggest that only views that endorse a particular theory of punishment (but which one?) could be considered "reasonable," and hence worthy of being contributors to

---

12. JOHN RAWLS, POLITICAL LIBERALISM 213 (1993).
13. *Id.* at 9.
14. *Id.*
15. *Id.* at 9-10.

the "shared public reason" on this issue. But I see no way to develop his notion of the 'reasonable' to get this result without violating the attitude of moral and religious tolerance that Rawls is aiming to defend—nor do I know, just from the term, which of the many theories of punishment that have been defended over the years, counts as "reasonable."

I see only one way out for Rawls. He might maintain that implicit in our legal system right now, as it happens, is a shared conception of what the state is doing when it punishes criminals, that includes a (perhaps not very well-worked-out) amalgam of deterrence, educative and retributive ideas, to which all reasonable citizens (more or less) are committed. Certainly it is true that our punishment system *is* a mishmash of all these ideas. It is arguable whether this inconsistent mishmash can qualify as something called "public reason," and even more arguable (indeed, to many, alarming) whether something called "public reason" could simply be defined by looking at status quo thinking. But aside from these worries, note that this way of interpreting Rawls would show him to be endorsing at least a partial retributive role for our liberal society, consistent with his commitment to the neutrality of liberal states.

These remarks should illustrate that those who push the idea of neutrality liberalism will only have a plausible and coherent view if they develop a conception of what it is to be "neutral" in a pluralist society that allows the state to act in *some* way. To disallow the state to take sides is to effectively paralyze it. The trick is to figure out how the state can take sides, and yet remain tolerant. I suggest that if this trick can be pulled off, there is no particular reason to think that retribution should not be endorsed as at least part of the neutral liberal's state's justifying aim.

The issue of the coherence of neutrality liberalism is raised in a different way by reflecting on whether the neutral liberal state could recognize the harm principle. Initially it would seem that an advocate of this view could develop a criminal code using the harm principle as long as there was an overlapping consensus (in Rawls's case, among reasonable views) on this principle. However, in order to be effective in guiding legislation, that consensus could not be general and unspecified, but would have to be detailed enough to cover the meaning of "harm" so as to serve as a theoretical guide for criminal legislation. But it would be difficult—nay impossible—for a pluralist society such as ours to generate any overlapping consensus on what counts as harm. For example, understanding harms as defined by reference to "rights"—a term that Murphy himself uses in one of his formulations of the harm principle— seems precluded, because the existence and content of rights is hotly contested in our pluralist society. Moreover, what counts as a harm is fraught with moral controversy. To illustrate this with a contemporary example: consider some feminists' claim that pornography "harms" women, even if it does not cause violence against their person, because it lowers their self-esteem and fosters demeaning treatments of them. This claim is countered by those who believe that the only relevant sense of "harm" that should govern legislation is damage or threatened damage to the body or possessions of people. Is a society that, noting the controversy surrounding this question, decides to allow the pornography insofar as it has not been shown to encourage violence against women's persons or property, being "neutral"? While it is true that virtually everyone agrees that damage to a person or property counts as harm, if

legislation could only be based on this commonly held belief, it would *not* be neutral between the moral views of feminists and their pornography-tolerant opponents.

So again we see that this version of neutrality liberalism may not be a coherent position, in which case its failure to accommodate both retribution and the harm principle is philosophically unimportant since it would be a bankrupt conception of liberalism.

However, I also wonder whether it is even possible for radically neutral liberals or Rawlsian neutrality liberals to endorse the harm principle (even one that is only vaguely stated). If Murphy is right to understand Lord Devlin's legal moralism as endorsing the legitimacy of legislating against certain actions based upon people's *shared moral beliefs*, then both these versions of neutrality liberalism (but again, *not* Dworkin's brand of neutrality liberalism) would seem to endorse a variant of such legal moralism, namely, the view that it is legitimate to legislate against certain actions based upon people's shared (and, in Rawls's case, reasonable) moral beliefs in an overlapping consensus. So, if that consensus included, say, the belief that homosexual acts were so intrinsically immoral that they should be coercively suppressed, then they would seem to be an appropriate subject in *this* society for criminal legislation. Rawls might want to argue that any belief that homosexual acts were "intrinsically immoral" could not be reasonable, but I see no substantive conception of the reasonable in his work on which that argument could be based, and certainly history gives us plenty of examples of societies in which the belief that actions of this sort were intrinsically immoral was widely (sometimes nearly universally) shared.

