

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

The Reactive Theory of Punishment

Permalink

<https://escholarship.org/uc/item/8d55r6mx>

Author

Engen, Andrew Dean

Publication Date

2011

Peer reviewed|Thesis/dissertation

The Reactive Theory of Punishment

By

Andrew Dean Engen

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Philosophy

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Niko Kolodny, Co-Chair
Professor R. Jay Wallace, Co-Chair
Professor Christopher Kutz

Fall 2011

Abstract

The Reactive Theory of Punishment

by

Andrew Dean Engen

Doctor of Philosophy in Philosophy

University of California, Berkeley

Professor Niko Kolodny, Co-Chair

Professor R. Jay Wallace, Co-Chair

The widespread practice of state punishment is somewhat puzzling from a moral point of view. Why should the state devote its limited resources to harming its own citizens? For that matter, why is the state even morally permitted to punish? After all, punishment deprives criminals of goods to which people are normally entitled, such as their liberty and property.

Traditional justifications of punishment do not satisfactorily account for the reason the state has to punish serious crime. Deterrence theory justifies punishment on the grounds that it reduces crime in society. Retributivism justifies punishment on the grounds that it inflicts deserved suffering on criminals. Both views are susceptible to counterexamples. Punishment intuitively seems justified in cases where it does not appreciably deter crime: where, for example, someone has committed a hate crime in the past in order to achieve a political goal that no one in society views as achievable anymore. Punishment also intuitively seems justified in cases where it does not inflict suffering: where, for example, someone does not suffer through his community service.

I offer an alternative view: the reactive theory of punishment. According to this theory, punishment is justified because the act of punishment expresses indignation that appropriately blames criminals for serious wrongdoing. In developing the account, I draw on P.F. Strawson's idea that when we feel the reactive sentiments of resentment, indignation, and guilt, we hold those whom we feel them toward responsible for their actions. From Joel Feinberg I take the insight that where political institutions reflect the citizens' will, punishment gives expression to the emotional reactions deemed appropriate by members of that community.

I argue that the same values that make blame morally important also give the state reason to express blame through punishment. Feeling the reactive sentiments when people have been wronged shows that we value the victims of wrongdoing in a morally significant respect. These emotions are partially constituted by the judgment that the person who has been mistreated is owed moral consideration. By establishing and maintaining egalitarian institutions that express indignation in response to serious crimes, a society demonstrates that it takes seriously the idea that all its citizens have rights about whose violation it would be appropriate to get emotionally exercised. In order to signal to all members of the community the value of those who have been victimized, the form that blame takes must be accessible to everyone in society as expressive of an unambiguously high degree of blame. Punishment is well-suited for this role because it denies goods that we all recognize to be important. I contend that this public blaming function gives us reason to punish in all cases in which punishment is intuitively justified.

Setting aside the question of why the state has reason to punish, the reactive theory also provides a better answer than deterrence theory or retributivism to the question of why punishment is morally permissible. The state is not normally permitted to violate rights just because doing so has social benefits, and it is hard to rationally defend the idea that somehow it is intrinsically good that wrongdoers suffer. I argue that punishment is permissible because we are justified in treating people in ways that would otherwise be impermissible when that treatment expresses a proportionate degree of blame. It is clear in the interpersonal context that the targets of appropriate blame lose their standing to complain about the loss of social goods. For instance, someone who has recklessly betrayed your trust cannot reasonably object to your proportionate expression of resentment on the grounds that it would cause him distress and the loss of social regard. State punishment is continuous with these interpersonal responses on the reactive theory. Those who commit serious crimes do not have the standing to object to deprivations of goods to which they are normally entitled when those deprivations are expressive of the community's appropriately high degree of indignation.

For Mom and Dad

Table of Contents

| | |
|---|-----------|
| Chapter 1: Traditional Justifications..... | 1 |
| 1 Consequentialist Justifications of Punishment..... | 3 |
| 1.1 The Utilitarian Justification of Punishment..... | 3 |
| 1.2 Changing the Consequences..... | 8 |
| 2 Retributivism..... | 10 |
| 2.1 The Retributive Idea..... | 11 |
| 2.1.1 Morris’s Retributivism..... | 13 |
| 2.1.2 Hampton’s Retributivism..... | 16 |
| 2.2 Other Problems for Retributivism..... | 18 |
| 3 Conclusion..... | 23 |
| | |
| Chapter 2: The Reactive Theory..... | 24 |
| 1 Punishment as an Expression of the Indignation of the Community..... | 26 |
| 2 Are the Reactive Sentiments Morally Problematic?..... | 31 |
| 3 Punishment and the Value of Victims of Crime..... | 33 |
| 4 Why Grant the State the Authority to Express Our Indignation?..... | 37 |
| 5 Conclusion..... | 41 |
| | |
| Chapter 3: Reasons to Express Indignation with Punishment..... | 42 |
| 1 Why Not Say it with Weeds..... | 46 |
| 1.1.Punishment as a Constructive Expression of Indignation..... | 47 |
| 1.2 Punishment as a Proportionate Expression of Indignation..... | 48 |
| 2 Alternative Societies..... | 51 |
| 3 Acts which Call for a High Degree of Indignation, but not Punishment..... | 57 |
| 4 Conclusion..... | 60 |
| | |
| Chapter 4: The Blame Fits the Crime..... | 64 |
| 1 The Reactive Sentiments and the Seriousness of Crime..... | 65 |
| 2 Reactive Proportionality and the Traditional Alternatives..... | 67 |
| 3 Conclusion..... | 71 |
| | |
| Chapter 5: Blame and the Standing to Object..... | 73 |
| 1 Forfeiture..... | 73 |
| 2 Valid Objections to Interpersonal Blame..... | 76 |
| 3 Reactive Punishment and the Standing to Object..... | 79 |
| 4 Conclusion..... | 84 |

Acknowledgements

This dissertation has greatly benefitted from feedback from a number of different audiences, including those at Berkeley-London conferences, Berkeley-Stanford-Davis conferences, dissertation seminars in the University of California-Berkeley, my dissertation writing group, and my job talk. Special thanks to Joseph Barnes, Erin Beehly, Brian Berkey, Tony Bezsylko, Benjamin Boudreaux, Lara Buchak, John Campbell, Alan Coffee, David Crane, Tim Cooper, Amanda R. Greene, Erica Klempner, Markus Kohl, Christopher Kutz, Jessica Gelber, Kristina Gehrman, Hannah Ginsborg, Araceli Hernandez, Michael Martin, John MacFarlane, Julia Nefsky, Veronique Munoz-Darde, Alva Noë, Matthew Parrott, Luis Ramos, Roisin Reardon, Sherrilyn Roush, Esperanza Sanchez, Ian Schnee, John Schwenkler, Orna Shaughnessy, Josh Sheptow, Sabrina Soracco, Hans Sluga, Kalbir Sohi, James Stazicker, Barry Stroud, and Ari Weinstein.

There are not words that express how grateful I am to my advisors, Niko Kolodny, Sam Scheffler, and Jay Wallace. Each was incredibly generous with time and insight that sustained me in this project. I am forever in their debt.

Chapter 1 Traditional Justifications of Punishment

Secure in our moral thinking is the conviction that those who commit serious acts of wrongdoing should be punished. Public debate about whether an accused murderer or rapist should be punished focuses on his guilt or innocence, not whether he should be punished if he is guilty.¹ Certain modes of punishment, such as the death penalty, are sometimes the subject of vigorous debate, but the idea that murderers and rapists should be punished is not. A theory of punishment that takes commonsense morality seriously must give an account of the positive reason the state has to punish those who commit serious crimes in our society. Without such a reason, the state would be devoting its limited resources to pointlessly harming its citizens. In order to make sense of our moral thinking, this account should be sufficiently general.

Positive reason desideratum: A theory of punishment should explain why we always, or almost always, have positive reason to punish serious crimes.²

Not only does commonsense morality hold that we have reason to punish serious criminals, it holds that we have reason to inflict punishment that is proportional to the seriousness of the crime. Though there is disagreement about what mode of punishment is appropriate for certain crimes—again, capital punishment has its staunch defenders and critics—there is widespread agreement that the severity of punishment should largely be determined by the seriousness of the crime. With respect to incarceration, we think those who commit the more serious crimes should be sentenced to more time in prison than those who commit less serious crimes, and those who commit the same crimes should be sentenced to roughly equal amounts of prison time. A theory of punishment should explain why we generally think the severity of punishment should be a function of the seriousness of crime.

Comparative proportionality desideratum: A theory of punishment should explain why we always, or almost always, have reason to punish more severely those who commit more serious crimes.

The security in our moral thought of the idea that serious wrongdoers should be punished is remarkable given another firmly held moral conviction: that people have moral rights to goods such as the liberty of bodily movement. Our conventional modes of punishment, such as imprisonment, seem to violate this set of rights. A theory of punishment should explain how we are ever permitted to punish any persons who normally bear rights to liberty and property.

¹This is not simply out of respect for the rule of law. The applications of laws that are controversial are themselves often the subject of public debate, in many cases to highlight the injustice of those laws.

²It need not be that these reasons are conclusive: there may be cases in which all things considered we have most reason not to punish some serious crime(s).

No valid objection from rights desideratum: A theory of punishment should explain why the punished cannot reasonably object to their punishment, even though it deprives them of goods to which they normally have a right.

I hope it is self-evident that the *positive reason*, *comparative proportionality*, and *no valid objection from rights desiderata* capture central features of commonsense moral thinking about punishment. Yet, I believe that traditional theories of punishment have difficulty satisfying them.

A different approach to the justification of punishment is needed. In this dissertation I propose and develop the “reactive theory” of punishment. According to this theory, punishment is justified because the act of punishment expresses indignation that appropriately blames criminals for serious wrongdoing. In developing the account, I draw on P.F. Strawson's idea that when we feel the reactive sentiments of resentment, indignation, and guilt, we hold those whom we feel them toward responsible for their actions.³ From Joel Feinberg I take the insight that where political institutions reflect the citizens' will, punishment gives expression to the emotional reactions deemed appropriate by members of that community.⁴

In this chapter I argue that the traditional theories of punishment, consequentialism and retributivism, cannot make sense of our moral thinking. In the next, I sketch the reactive theory. In Chapter 3, “Reasons to Express Indignation with Punishment” I explain why that the reactive theory satisfies the *positive reason desideratum*. I take up what many have considered to be a knockdown challenge to expressive justifications of punishment, that punishment cannot be justified by its expressive function because punishment's expressive content could be expressed through other means. I argue that in order to signal to all members of the community the value of those who have been victimized, the form that blame takes must be accessible to everyone in society as expressive of an unambiguously high degree of blame. Punishment is well-suited for this role because it denies goods that we all recognize to be important. In Chapter 4, “The Blame Fits the Crime,” I maintain that reactive proportionality satisfies the *comparative proportionality desideratum*. I argue that we have reason to punish those who commit more serious crimes more severely because we have reason to express a higher degree of indignation in response to their crimes and more severe punishment expresses a higher degree of indignation. In Chapter 5, “Blame and the Standing to Object,” I contend that the reactive theory satisfies the *no valid objection from rights desideratum*. I claim that punishment is permissible because we are justified in treating people in ways that would otherwise be impermissible when that treatment expresses a proportionate degree of blame.

³ Strawson, P. F., 1962, "Freedom and Resentment"; reprinted in Watson, Gary, 1982, *Free Will*. New York: Oxford University Press.

⁴ Feinberg, Joel, 1965, “The Expressive Function of Punishment”; reprinted in Feinberg, Joel, 1970, *Doing and Deserving*. Princeton: Princeton University Press.

1) Consequentialist Justifications of Punishment

When pressed for a reason that punishment is justified, it is common to appeal to the consequences of punishment. “It keeps dangerous people off the streets.” “It shows potential criminals that crime doesn’t pay.” “It discourages criminals from repeating their offenses.” Consequentialist theories justify punishment based on the consequences that such features as its incapacitating, deterrent, and rehabilitative features engender. Such factors undoubtedly play a role in why we think punishment is justified. Consequentialists hold that punishment can be justified solely by appealing to its good consequences.

Consequentialist Principle: The state is justified in punishing someone if punishing him will lead to better consequences than not punishing him.

Note that “justification” here is ambiguous: it might mean that punishment is morally permissible or it may mean that we should punish. On consequentialist theories these elements come together. The good consequences of punishment both make it permissible and give us reason to punish.

1.1 The Utilitarian Justification of Punishment

Different consequentialist theories propose different ways of ranking the potential outcomes to measure whether punishment is justified. Typically theories do not simply appeal to the absence of crime in their rankings. After all, though crime would seemingly be minimized if all citizens were locked up, that would not be a good state of affairs. The most famous consequentialist theory of punishment is Jeremy Bentham’s utilitarian theory, upon which outcomes are ranked by the aggregate sum of happiness of each sentient being where happiness is understood as pleasure and the absence of pain. On Bentham’s view, the suffering that punishment inflicts on the punished is a necessary evil. Punishment is only justified if the pain it causes the punished is outweighed by its promotion of the aggregate sum of happiness of others.⁵ We have reason to punish serious wrongdoing because the performance and prospect of wrongdoing cause pain. By disincentivizing serious wrongdoing with the threat of punishment, we reduce the likelihood that such crimes will be performed and promote aggregate well-being.

⁵ See his *Principles of Moral and Legislation*, Chapter XIII, Section 1: “I. *The end of the law is, to augment happiness.* The general object which all laws have, or ought to have, in common, is to augment the total happiness of the community; and therefore, in the first place, to exclude, as far as may be, every thing that tends to subtract from that happiness: in other words to exclude mischief. II. *But punishment is an evil.* But all punishment is mischief: all punishment is in itself evil. Upon the principle of utility, if it ought at all to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil.”

Utilitarian Principle: The state is justified in punishing someone if punishing him will lead to a greater aggregate sum of pleasure minus pain than not punishing him.

The utilitarian justification of punishment inherits the plausibility of the ethical theory of utilitarianism. It appeals to considerations that are of undeniable appeal: the goodness of pleasure and the badness of pain, the equality of everyone's interests, and the idea that we should try to make the world a better place. We obviously care about pleasure and pain: in our practical reasoning we often seek pleasure and avoid pain. A world with more pleasure and less pain than the one in which we live is a better world so that we have reason to pursue it. The Utilitarian Principle treats everyone equally by counting everyone's pleasure and pain equally in measuring outcomes. So, the Utilitarian Principle justifies punishment when it makes the world a better place and takes interests that we all have equally in determining what's better.

Act utilitarianism is often subject to the criticism that in particular cases it recommends courses of action that are at odds with commonsense morality. In this subsection, I will argue that the Utilitarian Principle is susceptible to such counterexamples, so it is unable to satisfy the *positive reason*, *comparative proportionality*, and *no valid objection from rights desiderata*.

Consider first the *no valid objection from rights desideratum*. It is a familiar objection to utilitarian theories that they do not rule out the punishment of the innocent. If happiness in society were maximized in some situation by punishing an innocent person, the Utilitarian Principle would seem to endorse his punishment. H. J. McCloskey describes a case in which a local sheriff would be able to forestall a series of lynchings in retaliation for some unsolved crime by framing someone he knows to be innocent.⁶ We think that an innocent person, such as the scapegoat in McCloskey's case, can reasonably object to being harmed in the ways characteristic of punishment as a means to the general security, and therefore happiness, of society. The good utilitarian consequences of punishment do not explain why a criminal does not have a reasonable objection to being punished, because in other cases potential good utilitarian consequences alone do not justify it.

The Utilitarian Principle also fails to satisfy the *comparative proportionality desideratum*. It might justify severe punishment for minor crimes and relatively light sentences for serious crimes, depending on the circumstances. For example, if authorities can dramatically decrease incidents of shoplifting in the community by making an example out of some celebrity who shoplifts by putting him in prison for an extended period of time, such a draconian punishment may be justified by the Consequentialist Principle. According to the principle, the evil of the suffering endured by the famous shoplifter will count in the ranking of outcomes, of course, but so will the aggregate happiness of all the merchants and customers who would reap the benefits of the decrease in shoplifting. On the other hand, we can imagine cases in which the Utilitarian Principle

⁶ McCloskey, H.J., 1965 "A Non-Utilitarian Approach to Punishment," *Inquiry* 8: 239-263.

would only justify a punishment that is much less severe than would be justified in other cases in which a crime of similar seriousness has been committed. Imagine that some armed robber terrorizes convenience store clerks so that she can have enough money to live the lifestyle she prefers, though the armed robbery itself is not part of her preferred lifestyle. She inherits a great deal of money at some point after she commits her crimes but before she is apprehended.⁷ She now has enough money so she does not need to steal so that there is no reason to punish her with the aim of keeping her from committing robbery again. If we publicize why we have exempted her from the regular punishment, we might also conclude that there is not much deterrent reason to punish her, because very few people will find themselves in these circumstances. We might on utilitarian grounds make her repay what she has stolen, however, so that we do not give an incentive to those who know they are soon to receive an inheritance to commit armed robbery without threat of punishment. But it is not clear that we would have utilitarian grounds for punishing her any more severely.

I suppose it is unsurprising that the Utilitarian Principle justifies punishments that are too harsh in some cases, given that it justifies punishments of innocents in others. Contrariwise, it is unsurprising that, given that the Utilitarian Principle justifies punishments that don't seem severe enough in some cases, it might not justify punishing some serious criminals at all. Consider someone who has assassinated a civil rights leader in the distant past. Imagine that he is not apprehended until years later, after social progress has been made, and he and others are no longer disposed to commit similar crimes because the political ends they were intended to achieve are no longer achievable. In such a case, we can suppose that the act of punishing the assassin will not make society any more secure. Though it might please some people in society, it is not obvious that their pleasure would outweigh the pain the civil rights leader would feel at his punishment. We can imagine it will not, anyhow. Commonsense morality, however, holds that we have positive reason to punish him because he has committed an atrocious wrong. This example shows that the utilitarian theory does not satisfy the *positive reason desideratum*.

Presented with these counterexamples, a utilitarian might simply accept that her view sometimes justifies acting in ways that are contrary to commonsense morality, but maintain “so much the worse for commonsense morality.”⁸ Such a radical utilitarian will not be concerned with arguing that her theory satisfies the *positive reason*,

⁷ This example is adapted from Moore, Michael S., 1997, *Placing Blame: A Theory of Criminal Law*. Oxford: Oxford University Press.

⁸ McCloskey attributes this view to the utilitarian J.J.C. Smart. But Smart writes that he, in fact, finds unsatisfying the implication that utilitarianism sometimes justifies the punishment of the innocent. Smart, J. J. C., 1973, “An Outline of a System of Utilitarian Ethics,” In Smart & B Williams, *Utilitarianism: For and Against*. Cambridge: Cambridge University Press, p. 71: “McCloskey then remarks: ‘But as far as I know, only J.J.C. Smart among the contemporary utilitarians, is happy to adopt this ‘solution.’” Here I must lodge a mild protest. McCloskey’s use of the word ‘happy’ surely makes me look like a reprehensible person. Even in my most utilitarian moods I am not *happy* about this consequence of utilitarianism. Nevertheless, however unhappy about it he may be, the utilitarian must admit that he draws the consequence that he might find himself in circumstances where he ought to be unjust.”

proportionality, and *no valid objection from rights desiderata*. Other utilitarians, however, do argue that the utilitarian theory of punishment is consistent with commonsense morality. In the remainder of this subsection, I will consider and reject two strategies that utilitarians employ against objections of the sort I raise. My discussion will focus on the case of the punishment of innocents, since that is the objection to utilitarian theories of punishment that has received the most attention in the literature. But the strategies here apply to cases in which the principle seems to endorse disproportionate punishment or not punishing those who commit serious crimes, insofar as they are general strategies for addressing particular cases in which utilitarianism seems to recommend courses of action at odds with commonsense morality.

In “A Utilitarian Reply to Dr. McCloskey,” T.L.S Sprigge criticizes McCloskey’s example purporting to show the divergence between utilitarianism and commonsense morality. Sprigge argues that when more details are filled in about the example of the sheriff who could frame an innocent man to prevent a series of lynchings, it is unlikely that utilitarianism will justify the punishment. One must consider how much we should trust the sheriff’s knowledge of the supposed consequences. Given epistemic limitations, it will be less certain to the sheriff that lynchings will occur than that framing the innocent man will bring him great unhappiness. One must consider the likelihood that the sheriff will be found out. If found out, punishing the innocent man will not forestall the lynchings and will undermine the rule of law in the community. One must consider how punishment of the innocent causes more unhappiness than punishment of the guilty

An innocent man is liable to suffer more shock at being thus punished. He will suffer from an indignant fury as the guilty man will not. Whatever the utilitarian thinks of the appropriateness of such indignation (a matter too complicated for comment here) he must take it into account as a fact. But apart from the indignation, the punishment will come on him as much more of a surprise, and thus be something he is less able to cope with psychologically (accept) or even practically. The distress caused to his relatives will also probably be greater, since it is likely to come as more of a shock to them also.⁹

Such considerations might go some ways towards closing the gap between what the Utilitarian Principle and commonsense morality would say about individual cases. Nevertheless, it seems that our judgment about whether an innocent person should be punished does not hang in the balance as we weigh such considerations. The fact that the punishment of an innocent man would shock his grandmother more than the grandmother of a guilty man seems rather beside the point. No matter how careful our calculation, if all that we can appeal to is the pleasure or pain that punishing him would cause across society, it seems that the innocent man can reasonably appeal to his right to liberty to object to his punishment. Likewise, our conviction that those who commit serious crimes should be punished, and that the severity of that punishment should be a function of the

⁹ Sprigge, T. L. S., 1965, “A Utilitarian Reply to Dr. McCloskey”, *Inquiry* 8: 264-91, p. 278.

seriousness of their crimes does not seem to require faith that in individual cases punishing serious criminals to a proportionate degree will maximize happiness.

An alternative way of responding to cases in which utilitarianism endorses courses of action that run counter to commonsense morality is to argue that making exceptions in particular cases leads to worse consequences overall. The maximization of happiness across society is achieved when those who commit serious crimes are punished proportionally and those who do not commit crimes are not punished. Aggregate happiness is better promoted, for instance, if rogue sheriffs do not think that they should go about framing innocent people whenever they think that that would lead to the best consequences, so there should be a rule against punishing the innocent. This suggests a new candidate principle

Rule-Utilitarian Principle: The state is justified in punishing someone if his punishment is prescribed by a rule that, when universally followed (or internalized), leads to a greater aggregate sum of pleasure minus pain than not punishing him.

There are at least two worries about this way of dealing with counterexamples to act-utilitarianism. First, many, including the utilitarian Sprigge, brand rule-utilitarianism as an unstable view.¹⁰ Rule-utilitarianism instructs people to follow rules that generally promote the greatest happiness even in situations in which following those rules would not have that result. But what could be the justification of following the rule in cases in which it does not produce the best outcome, given that the rule itself is justified by appeal to the best outcome? If the sheriff is a rule-utilitarian because he is committed to maximizing aggregate happiness, it seems to be incoherent for him to refuse to frame the innocent man based on a rule that will not maximize happiness in the circumstances in which he finds himself. If it is subject to this incoherence, the Rule-Utilitarian Principle cannot satisfy the desiderata of a theory of punishment.¹¹

The second worry about the Rule-Utilitarian Principle is that it remains an open question whether we will arrive at a set of rules that is consistent with commonsense

¹⁰ Sprigge, T. L. S., 1965, "A Utilitarian Reply to Dr. McCloskey", *Inquiry* 8: 264–91, p. 286: "There is indeed something absurd about rule-utilitarianism if it is thought to be in conflict with act-utilitarianism. If a rule is good precisely because action in accordance with it usually advances the general good, and this is the only reason, it is very odd to think that any value should be attached to action according to the rule in those exceptional cases where it hinders rather than advances the general good."

¹¹ Some rule-utilitarians have suggested that this objection only works against versions of the theory that are based solely on a commitment to maximizing happiness. They suggest that there are other grounds for endorsing rule-utilitarianism that diverge from reasons for endorsing act-utilitarianism. See, for example Hooker, Brad, 2000, *Ideal Code, Real World*. Oxford: Clarendon Press, p. 101: "[T]he best argument for rule-consequentialism is *not* that it derives from an overarching commitment to maximize the good. The best argument for rule-consequentialism is that it does a better job than its rivals of matching and tying together our moral convictions, as well as offering us help with our moral disagreements and uncertainties." Even if it avoids the charge of instability, it is difficult to see how such a view offers the right kind of explanation of what is wrong with punishing the innocent. I believe the reactive theory offers an intuitive explanation: the innocent man should not be punished because he is not blameworthy.

morality when we calculate which rules will maximize happiness. Consider the legal status of the severely mentally disabled, who we think are not morally responsible for their actions. Bentham holds that there are utilitarian reasons not to punish severely mentally disabled people. Punishing them does not fulfill a deterrent purpose because they do not have the kind of will that can be deterred.¹² H.L.A. Hart argues that this consideration does not establish a utilitarian conclusion in favor of having a rule against punishing the severely mentally disabled because having such an excuse in the legal system encourages criminals to try to deceive juries into thinking they are not responsible for their crimes, which could undermine utilitarian aims. “The belief that such deception is feasible may embolden persons who would not otherwise risk punishment of deceiving a jury in this way.... [A] criminal who actually succeeds in this deception will be left at large, though belonging to the class which the law is concerned to incapacitate.”¹³ Further arguments could be made on either side of the issue of whether a prohibition on punishing the severely mentally disabled actually leads to the greatest amount of aggregate future happiness. But from the point of view of commonsense morality, such arguments do not concern us when considering whether the severely mentally handicapped ought to be punished. Even if it would slightly increase aggregate happiness to abolish the insanity defense, it would be wrong to punish those who were really not responsible for their actions due to insanity. I believe that similar lines of argument can be made with respect to the desiderata that a theory that takes seriously commonsense morality must satisfy. Perhaps happiness across society would be maximized by exempting some group of serious criminals from prosecution; or systematic draconian punishments of some minor crimes; or policies of punishing relatives of the most serious criminals. The worry in each case is not that the Rule-Utilitarian Principle would in fact recommend rules that are at odds with the desiderata; the worry, rather, is that they leave open the possibility in a way that commonsense morality does not.

