

HARM AND MORALITY REVISITED: A RECONCEPTUALIZATION OF THE HARM PRINCIPLE

The debate between legal moralists and liberals has typically centered on a distinction between “harm arguments” and “morals arguments.” Liberals argue that the only acceptable reason to criminalize conduct is that it causes someone harm. Legal moralists agree that harm is *one* reason to criminalize conduct, but they also argue that if an act is immoral, this fact may be a reason to criminalize the conduct. In the first part of this paper, I argue that this traditional distinction between harm and morality is flawed and suggest a new way of conceptualizing the distinction between harm and morality. I argue that harm arguments and morals arguments are not different categories of argument. Rather, a harm argument is a *type* of moral argument, and the concept of harm is a rhetorical device for making a specific type of morals argument. If I am correct, legal liberals should not base their arguments for the “harm principle” on the notion that “harm arguments” are of a different category than moral arguments. They are not.

In the second part of this paper I will explore possible justifications for the harm principle once “harm arguments” have been reconceptualized as morals arguments. I will argue that three ideas best explain the appeal of the harm principle: First, the “harm principle” does moral work similar to John Rawls’s “thin theory of the good,” producing arguments that are likely part of an “overlapping consensus.” Thus, the notion of harm acts as a sort of “filter,” allowing some concepts of morality into political debate, while keeping out others. Second, the harm principle allows citizens to justify laws to each other using terms that everyone can accept. Using this rhetoric lets citizens respect each other as moral equals, an idea at the heart of liberalism. It also tracks Amy Gutmann’s and Dennis Thompson’s notion of reciprocity, which they argue is central to “deliberative democ-

racy.”¹ Finally, the harm principle is a useful *linguistic* device because it allows citizens in a democracy to express moral ideas without using contentious moral language. In my conclusion, I raise questions about the viability of these three proposed justifications.

I. HARM VERSUS MORALITY

There is no general consensus on the meaning of morality, and precisely what it means to call an act *immoral* is beyond the scope of this paper. My goal in this section is to show that no matter what a person means when he calls an act *immoral*, his concept of harm will depend—at least in part—on his concept of morality. Specifically, a “harm” argument is not a different *category* of argument; it is merely a *type* of morals argument.

To show why I come to this conclusion, I offer an example. Imagine that I feed a person poison that causes him to be unable to hear. Assuming that the poison does not also have some countervailing beneficial effect, everyone would agree that he has been harmed. But why? I see two plausible explanations. The first would go like this: Most people prefer to be able to hear over not being able to hear. Therefore, we assume that the poisoned person also preferred to be able to hear, and since I gave my victim results he did not like, I have harmed him. Notice that in this explanation, to be harmed is simply to receive a result that one dislikes. A second explanation would go like this: Most people would say that other things equal, it is better to be able to hear than not to be able to hear. Being able to hear enables us to accomplish ends that are morally worthwhile. For example, it allows us to carry on conversations with others, which allows us to build friendships, and which most of us believe to be a morally worthwhile activity. By causing the man to be unable to achieve morally worthwhile goals that he would otherwise have achieved, I have harmed him.

Although it is initially plausible, the first explanation cannot adequately describe what legal liberals mean by “harm” because it links harm too closely with emotions. Legal liberals typically argue that the state should not ban conduct unless it is “harmful,” but they refuse to count the community’s negative feelings (like community outrage or disgust) as harm. For example, liberals typically argue that the state should not criminalize homosexual sodomy because it does not harm the community, even though many in the community strongly desire to live in a community free from homosexual sodomy. If harm were merely a matter of not getting what one desired (or having negative feelings about the results of a particular act), homosexual sodomy would undoubtedly harm people who take offense at it. Since the liberal notion of harm assumes that outrage and disgust are not harm, one cannot be harmed simply because one does not get what one wants.

The second explanation, however, does explain why liberals do not count outrage and disgust as harm. Under this theory, one starts with a model of human flourishing, a picture of the morally worthwhile goals that make up an ideal human life. We might call these goals collectively “flourishing.” If a state of affairs is incompatible with flourishing, it causes harm - regardless of how other people feel about it. And while some people may feel disgust at certain conduct, they are not necessarily harmed by it.