So ironically, radical neutrality liberalism and Rawlsian political liberalism, may be, contra Murphy, better at accommodating retribution than the harm principle! Fortunately for advocates of both the harm principle and of retribution, there are other forms of liberalism! How do they fare in consistently accommodating both principles?

III. VARIETIES OF LIBERALISM

In his paper Murphy mentions another version of liberalism, which he calls, after Joseph Raz, perfectionist liberalism, and which he argues, is able to accommodate retribution, but probably not the harm principle. But another form of liberalism whose foundations are sufficient to motivate both the harm principle and retribution is the sort endorsed by Joel Feinberg.[16] I call it "rights-based liberalism". This view invokes the moral idea of "rights" in order to legitimate and define the harm principle, arguing that state power must be used to protect rights of a certain kind, which I will henceforth call autonomy rights. The harm principle requires that laws can only be enacted which prohibit behavior that violates these, and only these, rights. Feinberg himself argues that because a rights-based liberalism is unafraid of incorporating explicit moral ideas in its foundation, and indeed, welcomes the idea that the liberal state has a certain kind of moral role to play in the community, it can include a retributive conception of punishment that grants the state a moral role in its punishment activity. Starting with the (by now) familiar distinction between the general justifying aim of an institution and the rules of fair

---

16. Feinberg, *supra* note 4, at 249-275.

procedure governing its actions, Feinberg maintains that when we apply this distinction to the criminal law we find that

> it is a misrepresentation of the liberal position (at least as I have tried to formulate it) to say that it ascribes an entirely nonmoral justifying aim to the criminal law. There is a clear respect in which the liberal's liberty-limiting principle *is* a moral one. The justifying aim of the system of criminal law, on his view, is not merely to minimize harms, in the sense of setback interests, all round. If that is what he advocated he would have no quarrel with the legal paternalist. In fact, his principle permits prohibitory statutes only when necessary to prevent those harms (and offenses) that are also *wrongs*: those that are unconsented to, involuntarily suffered, and neither justified nor excused. The criminal law, he insists, must serve a profoundly moral purpose, namely the protection of individual's [sic] moral rights."[17]

Whereas a neutrality liberal has to attribute to any modern pluralist liberal state a non-moral understanding not only of its punishment actions but also of the offenders' acts which warrant it, Feinberg appreciates that a rights-based liberal is free to appropriate moral language to characterize criminal actions not merely as harms but as wrongs, insofar as these acts are violations of rights. Hence Feinberg's liberal state is unabashedly taking ethical positions: some conceptions of right and wrong are rejected out of hand, and lawbreakers who perform criminal actions that demonstrate they subscribe to what this state determines to be false moral views would be told, via punishment, that both their views and their actions following from these views, are unacceptable. As Feinberg puts it, "even a penal code based exclusively on the harm principle (and any penal code will be *largely* based on that principle) is meant to do more than merely prevent harm. In so protecting people, it also means to vindicate the morality of preventing harm and respecting autonomy."[18] Feinberg goes on to attribute to his liberal state a retributive response to wrongdoers, which involves morally condemning the criminal through its punishment, "expressing public reprobation and moral censure of the harm-causing wrongdoer."[19] And as we have already noted, this advocate of the harm principle regards the state as a great "moral machine."

Those attracted to radical neutrality liberalism or political liberalism will object to the fact that such partisan moral positions are taken by Feinberg's liberal state (but note again that Dworkin is one liberal who will not so object). Aside from the charge of partisanship, these liberals may also charge that Feinberg's position makes the state *paternalistic*, in a way that violates the value of freedom fundamental to liberalism. A rights-based liberal state not only preaches but also enforces a moral code, in just the way that parents do as they raise children. Once he attributes such a role to the state, hasn't the rights-based liberal fatally compromised the idea that the state should promote the autonomy of its citizens and behave in a way that is tolerant of their competing views?