1.1 Changing the Consequences

Utilitarian justifications of punishment are attractive because they appeal to the way in which punishment can make the world a happier place. I have suggested that the utilitarian view is at odds with commonsense morality, however. One way of responding to this line of objection is to maintain that utilitarianism is correct that justified punishment makes the world better, but wrong about how we should measure potential outcomes. While pleasure and pain are of obvious moral significance, they are not *the only* factors of moral significance when we rank states of the world. Remember the original Consequentialist Principle did not give criteria for measuring outcomes

Consequentialist Principle: The state is justified in punishing someone if punishing him will lead to better consequences than not punishing him.

¹² *Principles of Morals and Legislation*, XIII.3.IX.

¹³ Hart, H. L. A., 1968, *Punishment and Responsibility*. Oxford: Oxford University Press, pp. 19–20.

We might think that some actions are intrinsically better than others or that more equal distributions are better than less equal ones. Allowing more considerations that we consider morally important to influence the ranking of outcomes will potentially bring the Consequentialist Principle closer to our commonsense moral thinking about punishment. For instance, consider a version of the principle that favors fairer outcomes that treat like cases like. Such a principle has resources to explain why those who commit similar crimes should be punished similarly and innocent people should not be punished that the Utilitarian Principle lacks.

Nevertheless it is difficult to see how any Consequentialist Principle will satisfy the *no valid objection from rights desideratum* because of the maximizing structure of consequentialism. Whatever the moral considerations from which we draw, consequentialism introduces them in the wrong way to satisfy this desideratum. In order to demonstrate this, it is instructive to consider a toy version of the Consequentialist Principle according to which punishing innocent people is intrinsically bad and nothing else matters morally. Such a principle would still recommend punishing the innocent in particular cases. Imagine that some judge believes that he can best prevent the punishment of innocents by sentencing a man he knows to be innocent to capital punishment. After the punishment, facts will come out to exonerate the punished and raise public consciousness about the importance of having mechanisms in place in the legal system to ensure that the innocent are not punished. No other avenue for achieving this goal would be as effective in minimizing the number of innocents punished. If the punishment of innocents is what makes states of affairs bad, following the Consequentialist Principle, it seems as though the capital punishment of the innocent man is justified because, even if the badness of that act as an instance of punishing the innocent is included in the outcome of the punishment, the punishment would be justified by the future punishment of innocents that it would prevent. This seems to me at odds with commonsense morality; we think that the punishment of innocents is not justified even if it reduces the number of innocents who are punished. The status of the action of punishing the innocent in commonsense morality makes it an action not to be permitted rather than one whose occurrence is to be minimized. If the potential consequences of punishment cannot explain why innocents should not be punished, appealing to the consequences of punishment alone will not explain why those who commit serious crimes cannot object to their punishment.¹⁴

¹⁴ I am uncertain whether the same line of argument works against broadly consequentialist treatments of the other desiderata. It seems that we should not excessively punish one person, even if that punishment would somehow lead to many more proportional punishments (even if, say, prolonging his detention would eventually cause him to implicate others). On the other hand, it seems that we are sometimes willing to let some criminals get off with lesser or no punishment when doing so leads to other criminals being punished. We think that prosecutors sometimes correctly exercise their discretion in granting immunity to some who have committed crimes in order that they are witnesses against multiple other criminals. But are we comfortable granting immunity to those who commit the most horrific crimes in cases in which doing so would lead to the punishment of others who have committed horrific crimes? If releasing some serious criminal would be likely to cause other serious criminals who fear him to turn themselves in, should we set him free? I do not have firm intuitions about these cases.

Kantian philosophers have paid special attention to consequentialism's failure to satisfy the *no objection from rights desideratum*. They claim that consequentialist theories justify punishment by treating criminals as "mere means, rather than ends in themselves."¹⁵ The idea that it is never permissible to treat people as mere means is not uncontroversial.¹⁶ Nevertheless, if we punish criminals simply in order to achieve some social goal, it seems that we do treat them as a means in an objectionable sense, insofar as we treat their fundamental rights as instrumental. Were fundamental rights instrumental, simply because of the desirable social consequences of punishment, criminals could not reasonably complain about being denied goods to which they are otherwise entitled. But a purely instrumental view stretches the concept of rights beyond recognition. On consequentialist justifications, the punished are treated as causal levers in a system that aims to achieve social goods. Rights are commonly thought to put limits on how people may treat one another independently of the consequences of that treatment. The state is not normally justified in violating people's rights simply on the basis that doing so would lead to good social consequences. Although consequentialist theories might draw upon considerations of obvious moral importance, they cannot make sense of our commonsense moral thinking about punishment.

2) Retributivism

Another way in which people often answer the question of why punishment is justified does not refer to its social consequences. They might say of someone who has committed a serious wrong, "He deserves to suffer for what he did." The rival justificatory tradition to deterrence theory, retributivism, is based on the "retributive idea" that wrongdoers deserve to suffer or be harmed. Punishment fulfills its retributive function by inflicting this deserved harm or suffering on the punished. Many philosophers have been attracted to some form of retributivism because unlike consequentialism it does not treat criminals merely as means to a further social good.

Before assessing whether retributivism can meet the desiderata for a theory of punishment that makes sense of commonsense morality, it will be useful to disambiguate a couple of different general versions of retributivism. According to one classical version of retributivism the retributive idea is defended based on the idea that people deserve happiness in proportion to their character so that a better world is realized in which we make bad people suffer by punishing them.¹⁷ I think it is clear that this form of retributivism will fail to satisfy the desiderata. Punishment often causes unintended

¹⁵ See, for example, Murphy, Jeffrie, 1973, "Marxism and Retribution," *Philosophy and Public Affairs* 2: 217–43.

¹⁶ For an argument that it is permissible to treat people as a means in a number of cases see forthcoming Parfit.

¹⁷ This classical view is a sort of consequentialism, on which the outcome maximized is where deserved happiness minus undeserved happiness is maximized. In my following discussion of retributivism, I will understand it non-consequentially: punishment is justified by the fittingness of the infliction of suffering or harm itself rather than its further consequences.

suffering for those with good character who care about the punished. This version of retributivism fails to satisfy the *positive reason desideratum* because it will only justify punishing serious criminals in those cases in which the deserved suffering of the criminal is not counterbalanced by the undeserved suffering of those who care about him.¹⁸ Moreover, this classical account makes the amount of suffering people deserve a function of their character and the amount of expected happiness of their lives rather than the seriousness of their crimes, so that it does not satisfy the *comparative proportionality desideratum*. A more promising version of retributivism says that wrong *actions* deserve harm or suffering.

Retributive Principle: The state is justified in punishing someone if punishing him will inflict upon him the harm or suffering he deserves in virtue of his wrongdoing.

I will critically analyze this version of the Retributive Principle in this section. As with the Consequentialist Principle, I will understand the sense of justification that the principle employs to both explain why punishment is permissible, and why we should punish.¹⁹ It is worth noting that the “harm” and “suffering” are vague terms and while some retributivists use them interchangeably, others do not. For example, some have physical pain in mind when they use the word “suffering” and some have something distinct from physical pain in mind when they use the word “harm.” I think that most of my critical points do not depend on particular notions of “suffering” or “harm,” but at some points I will consider whether different understandings of “suffering” or “harm” might have more compelling responses to my criticisms.

2.1 The Retributive Idea

Though many people have the intuition that wrongdoers deserve to suffer or to be harmed, many philosophers have claimed that the retributive idea is difficult to rationally defend. If there are good reasons to be skeptical about the retributive idea, it is difficult to see how retributivism could satisfy any of the desiderata of a theory of punishment that

¹⁸ A. C. Ewing emphasizes this point in an objection to retributivism. Ewing, A.C., 1929, *The Morality of Punishment*. London: Kegan Paul, Trench, Trubner & Co., Ltd, 1929, p. 43.

¹⁹ One could subscribe to a retributive explanation of why punishment is permissible and a consequentialist explanation of the reason we have to punish. For a simple example see the “Compromise View” in Ten, C. L., 1987, *Crime, Guilt and Punishment*. Oxford: Oxford University Press. I believe that such accounts fail to satisfy the *no valid objection from rights desideratum* for the reasons that retributive theories fail and the *positive reason* and *comparative proportionality desiderata* for the reasons that consequentialist theories fail. See, for instance, Igor Primoratz’s criticism of H. L. A Hart’s account on which utilitarianism is the “General Justifying Aim” of punishment, but does not settle the matter of whether we are permitted to punish in individual cases. Primoratz argues that Hart’s account does not explain why we have general positive reason to punish serious crimes. Hart, H. L. A., 1959, “Prolegomenon to the Principles of Punishment,” reprinted in Hart, H.L.A., 1968, *Punishment and Responsibility*, Oxford University Press, pp. 1–27. Primoratz, Igor, 1989, *Justifying Punishment*. Atlantic Highlands, N.J.: Humanities Press International, Inc., p. 142–143.

takes commonsense morality seriously. Retributivism would not explain why we have any reason at all to punish wrongdoers, proportionately or otherwise. If criminals were punished in order to receive their purportedly deserved suffering it seems as though the punished would have a valid objection to that punishment. In this section I will argue that there is good, if not dispositive, reason to be skeptical of the retributive idea. I will examine and reject a couple of prominent attempts to defend the retributive idea by Herbert Morris and Jean Hampton. Finally, I will argue that even if the retributive idea were true, it is not clear that the Retributive Principle would satisfy the desiderata for a theory of punishment that takes commonsense morality seriously.

In the evocative phrase of H. L. A. Hart, the retributive idea appears to its critics to be “a mysterious piece of moral alchemy in which the combination of the two evils of moral wickedness and suffering are transmuted into good.”²⁰ The line of skepticism I will pursue against the retributive idea is based on two plausible claims. First, the retributive idea seems in tension with the rest of our moral commitments. So it is difficult to accept without further grounding. Second, attempts to ground the retributive idea are notoriously obscure.

The retributive idea seems in tension with the rest of our moral commitments because we generally think that it is morally suspect to hold any one’s suffering to be intrinsically good. Ewing argues for this view in *The Morality of Punishment*

In every other instance the deliberate infliction of pain is wrong, except where necessary as a means to happiness or ethical improvement, in every other instance our primary duty is to abstain from bringing evil on our fellow-men.... Yet here we are asked to inflict pain for pain’s sake. It seems strange that a kind of action which under ordinary conditions is regarded as the very extreme of moral depravity should become a virtue in the case of punishment.²¹

In response to this charge, a defender of the retributive idea might simply maintain that it is a feature of commonsense morality that acts of wrongdoing are unique in rendering the infliction of suffering or harm intrinsically good. However, I think that there is a class of cases in which commonsense morality seems to reject the retributive idea: our reactions to minor wrongs in interpersonal relationships. Imagine a friend wrongs you in a minor way. He fails to keep a meeting time that he does not have a good excuse for missing, for example. You will probably be upset. It seems inappropriate, however, to think it is good that your friend should suffer, or that he should be harmed in some way that matched his crime. It would be inappropriate, for instance, for you to schedule another meeting with him, knowing that you plan not to show up in order that he gets his just deserts. Such thoughts and behavior seem inappropriate in the context of friendship; they seem to involve a sort of vindictiveness that friendship precludes. In fact, it seems petty to me to act in this way toward anyone in response to a minor wrong, regardless of

²⁰ Hart, H. L. A., 1968, *Punishment and Responsibility*. Oxford: Oxford University Press, p.234.

²¹ Ewing, A. C., 1929, *The Morality of Punishment*. London: Kegan Paul, Trench, Trubner & Co., Ltd, 1929, p. 29.

whether he is a friend. Such cases place the retributive idea in tension with our lived convictions about how we ought to respond to wrongdoing. They make it difficult to believe that we should simply take the retributive idea for granted. The idea that wrongdoers deserve to suffer or be harmed stands in need of rational explanation.

According to the second claim of line of skepticism, such explanations have proven difficult. Upon reflection, some philosophers have wondered whether there could be any grounds for the retributive idea. Warren Quinn writes, “[I]t seems clear that retributivism is burdened with a *prima facie* mystery. The idea that it is just (and, therefore, in some sense morally good) to harm someone else’s interests simply because he has harmed someone else’s interests is, when considered in the cold light of reason, hard to understand.”²² The metaphors traditionally used to explain the retributive idea seem to be no step toward clarity by themselves. It is unclear, for example, how to make out the ideas that “retributive punishment annuls the crime” or through retributive punishment offenders “pay back their debt to society” without offending commonsense.²³ This is not to say that with further elucidation the metaphors could be used in a justification of punishment—it is that they themselves are seemingly as mysterious as, or more so than, the retributive idea itself. Making these metaphors illuminating has proven difficult. This verdict is often reached by critics of retributivism, as well as some of its defenders.²⁴ In order to lend credence to the view that retributive theories do not provide clear explanations of the idea that wrongdoers deserve to suffer or be harmed, I will examine contemporary retributive theories proposed by Herbert Morris and Jean Hampton.

2.1.1 Morris's Retributivism

In “Persons and Punishment,” Herbert Morris lays out what has recently been called “probably the most influential defense of retributivism.”²⁵ Morris sets up his

²² Quinn, Warren, 1985, “The Right to Threaten and the Right to Punish,” *Philosophy and Public Affairs* 14: 327–73, p. 334.

²³ Nigel Walker is dubious of the notion that “punishment annuls the crime.” Walker, Nigel, 1991, *Why Punish?* Oxford: Oxford University Press, p. 74: “Even the dullest capacity is aware that ‘dead men rise up never’, even when you hang their murderers.”

²⁴ In a section entitled “The End of All Retributivism,” Honderich argues, “[Y]ou can suppose, no doubt, that the obscurity etc. of [retributive] theories is part of the explanation of their failing to offer a substantial reason for punishment. You can also suppose, no doubt, exactly oppositely, that the lack of a real reason for punishment is part of the explanation of the obscurity. You can indeed be suspicious of the obscurity. You can wonder about a cover-up, whether or not a self-aware or deliberate one.” Honderich, Ted, 2005, *Punishment: The Supposed Justifications Revisited*. London: Pluto Publishing, p. 196. Before presenting his own retributive view, George Sher remarks, “[A] cursory historical survey makes it clear that although retributivism has been, and still is, a live philosophical option, its theoretical base has not been clearly articulated.” Sher, George, 1987, *Desert*. Princeton: Princeton University Press, p. 69.

²⁵ Morris, Herbert, 1968, “Persons and Punishment,” *The Monist* 52: 475–501, characterized by Dolinko, David, 1991, “Some Thoughts about Retributivism,” *Ethics* 101: 537–59, p. 545. For interesting developments of Morris’s view see Murphy, Jeffrie, 1973, “Marxism and Retribution”. *Philosophy and Public Affairs* 2: 217–43, and Sher, George, 1987, *Desert*. Princeton: Princeton University Press.

theory in recognizably Hobbesian terms. We are all capable of doing harm to one another and all better off with a system of criminal laws which prohibits everyone from doing this harm. We all take on a slight burden; we restrain ourselves from harming others in the prohibited ways. But by taking on this burden we receive a substantial benefit; we can go about the business of our lives with a significantly reduced risk of being harmed by others. Fairness requires that the burdens and benefits are distributed equally across people in society. To maintain this distribution, those who break the laws are subject to having burdens inflicted on them through punishment.

A person who violates the rules has something others have—the benefits of the system—but by renouncing what others have assumed, the burdens of self-restraint, he has acquired an unfair advantage. Matters are not even until the advantage is in some way erased. Another way of putting it is that he owes something that does not rightfully belong to him. Justice—that is punishing such individuals—restores the equilibrium of benefits and burdens by taking from the individual what he owes, that is, exacting the debt.²⁶

Morris's theory may be understood to be making sense of the retributive metaphor that through punishment, wrongdoers pay their debt to society. In violating the law they gain an unfair advantage over the law-abiding members of society—they have the benefits of being protected by the system of law and without the burden of restraining their own conduct. Through the harm of punishment we remove the unfair benefit, making that harm a morally fitting response to wrongdoing.

Despite its apparent simplicity, it can be difficult to understand what the relevant burdens and benefits to be equalized are on Morris's view. When considering what it is that the criminal has that does not rightfully belong to him, we naturally think of cases such as theft where the criminal actually has things that belong to someone else and it does seem morally fitting that we take those things away from the criminal. But this cannot be the unfair advantage that Morris has in mind because the model is not easily applied to other crimes. We cannot point to any ill-gotten material goods that the murderer has in virtue of his crime that we take away by punishing him. Another way of understanding the unfair advantage is according to the notion that criminals have received the benefit that protection of the law they have broken provides, without bearing the cost that prevents others from doing to him what he has done to others. But we think that we are justified in punishing people who break laws that they are not themselves protected by. In "Do the Guilty Deserve Punishment?" Richard Burgh proposes a couple of cases in which this seems to be the case: punishment of an embezzler is justified even if, being poor, he is not someone who receives any real protection from laws prohibiting embezzlement, as is punishment of a human for cruelty toward non-human animals.²⁷ A

²⁶ Morris, Herbert, 1968, "Persons and Punishment," *The Monist* 52: 475–501, p. 478.

²⁷ Burgh, Richard, 1982, "Do the Guilty Deserve Punishment?" *Journal of Philosophy* 79: 193–210. Burgh acknowledges that a defender of this interpretation of the unfair advantage could argue that the embezzler and animal torturer have received the advantage of protection by the whole system of criminal laws without

third reading of what the criminal has that does not belong to him is some sort of freedom from the restraint of the law that the rest of us do not have because we have followed the law.²⁸ But it is strange to talk of this freedom as a benefit with regard to most serious crimes. We do not generally consider whatever freedom from restraint is involved in committing murder or rape to be an advantage: who among us would want that kind of freedom? Such an interpretation makes the theory, in the words of Jean Hampton, “too ‘creepy’ to be right.”²⁹ This last point suggests that, at least with regard to a number of serious crimes, Morris’s theory is not consistent with commonsense morality. Regardless of how we understand the benefits that criminals acquire through their crimes, we all take on the “burdens of self-restraint” in obeying the law. But for most of us, refraining from committing acts of rape and murder is no burden at all. Clearly making out the unfair advantage of Morris’s account in a plausible way proves difficult.

Aside from difficulties with interpreting the “burdens” and “benefits” on Morris’s view, one might worry whether the theory he sketches is actually correctly characterized as retributivism. Hampton claims that Morris’s theory does not give a plausible account of the retributive idea, because “it fails in a fully adequate way to link our condemnation of a wrongdoer *to that which makes his conduct wrong*.”³⁰ Those who find the idea that it is morally fitting that wrongdoers suffer in many cases do not think the reason they should suffer is because they have gained an unfair advantage over the rest of us, except perhaps in special cases such as tax evasion. In most cases the idea is that wrongdoers deserve to suffer because they have done something terrible to someone. Rapists deserve to suffer because of the way they treat their victims not because of some advantage they gain over the law-abiding through the rape. The moral fittingness of economic justice might be naturally understood on the distributive model of rectifying unfair advantages in society but the moral fittingness of retributive justice seems fundamentally different. George Sher offers another reason to think that Morris’s view is not an analysis of the retributive idea: those who accept the retributive idea believe it applies pre-institutionally.³¹ They are likely to think that those who commit mass genocide deserve to suffer, even if genocide is not outlawed in the society in which it takes place. These criticisms, of course, are not knockdown criticisms of Morris’s view *per se*. But it would be striking if “the most influential defense of retributivism” is not really retributive. It

taking on the cost of obeying those laws. Burgh argues that this is an untenable response, because it makes the view unable to distinguish between the seriousness of crimes. “Under this analysis, the violation of any particular law creates an equivalent degree of dis-equilibrium in the distribution of benefits and burdens in the sense that the undeserved benefit is the same for all offenders, namely, the benefits that result from law in general. If the restoration of this equilibrium, by removing the unentitled benefit, is the basis upon which punishment is deserved, it follows that all offenders are, regardless of the offense they committed, deserving of the same punishment”(206).

²⁸ See Finnis, John, 1972, “The Restoration of Retribution,” *Analysis* 32: 131–135, and Sher, George, *Desert*.

²⁹ Murphy, Jeffrie, & Hampton, Jean, 1988, *Forgiveness and Mercy*. Cambridge: Cambridge University Press, p. 116, see also Duff, R. A., 1986, *Trials and Punishments*. Cambridge: Cambridge University Press, pp. 210–214.

³⁰ Murphy, Jeffrie, & Hampton, Jean, 1988, *Forgiveness and Mercy*, pp. 116–117.

³¹ Sher, George, 1987, *Desert*. Princeton: Princeton University Press, p. 76.

provides a kind of support for the suggestion that it is difficult to give a rational justification of the retributive idea.

2.1.2 Hampton's Retributivism

Hampton offers an alternative account of the moral fittingness of suffering as a response to wrongdoing. According to Hampton's theory, in committing a crime a wrongdoer seems to provide false evidence about his relation to his victim because he treats the victim as someone of lower rank than him. When we make the wrongdoer suffer through punishment, he experiences that punishment as inflicted against his will so that he is made to submit to the will of the punisher. His punishment shows that the evidence provided by his wrongdoing is false because in defeating the criminal we show that the victim is as valuable as he is.

By victimizing me, the wrongdoer has declared himself elevated with respect to me, acting as a superior who is permitted to use me for his purposes. A false moral claim has been made. Moral reality has been denied. The retributivist demands the false claim be corrected. The lord must be humbled to show that he is not the lord of the victim. If I cause the wrongdoer to suffer in proportion to my suffering at his hands, his elevation over me is denied and moral reality is affirmed. I master the purported master, showing that he is my peer. So I am proposing that retributive punishment is the defeat of the wrongdoer at the hands of the victim (either indirectly or through an agent of the victim's, e.g. the state) that symbolizes the correct relative value of wrongdoer and victim.³²

Hampton associates this view with the retributive metaphor of "annulling the crime," because through the punishment she describes we annul the seeming evidence of superiority of the wrongdoer over the victim.³³

It is not obvious how to understand the relation of superiority between wrongdoer and victim that Hampton has in mind in a way that makes punishment a morally fitting response to seeming evidence of this relation. She sometimes talks about the relation in terms of moral rank. But on this reading it is difficult to see how the act of wrongdoing provides any even seeming evidence of the criminal's superiority. We do not usually take violations of moral obligations as actual evidence that the violator, in fact, is of higher rank than his victims. Wrongdoing in many cases might be evidence that some wrongdoer has false beliefs about the value of his victim in relation to him.³⁴ But it is not clear why the most fitting response to such beliefs should be the hard treatment of punishment rather than, say, moral argument. Further, if false evidence of this relation of

³² Murphy, Jeffrie, & Hampton, Jean, 1988, *Forgiveness and Mercy*. Cambridge: Cambridge University, p. 125.

³³ Murphy, Jeffrie, & Hampton, Jean, 1988, *Forgiveness and Mercy*, p. 131.

³⁴ It's not clear this is always the case. A bank robber may not be making any claim of superiority; maybe he just wants the money.

moral value is what punishment is a response to, it does not seem that punishment links up with that which makes the action wrong. In a critique of Hampton's view, David Dolinko charges that Hampton makes the same mistake that she accuses Morris of because the wrongness of crime is not captured by the message it sends.

She connects punishment to the message which, she believes, the wrongful act expresses. But what makes such an act wrong is not that it expresses the wrongdoer's morally false message "I'm worth more than you are...." What is crucially more important is the *way* in which the wrongdoer goes about conveying his moral falsehood.... A rapist deserves punishment not because he has communicated his belief that he is of greater value than his victim but because he has done so by raping her.³⁵

Retributive punishment is a fitting response not merely to false claims of value superiority, according to the objection, but also to the morally objectionable ways that those claims are made.

The relation of superiority of wrongdoer and victim might alternatively be interpreted as one of domination of will. This is what is suggested by referring to the wrongdoer as a lord who must be humbled. Hampton writes, "punishment is an experience designed to 'humble the will' of the person being punished."³⁶ The evidence of superiority implicit in wrongdoing on this view is given by the fact that someone has humbled someone else's will. Punishment at the hands of the victim is appropriate because it shows that really the victim is as powerful as the wrongdoer; they can both humble each other. It is not clear on this interpretation, however, that Hampton's theory can explain why punishment is *an* appropriate response in some cases we think it is. For example, in cases of murder, we cannot make the victim lord over his murderer except in an exceedingly metaphorical sense, since he is no longer around to take part in this relation. Also this interpretation makes it difficult to understand how state punishment would be a fitting response to crime. After all, it is not the victim who dominates the wrongdoer in state punishment but the collective instruments of the state—and making the wrongdoer a victim of the state does not show why he is not as powerful as his victim.³⁷ It is also not obvious that the domination interpretation links up the deserved suffering to what makes the action to which it is a response wrong. A relationship of domination is part of what makes rape wrong, insofar as the rapist forces his victim to submit to his will. But one suspects that there are other wrong-making features of rape that trigger the idea that suffering would be a fitting response. Or consider someone who

³⁵ Dolinko, David, 1991, "Some Thoughts about Retributivism," *Ethics* 101: 537–59, p. 552.

³⁶ Murphy, Jeffrie, & Hampton, Jean, 1988, *Forgiveness and Mercy*. Cambridge: Cambridge University, p. 126.