The example demonstrates that inherent in the concept of harm is a concept of what is valuable in life: what kind of life we should pursue. Without knowing what is valuable, I cannot say whether an act causes harm. It is the job of moral theory to explain what is valuable in life (i.e. what constitutes a “good life”). Therefore, one cannot say anything about harm without first adopting a moral theory.

This conclusion does not presuppose any particular moral theory. For example, it is compatible with Christian moral theory, which says that serving God is the end of human life and explains why Christians believe that acts incompatible with serving God are harmful. It is compatible with utilitarianism, which posits that it is right to maximize utility and would conclude that a person is harmed if his utility decreases. And it is compatible with Kantian morality, which posits that human flourishing consists in fully expressing one’s autonomy. To the extent that an act interferes with a person’s autonomy, a Kantian theorist would say that the person has been harmed.

II. WORK DONE BY THE HARM PRINCIPLE

Harm as a “Thin Theory of Good”

If, as I have argued, harm depends on a notion of morality, it is puzzling why philosophers have so carefully distinguished “harm” arguments from “morals” arguments. In particular, it is puzzling why liberals dismiss certain arguments as “morals” arguments because they are premised on a notion of morality but would allow “harm” arguments that are just as dependent on a notion of morality. For example, when a politician says that banning guns would save lives, we conclude that he is making a “harm” argument. But when a philosopher argues that homosexual sodomy is incompatible with human flourishing² and that living in a society free from vice is in every person’s interest because it helps each person to flourish,³ he is typically considered to be making a *morals* argument, rather than a harm argument. Both arguments depend on a notion of morality. Even so, I think our intuition that this second argument is a *moral* one rather than a *harm* argument is strong. What accounts for this intuition?

I have argued that “harm” and “morals” arguments are quite similar in that both rest on a notion of what it means to live a good life, which is itself a moral idea. In this section, I want to argue that what distinguishes a harm argument from a morals argument is the breadth of this notion of flourishing. A morals argument can contain our full concept of what it means to flourish. A “harm” argument, by contrast, cannot rely on our full idea of what it means to flourish. Thus, the idea of harm acts as a filter: some concepts of flourishing are allowed in; some are not.

The idea that “harm” cannot rest on a full account of human flourishing is puzzling and requires some justification. I want to argue that the same liberal vision that animates John Rawls’s *Political Liberalism* and *A Theory of Justice* could supply the necessary justification for the “filtering” that the harm principle does.

In *Political Liberalism*, John Rawls tackles the problem of how citizens can share a common democratic government when reasonable people disagree about the basic principles of morality. He argues that the moral principles forming the basis of our government—which he calls the principles of justice—ought to be principles that each reasonable person can accept on the basis of his own (more comprehensive) moral theory. Since these principles would be at the “overlap” of many competing moral theories, Rawls calls them the product of an “overlapping consensus.” He then proposes a thought experiment—“the veil of ignorance”⁴—that he claims models two such principles: “the idea of society as a fair system of cooperation,” and the idea that citizens should treat each other as “free and equal.”⁵ Thus, Rawls’s theory allows moral principles to enter political debate, but not just any moral principle. The overlapping consensus acts as a filter to prevent just *any* moral idea from being used in politics.

At first approximation, this “filtering idea” can explain the work that the “harm principle” does in practice. When the government bans murder, alleging that human death is “harmful,” it indeed relies on a moral idea that death is incompatible with human flourishing, but this idea is one shared by all reasonable comprehensive moral views.⁶ By contrast, when Robert George alleges that homosexual sodomy is incompatible with the good, he bases his claim on a notion of human good that most people do not share. Thus if the notion of flourishing encompassed in “harm” must be part of an overlapping consensus, George’s idea does not count. It is a morals argument, but not a harm argument.

If my sketch is correct, the idea of “harm” serves in *real* political dialogue the same function that the “thin theory of the good” serves in Rawls’s veil of ignorance thought experiment. In Rawls’s theory the “thin theory of the good” consists of some “bare essentials” that all people could

agree are desirable—no matter what moral theory they embrace. Rawls allows this thin theory of the good to influence the principles of justice that form the basis of a just state. In real political discourse, the notion of harm seems to serve the same purpose: it excludes many notions of flourishing, allowing only the very basics into the debate, for example the idea that death, maiming, and human pain are incompatible with flourishing.