Feinberg has an interesting argument denying this contention. What makes a state "liberal" for him is not its rejection of any moral ideas as the foundation for its criminal law, but its rejection of the idea that any enforcement of moral behavior should include punishment of immoral behavior

---

17.   *Id.* at 257. *See also* FEINBERG, *supra* note 1 at 11-14.
18.   FEINBERG, *supra* note 1, at 12.
19.   *Id.* at 12.

which nonetheless has no victim other than the offender himself. So it is endorsement of the harm principle that is, for him, the marker of a state's liberalism, and not any eschewal of a moral role. Even though the liberal state is in the business of enforcing morality, according to Feinberg, it is not in the business of enforcing those aspects of morality which do not involve the rights of others. The term 'paternalistic', on his view, refers only to states that take on the latter role, not to states that take on the former role. As he explains:

> The liberal does *not* urge that the legislators of criminal law be unconcerned with "a man's morals." Indeed, everything about a person that the criminal law should be concerned with is included in his morals. But not everything in a person's morals should be the concern of the law, only his disposition to violate the rights of other parties. He may be morally blameworthy for his beliefs and desires, his taboo infractions, his tastes, his harmless exploitations, and other free-floating evils, but *these* moral judgments are not the business of the criminal law.[20]

So Feinberg believes that those who participate in and theorize about the liberal state are involved up to their necks in morality. But having said this, Feinberg wants to disassociate himself from the position that the moral role of liberal states includes enforcing victimless morality, or, as he puts it, "nongrievance" morality. So he tries to mark out a middle ground upon which to build his conception of liberalism, one which acknowledges that the state is in the business of enforcing morality—but only a "grievance" morality.

Most liberals in our history have endorsed something like Feinberg's rights-based conception of liberalism (including Locke, Rousseau, Kant, and Mill) although they have disagreed with him and with one another about the best moral theory to use to define and vindicate these rights (and in Mill's case, vigorously denying that rights should be taken to be in any way fundamental from a moral point of view, making it problematic to label his version of liberalism 'rights-based'). For any person who has this form of liberalism, justifying retributive punishment under a liberal regime is straightforward. And retributive ideas appear explicitly in the writings of Locke and Kant. Indeed, it is surely not an accident that the clearest proponent of retribution in philosophical history, namely Kant, is also one of the most ardent supporters of the liberal conception of the state.

Kant is also famous for arguing that retributive punishment is justified because it responds to wrongdoers in a way that honors their rights as persons, and expresses the fact that their criminal behavior is not only societally harmful, but more importantly violative of behavior which other individuals in the society have a right to expect. On this view, retributive punishment (perhaps along with morally educative punishment) is properly responsive to those individual rights that are also part of the moral foundation for the harm principle. Such a view makes the philosophical foundations of the harm principle and retribution not only consistent, but also mutually supportive.

IV. LIBERALISM

The preceding discussion has tried to say something about the relationship between retribution and the harm principle, and also about the nature of liberalism. There are lots of liberalisms—I have mentioned four in

---

20. Feinberg, *supra* note 4, at 259.

this paper (three forms of neutrality liberalism, perfectionist liberalism, and rights-based liberalism), and I could have mentioned more (and in particular, I could have differentiated a number of forms of "rights-based" liberalism). None of us should confuse certain (presently trendy) form of neutrality liberalism, which I have argued are beset with theoretical problems, with all the forms that have been developed in the last few hundred years. Whereas the question: can the harm principle be reconciled with the retributive theory of punishment is a well-formed question (with the answer "yes"), the question "can liberalism consistently accommodate retribution" is not, because the answer to the question depends on what one means by "liberalism."

These various kinds of liberalisms are importantly different from one another. So why do they all count as forms of something called "liberalism"? In another place, I've tried to pick out common features that all forms of liberalism have (such as a commitment to the idea that all human beings are free and equal),[21] and neither the harm principle nor retribution is amongst them. Indeed, it is relatively easy for us to think of liberal theories that endorse neither. Now perhaps I was right that there are unifying ideas in all these liberal theories, but lately I've been wondering whether even that is so, and whether the reason political philosophers are so apt to call their theories of legitimate state power "liberal" is that it is a success word of the first order among a considerable part of the philosophical community. Since, for many it is *de rigueur* to put forward a "liberal" view of state power, accordingly many people say that they do, and yet what they put forward varies enormously. I'm wondering, therefore, if this word 'liberal' is starting to get in the way, leading us to think that we are more unified than we are, and causing some of us to lose track of the real issue that we should be pursuing, which is determining what theoretical components a *good* theory of state power should have. Whether such a theory is "recognizably liberal" is, I would suggest, beside the point.

---

21.   *See* Hampton, *supra note* 10, at 259.