³⁷ See Dolinko, David, 1991, "Some Thoughts about Retributivism," p. 553: "How does it get across that the victim has evened the score and can now claim whatever mastery the wrongdoer can? It could seem, rather, that while the wrongdoer claimed 'superiority' by defeating the victim himself, a whole gang of partisans of the victim has now banded together and defeated the hopelessly outnumbered wrongdoer!"

is shot in the back and killed. It does not seem that this act of cowardice is a mastery of someone else's will; the victim's will is not involved in any straightforward way.

A third way to read the relation of superiority between the wrongdoer and victim which wrongdoing seems to give evidence of on Hampton's view is one of both domination and moral rank. On this reading, in wronging her victim a wrongdoer seemingly provides evidence that she is master over him *and* that she is of higher moral rank than him, because in wrongdoing she treats him as having a lower rank than her and thus acts as his master. Punishment is an appropriate response because only through punishment can we annul the evidence that she is of greater value than her victim because only punishment shows that she is not his master. Consider Hampton's response to the worry that we could vindicate the victim's relative worth to the one who wronged him through non-punitive means. "Still the fact that he had been mastered by the wrongdoer would stand. He would have lost to her, and no matter how much the community might contend that he was not her inferior, the loss counts as evidence that he is."³⁸ But by yoking the two accounts of superiority, the theory seems to rely on a dubious claim that the ability to dominate others is somehow evidence of moral superiority.³⁹ Or perhaps, less dubiously, domination stands as false evidence of moral superiority that no one really takes to be evidence. But then it is unclear why the suffering of the wrongdoer should be necessary to annul that evidence. It is unclear on Hampton's theory how to understand the relationship of superiority between wrongdoer and victim, for which wrongdoing provides evidence, in a way that makes it obvious why suffering which humbles the wrongdoers is a morally fitting response.

2.2 Other Problems for Retributivism

I hope that the discussion of the retributive theories of Morris and Hampton supports the claim that it is difficult to give a rational explanation of the retributive idea. By itself, this might not be problematic for retributivism; it might be that the retributive idea is a basic intuition that does not admit of further analysis. But in conjunction with worries I presented earlier about the consistency of the retributive idea with other commonsense moral beliefs, the retributive idea looks dubious. If this line of criticism of the retributive idea is successful, retributive theories cannot satisfy any of the three desiderata of a theory of punishment I proposed at the outset. Nevertheless, it is not a conclusive argument. I may have overlooked a retributive theory that clearly grounds the retributive idea, or one may yet be formulated. Perhaps an account could be given of why the basic intuition only applies to serious, and not minor, wrongs. I will conclude by

³⁸ Murphy, Jeffrie, & Hampton, Jean, 1988, *Forgiveness and Mercy*. Cambridge: Cambridge University, p. 128. Note that the wrongdoing counts as evidence of the wrongdoer's superiority over the victim even if the community does not think it does, appropriately so it would seem. Presumably in many cases neither the wrongdoer (who might not think much of his own worth) nor victim will take it as evidence, appropriately so. Still it somehow counts as evidence in some objective sense.

³⁹ For a fuller development of this last point, see Gert, Heather and Linda Raznik, Michael Hand, 2004, "Hampton on the Expressive Power of Punishment," *Journal of Social Philosophy* 35: 79–90.

arguing that *even if the retributive idea is rationally defensible*, there are reasons to think that the Retributive Principle does not straightforwardly satisfy the *positive reason*, *comparative proportionality*, and *no valid objection from rights desiderata*.

The Retributive Principle is not subject to the same Kantian criticism as the Consequentialist Principle with respect to the *no valid objection from rights desideratum*. People are often attracted to retributivism because it takes a noninstrumental view of criminals. If we deprive the punished of goods to which they normally have a right in order that they get their just deserts, we respond to them as moral agents and do not use them merely as tools for some further social goal. Nevertheless, some critics of retributivism have claimed that desert claims do not generally override rights claims. Dolinko motivates this claim by presenting a case in which a father leaves his fortune to an undeserving son over a deserving one. Though the inheritor does not deserve the fortune, he is entitled to it, and the legal system treats this as a conclusive reason to give it to him.⁴⁰ Even if wrongdoers deserve suffering or harm in virtue of their wrongdoing, we may not be justified in punishing them in order to give them their just deserts. Punishment involves treating them in a way that would normally violate their rights and people can be entitled to things they do not deserve. A retributive theory must explain what makes retributive desert different from other sorts of desert so that, unlike those sorts of desert, retributive desert makes permissible treatment that would otherwise be impermissible because it violates rights.⁴¹

It is also not clear that the Retributive Principle satisfies the *positive reason desideratum*. First, consider a version of the Retributive Principle that holds wrongdoing deserves suffering (rather than harm)

Suffering-Based Retributive Principle: The state is justified in punishing someone if punishing him will inflict upon him the suffering he deserves in virtue of his wrongdoing.

Intuitively, it seems that we have reason to inflict state punishment in cases in which it does not cause suffering in the punished. It is perhaps easiest to demonstrate this point with respect to less severe forms of punishment, such as community service. We do not think that the punishment of someone who finds that he enjoys community service fails to serve its purpose because it does not make him suffer. Of course, we do not punish the most serious criminals, such as murderers, with community service. However, I do not believe that this is because other forms of punishment, such as imprisonment, inflict deserved suffering. I grant that the life of someone in prison often involves a great deal of suffering; confinement in small spaces often takes a physical and an emotional toll. But imprisonment does not always inflict more suffering on inmates than they would have experienced on the outside. Sometimes imprisonment takes people out of

⁴⁰ Dolinko, David, 1991, "Some Thoughts about Retributivism," *Ethics* 101: 537–59, p. 544.

⁴¹ Some retributivists attempt to avoid this worry by appeal to "forfeiture," according to which those who violate the rights of others thereby forfeit their own rights. I critically analyze this view in Chapter 5.

dire and dangerous situations and provides them with three meals a day and a mattress to sleep on that they otherwise might not have had. At least some inmates do not endure a great deal of suffering day to day in prison. When punishing someone causes less suffering in a person's life than not punishing him, his punishment does not carry out its purported retributive function. Yet commonsense morality does not hold that such punishment thereby fails to be a fitting response to wrongdoing.

Harm-based retributivism might be a better candidate for satisfying the *positive reason desideratum*. If every punishment, by definition, denies the punished something to which he usually has a right, it might be thought that every punishment harms the punished. But not every intuitively justified punishment harms in some ordinary senses of harm that might appeal to retributivists. Some retributivists have an “eye for an eye” interpretation of the retributive idea, according to which we should do to wrongdoers what they have done to their victims.

Harm-Based Retributive Principle 1: The state is justified in punishing someone if punishing him will harm him in just the way that he has harmed his victims, as he deserves.

Retributivism of this sort is susceptible to a well-known criticism: the state cannot always do to the wrongdoer what he has done to his victim. In his *Commentaries on the Laws of England*, Sir William Blackstone asserts, “There are very many crimes that will in no way admit of these [matching] penalties, without manifest absurdity and wickedness. Theft cannot be punished with theft, defamation by defamation, forgery by forgery, adultery by adultery, and the like.”⁴² The principle would give us reason to commit forgery against the forger, but it is not clear that it would justify imprisoning him, though it seems that we may be justified in imprisoning him.

Another version of harm-based retributivism is based on the idea that the wrongdoer has acted in a way that typically makes the lives of people who are victims of that wrongdoing worse than they otherwise would be, so that it is fitting to make the wrongdoer's life worse in return.

Harm-Based Retributive Principle 2: The state is justified in punishing someone if punishing him will make his life worse than it otherwise would have been in the manner he deserves in virtue of his wrongdoing.

Commonsense morality holds that not all justified punishments are captured by this principle. We think that justified punishments sometimes *improve* the lives of the punished overall. We rightly welcome success stories, in which punishment causes the punished to reflect on their crimes and turn his lives around. Some who are justifiably punished might describe their punishment as “beneficial” to them. An alternative version

⁴² Blackstone, William, 1765–1769, *Commentaries on the Laws of England*. Book 4, Chapter 1.

of the principle could allow that the harm punishment inflicts is good in the long term, so long as it makes their life worse than it would have otherwise been in the short term.

Harm-Based Retributive Principle 3: The state is justified in punishing someone if punishing him will make a period of his life worse than it otherwise would have been in the manner he deserves in virtue of his wrongdoing.

But the line of reasoning I used to reject the Suffering-Based Retributive Principle also seems to undermine this version of the Harm-Based Retributive Principle. It might turn out that the life of a serious criminal in prison is not much worse than it would have been on the outside for the period that he is there. That does not seem to show that we did not have reason to imprison him.

A fourth version of the principle might maintain that what the criminal deserves in virtue of his crime is not that his life goes worse, but only that he is denied some objective good.

Harm-Based Retributive Principle 4: The state is justified in punishing someone if punishing him will deny him some objective good that he does not deserve in virtue of his wrongdoing.

I believe that this version of the principle could yield results in line with the *positive reason desideratum*. Every punishment denies its object some objective good insofar as it violates his right to liberty. It is a further question whether Harm-Based Retributive Principle 4 actually either captures the retributive idea as a basic intuition or is the retributive idea employed by developed retributive theories. It seems doubtful to me that the people who are in the grip of the intuition that those who do terrible things deserve to suffer or be harmed have in it mind that wrongdoers who at no time live a life that is worse than it otherwise would have been after their crimes have received their just deserts. Such a person need not disagree with commonsense morality about which punishments are justified. Even if there is such a thing as retributive desert, the person could agree that punishment can be justified if it does not deliver retributive deserts to criminals. On the other hand, at least one interpretation of Morris's account makes it consistent with the principle. Take the third interpretation, upon which the advantage gained by the wrongdoer is freedom from restraint by the law, and punishment removes this advantage. It is questionable whether freedom from restraint of the law is an objective good, but if it is, Morris's view seems to reflect Harm-Based Retributive Principle 4.

It is sometimes argued that retributivists can give a better account of comparative proportionality than consequentialists, because those who accept the retributive idea typically believe that the amount of harm or suffering deserved is a function of the wrongness of the crime, rather than the consequences of punishing it. For instance, a retributivist might hold that the suffering or harm a wrongdoer deserves is commensurate

with the expected disutility for others of his crime.⁴³ Retributive proportionality does not endorse inflicting an unusually high degree of harm or suffering on the celebrity shoplifter in order to deter others, because doing would inflict more suffering or harm than is fitting, given his wrongdoing. Nevertheless, the considerations that cast doubt on whether Suffering-Based Retributive Principle, Harm-Based Retributive Principle 2, and Harm-Based Retributive Principle 3 satisfy the *positive reason desideratum* also cast doubt on whether they satisfy the *comparative proportionality desideratum*. I claimed that, for those principles, whether the conventional modes of punishment inflict suffering or harm on the punished is partially a function of the individual circumstances of the punished. Because the degree to which someone's punishment harms him or causes him to suffer varies between people, sometimes the same mode of punishment can inflict a great deal more harm or suffering on one criminal than another, because of contingent features of the criminals' respective circumstances.⁴⁴ It might affect the employment status of one but not the other. It might cause one but not the other long term psychological distress.

The way that the amount of harm and suffering inflicted by punishment is determined by individual circumstances undermines retributivism's claim to satisfy the *comparative proportionality desideratum*.⁴⁵ Consider two criminals, A and B. A commits a much more serious crime than B, one that makes A twice as deserving as B according to some Suffering-Based Retributive Principle. Yet A is so constituted to experience twice as much suffering in prison as B. The Suffering-Based Retributive Principle recommends the same prison sentence for A and B, even though A has committed a much more serious crime. Corresponding cases can be formulated against the second and third versions of the Harm-Based Retributive Principle.⁴⁶ Again the fourth version seems to evade the worry. If the undeserved objective benefit that punishment denies wrongdoers is determined without reference to the individual

⁴³ See, for instance, Nozick, Robert, 1981, *Philosophical Explanations*, Cambridge, MA: Harvard University Press, p. 365: "[I]t seems appropriate to let the penalty be the maximum of the amount of disutility the victim reasonably could have been expected to undergo, and the amount of disutility the perpetrator would (reasonably be expected to?) undergo."

⁴⁴ This feature of punishment is highlighted by Nigel Walker in Walker, Nigel, 1991, *Why Punish?* Oxford: Oxford University Press, p. 99: "[T]he intensity of the suffering, hardship, or inconvenience which an individual offender: on sex, age, social position, and so on. Even if it is only a fine, and even when the amount of it is adjusted to take his means and his financial commitments into account, will it inflict hardship or mere inconvenience? What sacrifices will it force him to make? Or will he simply pass them on to his wife and children? If the penalty is imprisonment, how much does the loss of freedom mean to him? How much will he mind the squalor of the gaol?"

⁴⁵ Our commonsense morality might allow *some* of these circumstances to be taken into account. For instance, many people have the intuition that the amount of criminal fines should depend on how rich criminals are, so that if a rich and poor man commit the same crime, punishable by fine, they should be subject to different fines. Making sense of this intuition does not require a retributive understanding of proportionality, however. See my alternative account in Chapter 4.

⁴⁶ The "eye for an eye" Harm-Based Retributive Principle 1 does not give a plausible account of comparative proportionality because it does not seem to give any guidance when we cannot do to the wrongdoer what he has done to others.

circumstances of the punished, the degree of undeserved benefit that we are justified in denying the criminal could be a function of the seriousness of their crime without taking into account those circumstances. On such a view we would generally have reason to sentence those who commit more serious crimes to longer prison sentences.

3) Conclusion

Despite their place in the philosophical tradition, neither consequentialist nor retributive theories make sense of our commonsense thinking about the morality of punishment. In the next chapter, I propose an alternative theory that instead focuses on the blaming function of punishment. In Chapters 3, 4, and 5, I argue that this reactive theory satisfies the *positive reason*, *comparative proportionality*, and *no valid objection from rights desiderata*. This justification is able to capture some of what makes the traditional views appealing. Part of the explanation of why we are willing to contribute significant state resources to systems of punishment is undoubtedly that we think that crime reduction is an important goal. I will contend that the nature of the emotions that are expressed by blame gives us reason to express it in ways that contribute to that goal. On the other hand, one reason why people are sometimes reluctant to give up on retributivism, even if they are skeptical of the retributive idea, is that retributivism assigns agential status to the punished. The reactive theory offers an alternative, according to which we can treat the punished as responsible agents without being retributivists.

Chapter 2 The Reactive Theory of Punishment

After considering objections to consequentialism and retributivism, one begins to wonder whether we are morally justified in spending state resources to ensure that certain members of society are treated in the normally morally impermissible ways that constitute punishment.¹ But this skepticism is confronted by the nearly universal presence of systems of state punishment. In this chapter I will argue for an understanding of our systems of punishment in terms of our blaming practices, in particular the emotions essential to those practices. In response to serious wrongdoing, nearly all humans are subject to moral sentiments of resentment, indignation, and guilt. When we feel these emotions, we hold those whom we feel them toward responsible for their actions. In societies in which the political institutions are reflective of the will of the citizenry, the practice of punishment gives expression to the emotional reactions deemed appropriate by members of the community. In expressing the indignation of the community toward acts of serious wrongdoing through punishment, the community blames criminals for their crimes. This understanding helps to explain the pervasiveness of systems of punishment. I will further argue that the indignation that it expresses is morally significant in an important respect, so that an expressive interpretation of the practice of punishment helps us to understand not only its ubiquity, but also its moral justification.

In his seminal 1965 article, “The Expressive Function of Punishment,” Joel Feinberg criticizes contemporary philosophical discussions of punishment for ignoring the way in which punishment expresses both the community’s emotions and also the judgment that the criminal acted wrongly.² He points out a number of important social roles that punishment plays in virtue of its expressive function.³ However, he conveys unease about the emotions punishment expresses and questions whether the harms characteristic of punishment are actually required to carry out its expressive function.⁴ It

¹ See R.A. Duff’s characterization of one sort of grounds for an abolitionist position. Duff, R.A., 2000, *Punishment, Communication, and Community*. New York: Oxford University Press, p. 32: “If asked *why* we should abolish punishment, abolitionists might with some justice reply that the onus instead lies on their opponents to show why we should maintain it. Punishment is, after all, a practice that severely coerces those subjected to it and inflicts burdensome sufferings on them. It requires some persuasive justification; and abolitionists could simply argue that the justifications offered for it fail and that therefore we cannot justifiably maintain it.”

² Feinberg, Joel, 1965, “The Expressive Function of Punishment”; reprinted in Feinberg, Joel, 1970, *Doing and Deserving*. Princeton: Princeton University Press.

³ Feinberg, Joel, “The Expressive Function of Punishment,” pp. 101–105. The roles are authoritative disavowal, symbolic nonacquiescence (“speaking in the name of the people”), vindication of the law, and absolution of others.

⁴ Feinberg, Joel, “The Expressive Function of Punishment,” pp. 115–116: “One can imagine a public ritual, exploiting the most trustworthy devices of religion and mystery, music and drama, to express in the most solemn way the community’s condemnation of a criminal for his dastardly deed. Such a ritual might condemn so very emphatically that there could be no doubt of its genuineness, thus rendering symbolically superfluous any further physical hard treatment. Such a device would preserve the condemnatory function of punishment while dispensing with its usual physical media—incarceration and corporal mistreatment.

is a testament to Feinberg's influence that it is now largely taken for granted in the philosophical literature that punishment has an expressive function, but that the emotions it expresses do not play a key role in its justification.⁵

In this chapter I will give a rudimentary account of the "reactive theory" of punishment, which is constructed out of a Strawsonian account of blame that understands our blaming practices in terms of the reactive sentiments.⁶ I argue that punishment can be understood to express the indignation of the community in response to serious wrongdoing, and that we have reason to express this indignation because in doing so we show that we value our fellow citizens as appropriately safeguarded by certain moral obligations.

The Reactive Principle: The state is justified in punishing someone if punishing him expresses well the appropriate indignation of the community toward his crime.

In the first section, I argue that systems of punishment should be understood in terms of the reactive sentiments. In the next two sections, I argue for the moral appropriateness of these emotions. In the second section, I maintain that unlike some of the emotions Feinberg maintains that punishment sometimes expresses, the reactive sentiments are neither vindictive nor dehumanizing. In the third section, I explain why I take the reactive sentiments to be morally important. This account offers a preliminary

Perhaps this is only idle fantasy; or perhaps there is more to it. The question is surely open." In the next chapter, I argue that the values that give us reason to blame serious criminals also give us reason to punish them.

⁵ There actually have been a number of subsequent justifications of punishment that focus on the judgment, rather than the emotions, of the community that are expressed by punishment. See, for example Duff, R.A., 2000, *Punishment, Communication, and Community*. New York: Oxford University Press, Hampton, Jean, 1984, "The Moral Education Theory of Punishment," *Philosophy and Public Affairs*, 13, pp. 208–38, Nozick, Robert, 1981, *Philosophical Explanations*, Cambridge, Mass.: Harvard University Press. Views that mention, but are not built around, the blaming *emotions* that punishment expresses include Narayan, Uma, 1993, "Appropriate Responses and Preventive Benefits: Justifying Censure and Hard Treatment in Legal Punishment," *Oxford Journal of Legal Studies* 13: 166–82, and von Hirsch, Andrew, 1993, *Censure and Sanctions*. Oxford: Oxford University Press.

⁶ Christopher Bennett offers a recent, retributive theory of punishment that draws upon the Strawsonian reactive attitudes. Bennett argues that blaming punishment should express how much guilt the punished should feel in virtue of his wrongdoing in order that he can symbolically make amends for his wrongdoing. But it is not clear to me that the guilt required to make amends can be externalized in the way that the theory requires in order to justify punishing those who do not truly feel guilty about their wrongdoing. It seems to me that such criminals do not in fact make amends through the suffering their punishment inflicts upon them, though commonsense morality still holds their punishment justified. I believe that focusing on the normative indignation, rather than guilt, that punishment expresses is a more promising place to focus a theory. The resulting theory does not require that the punished suffer, so avoids worries that I raised in the previous section about whether suffering-based retributivism satisfies the *positive reason* desideratum. Bennett, Christopher, 2008, *The Apology Ritual: A Philosophical Theory of Punishment*. Cambridge: Cambridge University Press. An earlier retributive theory that also appeals to the guilt that wrongdoers ought to feel is found in Moore, Michael S., 1997, *Placing Blame: A Theory of Criminal Law*. Oxford: Oxford University Press.

explanation of why the reactive theory satisfies the *positive desideratum*. I claim that because the reactive sentiments are bound up with values at the center of morality, we always have reason to publicly blame serious criminals in order to demonstrate our commitment to those values. In the fourth section, I offer a number of reasons for the state to express this public blame in response to serious crimes, rather than leaving it to the individuals who feel the reactive sentiments. I go on in the next three chapters to argue that, unlike consequentialism and retributivism, the reactive theory satisfies the three desiderata of a theory of punishment that takes commonsense morality seriously.

1) Punishment as an Expression of the Indignation of the Community

In “Freedom and Resentment,”⁷ P. F. Strawson argues that by looking at our practice of holding one another responsible, we can see that ascriptions of moral responsibility would not be threatened by the truth of determinism. I will not assess Strawson’s argument as a response to incompatibilism about responsibility and determinism here, but will take from his famous lecture the account of moral responsibility and its connection to the emotions. According to Strawson, when we hold agents responsible we are subject to a range of emotions that are given to us with the structure of human relationships. We feel resentment when we are wronged, indignation on behalf of others when they are wronged, and guilt when we ourselves act wrongly. Strawson maintains that these emotions are constitutive of our practice of holding one another responsible: “the making of [a moral] demand *is* the proneness to such attitudes.”⁸

The connection of resentment, indignation, and guilt to moral responsibility suggested by Strawson’s article is given a fuller account by R. Jay Wallace’s in *Responsibility and the Moral Sentiments*. Wallace maintains that there is an essential connection between the reactive sentiments and holding someone to a demand or expectation, so that holding someone to a demand involves a susceptibility to the reactive sentiments and vice versa. The stance of holding someone to an expectation is taken distinctively with respect to normative as opposed to merely epistemic expectations (i.e. expectations not about what someone *ought* to do but only what she *will* do). For one to hold someone to an expectation “is to be susceptible to a certain range of emotions if the expectation is violated, or to believe that it would be appropriate for one to feel those emotions if the expectation is violated.”⁹ Guilt is tied up with the expectations we have of ourselves. Indignation and resentment are linked with those we have of others, the latter specifically in regards to other people’s conduct towards ourselves. The reactive sentiments can be understood to be *moral* sentiments when their associated demands are moral obligations. I find the reactive account intuitively plausible. It is difficult to

⁷ Strawson, P. F., 1962, “Freedom and Resentment”; reprinted in Watson, Gary, 1982, *Free Will*. New York: Oxford University Press.

⁸ Strawson, P. F., 1962, “Freedom and Resentment,” p. 77.

⁹ Wallace, R. Jay, 1994, *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, p. 23.

imagine a case in which we hold someone morally accountable but neither feel the reactive sentiments in response to violations of moral obligations nor think them appropriate.¹⁰ It might be that we have competing reasons not to express these attitudes toward someone we hold responsible—she might be a boss who is cheating colleagues out of money at a job which provides a source of income we need—but we do not think in such cases that the attitudes themselves are rendered inappropriate even if, all things considered, they are to be concealed.

Our practice of blame is closely related to our practice of holding people morally responsible. Wallace spells out the connection

[T]o blame someone is to be subject to one of the reactive emotions in terms of which the stance of holding people responsible is essentially defined, and these emotions are *expressed* by the sanctioning behavior to which the stance of holding people responsible naturally inclines us.¹¹

Wallace's reactive account of blame identifies the reactive sentiments as the common element of all forms that moral blame takes—from private, unexpressed blame to public censure. He argues that accounts of blame that ignore the reactive sentiments leave out the attitudinal quality of opprobrium that is characteristic of blame.¹² This quality sets blame apart from other negative assessments, such as criticisms of people's arithmetic or hairstyles, which are not typically accompanied by the reactive sentiments.

We can understand punishment as blame, on the reactive account of blame, if we can understand it to express the reactive sentiments. In the remainder of this section I will argue that acts of punishment can be understood as externalized expressions of the indignation of the community in societies where systems of punishment reflect the will of members of the community. The criminal law sets out a set of moral demands on members of the society in which it applies. In societies where the criminal laws are enacted by procedures that are answerable to the will of the community, the moral demands enshrined in the law are normative expectations to which citizens hold one another. According to the reactive account of moral responsibility, this stance of holding one another responsible is to be understood in terms of the reactive sentiments. When citizens hold all their fellow citizens to the normative expectation of following the laws, they believe it would be appropriate to feel the reactive sentiments when their fellow citizens violate those laws. The criminal law also delineates the punishment for those who violate its demands. In democratic societies, acts of punishment that deprive their

¹⁰ Forgiveness might be a case in which the connection between the reactive sentiments and attributions of moral responsibility are severed, insofar as forgiveness is thought to involve the forswearing of resentment but not the forswearing of accountability. But it seems to involve the recognition that the forgiven has done something which merits resentment or it becomes mysterious why it needs to be forsworn in the first place. Wallace makes a similar point in Wallace, R. Jay, 1994, *Responsibility and the Moral Sentiments*, Cambridge, MA: Harvard University Press pp. 72–73.

¹¹ Wallace, R. Jay, 1994, *Responsibility and the Moral Sentiments*, p. 52.

¹² Wallace, R. Jay, 1994, *Responsibility and the Moral Sentiments*, p. 81. See also his forthcoming “Dispassionate Opprobrium.”

objects of important goods, such as liberty and property, can be conceived of as expressing the indignation the community has endorsed as appropriate in response to the crime. Acts of punishment thereby blame criminals for their crimes.