The Language of Harm and Pursuit of Reciprocity

In *Democracy and Disagreement*, Amy Gutmann and Dennis Thompson argue that because “[p]olitical decisions are collectively binding,” “they should therefore be justifiable, as far as possible, to everyone bound by them.”⁷⁷ To this end, Gutmann and Thompson argue for a theory of “deliberative democracy,” in which citizens justify political decisions to each other using three principles: “reciprocity, publicity, and accountability.” Reciprocity “asks us to appeal to reasons that are shared or could come to be shared by our fellow citizens... . It enables us, for example, mutually to respect one another as moral agents who share the goal of reaching deliberative agreement even when we disagree with one another’s conclusions.”⁷⁸ Reciprocity “tells citizen to appeal to reasons that are “recognizably moral in form and mutually acceptable in content.”⁷⁹

If harm acts as the “filter” I have described, a second justification for making harm arguments is that a harm argument is the perfect embodiment of Gutmann and Thompson’s reciprocity: it forces citizens to make political arguments citing reasons that everyone can accept. Since harm limits the idea of “flourishing” to the bare necessities, ideas of flourishing can be part of the idea of harm only if they are part of an overlapping consensus of the competing moral theories held in our society. Thus, when citizens argue that a certain act causes harm, they are arguing that it causes a state of affairs that everyone *already* believes is less-than-ideal. By using the universal language of harm, citizens can practice reciprocity without having to make different arguments to different groups of people. Using moral language, a citizen might have to argue that murder reduces overall utility, violates autonomy, and violates the Ten Commandments. But using the language of harm, the citizen can simply say that murder causes harm, leaving groups with different moral theories to translate “harm” into their own moral jargon.

The Language of Harm and Political Expediency

Finally, if the harm principle acts as I have argued, it is a political expedient. Moral language can often be contentious, and while the harm principle does not rid politics of morality altogether, it does cut out the most contentious moral views from political dialogue. In addition, it

changes the language we use to express the moral views that are allowed in politics. A Christian who wants to ban theft need not argue publicly that theft is wrong because God has banned it. Rather, he can simply say that theft is harmful. The former argument would likely provoke heated debate, even though everyone would want to outlaw theft. The latter argument would be uncontroversial.

The harm principle thus requires citizens who feel strongly about issues to express their passionately-felt views using non-inflammatory language. Thus, the harm principle does not simply require reciprocity in the *content* of moral argument. It also requires reciprocity in *language*.

III. CONCLUSION

I have argued that the traditional reason for preferring a harm argument to a morals one in political debate is flawed, and I have proposed three better justifications for the harm principle. In conclusion, I want to note that each of these proposed justifications suffers from its own shortcomings—in many cases, I believe, fatal shortcomings. The future debate between legal moralists and legal liberals therefore ought to center on these shortcomings and whether liberal theorists can answer them.

Harm as a “Thin Theory” of the Good

One of the biggest potential problems with defining harm as a real-world embodiment of Rawls’s “thin theory of the good,” is that Rawls’s thin theory of the good is arbitrarily thin. Rawls’s reason for excluding full theories of the good from his theory centers on his notion of a “political theory.”¹⁰ A legitimate state, he argues, ought to be founded on a “political” theory of morality, which means one that is part of an overlapping consensus and therefore does not depend on a comprehensive moral theory for justification. Yet Rawls’s thin theory of the good is much thinner than this notion would require. He does not attempt to explore all ideas of good that could be part of an overlapping consensus. Rather, he allows citizens under the “veil of ignorance” to know only the “bare essentials”¹¹ necessary to allow them to choose principles of justice. This limitation is arbitrary, so Rawls’s thin theory of the good is arbitrarily thin—even if we accept Rawls’s overall framework.

The problem is similar for the harm principle. If the harm principle allows in only the most “basic” notions of flourishing—such as life and health—it is arbitrarily thin. In principle, there is no reason to disallow other notions of flourishing that could be part of an overlapping consensus.

In theory, we could allow the idea of harm to be based on *all* the notions of flourishing that could be part of an overlapping consensus in

our democracy. Then, if we accept a Rawlsian framework, the thin theory wouldn't be arbitrarily thin. However, doing so would expand the category of permissible "harm" arguments so far that it wouldn't be recognizable.