Punishment can be understood to express indignation, because the disposition to punish, or, more precisely, to approve of punishment, is among those that are characteristic of the reactive sentiments. Contemporary conceptual analyses of emotions typically count dispositions to action among their defining characteristics.¹³ All other things being equal, those who feel compassion are disposed to act in ways that relieve the suffering of those they feel compassionate toward, for example, and those who feel love are disposed to promote the interests of their beloved. The reactive sentiments, on the other hand, typically dispose those who feel them to act in ways that are objectively bad for the blamed. Punishment is continuous with such actions that we are disposed to perform in the interpersonal context, where those responses are not mediated through state institutions, in that they characteristically dispose those who feel them to act in ways that are objectively bad for the blamed.¹⁴ This limits the field of actions which express resentment and indignation. It explains why applause and gifts of large amounts of cash do not express of those emotions except in very exceptional cases. Resentment and indignation always deny their objects something objectively good, a certain sort of social regard. For most people, this makes it distressing to be the object of the reactive sentiments. Objects of blame often feel guilty and feel the sting of the attitudes others feel toward them. People who feel indignation and resentment are characteristically disposed to perform actions that are typically bad for the objects of blame, even when those who feel them do not desire that those objects suffer. Even when one is the target of a merely verbal expression of blame, one often becomes painfully aware of his standing with others. Expressions of the reactive sentiments sometimes take the form of the denials of further social goods that come with other people withdrawing goodwill that they would otherwise have toward you. Examples of these harms include being excluded from a social circle or not being able to receive aid with one's projects.

We can see the harms that punishment involves as continuous with these interpersonal responses. Though we are not typically disposed to deprive one another of liberty of bodily movement by our own hands, we are disposed to approve the state carrying out such punishment. This is a generalization, but I believe that it fits the way most people experience indignation when their minds are emotionally fixed on an act of

¹³ See, for example, Goldie, Peter, 2000, *The Emotions*. Oxford: Oxford University Press, pp. 12–13: “An emotion is complex in that it will typically involve many different elements: it involves episodes of emotional experience, including perceptions, thoughts, and feelings of various kinds, and bodily changes of various kinds; and it involves dispositions, including dispositions to experience further emotional episodes, to have further thoughts and feelings, and to behave in certain ways.”

¹⁴ This idea can be found in Strawson, P. F., 1962, “Freedom and Resentment”; reprinted in Watson, Gary, 1982, *Free Will*. New York: Oxford University Press, p. 77: “The partial withdrawal of goodwill which these attitudes entail, the modification they entail of the general demand that another should, if possible, be spared suffering, is, rather, the consequence of continuing to view him as a member of the moral community; only as one who has offended against its demands. So the preparedness to acquiesce in that infliction of suffering on the offender which is an essential part of punishment is all of a piece with this whole range of attitudes of which I have been speaking.”

serious wrongdoing. I invite the reader to take a moment to imagine some terrible crime with which they are familiar and muster up indignation toward the person who has committed that crime on behalf of the victim of that crime. While in the grip of indignation consider how you are disposed to go in for the punishment of the serious criminal and how naturally you are so disposed from within the emotional experience. That we are not generally disposed to punish with our own hands, I believe, can be explained by the function of modern states. Depriving the goods characteristic of punishment usually requires a credible threat of force. We can exclude someone else from our social circle without accounting for his will, but in order to limit his bodily movement we need him to decide to go along with the limitations or force him to follow them against his will.¹⁵ In the modern societies in which we are emotionally educated,¹⁶ states have monopolies on the legitimate use of force.¹⁷ The actions that our indignation in response dispose us toward are shaped by the recognition of this fact.

According to the reactive interpretation of systems of punishment I have been advocating, we can understand those that are the product of democratic procedures as expressive of the indignation in those communities, because most people are disposed to approve of punishment when they feel indignation in response to serious crimes. This account conceives of punishment as blame in a way that is consistent with Wallace's reactive account, which defines blame in terms of the reactive sentiments. One might worry, however, that this understanding faces a dilemma. Either systems of punishment express the indignation that people in society feel or it does not express any indignation at all. People feel indignation inconsistently. They do not feel indignation toward all crimes, and whether they feel indignation toward particular crimes is often determined by factors that do not, and certainly should not, determine whether someone who commits a serious crime is punished. For instance, media coverage can make a large difference in whether members of the community feel indignation toward some serious criminal. Even independent of media coverage, the extent to which people in the community feel indignant in response to various criminal actions might depend on whether they are partial to some criminals or crime victims. But systems of punishment typically prescribe the same punishment to those who commit the same crimes, so punishment does not express the actual indignation members of the community feel in response to crimes. Thus we must avoid the first horn of the dilemma. If it does not express the actual emotions that people feel, however, how can it be expressing any emotion at all? We do not usually think that actions that people perform without feeling emotions express emotions, even when people act in ways to which people feeling those emotions are

¹⁵ It seems that typical, interpersonal expressions of the reactive sentiments are sometimes *de facto* deprivations of property. When some community ensures, "he'll never work in this town again" this has social and financial consequences. Taking property that someone already possesses, however, usually does require the threat of force.

¹⁶ The idea that the actions to which our emotions dispose us is partly a consequence of the societies we are emotionally educated in is pursued further in the next chapter.

¹⁷ Max Weber famously defines the state in terms of this monopoly in Weber, Max, 1919, "Politics as a Vocation," in *From Max Weber: Essays in Sociology*, ed. by H.H. Gerth and C. Wright Mills, London: Routledge and Kegan Paul, 1977, pp. 77–128.

disposed. For instance, someone who is dreadfully bored at a party who smiles to give his friends the false impression that he is happy does not express joy.¹⁸

I believe that the reactive understanding of systems of punishment can evade the second horn of this dilemma because there is at least one sort of action which may express emotions that people do not actually feel: when they are expressed in someone else's name. Imagine that before leaving for Antarctic seclusion, I ask my neighbor to have flowers sent to my grandmother should she become sick. My grandmother falls ill and my neighbor follows through, sending the flowers with a sympathy card. The flowers express my sympathy, or the sympathy that I have endorsed as appropriate. It is in this way that political institutions express the emotions of community: in democratic societies, state actions speak on behalf of citizens. Consider a recent resolution proposed in the United States Congress, described by its authors as "a resolution calling on Congress to formally acknowledge and express regret for the passage of a series of laws during the turn of the 20th Century that violated the fundamental civil rights of Chinese-American settlers."¹⁹ If passed the resolution would serve to express an emotion of the community of American citizens for the treatment of early Chinese American settlers, even if a small percentage of Americans feel such regret at the time of passage. It does so in virtue of being the product of democratic procedures. In the representative system, elected officials speak on our behalf. When in their role as representatives, they authorize an action that would clearly express regret were it performed by an individual disposed to do so by his regret, like formally saying "I express regret," they express regret in our name.

Emotions can be expressed by the state, not only through official declarations, but also through the day to day operations of institutions. Consider, for example, the attitude of those in the community toward those who enlist in the armed forces and put themselves in harm's way for the interests of the community as a whole. Most of us think we have reason to feel gratitude toward those who serve in the military and there are laws in our society that prescribe actions to which those who feel gratitude are characteristically disposed targeting military members. For example, among other things, the G.I. Bill offers benefits to veterans who attend college and start businesses.²⁰ It is not necessary for members of the community to actually feel gratitude toward each individual military member in order for these benefits to express the gratitude of the community.

¹⁸ Sometimes actions can be expressive of an emotion that is not occurrently felt. A parent who spends his days running errands for his family might feel stress rather than love. It seems like these actions could still express love for his family. However, it does seem strange to understand these actions as expressions of love if he *never* experiences feelings of love.

¹⁹ From Congresswoman Judy Chu's press release. Chu, Judy, 2011, "Reps. Chu, Biggert & Coffman, Senators Feinstein & Brown Announce Resolution of Regret for Chinese Exclusion Laws." Retrieved in June 2011 from <http://chu.house.gov/2011/05/rep-chu-bigbert-coffman-senators-feinstein-brown-announce-resolution-of-regret-for-chinese-exclusio.shtml>

²⁰ It might be thought that such programs aim at recruitment not expressing gratitude. I think that such programs can consistently do *both*. The state has reason to express gratitude in a way that encourages others in society to pursue actions for which we are expressing gratitude. Following a similar line of thought, in the next chapter I will argue that we have reason to express indignation through forms of treatment that deter.

Members of the community do have a role in seeing that the emotions they find to be appropriate are expressed by institutions in holding accountable elected officials. When, as for instance happened in the case of the Walter Reed Army Medical Center, it comes out that men and women injured in combat get worse medical care than other members of society, it is a matter that politicians ignore at their own electoral peril.

I conclude that the indignation of the community expressed by punishment need not be the actual emotions that members of the community feel, but is instead a normative notion, the indignation that they have endorsed as appropriate through their (often implicit) endorsement of systems of state punishment. When we are indignant toward serious crimes, most of us are disposed to approve of state punishment, so acts of punishment can be understood to express the indignation members of the community have endorsed as appropriate. Thus, systems of punishment hold criminals accountable in a way that is consistent with the reactive account of blame.

2) Are the Reactive Sentiments Morally Problematic?

In highlighting the expressive function of punishment, Feinberg writes

At its best, in civilized and democratic countries, punishment surely expresses the community's strong *disapproval* of what the criminal did. Indeed, it can be said that punishment expresses the *judgment* (as distinct from any emotion) of the community that what the criminal did was wrong. I think it is fair to say of our community, however, that punishment generally expresses more than judgments of disapproval; it is also a symbolic way of getting back at the criminal, of expressing a kind of vindictive resentment. To any reader who has in fact spent time in a prison, I venture to say, even Professor Gardner's strong terms—"hatred, fear, or contempt for the convict"—will not seem too strong an account of what imprisonment is universally taken to express....²¹

In this quote, Feinberg separates the encomium of "at its best," from acts of punishment that express emotions. There are a number of reasons that one might hesitate to approve of the emotions that punishment expresses. A couple of important features of those emotions that seem to concern Feinberg here are that they are vindictive or dehumanizing. Vindictiveness is usually taken to be a negative character trait. In the previous chapter, I argued that our interpersonal responses to minor acts of wrongdoing seem to show that we reject the retributive idea in that context. We might criticize someone as vindictive, if he desires that others be harmed or suffered for the minor wrongs they have done to him. If the emotions expressed by punishment necessarily

²¹ Feinberg, Joel, 1965, "The Expressive Function of Punishment"; reprinted in Feinberg, Joel, 1970, *Doing and Deserving*. Princeton: Princeton University Press, pp. 100–101.

involve the desire that their object suffer or be harmed they are susceptible to the charge of vindictiveness, according to this line of criticism.²²

Though the reactive sentiments are marked by a quality of opprobrium, they are not constituted by the retributive desire that their objects deserve to suffer. Feeling and expressing the reactive sentiments is consistent with thinking it is never intrinsically good that anyone suffer. This seems clear from the interpersonal context. Imagine a friend fails to show up to drive you to the airport as scheduled. You have to call a taxi instead and almost miss your flight. You will probably feel some resentment at your friend for not keeping his word. You might also express it, perhaps by calling your friend and explaining that he forgot to pick you up. In this case, it does not seem that in feeling and expressing resentment, you must think that it would be a good thing for your friend to suffer. In fact this attitude seems inappropriate in the context of friendship; it seems to involve a sort of vindictiveness that friendship precludes. This example shows that the reactive sentiments need not involve a desire that their object should suffer. Those who do not endorse the retributive idea need not worry that the reactive sentiments, and our practice of holding one another morally responsible, are necessarily vindictive. Punishment that expresses the indignation of the community always reflects the judgment that blame is appropriate, but it need not reflect the judgment that it is intrinsically good that criminals suffer.

Another complaint that is sometimes raised against the emotions expressed by punishment is that they are dehumanizing because they judge their objects as less than human. Martha Nussbaum, for instance, has argued against disgust-based justifications of sodomy laws, in part based on the idea that disgust dehumanizes its object.²³ A worry about the contempt for criminals that Feinberg claims imprisonment expresses could be that, if it in fact expresses contempt, punishment treats the punished as subhuman. According to Robert Solomon's register of the emotions, someone who feels contempt views its object as "markedly inferior to oneself, even subhuman," and typical contempt metaphors describe the object as "a snake, a reptile, an insect, a worm, a spider, as slime or excrement, as degenerate or depraved."²⁴ The emotions punishment expresses are problematic if they deny the humanity of the punished.

There is reason to think that the reactive sentiments are not dehumanizing, however. Strawson and Wallace consider classes of cases in which it is not appropriate to feel the reactive sentiments toward someone. In some cases we *excuse* an agent who is generally an appropriate object of the reactive sentiments, when that agent unintentionally performs an action that would violate a moral obligation if he did it intentionally. Strawson uses the example of someone who painfully treads on my hand accidentally when trying to help me up after I have fallen. While he has an obligation not to willingly inflict great pain in me, if he does it so accidentally while trying to help me,

²² Of course, this line will not be compelling to those who endorse the retributive idea.

²³ Nussbaum, Martha, 1999, "'Secret Sewers of Vice': Disgust, Bodies and the Law" in Bandes, Susan A. (ed.) *The Passions of Law*. New York: New York University Press. She distinguishes disgust from indignation, the relevance of which she says the criminal law "may rightly admit".(22)

²⁴ Solomon, 1993, *The Passions: Emotions and the Meaning of Life*. Indianapolis: Hackett Publishing Company.

it is not appropriate for me to feel resentment because he has not really violated a moral demand to which he is appropriately held. In other cases we *exempt* someone from the reactive sentiments, either temporarily or permanently. We appropriately take Strawson's objective stance, suspending our reactive sentiments, toward those who are not suited for interpersonal relationships such as very young children and the severely mentally disabled. Wallace argues that the suspension of resentment and indignation is appropriate in such cases because young children and the severely mentally disabled are not fairly held to moral demands because they lack the powers of self-control, or "the general ability to grasp and apply moral reasons and to regulate their behavior by the light of such reasons."²⁵ When one is susceptible to the reactive sentiments in one's interactions with someone else, one treats that person as someone to whom moral obligations are properly addressed, or in Strawson's phrase "as a member of the moral community."²⁶ When we feel resentment or indignation toward someone else, we treat him as an agent who can understand moral obligations and regulate his behavior accordingly. We acknowledge his capacities for moral agency. So there is at least one important respect in which reactive punishment acknowledges the humanity of the punished. The reactive attitudes necessarily view their objects as responsible agents and not mere snakes or excrement.²⁷

3) Punishment and the Value of Victims of Crime

One might maintain that the same considerations that explain why blame is not dehumanizing explain why we have positive reason to punish. We have general reason to punish serious criminals, she might argue, because we have general reason to treat them as persons with capacities for moral action. This is one way to make out Hegel's famous, controversial doctrine that the criminal has a right to be punished.²⁸ But it does not seem

²⁵ Wallace, R. Jay, 1994, *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press *Sentiments*, p. 155.

²⁶ Strawson, P. F., 1962, "Freedom and Resentment"; reprinted in Watson, Gary, 1982, *Free Will*. New York: Oxford University Press, p. 73.

²⁷ Michelle Mason has argued that we should understand contempt as a reactive attitude so, like other reactive attitudes not necessarily dehumanizing. Mason, Michelle, 2003, "Contempt as a Moral Attitude," *Ethics*, 113/2 (January): 234–72, p. 263: "Regarding another with contempt does not thereby objectify another person; rather it is regarding him as beneath contempt that signals we have exiled him from moral community with us." On her view contempt involves a low appraisal of someone compared with an interpersonal ideal of a person, but still recognizes his personhood because contempt only makes sense as a reaction to persons who do not measure up to the ideal. If this is right, I take it that some of the central contempt metaphors that Solomon associates with contempt are problematic, unless the same emotion can appraise someone as excrement while respecting his personhood. Even if contempt is not dehumanizing, there seems to be another reason to think that punishment that expresses it is problematic insofar as contempt "takes a person as a whole" rather than an action as its object(257). Our legal system is best suited to discover whether people have committed crimes rather than making character judgments. Appraising criminals negatively as whole persons would be unjustified in at least some cases.

²⁸ Christopher Bennett appeals to the Strawsonian idea that when we are susceptible to the reactive sentiments we treat their objects as members of the moral community to argue that criminals have a right to

that if we fail to imprison a serious criminal we thereby fail to respect his capacities for agency. Given the choice between granting someone the same options for the sort of life he can live as those who are not imprisoned, and greatly limiting those options by imprisoning him, it is not obvious that the only way to respect him as an agent is to greatly limit his options for self-determination. Both imprisoning him, and letting him free, might be ways of acknowledging those capacities depending on the reasons for doing so.

It does not seem to me that our impulse to impose blaming punishment on those who commit serious crimes is, in the first place, an impulse to respect the agency of wrongdoers. That misplaces the focus of the moral motivation of our blame. I maintain that the positive reason we have to blame those who commit serious crimes is instead tied to our commitment to the value of the victims of crime. When we feel resentment, indignation, or guilt when people have been mistreated, we show that we value those people in a particularly important respect. These emotions are partially constituted by the value judgment that the person who has been wronged is owed moral consideration and so is properly protected by moral obligations. In this section I will argue that the moral impulse to impose blaming punishment on those who commit serious crimes is an impulse to stand up transparently for victims of crime.

The idea that our emotions generally reveal what we value has been prominent in recent work on value theory and the emotions. Some value theorists have maintained that when we value something we are susceptible to certain emotional reactions with regard to that thing.²⁹ Work on the emotions has emphasized that they are partially constituted by value judgments.³⁰ For instance, if I value athletic achievements I am disposed to feel pride if I accomplish them and admiration (or perhaps envy) when someone else does. These emotions are themselves partly constituted by the judgment that athletic achievements are good. Oftentimes when we express these emotions, acting in the ways in which they dispose us, we display our value judgments to others. If my admiration for athletic achievements disposes me to spend a great deal of money on Barry Bonds

blaming punishment insofar as they have a “right to be evaluated as a member of a moral relationship.” Bennett, Christopher, 2008, *The Apology Ritual: A Philosophical Theory of Punishment*. Cambridge: Cambridge University Press, p. 70.

²⁹ For examples see Anderson, Elizabeth, 1993, *Value in Ethics and Economics*. Cambridge, MA: President and Fellows of Harvard College. Kolodny, Niko, 2003, “Love as Valuing a Relationship,” *Philosophical Review*, 112:2, pp. 135–189. Scheffler, Samuel, 2004, “Projects, Relationships, and Reasons,” in Wallace, Jay, Phillip Pettit, Samuel Scheffler, and Michael Smith, *Reason and Value: Themes in the Philosophy of Joseph Raz*. Oxford: Oxford University Press.

³⁰ For examples see Stocker, Michael and Elizabeth Hegeman, 1996, *Valuing Emotions*. Cambridge: Press Syndicate of the University of Cambridge. Solomon, Robert, 2007, *True to Our Feelings*. New York: Oxford University Press. Stocker and others argue that this point can be found in Aristotle’s treatment of the emotions. In the *Rhetoric* Aristotle defines anger as “a desire accompanied by pain, for a conspicuous revenge for a conspicuous slight at the hands of men who have no call to slight oneself or one’s friends”(1378a31), where slighting is understood as “the actively entertained opinion of something as obviously of no importance”(1378b10). Thus, on Aristotle’s account it looks like the emotion of anger necessarily involves value judgments insofar as it involves taking oneself and one’s friends to have some importance. Translation of the *Rhetoric* by W. Rhys Roberts in Barnes, Johnathan, ed., 1984, *The Complete Works of Aristotle*. Princeton: Princeton University Press.

memorabilia and devote a room of my house to celebrating his career home run record, by acting on those dispositions I demonstrate to visitors to my home what I value (and, to an extent, what I do not).

The reactive sentiments are particularly morally important emotions because they show what we value as being protected by moral obligations. In *Forgiveness and Mercy*, Jeffrie Murphy highlights the way that we show that we respect ourselves by feeling resentment when we are wronged. When experiencing resentment at being wronged we show that we show a commitment to the idea that we are not to be treated that way. “Not to have what Peter Strawson calls the ‘reactive attitude’ of resentment when our rights are violated is to convey—emotionally—that we do not think we have rights or that we do not take our rights very seriously.”³¹ This point can be extended to indignation. When we feel indignation on behalf of those who have been wronged we show that we take them to be properly protected by the moral obligation which someone has violated with respect to them. Being emotionally exercised on behalf of the mistreated party reveals that we care about them in a way that simply intellectually recognizing their value does not. It shows that their status as persons who are owed moral consideration is something that really matters to us. When we express that indignation we often display to others our commitment to the idea that we care about those on behalf of whom we feel it. For instance, if a colleague is making another uncomfortable by saying disrespectful things, and you indignantly say “hey, knock it off,” you show your tormented colleague that you value him as appropriately being given moral consideration.

When we regard the reactive sentiments as appropriate consistently across cases of similar wrongdoing, we demonstrate not only that we value certain people as being owed moral consideration, but that we value all people and, because of this, the role that certain moral obligations play in human life. When we are susceptible to the reactive sentiments with respect to a moral obligation, we reveal that we care that human interactions are structured by it.³² For instance, imagine that Tom consistently feels indignant on behalf of musical artists when his friends tell him they are downloading albums for free through file sharing programs that are not endorsed by the artists. Sue is among his friends that download albums. She does not feel guilty about downloading free albums without the consent of their artists and does not feel indignant when her friends do it. Even if she acknowledges that downloading music for free without the consent of the artists is wrong, it seems clear Tom cares about the moral obligation prohibiting doing so in a way that she does not.

The role that the reactive sentiments play in showing that we value victims of wrongdoing and the moral obligations that ought to protect them is crucial to

³¹ Murphy, Jeffrie, & Hampton, Jean, 1988, *Forgiveness and Mercy*. Cambridge: Cambridge University Press, p. 17.

³² R. Jay Wallace suggests that the reactive sentiments reflect a valuing of a sort of relationship that moral norms make possible. See his forthcoming “Dispassionate Opprobrium,” p.35: “To internalize a concern for morality... is to care about relating to people in the distinctive way that is constituted through compliance with basic moral requirements. But people who care about this form of relationship naturally tend to hold themselves and others to the moral norms that are constitutive of it, where this in turn involves a susceptibility to the distinctively reactive sentiments.”

understanding the moral significance of punishment. In section 2 I maintained that, in democratic societies, we can view the criminal laws as a set of expectations to which citizens hold one another. Holding people to expectations involves a susceptibility to the reactive sentiments. According to considerations just adduced, when we are susceptible to indignation in response to some wrong it shows that we care that interactions between people are structured by those obligations. A society thereby identifies in its criminal laws a set of moral obligations that it takes to be especially important in governing interactions between citizens. In egalitarian societies, these are the obligations that protect everyone in society. In going in for a system of blame that expresses indignation on behalf of all victims of serious crimes, a society affirms that all of its members are owed moral consideration. When punishment expresses indignation on behalf of a victim of crime, a society thereby shows that it values that victim and refuses to countenance his mistreatment.

According to the reactive theory, this public blaming function serves as the basis for legitimate punishment. The values that underlie blame give us reason to express blame publicly in a way that speaks on behalf of the members of the community.³³ Failure to offer protection in the establishment and enforcement of laws to some members of the community signals a lack of respect for them. Part of what is objectionable about a society that punishes crimes against oppressed groups at a lower rate than other citizens is that it treats people in the oppressed groups as less valuable. When, on the other hand, a community aims to hold all of those who commit serious crimes to account, it shows that it takes seriously the idea that all of its citizens have rights, the violation of which it would be appropriate to be exercised about. It also reinforces the importance of those moral demands enshrined in criminal law as appropriately structuring relations between members of the community. Societies have reason to blame publicly in order to show that they value all their citizens and take certain moral standards to be inviolable.

We have reason to blame serious crime not only to speak on behalf of members of the community, but also in a sense to speak *to* them: we have reason to blame in a way that citizens understand as expressive of blame. Anyone in society with an interest in seeing that those who commit serious wrongs are held accountable by the community for their wrongdoing should be able to understand that they are. One important potential audience is the family and friends of victims of crime. These people have an interest in having the victim's value affirmed by the community.³⁴ Marginal groups in society are

³³ The way that it speaks on behalf of the entire community in rejecting wrongdoing is one of the social roles of punishment described by Feinberg, "symbolic nonacquiescence." He discusses a law in Texas according to which a man was not to be punished if he killed a man whom he caught having sex with his wife. Feinberg, Joel, 1965, "The Expressive Function of Punishment"; reprinted in Feinberg, Joel, 1970, *Doing and Deserving*. Princeton: Princeton University Press, p.103: "The demand for punishment in cases of this sort may represent the feeling that paramour killings deserve to be *condemned*, that the law in condoning, even approving of them, speaks for all citizens in expressing a wholly inappropriate attitude toward them. For in effect the law expresses the judgment of the 'people of Texas' in whose name it speaks, that the vindictive satisfaction in the mind of a cuckolded husband is a thing of greater value than the very life of his wife's lover."

³⁴ The way that punishment affirms the sense of victims of crime as being wronged is highlighted by T. M.

another important audience. Members of these groups have a special interest in seeing the affirmation of those members of their groups who are crime victims. People from these groups often have reasonable concerns about whether the institutions of society treat them as equals. These concerns feed feelings of alienation toward public institutions. By making good faith efforts to blame all those who commit serious crimes in ways that everyone in society can recognize as blame, institutions that express the indignation of the community can work against these feelings of alienation. A society can thereby show that it takes all its citizens to be owed the moral consideration that makes blaming responses on their behalf appropriate.³⁵

Punishment is justified on the reactive theory not simply because we are susceptible to certain emotions in response to wrongdoing, but rather due to the role those emotions play in our valuing practices. When we feel these emotions, we often reflectively endorse them as morally appropriate. We care about our fellow citizens so that we think it is appropriate to be emotionally exercised when they are the victims of serious wrongdoing. Systems of punishment not only give us a systematic way to hold criminals accountable, they provide a systematic way to evince that everyone in society is valued as being owed moral consideration.

4) Why Grant the State the Authority to Express Our Indignation?