An example will illustrate my point. In his book *Natural Law and Natural Rights*, John Finnis enumerates what he calls the seven "basic forms of human good": life, knowledge, play, aesthetic experience, sociability, practical reasonableness, and religion.¹² Although religion probably could not form part of an overlapping consensus in our country, Finnis's other six probably could. Damage to any of these "goods" ought be able to qualify as "harm" under the harm principle. If this is true, then Robert George's argument against homosexuality—along with just about any other "morals argument"—qualifies as a harm argument. George argues (for reasons that may or may not be convincing but are irrelevant here) that homosexual sodomy is detrimental to the good Finnis calls "sociability." To the extent that "sociability" is part of an overlapping consensus, his argument would be a permissible *harm* argument rather than an impermissible *morals* argument. Although the argument would be subject to empirical verification, it could not be dismissed out of hand as the inappropriate "type" of argument.

Interestingly, this expansion of "harm" arguments to encompass what were traditionally thought to be "morals" arguments is just what we see in practice. Bernard Harcourt has argued that once the harm principle triumphed in our political culture over legal moralism, moral "conservatives" switched gears by making harm arguments for what were traditionally seen as morals offenses.¹³ This observation suggests that practice is mirroring what my critique suggests: namely, that the idea of "harm" can accommodate a much thicker notion of the good than we traditionally think, and many or most morals arguments can also be cast as "harm" arguments.

Future debates on the viability of the harm principle, then, ought to focus on whether liberals can proffer a non-arbitrary limit for the thin theory of the good. If they can, the harm principle may turn out to be a real-world, workable embodiment of Rawls's theory. If they cannot, a Rawlsian justification for the harm principle probably cannot hold water.

Harm Arguments as Reciprocity

Earlier, I argued that the harm principle would allow citizens to practice the form of reciprocity that Gutmann and Thompson advocate. Yet this justification also raises a number of potentially troubling questions. Most important is a question of the overall coherence of the requirement of reciprocity. Behind many liberal theories of justice is what John Rawls calls the "fact of reasonable pluralism": the notion that people thinking

reasonably can disagree about basic moral ideas. Gutmann and Thompson argue that in deliberative democracy, arguments should be made using principles that anyone holding one of these “reasonable” viewpoints can accept.

The problem is that citizens in a deliberative democracy do not accept many of the arguments offered to them, and they will claim that the reasons supporting these arguments are not reasons that they “can accept.” This puts the deliberative democrat in the somewhat uncomfortable position of deciding whether these citizens are acting “reasonably.” If it were obvious which theories of morality are “reasonable” and which ones are not, the notion of reciprocity would be easy to implement. Unfortunately, however, the notion of what theories are “reasonable” looks different from person to person.

In Gutmann and Thompson’s argument, “reasonable” is both a procedural requirement and a requirement of substantive moral content.¹⁴ The procedural part of deliberative democracy is easy to implement: citizens must offer arguments to each other. The problem is implementing the requirement of substantive moral content. Reciprocity requires citizens to be reasonable—that is to offer *reasons* that “are shared, or could be shared, by their fellow citizens.”¹⁵ But no one can agree on which reasons count as “reasonable.” Amy Gutmann and Dennis Thompson argue that Christian arguments against teaching evolution in schools are not reasonable.¹⁶ Thus, these Christian arguments do not belong in the democratic process. Yet contrary to Gutmann and Thompson, millions of Jews and Christians believe that arguments for the authority of the Bible can be made using principles that anyone can accept. Gutmann and Thompson believe arguments embodying reciprocity can be made on both sides of the abortion debate.¹⁷ Yet contrary to Gutmann and Thompson, John Rawls, whose theory embodies a similar requirement of reasonableness¹⁸ denies that *any* reasonable argument could be made against legalized abortion.¹⁹

At its best, then, the requirement of reciprocity is difficult to implement because people will disagree about what sort of principles their fellow citizens can accept. At its worst, the requirement is impossible to implement: perhaps we cannot ever hope to achieve agreement over which theories are “reasonable.” Faced with these problems, a deliberative democrat must either abandon the notion of reciprocity as unworkable, or insist that a lot of people are “unreasonable” enough to have their views excluded from democratic deliberation. Neither choice looks particularly consistent with deliberative democracy.

“Harm Talk” as a Political Expedient

The final justification I suggested for the harm principle was that framing moral issues in terms of “harm” may sometimes avoid morally contentious language that would otherwise undermine the democratic process. Critics and proponents of the harm principle probably can both agree that avoiding unnecessary political confrontation is desirable.

Yet critics of the harm principle would probably attack the connection between this justification and the harm principle. While harm rhetoric is clearly *useful*, the harm principle claims that such rhetoric is morally *required*. To argue that this justification supports the harm principle is therefore to equate political expediency with morality, a connection that many advocates of the harm principle would not support. A proponent of this justification would, therefore, need to give an argument to support his claim that expediency and morality are the same. Absent such an argument, this justification supports the use of harm talk where possible, but it does not support the harm principle’s *requirement* that we use it.

Conclusion

I have argued that our normal characterization of “harm” and “morals” arguments as separate categories is flawed. A harm argument is a morals argument, although as I suggested in Part II, a special type of morals argument. Once we recognize this fact, much of modern anti-perfectionist liberal theory can be marshaled to support the harm principle. The success of the harm principle will thus turn on the success of a much wider body of liberal political theory.

I have only been able to gesture at the arguments that could be made for and against the harm principle once this wider connection to liberal theory is recognized. My goal has been to argue for a new framework in which to debate the harm principle. Within this framework, there is clearly still much theoretical work to do. I hope that future work can use the framework I have suggested to make more compelling arguments for or against the harm principle.

Conclusion

NOTES

- 1 AMY GUTMANN and DENNIS THOMPSON, *Democracy and Disagreement: Why Moral Conflict Cannot be Avoided in Politics, and What Should be Done About It* (Cambridge: Harvard University Press, 1996) 52.
- 2 See, ROBERT P. GEORGE, "Same Sex Marriage" and "Moral Neutrality," *The Clash of Orthodoxies* (Wilmington: ISI Books, 2001) 75.
- 3 ROBERT P. GEORGE, *Making Men Moral: Civil Liberties and Public Morality* (New York: Clarendon Press, 1993), 72.
- 4 JOHN RAWLS, *A Theory of Justice*, revised ed. (Cambridge: Harvard University Press, 1999), 118.
- 5 JOHN RAWLS, *Political Liberalism* (New York: Columbia University Press, 1996) 14.
- 6 Of course, the comprehensive moral theories held by some terrorists would not be part of this overlapping consensus. Rawls, would likely dismiss these theories as being "unreasonable," a term which, as he uses it, has moral content. Whether this view is convincing I will not consider here.
- 7 *Democracy and Disagreement* 13.
- 8 *Id.* at 14.
- 9 *Id.* at 57.
- 10 Contrary to popular belief, Rawls's claims that a theory of justice ought to be deontological and recognize the right as prior to the good do not require him to develop the "thin theory." Some versions of natural law theory, which is based on a full theory of the good, are deontological by Rawls's definition because they do not define justice in terms of "maximizing the good." (In fact, many of these theories argue for a plurality of basic human goods, which are incommensurable. Therefore, they argue, "maximizing the good" is an incoherent concept.) Similarly, Rawls's justification for "priority of right" simply collapses into his idea of a "political theory" and therefore does not independently require a thin theory of the good. For a fuller exposition, see Chapter III of Mark Davis, "Goodness and Justice: A Critique of Neutrality in Rawlsian Liberalism" (Princeton University Senior Thesis, 2002).
- 11 *A Theory of Justice* 348.
- 12 JOHN FINNIS, *Natural Law and Natural Rights* (New York: Clarendon Press, 1980) 86–89.
- 13 BERNARD HARCOURT, "The Collapse of the Harm Principle," 90 *The Journal of Criminal Law and Criminology* 109 (1999).
- 14 GUTMANN, AMY and DENNIS THOMPSON. "Deliberative Democracy Beyond Process." pg. 8. Prepared for the Conference on Deliberating about Deliberative Democracy. University of Texas, Austin. February 4–6, 2000. Available at: <http://www.la.utexas.edu/research/delpol/conf2000/papers/DelibDemBeyondProcess.pdf>.
- 15 *Democracy and Disagreement* 25.
- 16 *Id.* at 64.
- 17 *Id.* at 74–78.
- 18 That is a requirement that citizens treat each other as "free and equal persons whose relations with respect to one another were fair." *A Theory of Justice* 12.
- 19 See, *Political Liberalism* 243f.