I have argued for an understanding of state punishment in terms of the reactive sentiments, as expressive of the indignation of the community in response to serious crimes. Further I have argued that indignation is not morally problematic in the way that some other emotions punishment sometimes expresses and are rather is bound up with values at the center of morality. One might accept these conclusions, but still wonder why we give the state the authority to express our sentiments. We think the expression of most emotions is properly left to individuals, so why should we deputize the state to express our indignation toward the most serious crimes? In this section, I will offer a number of reasons to think that that state punishment is a superior way to express indignation toward serious crimes to punishment that is carried out by individuals.

John Locke's account of why we are better out of the state of nature provides an instructive jumping off point. According to Locke, punishment by individuals is legitimate in the state of nature: when someone transgresses the Law of Nature everyone

Scanlon, though he does not think that the expressive role of punishment is enough to justify its hard treatment for a reason that I address in the next section. See Scanlon, T. M., 1999, "Punishment and the Rule of Law"; reprinted in Scanlon, T. M., 2003, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press.

³⁵ The relationship between blaming-institutions and marginalized groups in society is very complex. In many cases their alienation from those institutions is grounded in reasonable concerns about the demographics of those blamed, rather than the demographics of victims affirmed. These concerns are sometimes legitimately directed at those institutions and can be alleviated by those institutions blaming more consistently. Other times, other institutions in the network in which the blaming-institutions are embedded are more legitimate targets of reform.

else in the state of nature has a right to punish him.³⁶ Were this not the case, we would not be able to explain why states can legitimately punish foreigners who have not consented to the laws as citizens have.³⁷ Though everyone has the right to punish violators in the state of nature, exercises of that right bring with them “inconveniences”:

I doubt not but it will be objected that it is unreasonable for men to be judges in their own cases, that self-love will make men partial to themselves and their friends; and, on the other side, ill-nature, passion, and revenge will carry them too far in punishing others, and hence nothing but confusion and disorder will follow, and that therefore God hath certainly appointed government to restrain the partiality and violence of men. I easily grant that civil government is the proper remedy for the inconveniences of the state of Nature....³⁸

From this description of the predicament of the state of nature, I will pull out three important shortcomings of the state of nature that give us reason to prefer state punishment. First, people are more likely to punish (and punish more severely) when they and those to whom they are close are victims of wrongdoing and less likely to punish (or punish less severely) when those to whom they are close commit wrongdoing. So whether, and extent to which, someone suspected of violating the Law of Nature will be punished is largely determined by the relations he stands in to others. Second, passion sometimes leads people to punish those whom they think have done wrong before being able to fully consider the evidence. Third, when people disagree about whether the Law of Nature has been properly enforced, this leads to further punishment and further disagreement. For example, imagine I believe that you have wronged my friend so I punish you. But your friends do not think you have done wrong so they punish me for the wrong I have done you. My friends now think you have wronged me. A hopeless spiral ensues.

According to Locke, as the quoted passage suggests, in civil society people give up their right to punish to the state in order to avoid the inconveniences of the state of nature

[A]ll private judgment of every particular member being excluded, the community comes to be umpire, and by understanding indifferent rules and men authorised by the community for their execution, decides all the differences that may happen between any members of that society concerning any matter of right,

³⁶ See the *Second Treatise of Government*, II.8: “every man hath a right to punish the offender, and be executioner of the law of Nature.”

³⁷ *Second Treatise of Government*, II.9: “Those who have the supreme power of making laws in England, France, or Holland are, to an Indian, but like the rest of the world—men without authority. And therefore, if by the law of Nature every man hath not a power to punish offences against it, as he soberly judges the case to require, I see not how the magistrates of any community can punish an alien of another country, since, in reference to him, they can have no more power than what every man naturally may have over another.”

³⁸ *Second Treatise of Government* II.13.

and punishes those offences which any member hath committed against the society with such penalties as the law has established....³⁹

When a system of criminal law and punishment replaces the anarchy of the state of nature, confusion and disorder are mitigated and the shortcomings of the state of nature are overcome. The state executes the law consistently so that whether or not people are punished is a function of whether they have violated an agreed upon conception of the Law of Nature rather than the relationships of partiality that they stand in with others. Procedures are put in place in order to determine whether those accused of breaking the law are guilty. Because only the state has the right to punish, spirals of punishment are not unleashed.

The reasons to turn over the personal right to punish to the state that Locke adduces can be utilized by the reactive theory of punishment in order to explain why we have reason to give the state, rather than individuals, the authority to express indignation toward the most serious crimes. Think of the state of nature as one in which people are susceptible to feeling the reactive sentiments, but where there is no agreed upon institutional means of expressing resentment and indignation toward the most serious crimes. People generally feel the sentiments most intensely in response to the most serious wrongs. These sentiments dispose them to act in ways that are bad for their objects, but they do so somewhat inconsistently. In response to murder some are disposed to inflict physical pain on the murderer, others are disposed to take his property, others are disposed to ostracize him from society, and still others are disposed to deny him his freedom of bodily movement. It is clear that blaming responses in this state of nature will be inconsistent, based upon the actions that people's reactive sentiments dispose them to perform. Those who have many friends will be more likely to have others blame on their behalf, and those with few friends might not have anyone stand up for them. On the other hand, when their friends are accused, people will often rally to their defense and be less willing to blame them. Other morally arbitrary factors, such as media attention and whether potential blamers have the material goods to carry out the actions to which their blaming emotions dispose them, will play a role in whether those who commit acts of wrongdoing are blamed. As in Locke's state of nature, some will rush to judgment. Sometimes those in the grip of resentment or indignation will act on the emotion before they have gathered any evidence or given it careful consideration. Also, as in Locke's state of nature, spirals of punishment are an ever-present possibility. When people believe they have been wrongly blamed, they and their friends will blame those who have wrongly blamed them.

The predicament of the state of nature in which everyone blames each other however they see fit gives everyone reason to deputize the state to blame serious wrongdoers by depriving them of goods whose successful deprivation usually requires the threat of force. People in this version of the state of nature have reason to come together to decide upon which wrongs are so serious that they merit criminalization and how those crimes should be punished and then give up the prerogative to punishment to

³⁹ *Second Treatise of Government* VII.87.

the state. When people forfeit the right to use force, spirals of personal punishment are avoided. Disputants agree that the court system will serve as the final word on whether individuals are blameworthy for serious wrongs. Another advantage of a state punishment over personal punishment is that state resources can be used to establish court systems which follow standards of evidence and proceed deliberately so that those who are accused of serious wrongdoing are given due process. The accused have a reasonable claim to careful consideration of the evidence, including evidence concerning whether they ought to be excused or exempted, before being denied the important goods characteristic of punishment. Lastly, it seems that there needs to be a regularization of our responses so that whether, and to what extent, people in society are the objects of blame which deprives them of important goods depends upon whether they have done wrong, and if so, the seriousness of that wrong. In authorizing the state to express our indignation for serious crimes through punishment we are able to operationalize the blame that we find appropriate and apply it consistently.

It is only through responding to serious crimes *consistently* that a community is able to show all citizens that it takes them to be owed moral consideration. A community's claim to value all its citizens would be undermined in a state of nature where whether anyone was blamed for a particular instance of wrongdoing was in large part a function of whether the victim of wrongdoing had many friends, wealth, and media support. In fact, without state action, it is difficult to see how punishment could reflect the commitment of a *community* to stand up for the victims of crime, and so to perform the public blaming function described in the previous section, as opposed to the commitment of the individual expressing his indignation. State punishment that is answerable to democratic procedures can be understood to express "the indignation of the community" in a way that would only occur in the state of nature when a blaming response is coordinated among most of its members. Even in those rare cases in the state of nature where the community comes together to express their indignation toward one case of wrongdoing, they only show that they value the person who has been wronged, not everyone in the community. Alternatively, when the community comes together to establish a system of public blame, which offers moral protection to all its members, the community shows that it cares about all its members.

In participating in the maintenance of a system of criminal justice, not only is a community as a whole able to show that it values all its members, but individuals are also able to do so in a way that is unavailable to them absent such a system. Even punishing one person in the state of nature, while at the same time giving him something like due process, would regularly take a great amount of time and resources for an individual. When we pool our time and resources in order to support a system of punishment, we are able to publicly blame all wrongdoers in a way that is not too demanding on individual citizens. The ways in which the average citizens participate in the criminal justice system—by, for instance, funding the system by paying taxes, participating with criminal investigations, serving jury duty, and electing representatives—do not place unreasonable demands on them. This leaves citizens free to pursue other valuable ends in their lives besides blaming wrongdoers, while still exhibiting a commitment to the value of all

members of society and the importance of certain standards in governing their interactions.

5) Conclusion

I have argued that we should understand systems of punishment as expressive of indignation of the community in response to serious wrongdoing. Punishment thereby blames criminals for their serious crimes. I maintained that we have a number of reasons to grant the authority to punish to the state. When it is done well, state punishment is executed consistently and impartially, gives due process to the accused, places disputes over blameworthiness over serious wrongdoing in the courts so that spirals of punishment are avoided, and speaks on behalf of the whole community in upholding the values that make public blame morally important. I began this chapter by alluding to Feinberg's skepticism about whether punishment can be justified by the emotions it expresses. I argued that the reactive sentiments are not problematic in some of the ways that the other emotions punishment expresses are. In the next chapter, I address his (and others') main source of skepticism about the justificatory force of the expressive function of punishment: the suspicion that it could be carried out just as well by expressions that do not involve the harms of punishment.

Chapter 3

Reasons to Express Indignation with Punishment

In Chapter 1, I claimed that a theory of punishment must satisfy the *positive reason desideratum*.

Positive reason desideratum: A theory of punishment should explain why we always, or almost always, have positive reason to punish serious crimes.

I argued that traditional theories of punishment fail to satisfy this desideratum. Consequentialist theories fail because we have reason to punish when doing so does not achieve the best possible state of affairs. Retributive theories fail for two reasons. First, they rely on the groundless idea that wrongdoers deserve to suffer or be harmed. Second, we think that punishment can be justified even if it does not fulfill its purported retributive function of inflicting harm or suffering on the punished.¹

I believe that the reactive theory of punishment described in the previous chapter offers a promising alternative account of the positive reason we have to punish serious crime that avoids the difficulties of the traditional alternatives. According to the reactive account, the state has positive reason to punish because doing so expresses the indignation of the community and thereby shows that members of the community value one another as being owed moral consideration. Instead of relying on the questionable retributive idea that wrongdoers deserve to suffer or be harmed, the reactive theory relies on a “reactive idea” that is widely shared: that reactive sentiments are appropriate responses to wrongdoing. Rather than being “a mysterious piece of moral alchemy,” the fittingness of these sentiments is a lived conviction for most of us, given to us with interpersonal relationships. The reactive idea, unlike the retributive alternative, is consistent with our interpersonal responses to minor wrongs. We think it is appropriate to feel the reactive sentiments in response to such wrongs, but not necessarily to desire that the person toward whom the sentiments are felt suffer or live a worse life in virtue of that wrongdoing.²

The reactive theory potentially offers a superior account to consequentialism and retributivism of the positive reason we have to punish, because every legitimate act of punishment expresses appropriate indignation toward a criminal for a serious act of wrongdoing and thereby carries out its primary justifying function. Thus, it is in line with our shared moral thinking about when punishment is justified. We tend to think that

¹ This second criticism does not apply to all versions of retributivism, a point I explained in Chapter 1 and revisit in the conclusion of this chapter.

² One might think that the appropriateness reactive sentiment of guilt shows that it *is* intrinsically good that wrongdoers suffer even in response to minor wrongs. As with the other reactive sentiments, someone who feels guilty shows that he values the person he has wronged. While guilt is generally an appropriate response to wrongdoing, it is unclear to me that its appropriateness is necessarily relevant to third-person responses to wrongdoing. Remember the case from Chapter 2, where your friend forgets to give you a ride to the airport. It is appropriate for him to feel guilty in that case. But it also seems it would be inappropriately vindictive for you to attempt to guilt trip him in order that he suffers appropriately, as opposed to, say, telling him he didn't give you a ride in order to stand up for yourself.

punishments of serious crimes are justified before considering whether those individual acts of punishment will produce the optimal consequences or inflict a significant amount of suffering on the punished. Commonsense morality does not seem to hold that punishment has to accomplish anything beyond itself in order to be justified; rather, we contemplate with equanimity a number of possibilities when it comes to justified punishment. This attitude can be explained on the reactive theory, because insofar as punishment is an appropriate expression of blame, it is justified. Nothing further is required.

One might object at this point that there are serious crimes for which punishment is called for but not indignation. Were this so, the reactive theory would not be able to generally explain the positive reason we have to punish serious crimes. I have in mind two classes of cases: those in which indignation is inappropriate because the criminal is not blameworthy, and those in which indignation is inappropriate because there is no victim on behalf of whom to be indignant. With respect to the first, consider strict liability offenses, for which neither criminal intent nor negligence is required to be proven for conviction. While strict liability is rare in the criminal law, in some jurisdictions serious crimes such as statutory rape are strict liability offenses, so that someone may be prosecuted for having sex with a minor even if the minor misrepresents his age through fake identification.³ Because strict liability laws punish those people who are not blameworthy for serious crimes, such as statutory rape, these laws undermine the reactive theory's claim to satisfy the *positive reason desideratum*. In response to this objection, I claim that it is not clear commonsense morality endorses such laws, even though they may contribute to the public good. Most people have the intuition that there is something unfair about punishing those who commit statutory rape despite taking all reasonable precautions in good faith, including checking for identification. I think underlying this intuition of unfairness is the belief that the person who has sex with a minor is not blameworthy for doing so, which gives us reason not to punish him. Ambivalent responses to strict liability laws in our moral thinking suggest that rather than being paradigm test cases for theories of punishment, strict liability offenses are the sort on which a theory of punishment might give us some guidance. If it turns out that the utilitarian theory of punishment is the most plausible justification, and such laws contribute to the overall happiness of society, they are defensible. If, on the other hand, it turns out that the reactive theory gives the most plausible view, we have reason to reconsider strict liability in criminal law.

³ For a description of such a case, see Carpenter, C. L., 2003, "On Statutory Rape, Strict Liability, and the Public Welfare Offense Model," *American University Law Review*, 53/2: 313–390, pp. 351–352: "The defendant in *Jenkins [vs State]* was convicted of statutory rape, called sexual seduction in Nevada, and sentenced to sixteen years in prison for having sexual relations with two girls under the age of sixteen. There was no dispute that Jenkins did believe, and could have reasonably believed that one of the victims, Sherry, was sixteen years of age and therefore able to consent. In fact, the court accepted as true that Sherry misrepresented her age to the defendant. Despite the victim's misrepresentation and the harsh sentence that awaited Jenkins, the court was not persuaded. Rejecting a mistake-of-age defense, the court determined that the crime of statutory sexual seduction was a strict liability offense."

Next consider cases in which indignation, as understood in the previous section, seems inappropriate because there are no crime victims, so no one on behalf of whom to feel indignant, but we think that there is positive reason to punish. A critic of the reactive theory might object that the importance of showing commitment to the idea that everyone in society is protected by certain moral obligations cannot explain why we have reason to punish in such cases.⁴ There are at least two classes of uncontroversial cases in which the reactive theory is able to explain why we have reason to punish serious crimes without actual victims. There are other sorts of cases of victimless crimes for which the reactive theory might not recommend punishment and so be at odds with some currently existing criminal laws. But these laws, like strict liability laws, are controversial, so that it is not obvious that those laws can be taken as counterexamples. Instead, it may be that the most plausible theory of punishment will explain why we do not have reason to reconsider those laws.

One class of uncontroversial cases in which we have reason to punish crimes without actual victims is those in which there are no victims because a criminal fails in his attempt to carry out his crime. These cases are unproblematic for the reactive theory because the reactive sentiments are appropriate responses to attempts to wrong others. If someone attempts to kill me, but fails due to incompetence, I can justifiably feel resentment in response. Another class where we seem to feel the reactive sentiments even when there are no victims are when people engage in behavior which puts others in grave danger even when they do not intend to wrong those they put in danger. For instance, if someone checks to see whether his gun works by shooting it into a crowd of people he has no intention of harming, even if they are not shot those in the crowd might justifiably feel resentment. There are other classes of criminalized actions without identifiable victims, actual or potential, for which it does not seem that the reactive theory would recommend punishment. Consider laws against consensual same-sex sodomy that were ruled unconstitutional by the United States Supreme Court in *Lawrence v. Texas* in 2003. Punishment of acts of consensual sodomy did not have the valuable public blaming function of expressing indignation on behalf of any victim.⁵ With respect to this victimless crime, I believe that the reactive principle yields a plausible result.

⁴ One potential response might maintain that we always have reason to be indignant when someone knowingly breaks the law because they have failed to give due respect to the community whose will is reflected in the law. But this would give us positive reason to punish someone who violated immoral laws, and it does not seem to me that commonsense morality maintains states have reason to imprison people who break laws that are themselves immoral. For instance, Nazi Germany did not have moral reason to punish those who violated the Nuremberg Laws (and would not have, even if those laws had been enacted by democratic procedures).

⁵ Martha Nussbaum notes that when Oscar Wilde was convicted of gross indecency for consensual homosexual behavior, at sentencing Justice Wills claimed his punishment “is consistent with the utmost sense of indignation at the horrible charges brought home to both of you.” Nussbaum claims that Wills is mischaracterizing his emotion, which is really disgust rather than indignation. Nussbaum, Martha, 1999, “‘Secret Sewers of Vice’: Disgust, Bodies and the Law” in Bandes, Susan A. (ed.) *The Passions of Law*. New York: New York University Press, p.46–47: “Disgust hides behind the screen of indignation, but it is clearly disgust, not indignation at harm that it driving the sentencing. What is really being said is: these are two slimy slugs who ought to be squashed before they insinuate themselves into our bodies.”

With respect to other so-called victimless crimes, it is not always as obvious whether there are victims on behalf of whom one might appropriately be indignant and this might explain why many of us have ambivalent feelings about criminalizing such behavior. I believe that whether or not we ought to consider such actions serious crimes depends in large part on the extent to which such crimes in fact violate moral obligations. Consider, for example, criminal laws against drug possession and use. While such laws are pervasive in the American criminal justice system, many critics believe that we do not have reason to spend state resources punishing drug crimes. I believe it is a question up for debate whether drug possession and use is an action that has victims that would make public blame important. A supporter of criminalization might identify victims of drug use. First, those who habitually take drugs sometimes do harm to themselves, so the supporter might argue that drug users violate some moral obligation they owe to themselves. He may argue that we have reason to express indignation toward drug users because in doing so we show that we value him as being owed moral consideration from himself. Further, the supporter of criminalization might point to the way that some drugs can diminish their users' capacities for moral agency so increase the likelihood they will act wrongly. He might claim we have reason to be indignant on behalf of those who might be victimized because of those diminished capacities. Finally, he might point out how dangerous the drug trade is. He might maintain that without demand from drug users the drug trade would cease and claim fewer victims.

I do not find these convincing arguments for the criminalization of drug use without further development.⁶ However, I hope they suggest some considerations that might be available to the reactive theory in determining whether so-called victimless crimes ought to be punished. If it turns out that in punishing some action the community cannot rightly be understood to express public blame on even potential victims, the reactive theory will not give us reason to punish that action. But it does not seem to me that in doing so the reactive theory fails to explain why we have reason to punish a *serious* crime. Consequentialism and retributivism, on the other hand, are susceptible to counterexamples for which they do not explain the positive reason we have to punish murder.

I have defended the reactive theory's claim that when we have reason to punish, we have reason to blame publicly in order to stand up for the victims of crime against a number of counterexamples. Whether the reactive theory is actually able to satisfy the *positive reason desideratum*, however, depends on defending the additional claim that punishment is an appropriate way to express indignation. One could accept that punishment does blame criminals for crimes, and that blaming those who commit serious crimes is morally important, but deny that we have reason to blame criminals with punishment because there are other, less costly ways for us to express the indignation of

⁶ With respect to the first argument, it does not seem to provide a general argument for punishing drug possession, since a significant number of drug users are not worse off for having taken drugs. Further, it is not clear that it is the state's business when members of the community violate obligations toward themselves, rather than others in the community. With respect to the first and second arguments, they would seem to prescribe criminalizing substances that we do not, such as alcohol. I find the third argument unconvincing, since much of the danger of the drug trade is a consequence of the criminalization of drugs.

the community. In the following section I argue that norms that govern emotions explain why punishment is a not only natural, but also good, mode of expression toward the most serious crimes, given the values that make public blame appropriate. In sections 2 and 3, I respond to some objections to this argument.

1) Why Not Say it with Weeds?

Many philosophers who acknowledge the expressive function of punishment are skeptical that that function gives us reason to deprive criminals of their liberty or property.⁷ T.M. Scanlon, for instance, writes

Pointing out “the expressive function of punishment” helps us to understand our reactions to punishing particular kinds of people, but what role if any does it have in the justification of punishment? It seems to have no positive role in justifying hard treatment of the legally blameworthy. Insofar as expression is our aim, we could just as well “say it with flowers” or, perhaps more appropriately, with weeds.⁸

Contra Scanlon, I will argue that when punishment is understood as an expression of indignation, the expressive function of punishment explains why we have reason to subject criminals to the characteristic deprivations of punishment.

Consider how expressions of another moral emotion, compassion, might be criticized. Imagine a very wealthy man is sitting in front of the television watching coverage of a natural disaster that has been devastating to some members of his community. He tells his personal assistant who is working nearby, “I feel a great deal of compassion for the victims of this disaster. Please send them flowers with a note expressing this compassion.” She replies, “If you were really compassionate, you could send them money to help provide for their basic needs and give them an opportunity to rebuild their lives. In fact, if you were really motivated, you could spend some time volunteering at the shelter that has been set up.”

There are at least two ways in which the wealthy man’s gift of flowers is deficient as an expression of compassion. First, the expression is not constructive in light of the

⁷ See for example, H. L. A. Hart’s objection to Lord Denning’s view that punishment is ultimately justified as an expression of condemnation. Hart, H. L. A., 1963, *Law, Liberty and Morality*. New York: Random House, p. 66: “What is meant by the claim that punishment of offenders is an appropriate way of expressing emphatic moral condemnation? The normal way in which moral condemnation is expressed is by *words*, and it is not clear, if denunciation is really required, why a solemn public statement of disapproval would not be the most ‘appropriate’ or ‘emphatic’ means of expressing this. Why should denunciation take the form of punishment?” More recent critics of the idea that punishment is justified by its expressive function include Boonin, David, 2008, *The Problem of Punishment*. New York: Cambridge University Press, pp. 176–179, and Hanna, Nathan, “Say What? A Critique of Expressive Retributivism,” *Law and Philosophy* 27: 325–349.

⁸ Scanlon, T. M., 1988, “The Significance of Choice.” In *The Tanner Lectures on Human Values*, Vol. 8. Salt Lake City: University of Utah Press.

values that partially constitute the compassion. When people feel compassion, they judge that the suffering of the objects of that compassion is bad. A purported expression of compassion that does nothing to alleviate this suffering when such alleviation is possible shows a lack of commitment to the value judgment that is characteristic of compassion. Second, a gift of flowers seems not to express a sufficient degree of compassion. Emotions have affective elements that lead to action. Compassion involves a disposition to engage in actions that typically alleviate suffering. Generally, the more intense the compassion is, the more one is willing to help. The wealthy man's refusal to do anything for the victims of the natural disaster beyond sending flowers belies his claim to feel a great deal of compassion. The suffering brought by the natural disaster in this example calls for a high degree of compassion, and merely sending flowers does not express enough of it. These mundane observations suggest that in expressing compassion, we cannot always just as well say it with flowers.

1.1 Punishment as a Constructive Expression of Indignation

When expressing indignation toward serious crimes, norms of constructiveness give us reasons to express our indignation with punishment rather than weeds. Constructive expressions of emotions are those in which the value judgment which partially constitutes the emotion is reflected in the very manner of its expression. Such expressions preserve and protect what the emotions value (or do the opposite to what the emotions disvalue). While I do not think that this notion of constructiveness perfectly tracks our emotional vocabulary, I think that it is intuitive that, when our emotions are appropriately felt, we have reason to express our emotions in ways that are constructive in this sense. This intuition can be reinforced with examples of expressions of emotions that might be evaluated as more or less constructive. I suggested that with respect to compassion, constructive expressions will aim to alleviate the suffering of their object. Next, consider admiration. Admiration is partially constituted by the value judgment that the object of admiration is good. Imagine the very wealthy man admires the practice of medicine. Setting up a scholarship which encourages low income undergraduates to study premed seems like a constructive way to express this admiration. It is likely to increase the number of people who practice medicine. On the other hand, simply telling his adult daughter who is a painter that he really admires medicine a great deal more than art is unlikely to do much to promote those values that inspire his admiration. Next, consider fear. When we feel fear, we typically anticipate something that we consider bad happening. Constructive expressions of fear keep what we consider bad from occurring. I more constructively express my fear of not being able to pay rent by looking for a job rather than lying in bed, "paralyzed" by fear. Finally, consider love. Among the value judgments love involves seem to be that you take the interests of your beloved to be important and that you value your relationship with her. In an example made famous by Bernard Williams, a man standing on the deck of a boat can only save the life of one of two people, one who is his wife and one who is a stranger.⁹ If, out of love, the man dives

⁹ In Williams, Bernard, 1976, "Persons, Character, and Morality"; reprinted in Williams, Bernard, 1981,

into the water and saves his wife's life, this seems to be a constructive expression: this action treats her interests as more important than the stranger's and preserves their relationship. If, instead, he saves the stranger's life, but yells "I love you," to his wife as he jumps off the boat, this expression is relatively unconstructive. In all these cases those feeling the emotions have *prima facie* reason to express them constructively. In doing so they show a commitment to what they value, and those values are realized in the world.

In chapter 1, I suggested that a plausible account of the justification of punishment could not disregard the moral significance of the social goal of crime reduction. The values that make blame important give us reason to express it in ways that contribute to this goal. While we characteristically perceive the significance of suffering when feeling compassion, when we are indignant we perceive the importance of the moral consideration that victims of wrongdoing are owed. Where possible, constructive expressions of indignation will reduce the likelihood that others are victimized. Such expressions demonstrate a commitment to the values that partially constitute appropriate indignation.

Punishment's deterrent function makes it a constructive way to express indignation in light of the values that appropriately inspire our indignation. The denial of goods to a criminal has the potential to deter crime, insofar as the punished and others in society who are aware of the punishment do not want to be denied the good in question. The particular goods that punishment denies can play further roles in crime reduction. Imprisonment, for instance, incapacitates those who are risks for committing crimes and places them in a situation in which they will have a lot of time to think about their wrongdoing. Prisons often (and should) have programs whereby criminals can earn college degrees and develop skills that will make it less likely that they will commit future crimes. In these ways, imprisonment increases the likelihood that moral obligations are not violated in the future and in doing so, protects the rights of members of the community. When we express indignation toward serious crimes through imprisonment, the values that explain the moral importance of that indignation are reflected in the manner of its expression. This is less obviously true if we express our indignation with weeds. Such expressions do not deter crime as effectively because people are not as likely to avoid receiving weeds as they are the characteristic harms of punishment. Weeds do not incapacitate and they do not as effectively encourage criminals to reflect on their crimes. We cannot just as well express indignation toward serious crimes with weeds as punishment, because such expressions do not show the same commitment to the values that make the reactive sentiments appropriate responses to wrongdoing.

1.2 Punishment as a Proportionate Expression of Indignation

One might be skeptical of whether norms of constructiveness can actually explain why we have general reason to punish serious crime. After all, I argued that consequentialist theories do not satisfy the *positive reason desideratum* because they do

not explain why we have reason to punish in cases in which the best outcome is not achieved by punishment. Similarly it might seem that considerations of constructiveness will only give us reason to punish in those cases in which punishing someone actually does deter crime, which not every act of justified punishment does. I believe that norms of constructiveness give us reason to punish in even these cases, but indirectly, because they inform our judgments about which forms of expression are proportionate expressions of indignation.

In this subsection I argue that norms of proportionality give us reason to express indignation toward the most serious crimes by means of punishment. We understand expressive behaviors that are objectively worse for their targets as signaling those sentiments in a greater degree. In order to signal to all members of the community the value of those who have been victimized, the form that blame takes must be accessible to everyone in society as expressive of an unambiguously high degree of blame. A society that consistently punishes serious crimes emphatically shows that it values the victims of those crimes as being owed moral consideration. Expressions that deprive their targets of something less important could legitimately be viewed by members of society as failing sufficiently to stand up for those victims.

The idea that expressions of emotion are governed by norms of proportionality so can be better or worse depending on whether they express an appropriate degree of indignation is familiar from the way that we think about the emotions. The wide range of actions to which emotions can dispose us is instrumental in indicating to others the degree to which those emotions are felt. The way we act sometimes shows that our emotions are felt to a deficient or excessive degree. When the very wealthy man expresses compassion in a way that does nothing to ease the suffering of those in his community, or the husband of the previous section shouts “I love you” to his wife as he saves the stranger, it shows that they feel those emotions to a deficient degree (if at all). People who feel those emotions to a higher degree would be disposed to behave differently. Conversely, if someone starts sobbing when his pen runs out of ink, in most circumstances his tears express an excessive degree of sadness. Acting in a way that expresses a proportionate degree of an emotion is especially important when that emotion is significant to a potential audience. If a friend of a friend with a pickup truck helps you to move across town, you have reason to express gratitude toward him, and certain ways of doing so would be deficient or excessive. A terse nod alone when the move is complete expresses too little gratitude, and designating him a benefactor in your will expresses too much.

Like other emotions, indignation disposes us to a wide range of actions and which actions it disposes us toward is partly a function of the degree to which we feel indignant. As a rough rule of thumb, we understand actions that are objectively worse for their target to express more indignation. Separate into three categories typical modes of expressing indignation I described in Chapter 2: merely verbal expressions, denials of social goods, and punishment. Actions in each category typically express progressively more indignation because they are objectively worse for their targets. The gap between how objectively bad they are for their targets are wide enough that we can express a wide variance of degrees of indignation across these categories. Actions in each category

bring with them the objective bad for their targets of the awareness that someone feels a reactive sentiment toward you, so denies you social regard. Actions in the second and third categories deny further objective goods beyond social regard. Typical expressions in the second category like exclusion from social circles and refusal to help with projects deny their targets goods that typically contribute to a good life. Punishment often involves the denial of such goods. When we imprison someone, for instance, we usually separate him from his social circles. When we fine someone heavily we usually make it more difficult for him to carry out his projects. But goods denied by punishment also go beyond the other categories. Unlike those denied by the other categories they take away goods, such as liberty and property, whose denial requires the threat of force. Holding fixed our relation to the object of our reactive sentiments, it seems that the category of good the denial of which our reactive sentiments dispose us to countenance roughly corresponds to the degree to which we feel those sentiments. Blaming serious criminals by denying them goods in the third category shows that we endorse a high degree of indignation in response to their crimes.

Given considerations of proportionality and blame's public role as described in the previous section, we have reason to express blame toward those who commit serious crimes by depriving them of goods in the third category. Such expressions are accessible to everyone in society as expressive of a high degree of indignation. We cannot just as well express blame with weeds; weeds would not express a sufficient degree of indignation for this task. Imagine that our government suddenly starts to say it with weeds, sending weeds to convicted murderers rather than punishing them. We and our fellow citizens are likely to have trouble conceptualizing such a procedure as one for doling out blame. Giving someone weeds is not an action that is typically objectively bad for its target and thus is difficult to understand as even a candidate expression of indignation.

But imagine the government makes clear the expressive meaning of these weeds and claims that they are to represent the highest degree of indignation. Those who receive the weeds would presumably be deprived of some benefit of social regard. There may be some in society who would accept this system as expressive of the appropriate amount of blame for murderers, but I suspect most people, given their own emotional experience, would find the expression insufficient. Many victims of serious crimes, and those who care about them, will still reflectively endorse their own reactive sentiments that dispose them to approve depriving their perpetrators of goods beyond social regard. If there is not collective acceptance of the weeds-dispensing system as expressing a sufficient degree of blame, the system will not function in a way that can be seen by everyone in society as expressive of appropriate blame. It will seem to many to fail to stand up adequately for the victims of those crimes, and therefore fail to fulfill its public blaming function. Giving weeds to murderers in response to their wrongdoing would neither speak for the community in expressing the degree of indignation they find appropriate in response to murder nor speak to victims of crime and those with an interest in seeing that their victimizers are held responsible in a way that makes clear that the community takes sufficiently seriously the idea that they are owed moral consideration.

We are now in a position to see how considerations of constructiveness inform our views about proportionality. While it seems to be in the nature of indignation that when we are more indignant we are disposed to act in ways that are objectively worse for the object of our indignation, the actions we are disposed to perform are partially molded by the societies in which we receive our emotional educations.¹⁰ Having grown up in societies in which imprisonment is thought to contribute to crime reduction surely is part of the reason why we are generally disposed to express indignation toward murderers through imprisonment. The disposition to punish when we feel indignation towards serious crimes that we develop is general, however, and is not sensitive to cases in which considerations of constructiveness would not give us reason to punish. In such cases if the state were to express indignation in ways that were not as objectively bad for their targets, it would seem to many in society as though it were expressing a lower degree of indignation than is appropriate on behalf of victims of those crimes. According to the reactive theory, we have reason to punish serious crime primarily because we have reason to respond to wrongdoing in a way that is accessible to everyone in society as expressing a high degree of blame. Punishment always does this. The reason why it takes the form of denying its objects liberty and property in particular is partially provided by the role that such deprivations play in deterring crime.

2) Alternative Societies

By subjecting criminals to punishment, we publicly acknowledge the moral standing of their victims, and show that we are committed as a society to the importance of certain moral standards in regulating our interactions with one another. I have argued that the public blaming role that punishment plays in our society explains why the state always has reason to punish. In section 3, I consider whether that argument commits the reactive theorist to the criminalization of behavior that we generally do not consider punishable. In this section I look at challenges to the argument that might be posed by societies with different blaming practices than our own. First, I will examine whether difficulties are raised for the reactive theory by the possibility of societies in which people do understand non-punitive expressions as indicating a high degree of blame. Then I will take up the possibility of societies in which cruel treatment would be required to play the blaming function of punishment, and whether the reactive theory must endorse cruel punishment in those societies.

¹⁰ For a view that the actions to which our emotions dispose us are both biologically and culturally determined, see de Sousa, Ronald, 1987, *The Rationality of Emotion*. Cambridge, MA: MIT Press, p. 182: “My hypothesis is this: We are made familiar with the vocabulary of emotion by association with *paradigm scenarios*. These are drawn first from our daily life as small children and later reinforced by the stories, art, and culture to which we are exposed. Later still, in literate cultures, they are supplemented and refined by literature. Paradigm scenarios involve two aspects: first, a situation type providing the characteristic *objects* of the specific emotion-type, and second, a set of characteristic or ‘normal’ *responses* to the situation, where normality is first a biological matter and then very quickly becomes a cultural one.”

I have argued that considerations of constructiveness inform which actions we think express a proportionate degree of indignation because the actions to which our emotions dispose us are partially a result of the societies in which we grow up. One might wonder whether this opens up the possibility that there might be societies in which the public blaming role of punishment could be expressed by weeds if people in those societies are given the right emotional education so that the highest degree of indignation disposed them to approve of the dispensing of weeds to serious criminals. Wouldn't this show that we have reason to retrain our emotions so that we could express our indignation toward serious crimes in less harmful ways that required fewer state resources?

I concede that it is possible that a society whose citizens are not disposed to punish by indignation is possible. Nevertheless, before considering the implications of this possibility it is worth remarking that it is difficult to find a society that actually fits this description. This reinforces the idea that there is a biological basis for the connection between indignation and the disposition to punish, even if we can dream up possible cultures in which the connection is severed. In *The Case Against Punishment*, Deirdre Golash recognizes that what she calls anger (and I have been calling resentment and indignation)¹¹ is an appropriate response to wrongdoing, but questions whether punishment is a good way to express it. She acknowledges that those of us raised in cultures with systems of punishment are likely to think that punishment is a good way (in fact the only way) to express justified anger toward those who commit crimes. But she claims that in other cultures anger is expressed in ways that are not harmful, and that gives us reason to reconsider the way that we express anger in our own. Golash cites anthropologist Catherine's Lutz's rich account of the emotional life of inhabitants of a Micronesian atoll.

Anger has, in other times and other places, been directed toward ends other than harm of the wrongdoer.... In the South Pacific atoll of the Ifaluk, justified anger is acknowledged through a stylized speech that a designated person will give to the angry person, recognizing the angry person's desire to fight, but appealing to his compassion, his desire for the respect of the community, and the bad consequences of violence to urge him to exercise self-control. A decision to exercise self-control is the expected, and almost universally attained outcome.¹²

The case, however, does not provide an example of how the reactive sentiments in other cultures are characteristically expressed in harmless ways; it seems instead to be a case in which anger is not expressed at all. The angry person is counseled not to hold the object of his anger accountable for his wrongdoing but to "forget it," according to Lutz. The

¹¹ In common language usage "anger" seems to have much wider application than resentment and indignation. It seems, for instance, that I can be angry that a mosquito is keeping me up at night, without feeling resentment.

¹² Golash, Deirdre, 2005, *The Case against Punishment: Retribution, Crime Prevention, and the Law*. New York: New York University Press, p. 69.

counseling involves “frequent statements that the person at whom the justifiable anger is directed is in fact ‘crazy and confused.’”¹³ The adviser characterizes the object of anger in terms that we would use to exempt someone from moral responsibility. She gives reasons not to blame at all, rather than reasons to channel blaming responses to ends that don’t involve harm.

Further investigation into the practices of the Ifaluk reveals that there are times when indignation is deemed to be appropriately expressed by hard treatment. Lutz describes the reaction of a village in Ifaluk toward news that some of its members had disturbed the peace late into the night in another village.

[M]ost emotional talk centered on the certainty that “the chiefs will be justifiably angry [*song*]” at the young men. The chiefs, as it turned out, were in fact justifiably angry, as evidenced by the fact that they soon met and decided to levy a fine of 200 yards of rope on the men.... This amount of rope, which represents hundreds of hours of work for the men who make it from coconut husk fiber, was a measure of the *song* of the chiefs. This episode epitomizes one of the most common ways that the concept of justified anger is used—a taboo is violated, a traditional law is disobeyed, and the people point not directly to the law but to the *song* of the chiefs.¹⁴

This episode, and its description in terms of justified anger, demonstrates that sometimes indignation is expressed on Ifaluk through hard treatment, a heavy fine in this case. The sentence of hard treatment is delivered by chiefs, whom everyone on the atoll recognizes as legitimate, and it is understood as an expression of indignation. In Ifaluk society, like in ours, there are official social channels in place to express indignation toward the most serious offenses through hard treatment. The exact form that punishment takes is different, but that should be expected given differences between our cultural circumstances. For instance, in small communities social ostracism might be understood to express a high degree of indignation and its prospect might have a significant deterrent effect. In large modern communities where people are mobile and not as aware of the wrongdoing of their fellow citizens, ostracism would not have the same social meaning and consequences.

I suspect that the connection between indignation and the disposition to punish is grounded in deep facts about our biological nature and the societies we naturally form. But we can certainly imagine societies in which weeds would be as good a way as punishment to express indignation toward serious crime. In light of the argument of the previous section, weeds would have to be a constructive and proportionate expression of indignation in those societies. Weeds will not be a constructive way to express indignation in response to murder in societies where imprisonment would much more effectively deter murders than weeds. If murder were prevalent in that society, with

¹³ Lutz, Catherine, A., 1988, *Unnatural Emotions: Everyday Sentiments on a Micronesian Atoll & Their Challenge to Western Theory*. Chicago: The University of Chicago Press, pp. 175–6.

¹⁴ Lutz, Catherine, A., 1988, *Unnatural Emotions: Everyday Sentiments on a Micronesian Atoll & Their Challenge to Western Theory*, p. 158

many repeat offenders, we would question the commitment of this society to the idea that all of its members have a value that would make murdering them a terrible wrong. Perhaps if weeds did have the social meaning of expressing the highest degree of indignation, however, they would have significant deterrent effect for those who have been raised in this society and educated to desire to avoid precisely this sort of blame. So imagine that in the alternative society giving weeds to serious criminals expresses a high degree of indignation and has an effective deterrent function. In this society, punishment would not be required to express blame in a way that shows a commitment to the value of all citizens. Thus the reactive theory would not explain why that society has positive reason to punish serious crimes.

I do not think that the possibility of a society in which the reactive theory does not justify punishment by itself is problematic for the theory. Commonsense morality is not committed to the view that every society for all time will have reason to employ our particular institutions of punishment. Those institutions would be ill-suited for Ifaluk society. An alternative system that did not require the threat of force and therefore did not require as many of the resources of society would thereby be appealing. But the positive reason we have to punish is relativized to the actual circumstances we find ourselves in, so that insofar as a significantly less costly non-punitive system of public blame is not really a live option for us given our cultural and historical circumstances,¹⁵ especially given the cost of retraining our emotions that it would require as well as the uncertainty about whether that retraining would be successful and whether those retrained responses would be as constructive.

Another possible alternative society seems to me more problematic for the reactive theory: one in which in order to be accessible to members of society, blame toward serious crimes must be expressed in ways that we would consider cruel. Unlike societies in which indignant people are not disposed to punish, it is not difficult to find a society that fits this description. Any democratic society that punishes cruelly will do. The same norms of proportionality that I argued give us reason to punish seem to give such societies positive reason to punish cruelly. In the remainder of this section, I will argue that norms of constructiveness explain why some cruel punishments are not good expressions of indignation. Nevertheless, I will conclude that in some societies the reactive theory will need to rely on resources outside of the public blaming function of punishment in order to rule out cruel treatment.

The values that partially constitute the reactive sentiments put limits on the form that legitimate modes of expression can take, ruling out some candidate modes of punishment. My argument that punishment is a constructive expression of indignation focused on two value judgments that the reactive sentiments involve: that a moral obligation ought to govern behavior, and that the person with respect to whom the obligation has been violated is owed moral consideration. There is a third value judgment associated with the reactive sentiments that pays a sort of respect to the person they are felt *toward*. It is this value judgment I appealed to in the previous section to

¹⁵ Of course, we might have reason to reform our expensive private prison systems in ways that do not without undermine the public blaming role that punishment plays.

argue that the reactive sentiments are not essentially dehumanizing. When one is susceptible to the reactive sentiments in one's interactions with someone else, one treats that person as someone to whom moral obligations are properly addressed, or in Strawson's phrase "as a member of the moral community."¹⁶ When we feel resentment or indignation toward someone else, we treat him as an agent who can understand moral obligations and regulate his behavior accordingly. We acknowledge his capacities for moral agency. Expressions of indignation that treat these capacities as if they are of no moral significance violate norms of constructiveness.

Consider, for instance, torture as a candidate way of blaming criminals. Because torture breaks the will of its target, it is not a fitting expression of emotions that are partially constituted by the recognition of the capacities for agency as objects of respect. It is not a constructive way to express blame because it fails to preserve those capacities that blame respects. In a recent article, David Sussman describes the way in which being the target of torture dismantles those capacities.¹⁷ By applying pain, the torturer treats the body of the tortured in ways that make the tortured relinquish control of his capacities for agency; they are no longer his and are used against him. It is clear how this operates in the most familiar cases of torture: specifically when torture is performed in order to extract information from someone, to put him in a position where he is no longer in control of whether or not he reveals that information. Even if torture is strictly punitive it has a similar effect, Sussman maintains, in that "the very basic structure of embodied agency is perverted or turned against itself."¹⁸

One might wonder if punishment is a constructive way to express indignation in light of the value of the moral agency of the objects of indignation. After all, in nearly all cases of punishment, we treat the punished in ways that they would not will for themselves. Acting contrary to someone's will, however, is quite different from destroying it. In fact it is an inevitable consequence of living with people with competing interests that they sometimes act against our will, and we sometimes act against theirs. They and we may do so while still respecting one another as responsible agents. Even imprisonment, which significantly limits prisoners' freedom, does not take away their capacities for moral agency. Torture, on the other hand, is fundamentally inconsistent

¹⁶ Strawson, P. F., 1962, "Freedom and Resentment"; reprinted in Watson, Gary, 1982, *Free Will*. New York: Oxford University Press, p. 73.

¹⁷ Sussman, David, 2005, "What's Wrong with Torture?" *Philosophy & Public Affairs* 33:1, pp. 1-33, p.31: "In normal action, a person's movements and feelings are expressive of his prior attitudes, desires, or intentions, which then have their significance recognized by others. But in ordeals, this expressive relation between the public and the private is reversed. Another person or institution has some interest, project, or attitude toward the victim, realized immediately in the body and bodily affects of the victim, who experiences himself as merely the ultimate passive recipient of these acts. His thoughts, his attention, his interest assume whatever form his body demands. In a sense, his body ceases to be his, to be the substance in which he expresses his own attitudes, intentions, and feelings in a way that can be meaningful for others as a form of self-expression. Since the victim cannot effectively reassert himself physically against the assault (by fighting, fleeing, or shielding himself), his body becomes the medium in which someone else realizes or expresses his agency. The victim here becomes little more than a point of pure receptivity, having the most basic forms of agency effectively at the command of another." .

¹⁸ Sussman, David, 2005, "What's Wrong with Torture?" p. 31.

with treating prisoners as moral agents.¹⁹ By destroying those capacities that make its target an appropriate object of blame in the first place, torture undermines itself as a way of holding people morally responsible.

Considerations of constructiveness are able to explain why we do not have reason to punish in many ways that we consider cruel. For instance, extended solitary confinement tends to have deleterious psychological effects on prisoners.²⁰ Insofar as it has these effects it does not preserve the capacities for moral action in the punished the way a constructive expression of indignation does. These considerations also seem to apply to capital punishment, insofar as it destroys those capacities that blame respects. I take it that commonsense morality is not committed to the idea that capital punishment is morally defensible and the reactive theory gives us reason to think it is not.

Still, there may be some forms of punishment that are cruel, but do not damage their objects' agential capacities. For instance, some forms of corporal punishment might be consistent with acknowledging the value of those capacities, if the amount of pain caused by the punishment does not destroy, disable, or do long term damage to the rational capacities of the punished. We might nevertheless think that it is cruel to intentionally cause physical pain. We can imagine societies whose members feel indignation that disposes them to treat criminals in those other cruel ways and punishment's public blaming function seems to provide reasons for the state to punish cruelly in such societies.

With respect to this deficiency, the blaming function has partners in guilt in consequentialism and retributivism. In fact, it is not obvious that consequentialism and retributivism can explain why we do not have reason to *torture* criminals *in our society*. The Utilitarian Principle will endorse torturing criminals whenever doing so maximizes happiness across society—and torture may be a very effective deterrent. Even versions of consequentialism that count acts of torture among bad outcomes will endorse torturing some criminals when doing so would serve to minimize the amount of torture that occurs across society. Those versions of retributivism that recommend retributive matching penalties, where possible, will recommend torturing torturers. With respect to those that do not specify the harm or suffering that criminals deserve, it is not obvious why this harm or suffering cannot take the form of torture. Perhaps retributivists could pursue an argument similar to the one I have offered that there is something ill-fitting between culpable wrongdoing and torture because of the rational capacities one needs to be deserving of harm or suffering. But such a strategy would still not explain why we do not have reason to punish cruelly when doing so does not undermine those capacities. In

¹⁹ See Sussman, David, 2005, "What's Wrong with Torture?" *Philosophy & Public Affairs* 33:1, pp. 1-33, p.14: "Unlike other kinds of unwanted imposition, pain characteristically compromises or undermines the very capacities constitutive of autonomous agency itself. It is almost impossible to reflect, deliberate, or even think straight when one is in agony. When sufficiently intense, pain becomes a person's entire universe and his entire self, crowding out every other aspect of his mental life. Unlike other harms, pain takes its victim's agency apart 'from the inside,' such that the agent may never be able to reconstitute himself fully.... [Torture is] a violation that is accomplished through the very annihilation of such agency itself, if only temporarily or incompletely."

²⁰ For a summary of the findings of a study of solitary confinement, see Grassian, S., 1983, "Psychopathological Effects of Solitary Confinement." *American Journal of Psychiatry*, 140: 1450–1454.

order to be consistent with the commonsense moral view that societies do not have reason to punish cruelly, consequentialism, retributivism, and the reactive theory must appeal to norms independent of what they argue generally justifies punishment.

3) Acts which Call for a High Degree of Indignation, but not Punishment

In response to the argument of section 1, there is another possible line of objection that does not require thinking of alternative societies, but instead can be motivated by thinking about responses to wrongdoing in our society. Above I argued that The Reactive Principle gives us positive reason to punish in all cases in which we think punishment is called for. One might argue that it would justify punishment in too many cases, however. It seems as though there are some acts of wrongdoing that call for a high degree of reactive sentiment in response, but are not criminalized. For instance, if someone recklessly cheats on his wife, she may appropriately feel a very high degree of resentment. Likewise someone who publishes racist material that contributes to a hostile environment in which some minority group are the victims of violence and discrimination might legitimately be the object of a high degree of indignation. Insofar as a high degree of indignation is called for in response to these wrongs, it seems that the Reactive Principle endorses punishing such wrongs, but many of us believe that such acts should not be criminalized. In the rest of this section, I will argue that the reactive theory can explain why the state does not have reason to punish in the proposed counterexamples, because, respectively, the state does not have reason to punish those wrongs regarded by citizens as private rather than public wrongs, or those wrongs whose criminalization would fail to respect the capacities for moral agency of the citizenry.

There seem to be some types of wrongdoing for which a high degree of indignation is appropriate, but for which it would only be appropriate for a small percentage of the community to feel this indignation. It would be inappropriate for the community to express its indignation toward these types of wrongdoing. Blame in these cases should be left to those who appropriately feel indignation. R. A. Duff maintains that crimes are distinct from torts because they are public wrongs. They are “wrongs in which ‘the public,’ the community as a whole, is properly interested.”²¹ One might say that the reactive theory does not give us reason to punish adultery because adultery is not a public wrong in which the community is properly interested, so public blame of adultery is inappropriate. While a spouse who is cheated upon, and perhaps some friends and family of the couple, might properly feel indignation, the state is not properly interested in adultery. One might wonder whether adultery is in fact a private wrong in our society. After all, indignation on behalf of the wronged members of celebrity infidelity that is documented by gossip magazines and websites is probably felt by a larger segment of society than is felt on behalf of the victims of most crimes. It is not clear, however, that we believe that it is *properly* the place of the wider community to

²¹ Duff, R.A., 2000, *Punishment, Communication, and Community*. New York: Oxford University Press, p. 61.

feel this indignation and the fact that adultery is not criminalized provides some evidence that we do not believe that it is.

The idea the reactive theory does not give us reason to punish some serious wrongs because they are not properly the interest of community and so do not call for a blaming response that speaks on behalf of all of us raises a couple of potential worries that are worth addressing. First, it again seems to leave the reactive theory open to a worry that arises when we look at alternative societies. Many societies in the past have criminalized adultery (and some still do). Such societies take adultery to be a public wrong, so the reactive theory could not rely on the distinction between public and private wrongs to explain why such societies do not have reason to punish adultery without an account of what makes some wrongs public and others private. It is difficult to give such an account,²² and this account is likely to rely on norms external to the justification of punishment about the legitimate role of the liberal state. It is not clear, however, that our belief that adultery should not be criminalized is one we think is universalizable across cultures. Different societies may reasonably hold different views about which wrongs are and are not of interest to the community generally, depending to some extent on contingent cultural factors. Second, one might worry that those who appropriately feel a high degree of the reactive sentiments toward private wrongs do not have an avenue to express those sentiments. If the state has a monopoly on the legitimate use of force, those in a position to feel resentment and indignation cannot express their resentment and indignation through private punishment. In responses to private wrongs, though, where the community does not have an interest in seeing that wrongdoers are blamed, the audiences for which expressions need to be transparent are much smaller. Actions which express well a high degree of blame will be accessible as such to the wrongdoer and victim but need not be to the community at large. A complete withdrawal of affection in a romantic relationship might sufficiently express resentment toward a partner for reckless infidelity, if it is appropriately regarded as private wrong, even if the expression is inaccessible to anyone outside the relationship.

It's not clear that the distinction between public and private wrongs can explain why those who publish offensive materials should not be punished. Publication, by definition, makes the materials available for public consumption. Nevertheless, it seems that, in general, offensive expressions do not call for a high degree of indignation. Though some indignation might be called for in such cases, we should reserve our highest degree of indignation for actions that violate natural rights. While many people feel a high degree of indignation toward those who express political views contrary to theirs, for instance, they can admit that really the highest degree of indignation is appropriate to murderers and rapists, not political opponents. But when someone publishes racist material that will predictably lead people who are exposed to act in ways

²² Duff, R.A., 2000, *Punishment, Communication, and Community*. New York: Oxford University Press, p. 63: "I do not think we could find general criteria for determining just which wrongs should be thus seen as 'public.' One could, for instance, imagine a community whose valuations of reputation and private property were such that libel was seen as a public wrong and theft as only a private wrong. However, any political community will have some set of shared values given which some wrongs will count as public wrongs...."

that violate the rights of members of minority groups in society, it might seem that such an action is rightly the object of the indignation of the community. Would not the reactive theory recommend punishment of the publisher? Through punishing the racist publisher the community would emphatically show that it takes all members of society to be owed moral consideration.

Many societies have laws against hate speech and perhaps they can be justified based upon the way that society stands up for the protected groups in response to hate speech. However, many of us are uncomfortable with the criminalization of expression on the grounds that it is likely to cause others to violate important moral obligations. Any theory of punishment must be embedded in a larger theory of democratic legitimacy that will articulate limits on state power which may explain why certain actions cannot legitimately be criminalized. I believe that plausible accounts of the limits of state power to punish wrongdoing such as publishing material that encourages wrongdoing accord well with the values that make up a system of reactive punishment. In “A Theory of Freedom of Expression,” Scanlon argues that the state cannot legitimately restrict speech on the basis that it purports to inform people of reasons to act in ways that are illegal, because “the powers of a state are limited to those the citizens could recognize while still regarding themselves as equal, autonomous, rational agents.”²³ In criminalizing speech because it advocates illegal activity, a state would treat citizens as if they do not have the autonomy to decide for themselves whether the purported reasons for action are good reasons. For an autonomous agent “in any given case his mere recognition that an action is required by law does not settle the question of whether he will do it.”²⁴ In order to regard themselves as autonomous agents, citizens must see that whether or not to follow laws is up to their own practical reasoning, so a legitimate state ought not censor speech that purports to give good reasons to break the law.

According to the reactive understanding of systems of punishment, an egalitarian system of criminal law both holds all citizens to and protects all citizens by those moral obligations codified into law. It thereby treats all citizens with capacities for moral agency as responsible for their actions. If criminalizing some wrongdoing undermines the ability for members of the community with these capacities to regard themselves as autonomous moral agents, it undermines the claim of the system of punishment which criminalizes that wrongdoing to treating all citizens as responsible agents. This gives the state reason not to criminalize the publication of material that encourages others to act immorally. Doing so would show a lack of respect for the reasoning capabilities of those citizens who might be a potential audience of that material.

²³ Scanlon, T.M., 1972, “A Theory of Freedom of Expression” *Philosophy & Public Affairs* 1: 204–26, p. 215.

²⁴ Scanlon, T.M., 1972, “A Theory of Freedom of Expression”, p. 216.

4) Conclusion

I have argued that the reactive theory of punishment satisfies the *positive reason desideratum*. We have general reason to blame those who commit serious crimes to show that we value the victims of such wrongdoing. Blaming such wrongs with punishment indicates to everyone in society that we take the wrong done to the victim seriously and are committed to a certain way of citizens getting along with one another. There may be other societies in which the values of public blame can be expressed without punishment, but given our emotional experience and the deterrent function of punishment, we have good reason to establish and maintain systems of punishment to express our indignation toward serious crimes. We do not have reason to punish some serious wrongs when those wrongs are not properly the interest of the whole community or when criminalizing them is inconsistent with members of society regarding themselves and each other as possessing capacities for moral agency.

The public blaming function of punishment is able to explain why we have reason to punish in cases of serious crimes where standard versions of traditional theories seem not to give us a reason. In chapter 1, I argued that the Utilitarian Principle does not recommend punishing in some cases in which commonsense morality holds that punishment is justified. I maintained that we believe we have positive reason to punish the assassin of a civil rights leader even if his punishment will not maximize aggregate happiness across society. The reactive theory can explain why we have reason to punish the assassin. We have reason to publicly blame him for what he has done in a way that shows that we take seriously what he has done. Mere verbal public condemnation without imprisonment would not be accessible to members of the community as expressing a high degree of indignation.

The reactive theory also explains why we have reason to punish serious crimes in cases in which retributivism does not. I argued in chapter 1 that we think we have positive reason to punish criminals even when their punishment does not inflict the suffering or harm on its object that many versions of retributivism claim give us reason to punish wrongdoers. The reactive theory is not susceptible to the same counterexamples. I argued that the following “eye for eye” version of retributivism does not explain the positive reason to punish because we think that punishment can be justified when it does not do to a criminal what he has done to others.

Harm-Based Retributive Principle 1: The state is justified in punishing someone if punishing him will harm him in just the way that he has harmed his victims, as he deserves.

Successfully expressing blame toward wrongdoers does not require doing to them what they have done to others. In the interpersonal context, a verbal expression of resentment might be an appropriate response to someone failing to show up for a meeting even though such a response does not do to its target what he has done to its expresser. Reactive punishment likewise does not need to do to a serious criminal what he has done

to others with his crime because conventional modes of hard treatment clearly express public blame.

Another version of retributivism holds that wrongdoers should not have their lives go as well as they otherwise would have gone in virtue of that wrongdoing.

Harm-Based Retributive Principle 2: The state is justified in punishing someone if punishing him will make his life worse than it otherwise would have been in the manner he deserves in virtue of his wrongdoing.

I argued that this version of retributivism is implausible because we are happy when punishment causes a criminal to turn his life around and live better than he otherwise would have. Our expressions of blame generally reflect this attitude. In the interpersonal context, if we express indignation at some friend for betraying another and the target of this expression reflects on his wrongdoing and treats others better in the future, the beneficial effects our expression has on his life are welcome. These effects are not required for our expression of blame to be successful, however. If the target of our indignation dismisses it we have nonetheless stood up for our friend. As an extension of our interpersonal blaming practices, reactive punishment is justified even if punishment does not improve the moral characters of the punished, but at the same time when their characters are improved in a way that allows them to live better, members of the community can welcome this result.

Other versions of retributivism rely on the idea that wrongdoers deserve to either suffer or have their lives made worse for some period of time because of their wrong.

Suffering-Based Retributive Principle: The state is justified in punishing someone if punishing him will inflict upon him the suffering he deserves in virtue of his wrongdoing.

Harm-Based Retributive Principle 3: The state is justified in punishing someone if punishing him will make a period of his life worse than it otherwise would have been in the manner he deserves in virtue of his wrongdoing.

I argued that these views do not explain the positive reason we have to punish in cases in which punishment does not inflict more suffering or harm on the punished than they would have been subject to without punishment. The reactive sentiments do not necessarily involve the idea that it would be intrinsically good for their object to suffer, so blame can be well expressed in ways that do not cause its object to suffer. According to the reactive theory, we have reason to express the community's indignation toward serious crime through punishment because it publicly shows that we take certain violations of moral obligations seriously and take the victims of crime to be owed moral consideration. This expression can be successful regardless of whether it causes its target to suffer or makes him worse off than he would have been for the time that he carries out his sentence. Sending a criminal to prison expresses a high degree of indignation

whether or not he suffers a great deal day to day as a result or his imprisonment makes his life worse than it otherwise would have been.

Finally, I claimed that there was a version of retributivism that is not subject to the same counterexamples with respect to the *positive reason* desideratum as other versions

Harm-Based Retributive Principle 4: The state is justified in punishing someone if punishing him will deny him some objective good that he does not deserve in virtue of his wrongdoing.

I argued that in order to explain the positive reason that we have to punish serious crime such a version of Harm-Based Retributivism would have to give a plausible account of why it is that someone does not deserve an objective good in virtue of his wrongdoing. It might be thought that the reactive theory provides such an account and so is actually a retributive theory. According to the reactive theory, when someone commits a wrong the reactive sentiments are appropriate in response. Those sentiments themselves deny their object of social regard so it might be thought that the reactive idea is really a retributive idea. Further, according to the theory, when someone commits a serious crime he makes himself an appropriate target for the expression of a high degree of blame that deprives its object of a good that punishment characteristically denies. It can look as though criminals deserve to be harmed in the sense that they deserve to be deprived of an objective good and punishment is justified because it deprives them of this good.

This understanding of the way in which punishment is deserved is suggested by Joel Feinberg in another of his landmark articles, “Justice and Personal Desert.” In that article he notes that, as responses to wrongdoing, the reactive sentiments are “felt as deserved.”²⁵ These emotional responses are what wrongdoers deserve and we can understand punishment to express them.²⁶ In that article, as in “The Expressive Function of Punishment,” Feinberg does not really argue that these emotions and their expression are morally important. He writes, “I am not sure how, if at all, these judgments of moral appropriateness are to be verified; but I suspect that they resemble certain aesthetic judgments—for example, that crimson and orange are clashing colors...”²⁷ But it seems that the reasons we have to mete out deserved punishment are significantly stronger than those we have to wear clothes that match.²⁸ The analogy

²⁵ Feinberg, Joel, 1970, “Justice and Personal Desert” in *Doing and Deserving*. Princeton, NJ: Princeton University Press, pp. 70–71.

²⁶ Feinberg, Joel, “Justice and Personal Desert,” p. 82: “This view suggests in turn that responsive attitudes are the basic things persons deserve and that “modes of treatment” are deserved only in a derivative way, insofar as they are natural or conventional means of expressing the morally fitting attitude. Thus punishment, for example, might be deserved by the criminal only because it is the customary way of expressing the resentment or reprobation he ‘has coming.’”

²⁷ Feinberg, Joel, 1970, “Justice and Personal Desert,” p. 82.

²⁸ George Sher uses this passage from Feinberg to argue that the reactive attitudes cannot ground desert claims about merit. Sher, George, 1987, *Desert*. Princeton: Princeton University Press, p. 114: “Even when we allow for differences in strength among aesthetic reasons, this implication seems to amount to a concession that the desert of the meritorious lacks serious normative force. In this context, as elsewhere, to

between the fittingness of the reactive sentiments and the fittingness of aesthetic judgments casts doubt on whether this fittingness can explain why states have general reason to use as many of their limited resources to punish as they do. By highlighting the value judgments that make up the reactive sentiments and that punishment expresses, the reactive theory provides an explanation of why the reactive sentiments are more morally significant in a way that ought to motivate action than aesthetic judgments.

So is the reactive theory best understood as a desert-based retributive theory? It has a number of features that place it at odds with much of the retributive tradition. On the reactive theory what wrongdoers deserve is not primarily suffering or harm but an emotional reaction. This emotional reaction does not necessarily involve the judgment that it would be intrinsically good that its object suffer or be harmed. It may be successfully expressed in a way that does not cause its object to suffer and, in fact, improves his life. Whether it ought in fact be expressed in punishment is partly based on contingent features of the society in which the expression takes place including deterrence considerations.

Chapter 4 The Blame Fits the Crime

I have argued that the reactive theory of punishment satisfies the *positive reason desideratum*. That argument relied on an account of proportionality according to which modes of treatment that denied their objects more important goods had the social meaning of expressing a greater degree of blame. I believe that this reactive account is able to satisfy the *comparative proportionality desideratum*

Comparative proportionality desideratum: A theory of punishment should explain why we always, or almost always, have reason to punish more severely those who commit more serious crimes.

In this chapter, I elaborate on the reactive account of proportionality. In the first section I suggest that the seriousness of crime can be understood in terms of the indignation that is called for in response. The seriousness of crimes is a function of their wrongness and more serious wrongs call for a higher degree of indignation. This gives the state reason to give longer sentences or heftier fines to those who commit more serious crimes because doing so denies them more of their liberty and property and thereby transparently expresses a higher degree of externalized indignation from the community. Then I will argue that the reactive theory offers a better account of proportionality than retributivism and consequentialism.

Before elaborating on the reactive account of proportionality, I will first dismiss a potential counterexample to it. I conceded in the previous chapter that the reactive theory might not endorse criminalizing all the actions that our system of criminal justice punishes. But I maintained that it is not obvious that these actions should be criminalized from the point of view of commonsense morality. Likewise, reactive proportionality may sometimes recommend sentences that diverge from those issued by our criminal justice system, yet it is not obvious that such sentences are compelling counterexamples to reactive proportionality. Consider, for example, California's Three Strikes law, according to which people who commit three felonies may be given extended (oftentimes life) sentences in prison. Given a reactive understanding of systems of punishment, the punishment for the third offense licensed by the Three Strikes law seems to express a much higher degree of indignation from the community than would be appropriate for that wrongdoing compared with the indignation we typically express toward those crimes. Even if it is appropriate to express marginally more blame toward a third time offender than a first time offender, this does not seem to warrant that we respond to his minor felony the same way that we would toward the most serious crimes, such as multiple first degree murder. Such a response symbolically expresses a very high degree of indignation inappropriate to relatively minor wrongdoing. Though the reactive account of proportionality does not seem to endorse the Three Strikes Law, I do not believe this is a knockdown objection to the account. In fact, some applications of the law seem problematic and reactive proportionality offers an explanation why. Consider an application of the law

Kevin Weber, who had been convicted of two residential burglaries, was apprehended after burglarizing a restaurant at night. His total gain from the burglary was four cookies, which had been shoved in his pockets. His sentence under the Three Strikes law (plus an additional one-year enhancement) was 26 years to life.¹

A theory that does not endorse Weber's punishment seems to get it right from the point of view of commonsense morality. In fact, it is difficult to see how any reading of "more serious crime" in the *comparative proportionality desideratum* would endorse his punishment. The reactive theory has an explanation of this. Weber's sentence symbolically expresses a much higher degree of indignation than his crime calls for.

1) The Reactive Sentiments and the Seriousness of Crime

As I mentioned at the beginning of Chapter 2, the seminal article in the discussion of the punishment's expressive character is Joel Feinberg's "The Expressive Function of Punishment." While he is dubious of the moral status of the emotions that punishment expresses and raises doubts about whether the expressive function of punishment can justify its deprivations of liberty and property, he does maintain that we ought to understand the severity of punishment that crime calls for in expressive, rather than retributive, terms.

What justice demands is that the *condemnatory aspect* of the punishment suit the crime, that the crime be of a kind that is truly worthy of reprobation. Further, the degree of disapproval expressed by the punishment should "fit" the crime only in the unproblematic sense that the more serious crimes should receive stronger disapproval than the less serious ones, the seriousness of the crime being determined by the harm it generally causes and the degree to which people are disposed to commit it.²

The reactive theory takes aboard Feinberg's idea that the punishment should fit the crime in the sense that more serious crimes should receive stronger disapproval. It understands this stronger disapproval in terms of the indignation from the community that would be appropriate in response to the crime.

One might raise worries about the factors that Feinberg claims determine the seriousness of crime. The first factor he cites is the harm that a crime generally causes. Intuitively, the amount of harm caused plays a large role in our determinations of the seriousness of crimes. Nevertheless, it is not clear that the various harms that crimes

¹ Walsh, Jennifer Edwards, 2007, *Three Strikes Laws*. Westport, CT: Greenwood Publishing Group, pp. 68–69.

² Feinberg, Joel, 1965, "The Expressive Function of Punishment"; reprinted in Feinberg, Joel, 1970, *Doing and Deserving*. Princeton: Princeton University Press, p. 118.

generally cause are commensurable,³ so that it is not clear how we could measure them on the same scale. How are we to weigh physical pain against invasions of privacy, or economic hardship? I do not think that the second factor Feinberg references, the degree to which people are disposed to commit crimes, has a significant influence on our intuitions about the seriousness of crime. If the claim is about the number of people in society who are disposed to commit the crime, it would seem to imply that tax evasion is a much more serious crime than murder in an important respect. If instead the claim is about the strength of the desire to commit the act of those who are disposed to do so, it would seem to imply a way in which all sex crimes were more serious than others.⁴ These implications do not seem to reflect an important criterion for the seriousness of crime, but rather something like the degree to which certain actions might be good candidates for deterrent punishment.

I believe that the reactive theory offers a promising alternative way to understand the seriousness of crime that is not determined by some combination of harm caused and propensity of people to perform the act, but the *wrongness* of the act. The reactive sentiments give us epistemic access to a scale on which to compare the wrongness of acts: those that are more wrong call for a higher degree of the reactive sentiments. We can measure the wrongness of various violations of moral obligations by considering the degree of reactive sentiment that would be an appropriate response to those violations. It is more appropriate to feel indignant about a friend being murdered than being stood up for a date because murder wrongs your friend much more seriously than standing him up for a date does. Of course, any plausible analysis of the seriousness of crime in terms of the harm it causes will be able to explain why murder is a serious crime. Consider a framework for determining the seriousness of crimes offered by Andrew von Hirsch and Nils Jareborg. They propose that seriousness of crimes should be determined according to how much they typically negatively affect the living standard of those who are victims of them, where the living standard is understood as “the *means* or *capabilities* for achieving a certain quality of life.”⁵ The seriousness of a crime on this account is based on the negative consequences that certain crimes typically involve for their victims and “the value-judgment about how significant those consequences are in reducing a person’s standard of living”⁶ Murder is a serious crime on this view because it makes achieving

³ In fact some crimes that wrong victims might not harm them, actually or potentially. For an example see Ripstein, Arthur, 2006, “Beyond the Harm Principle,” *Philosophy and Public Affairs*, 34 (3): 216–246, p. 218: “Suppose that, as you are reading this in your office or of the library, I let myself into your home, using burglary tools that do no damage to your locks, and take a nap in your bed. I make sure everything is clean. I bring hypoallergenic and lint-free pajamas and a hairnet. I put my own sheets and pillowcase down over yours. I do not weigh very much, so the wear and tear on your mattress is nonexistent. By any ordinary understanding of harm, I do you no harm. If I had the same effects on your home in some other way, nobody would suppose you had a grievance against me, let alone that you should be able to call the law to your aid. Your objection is to my deed, my trespass against your home, not to its effects.”

⁴ Or perhaps theft of food by hungry people would be more serious than sex crimes.

⁵ von Hirsch, Andrew and Nils Jareborg, 1991, “Gauging Criminal Harm: A Living Standard Analysis,” *Oxford Journal of Legal Studies* 11: 1–38, p. 10.

⁶ von Hirsch, Andrew and Nils Jareborg, 1991, “Gauging Criminal Harm: A Living Standard Analysis,” p. 14.

any quality of life impossible. But the living standard account of the seriousness of crime seems fraught with some of the same difficulties that made degree of harm an untenable standard for seriousness. It is not obvious that those things that contribute to a quality of life are commensurable, so it is not clear how we should compare them on a uniform scale.

On the other hand, I have suggested that the wrongness of actions can be compared on a uniform scale: the degree to which the reactive sentiments are appropriately felt by those in a proper position to feel them in response. Without reference to the reactive sentiments, the wrongness of actions is not any more commensurable than harms or components of standards of living, but the degree to which we believe the reactive sentiments are appropriate in response to various wrongs gives us a special epistemic access to how wrong we regard those actions to be. According to the reactive understanding of systems of punishment, when we punish crimes more severely that punishment expresses a higher degree of indignation from the community. The more severely we punish a crime, the more serious a wrong we demonstrate that we regard it to be. This analysis does not ignore the intuitive idea that the seriousness of crime is related to the harm it generally causes. The degree of reactive sentiment appropriate to a given action is sensitive, among other things, to the harm actions of that kind generally cause and how those harms typically bear on our standard of living.

2) Reactive Proportionality and the Traditional Alternatives

The retributivist might claim that the reactive theory has no advantage over retributivism on this particular score. Retributivism offers a uniform scale on which the wrongness of acts can be compared: the amount of harm or suffering those acts deserve. I believe that there are reasons to prefer reactive to retributive proportionality, however. As I argued in Chapter 1, I do not think that the retributive idea that wrongdoers deserve harm or suffering in virtue of their wrongdoing is well-grounded. By comparison most of us implicitly accept the appropriateness of the reactive idea that the reactive sentiments are appropriate responses to wrongdoing. The further judgment that more serious wrongs call for a higher degree of the reactive sentiments is also widely shared. Though most retributivists hold that more serious wrongs call for more harm or suffering, to the extent that the retributive idea is groundless, the amount of harm or suffering that people who commit various acts supposedly deserve cannot determine the seriousness of crime.

Another reason to prefer the reactive account of proportionality is that on most versions of retributive proportionality the severity of punishment that is called for is a function not only of the seriousness of the wrongdoing. According to the Retributive Principle, the severity of punishment that is justified is the sentence that visits upon the criminal his deserved harm or suffering. But conventionally more severe forms of punishment do not always inflict more suffering on the punished or make them worse off, so the severity of punishment that the principle justifies is a function not only of the seriousness of the wrong, but of how much suffering or harm punishment causes the

punished, given his individual circumstances.⁷ However, we typically do not take into account these circumstances in determining the extent to which criminals ought to be punished. If two people commit the same serious crime, the criminal justice system sentences them to roughly equal prison sentences, even if one is constituted such that he will suffer much more in prison than the other. According to the reactive account of proportionality sentencing two people to the same prison sentence expresses the same degree of indignation because it deprives both of them the same objective good. Even if one criminal suffers more in prison than another, if they are given the same sentence, they are blamed equally.

A retributivist may object here that courts sometimes do (or ought to) take individual circumstances into account in sentencing and this is evidence that we believe that suffering or harm should fit the crime rather than the denial of an objective good that symbolically expresses a degree of blame. Many European countries, and some American states, punish crimes with day fines where the amount of the fine is determined by the daily income of the punished and number of dependents. Insofar as we endorse these sentencing procedures, the retributivist could argue that we take into account the circumstances of the punished in order to inflict a similar amount of harm or suffering on them as others who commit the same crime. I grant that the retributivism gives a possible explanation of the endorsement of day fines, but I believe that the reactive theory offers a possible alternative explanation that accords better with our attitude toward other modes of punishment.

In some societies, it may be that imposing the same fine on people who have committed the same crime is not understood by members of the community to express the same degree of indignation depending on the material circumstances of the punished. In such societies, imposing the same fine on the poor and rich for committing the same crime would not actually be understood to deny them the same objective benefit. Instead an equal objective benefit would be understood as the same percentage of one's income. I think that this view of fines has some intuitive force: if Bill Gates and I both commit the same crime it is not clear that the community would understand fining us both \$25,000 as expressing the same degree of indignation. In the eyes of members of the community, Gates might be seen as getting away with his crime without a blaming response on behalf of his victim that takes it seriously. In addition, considerations of constructiveness might influence our emotional reactions in a way that supports this understanding of the objective good fines deprive the punished. Lump sum fines might be no burden at all for the rich so not deter them from criminal activity.

In a society in which people's indignation disposes them to deprive rich criminals of more of their property than poor for the same crimes, higher fines of rich people can be understood to express the same degree of indignation as lower fines for poor people. On this reactive explanation, day fines are indexed to income so that people who commit the same crimes are targets of an expression of the same degree of indignation from the community, not so they both receive the same amount of deserved suffering. Unlike the

⁷ I explained in Chapter 1 that this criticism does not apply to a version of retributivism on which wrongdoing deserves the deprivation of an objective good.

retributive explanation, it is consistent with ignoring a number of factors during sentencing which contribute to the amount of harm and suffering criminals undergo when they are punished, such as their psychological makeup or whether they have rich relatives who can cover their fines,⁸ —or, as many societies do, their income!

Setting aside these criticisms, some retributivists might claim that they offer a superior account of proportionality because their accounts offer guidance not only on how punishments of different offenses ought to compare but also on how the punishment ought to compare to the offense: it ought to be proportionate in the noncomparative sense that the severity of the punishment should match the seriousness of the crime considered in isolation from other punishments.⁹ Retributive principles of proportionality often claim to be able to offer a substantive account of this offense-punishment proportionality insofar as they recommend matching penalties: that the degree of punishment criminals deserve is the suffering or harm they have inflicted on their victims. The idea that we should treat criminals in ways that express an appropriate degree of indignation may strike some as vacuous by comparison.

The retributive matching account of proportionality has shortcomings, however, that tell in favor of allowing for some indeterminacy in offense-punishment proportionality. I argued in the previous section that the harm a wrong causes does not clearly determine its seriousness because harms are incommensurable. The incommensurability of harms tells against both ranking the seriousness of crime by their harms, and matching the harms that people commit by their punishment because we cannot always harm them in the ways that they harm others.¹⁰ On some understandings of suffering, such as physical pain, suffering is commensurable in principle, so that it could ground a uniform scale in which to measure the seriousness of wrongdoings. But the seriousness of wrongdoing is not always a result of the pain it causes. If a cyber-thief hacks into the bank account of an oblivious, old miser, who is unaware that he is being robbed, the thief's action is wrong even if it causes the miser no pain. Further, it seems that commonsense morality allows for indeterminacy in morally defensible systems of punishment. If the legal system of one state punishes a crime with one year in prison and a different state punishes the same crime with fourteen months in prison, we do not think

⁸ David Lewis points to this feature of fines in provocatively arguing that even non-Christians are of two minds about penal substitution, since they accept penal substitution of fines but not of imprisonment. For this reason, the penal substitution required for atonement does not pose a special problem for Christians. Lewis, David, 1997, "Do we believe in penal substitution?" reprinted in Lewis, David, 2000, *Papers on Ethics and Social Philosophy*. Cambridge University Press, Cambridge, p. 133: "We do not believe that the offender's friend can serve the offender's prison sentence of flogging, transportation, or hard labour. But we do believe — do we not? — that the friend can pay the offender's fine. (At least, if the offender consents.) Yet this is just as much a case of penal substitution as the others."

⁹ On the importance of this other kind of proportionality see Primoratz, Igor, 1989, *Justifying Punishment*. Atlantic Highlands, N.J.: Humanities Press International, Inc., pp. 89–90: "If murder is punished by a year in prison, because that is the most severe punishment the law provides for, we should say that the punishment is proportionate to the offense in a way, but only in a way, and in a very formal and unsatisfactory way at that, and that it is at the same time disproportionate in another, more important, substantive sense: that it is too mild, given the absolute gravity of the crime."

¹⁰ This is the reason I rejected the "eye for eye" version of retributivism in Chapter 1.

that at least one of the states must have gotten the sentence wrong. Both sentences might strike us as roughly appropriate responses to the crime, without consideration of whether social conditions are such in the society with the longer sentence that either the criminal act typically inflicts more harm or suffering on its victims or prison inflicts less harm or suffering on the imprisoned than in the other society. The reactive view can explain our lack of concern about the disparity in sentences. Different communities can reasonably come to different conclusions about what treatment expresses the degree of indignation that they find appropriate in response to the crime. Reactive proportionality's comparative vacuity is to its advantage.¹¹

One of the reasons why reactive proportionality does not definitely determine the severity of punishment that each offense calls for is that on the reactive theory the severity of punishment we have reason to inflict is partly determined by consequentialist considerations. On the reactive theory, these considerations are relevant to offense-punishment proportionality, but not comparative proportionality.¹² In Chapter 1, I pointed out some difficulties that consequentialist theories have satisfying the *comparative proportionality desideratum*. Following the Consequentialist Principle, the severity of justified punishment in every case is a function of the good consequences of the act of punishment. It is subject to counterexamples in which punishing a less serious crime more severely than a more serious one would have better consequences. If by subjecting a celebrity shoplifter to the same punishment we typically would a murderer we could greatly reduce the incidence of shoplifting and increase overall happiness, the Utilitarian Principle would endorse it. The reactive theory explains why we do not have reason to give the shoplifter the same punishment as the murderer. Given a reactive understanding of our systems of punishment, such a sentence would express a much higher degree than is appropriate in response to shoplifting. Conversely, utilitarianism will not endorse punishing beyond what will maximize happiness, so that sometimes we will not have reason to punish the most serious crimes severely. But in doing so, according to the reactive understanding of systems of punishment, we would inappropriately indicate that the crime does not call for a high degree of indignation.

As I suggested above, I do not think the second determining factor of the seriousness of crime identified by Feinberg, the degree to which people are disposed to commit the crime, actually determines how serious crimes are compared with other crimes, and so is not relevant to comparative proportionality. I do believe it is relevant to offense-punishment proportionality, however. While tax evasion is not a serious crime compared with murder, we might think it is a particularly good candidate for deterrent punishment insofar as many people are disposed to commit it. On the reactive theory, whether deterring an act with certain modes of punishment would be effective informs the ways that our indignation disposes us. These dispositions are embedded in a wider

¹¹ Versions of retributivism do not call for matching penalties, but simply more suffering and harm for more serious crime share this advantage.

¹² Uma Narayan proposes a similar view of the role the way that consequentialist considerations ought to fit with proportionality. Narayan, Uma, 1993, "Appropriate Responses and Preventative Benefits: Justifying Censure and Hard Treatment in Legal Punishment," *Oxford Journal of Legal Studies* 13: 166–182, p. 180.

set that disposes us to feel a higher degree of indignation toward serious crimes, and so approve of longer sentences toward those crimes. The disposition to approve of short prison terms for tax cheats is partly based on the deterrent function of punishment and those deterrence considerations indirectly contribute to the explanation of why people are disposed to give more serious criminals, such as murderers, long prison sentences. Considerations of constructiveness help explain why we punish certain crimes to the extent we do, but they do not give us reason to punish less serious wrongs more severely than more serious ones.

3) Conclusion

In the conclusion to Chapter 3, I considered whether the reactive theory was actually a retributive theory. I think that given the role considerations of constructiveness play in informing our emotional responses, one might claim instead that the view is, in essence, a version of rule consequentialism. The reactive theory does share some features with rule consequentialism. According to the reactive theory, the explanation of why we take ourselves to have reason to punish crimes to the degree we do is partly due to the fact that we believe that punishment generally has a deterrent function (and so shows we are committed to the idea that everyone in society is owed moral consideration). As with rule consequentialism, though the potential good consequences of punishment are relevant to justifying our laws, those consequences need not be realized on every occasion for reactive punishment to be justified.

While the reactive theory draws upon consequentialist considerations, I believe it puts them in their proper place so that it is able to avoid some of the difficulties that plague rule-consequentialist theories. One difficulty for rule-consequentialism emerges from the discussion which concluded the previous section. Because the severity of punishment we have reason to inflict is a function of the good consequences of doing so according to consequentialism, it leaves open the possibility that the best system of rules will punish intuitively less serious crimes more severely than more serious ones in a way that reactive proportionality does not. The reactive theory has features that allow it to avoid some other objections to rule utilitarianism raised in Chapter 1. First, I appealed to the familiar criticism that rule-utilitarianism is an unstable view. If only the consequences of our actions are of moral significance, it is unclear why we should follow the rule to punish serious criminals in cases in which doing so would not produce the best consequences. The reactive theory explains why we have reason to punish in these cases because punishment performs its public blaming function even when it does not achieve any further good outcomes, and this function is the primary reason the state has to punish. Second, I argued that there are a number of features of commonsense morality besides comparative proportionality that might not be captured by the set of rules that maximizes happiness across society. For instance, it is not clear that rule-utilitarianism will support a rule prohibiting the punishment of the severely mentally disabled. The reactive theory supports such a law because those without capacities for moral agency are not appropriate objects of blame. Finally, rule-consequentialism is open to the charge that it

treats criminals in a way that is not morally permitted, depriving them of good to which they have a right because of the good social consequences of doing so. In the next section I argue that the reactive theory is able to avoid this charge by seeing those deprivations as constitutive of appropriate expressions of blame.

Chapter 5 Blame and The Standing to Object

I have argued that the reactive theory of punishment explains why the state has reason to punish serious crime and explains why it has reason to punish more serious crimes more severely. But I have not yet explained why punishment is morally permissible on the reactive theory. This is the central issue that a theory of punishment must address: how are we permitted to treat criminals in ways that would normally be impermissible, denying them of goods to which they normally have a right? In Chapter 1, I argued that neither traditional justification satisfies the *no valid objection from rights desideratum*.

No valid objection from rights desideratum: A theory of punishment should explain why the punished cannot reasonably object to their punishment, even though it deprives them of goods to which they normally have a right.

Appealing to the good consequences of punishment will not suffice, because the state is not normally justified in violating citizens' rights just because doing so has good social consequences. On the other hand, even if the idea that wrongdoers deserve to suffer or be harmed can be grounded, it is not clear that the state would be justified in violating criminals' rights in order to give them their just desserts, because the state is not generally justified in violating citizens' rights in order to give people what they deserve.

In this chapter, I argue that the reactive account is able to satisfy the *no valid objection from rights desideratum*. This argument draws upon the account of previous chapters in which I argue that punishment is an appropriate expression of indignation. I claim that punishment is permissible because we are justified in treating people in ways that would otherwise be impermissible when that treatment expresses appropriate blame. Those who are the targets of appropriate expressions of blame do not have the moral standing to object to them on moral grounds. I begin this argument by rejecting a similar explanation for why serious criminals cannot reasonably object to their punishment, the idea that they forfeit their rights when they violate the rights of others. I raise a couple of objections to this view. Then I explain the moral standing alternative, focusing on how objections to blame function in the interpersonal context, where our blaming responses are not mediated through political institutions, and then extend the account to state punishment. I maintain that this rival explanation is not subject to the same objections as the forfeiture view. Finally, I compare the moral standing explanation of why we are permitted to punish criminals to the consequentialist and retributivist alternatives.

1) Forfeiture

The notion of forfeiture offers a potential explanation of why those who violate the rights of others cannot object to the state treating them in ways that would normally violate those rights. According to forfeiture criminals forfeit their rights when they

violate the rights of others. Forfeiture, so understood, is not a claim about the positive reason we have to punish wrongdoers, but an explanation of why violation of their rights is permissible.¹ While a theory of punishment that accepts forfeiture has a straightforward way to satisfy the *no objection from rights desideratum*, the idea of forfeiture itself has been subject to criticism. Some philosophers have claimed that it makes permissible harms that are intuitively impermissible. Others have argued that it would not be able to explain why the harms of punishment are permissible in a wide range of cases in which punishment is justified. In this section I will explain these criticisms.

In Warren Quinn's "The Right to Threaten and The Right to Punish," he presents a case in which it looks like a simple account of forfeiture would justify harm that is in fact impermissible:

The proper authorities are entitled to punish Jones, a generally decent young man who has foolishly stolen Smith's car, by depriving him of up to the amount of liberty forfeited in the theft. But suppose that before any such punishment takes place, Smith, for reasons having nothing whatever to do with the theft, kidnaps Jones and deprives him of exactly that amount of liberty. In this situation it is natural to suppose that Smith not only wrongs Jones but specifically violates his right to liberty.²

If forfeiture entails that Jones could not reasonably complain about his treatment by Smith on the grounds that it violates his right to liberty, this case is a counterexample to forfeiture. A defender of the forfeiture view might maintain that Jones can still complain about his treatment because he has only forfeited his right to liberty to the state, which has a monopoly on the legitimate use of force. According to this defense, only the state is justified in depriving goods, the deprivation of which usually requires the threat of force. Through his wrongdoing Jones has forfeited his right of liberty to the state, but not to all of his fellow citizens. Yet Smith's position as a private citizen does not seem to be the only grounds on which Jones can object to his treatment. The fact that Smith's reasons for kidnapping him have nothing to do with his wrongdoing gives Jones grounds for objecting. Jones does not generally forfeit his right to liberty in virtue of his wrongdoing, rather there seems to be something special about the circumscribed situation in which he is being punished *for that wrongdoing* that makes his appeal to his rights unreasonable.

While Quinn argues that there are cases in which forfeiture would justify morally impermissible acts, another line of criticism of forfeiture claims that it is unable to justify the characteristic harms of punishment in a wide range of cases in which we think that punishment is justified. In "The Paradox of Punishment," Alvin Goldman makes this case based on the premise that we do not think that those who violate the rights of others

¹ Thus, it differs from retributivism as I defined it in Chapter 1, because it does not say anything about what those who commit serious wrongdoing positively deserve.

² Quinn, Warren, 1985, "The Right to Threaten and the Right to Punish," *Philosophy and Public Affairs* 14: 327-73, p. 332.

thereby forfeit all their rights. He argues that under the most plausible interpretation of forfeiture, when one violates the rights of others, one forfeits those rights that one violates but not one's complete set of rights.

If A steals fifty dollars from B, this does not give B or anyone else, official or not, the right to impose all and any conceivable harms upon A in return. Nor does A thereby become available for any use to which the community then wants to put him.... [I]f we ask which rights are forfeited in violating rights of others, it is plausible to answer just those rights one violates (or an equivalent set). One continues to enjoy rights only as long as one respects those rights in others: violation constitutes forfeiture.³

Though Goldman endorses this account of forfeiture, he thinks that modern systems of punishment cannot be justified by it, because, on deterrence grounds, they often deprive criminals of protection by rights that are not equivalent to those of others that the criminals have violated.

[O]ur society does not limit punishment to deprivation of rights forfeited, that is, rights of others which have been violated by the criminal. Especially in regard to crimes against property, punishments by imprisonment are far more severe, on the average, than the harm caused to victims of these crimes.... The disproportion between violated or deprived rights of the victims and those of the criminals in these crimes is obvious.⁴

If those who violate the rights of others forfeit only those rights, forfeiture does not explain why they cannot reasonably object to punishment that deprives them of rights outside that set. Even if we can make sense of the deprivation of liberty and property that punishment involves as being commensurable to criminals' mistreatment of their victims, forfeiture cannot explain why we are sometimes permitted to deny them greater goods than they have deprived others. If, on the other hand, those who violate any of the rights of others thereby forfeit all their rights, those who commit the most minor rights violations could not reasonably object to any treatment on the grounds that it violates their rights. A theory of punishment that satisfies the *no valid objection from rights desideratum* in a way that accords with commonsense morality must explain why criminals cannot reasonably object to being treated in ways that would otherwise violate their rights, even when those rights violations are not equivalent to those they have violated in others, but this explanation must not imply that criminals can never object to any treatment on the grounds that it violates their rights.

³ Goldman, Alvin, 1979, "The Paradox of Punishment," *Philosophy and Public Affairs* 9: 42–58, p. 46.

⁴ Goldman, Alvin, 1979, "The Paradox of Punishment," p. 49.

2) Valid Objections to Interpersonal Blame

I believe that the reactive theory can explain why serious criminals cannot reasonably object to their punishment on the grounds that it deprives them of a good to which they have a right in a way that avoids the objections faced by forfeiture. In this section, I argue that, in the interpersonal context the targets of appropriate blame lose their standing to complain about the loss of social goods that appropriately express that blame. In the next, I argue that the reactive theory can satisfy the *no objection from rights desideratum* because state punishment is continuous with these interpersonal responses. Punishment is permissible because we are justified in treating people in ways that would otherwise be impermissible when that treatment expresses a proportionate degree of blame. Those who commit serious crimes do not have the standing to object to deprivations of goods to which they are normally entitled when those deprivations are expressive of the community's appropriately high degree of indignation.

It is evident from the interpersonal context that we accept the idea that targets of appropriate blame do not have the standing to object to that blame. Someone might object to being blamed because he feels distress when the reactive sentiments of others are directed toward him. This complaint is reasonable if he was justified in acting the way he did or has an excuse for acting in that way. In these cases it is inappropriate to blame the person, because he has not really violated a moral obligation or there is some extenuating circumstance that renders blame inappropriate. In other cases his objection is unreasonable unless he is properly exempted from blame.⁵ Those without capacities for moral reasoning or guiding their behavior in light of moral reasons are appropriately exempted from being the objects of reactive sentiments. But if someone commits an unexcused wrong while in the possession of these capacities, he cannot reasonably object to others, appropriately situated, feeling the reactive sentiments toward him to a proportionate degree.

This point is about the standing of those who violate moral obligations to object to blame in the interpersonal context, and *not* about whether those in a position to blame always have conclusive reason to do so. There are other ways of responding to wrongdoing that are consistent with treating those who commit wrongdoing as responsible agents. For example, the swearing off of the resentment that is characteristic of forgiveness is a way of respecting the forgiven as a moral agent. But someone who has committed a wrong is not in a position to demand forgiveness when he is appropriately the object of resentment; he is rightly at the mercy of the person he has wronged. In another example, I might fail to feel resentment toward some minor wrong because I am preoccupied with another matter. This need not mean that I fail to treat the person who has wronged me as a responsible agent. Nor is that person in a position to reasonably demand that I should spend my cognitive energy on other matters, however.

⁵ For an extensive discussion of excusing and exempting conditions that make blame inappropriate see Wallace, R. Jay, 1994, *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, Chapters 5–6.

Not only is it unreasonable for someone who has violated a moral obligation to object to others, appropriately situated, feeling the reactive sentiments to a proportionate degree toward him, he also does not have standing to object to their appropriate expression. Imagine that Jim has betrayed Susie's confidence in a matter of some importance, and she expresses her resentment toward him, "I can't believe you did that! You betrayed my confidence." It would be peculiar for Jim to reply by saying, "You ought not express your resentment like that, because it really hurts my feelings." Hurt feelings sometimes provide grounds for objecting to an action, and if Jim had not actually betrayed her confidence, he could legitimately appeal to these feelings in objecting. What makes the response unsuitable, I maintain, is that Jim does not have the standing to object to the negative impact an appropriate expression of blame has on him.

One might question this diagnosis. It is unreasonable to object to a wide range of actions performed by others on the basis of hurt feelings. It would be unreasonable, for example, for me to demand that my unrequited love date me, or my teacher give me a good grade, because failure to do so would hurt my feelings. In light of such examples, it might be thought—too quickly, no doubt—that hurt feelings do not constitute much of a basis for objecting to the actions of others. Consider instead, then, expressions of blame in the interpersonal context that take the form of deprivations of goods beyond social regard. Imagine that Susie withdraws from Jim an invitation to a party she is throwing in response to his betrayal. He responds, "You ought not express your resentment like that, because I had a legitimate expectation to go to that party and didn't make other plans that night." This is an unreasonable response, not because it is unreasonable to object to having one's legitimate expectations thwarted, but rather because Jim is not in a position to object to the negative effects that his own wrongdoing has on him when they are a result of his being appropriately blamed for that wrong. In response to Jim's betrayal, Susie's father might express his indignation by breaking off a mentoring relationship with Jim. Assuming the betrayal was serious enough to make the severing of this relationship appropriate, Jim could not reasonably object that Susie's father has broken a promise to mentor him. Though one can usually reasonably complain about a broken promise, one cannot reasonably demand those who are blaming them honor their promises when appropriately blaming one involves not honoring those promises.⁶

Jim's standing to object to similar treatment is not undermined when that treatment is not a blaming response or when the person treating him that way is not in a position to blame him, however. For example, were the rescinding of the invitation not a blaming response—say a friend of Susie's who is helping throw the party and unaware of

⁶ In his recent work on blame, T. M. Scanlon makes a claim that seems to be in tension with the point I am making here. Scanlon, T. M., 2008, *Moral Dimensions: Permissibility, Meaning, and Blame*, Cambridge: Belknap Press of Harvard University Press, p. 142: "Even those who have no regard for the justifiability of their actions toward others retain their basic moral rights—they still have claims on us not to be hurt or killed, to be helped when they are in dire need, and to have us honor promises we have made to them." However, it seems to me uncontroversial that Susie's father is permitted to break his promise to Jim in response to his wrongdoing. This need not be because that promise is somehow conditional on Jim's relationship with Susie. He might have made the promise to mentor before Jim met Susie. He would still be permitted to not to honor it in these circumstances.

Jim's betrayal randomly chooses to disinvite him—Jim could reasonably object to it. Jim also retains the standing to object to blaming responses from those who are not in an appropriate position to blame. I claimed in Chapter 2 that it is difficult to give a complete analysis of how we should distinguish public from private wrongs. But stipulate that Jim's betrayal of Susie is properly considered a private wrong. Imagine that someone who has promised to mentor Jim and has no relationship to Susie finds out about his wrongdoing from the grapevine. She indignantly refuses to fulfill her promise. If she is not in a position to blame him, he can reasonably object that breaking her promise to him violates an obligation that he is owed.

Jim also retains the standing to object to treatment that expresses blame because it is disproportionate or tampers with his capacities for moral agency. If Susie conspires to get him fired from his job as an expression of her resentment, he could reasonably complain. Such a course of action is not an appropriate expression of blame; it is an excessive response to his wrongdoing, expressing a higher degree of resentment than is appropriate. Jim would be in a position to object to Susie's behavior in these cases, because he does not lose the standing to object to treatment that negatively impacts him in ways that go beyond what expresses the appropriate degree of blame. Or, imagine that Susie is trained in psychological torture. Jim could reasonably object to her using these techniques on him in order to express her blame because they fail to respect those capacities that make blame appropriate (in the manner described in Chapter 3).⁷

I have identified some grounds for validly objecting to expressions of resentment and indignation and their expression in the interpersonal context. One can reasonably object to being blamed for a wrong that one has not actually committed, or for which one has an excuse for having committed. One can reasonably object if one is properly exempted from blame for the wrong because one was not in possession of the capacities for moral agency at the time of the wrongdoing. One can reasonably object to blame from those who are not in a position to blame one for a wrong, as members of the general public are not in a position to blame a private wrong. One can reasonably object to expressions of blame that express a higher degree of blame than is appropriate in response to his wrongdoing. Finally, one can object to expressions of blame that fail to preserve one's capacities for moral agency. I maintained, however, that one cannot object to expressions of blame simply on the grounds that those expressions deny one the sorts of goods that proportionate, constructive expressions of resentment and indignation characteristically deny their objects. Even if one typically has a valid objection to being denied these goods, one loses the standing to make that objection when that denial constitutive of, or a consequence of, appropriately expressed blame.

⁷ While the targets of blaming expressions can reasonably object to expressions that are unconstructive because they do not respect their reasoning capabilities, it does not seem that they can reasonably complain on the grounds that those expressions are not constructive because they do not deter.

3) Reactive Punishment and the Standing to Object

These observations about the way blame works in the interpersonal context can be extended to explain why serious criminals do not have moral standing to object to punishment that denies them their liberty or property. I have described how in the interpersonal context, the targets of appropriate blame cannot object to treatment that would otherwise be morally problematic (such as the failure to honor a promise) when that treatment is bound up with an appropriate expression of blame. Likewise, the punished cannot object to treatment that would otherwise violate their rights when that treatment appropriately expresses the indignation of the community.

I argued in Chapter 3 that communities have reason to express indignation toward those who commit serious crime by depriving them of important goods. Such deprivations signal an appropriately high degree of indignation. It shows that members of the community take such crimes seriously and are committed to reducing their occurrence. Because punishment is an appropriate way to express indignation toward serious crimes, those who commit such crimes cannot reasonably object to punishment that blames them, even if it deprives them of goods to which they would otherwise have a right. Those who have committed serious crimes and been fairly convicted do not have the same standing to appeal to their rights to liberty and property that other people have, when the deprivation of their liberty or property partly constitutes proportionate expressions of blame. It would be unreasonable for a fairly convicted murderer to object at sentencing, “You cannot imprison me; it violates my right to freedom.” The murderer cannot appeal to his right to liberty, not because this is not a right that people are usually protected by, but because convicted murderers do not have the standing to appeal to that right when members of the community have appropriately endorsed imprisonment as the way to blame those who commit murder.

There are a number of valid grounds for objecting to reactive punishment with appeal to one’s rights. If one has committed no crime or has a valid excuse for acting in a way that would otherwise be criminal, one can reasonably object to being punished. One can also reasonably object to reactive punishment if one was properly exempted from reactive blame at the time of one’s wrongdoing, because one was not in possession of the capacities for moral agency.⁸ Given the importance of the goods that punishment deprives the punished, an accused criminal can reasonably demand that those who would blame him believe with some certainty that he has in fact committed a crime, that he does not have a valid excuse or justification for having done so, and that he was not in a condition at the time of the crime that would exempt him from moral assessment. In

⁸ Those who permanently lack the moral capacities may not have the standing to appeal to their rights, if they retain those rights in virtue of possessing capacities for moral agency. See, Strawson, P. F., 1962, “Freedom and Resentment”; reprinted in Watson, Gary, 1982, *Free Will*. New York: Oxford University Press, p. 66: “To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained.”

criminal justice systems in which there are procedures in place for establishing guilt, a criminal can reasonably object to his punishment if these procedures are not followed.

As in the interpersonal case, one who has committed a serious wrong can reasonably object to blaming expressions from those who are not in an appropriate position to blame him. So if some wrong is properly regarded as private in a community, someone who commits that wrong can object to reactive punishment in response to it. Because punishment is understood to express the indignation of the community, if infidelity is not properly in the interest of the community as a whole, an adulterer can object to reactive punishment. He retains the standing to object to state punishment in this case, but not the standing to object to interpersonal blaming responses that appropriately express the reactive sentiments of those who are in an appropriate position to blame him.

Even if one is worthy of public blame and is fairly convicted, one can reasonably object to an expression of the indignation of the community that expresses a higher degree of indignation than is appropriate. The celebrity shoplifter could reasonably object to a sentence that imprisons him for life, even if it would have good social consequences, because it would express an excessive degree of blame for his crime. A target of reactive punishment can also reasonably object to punishment that undermines its claim to respecting his capacities for moral agency with the way it treats those capacities. Even if one has been fairly convicted of a serious crime he can reasonably object to the state torturing him in response, on the grounds that torture characteristically destroys those capacities that make blame appropriate.

Though there are a number of situations in which someone who has committed a serious act of wrongdoing can reasonably object to treatment which deprives him of goods to which he has a right, I have maintained that there is a very circumscribed situation in which he does not have the standing to appeal to those rights: when denying him that good appropriately expresses the indignation of the community in response to his crime. This moral standing explanation of why punishment is permissible differs from the forfeiture view in a way that makes it immune to Quinn's and Goldman's criticisms of forfeiture. It explains why the state can permissibly deprive Jones of his liberty while Smith who is ignorant of Jones's wrongdoing cannot. There are at least two reasons to think that Jones has a valid objection to his kidnapping. First, in societies in which the state has a monopoly on the legitimate use of force, Jones can reasonably object to a fellow citizen holding him against his will, whether or not it expresses an appropriate degree of blame. Second, and more importantly for distinguishing moral standing from forfeiture, Jones's kidnapping cannot be understood as a blaming response because Smith is unaware of Jones's crime. According to the moral standing view, criminals do not forfeit any of their rights without qualification when they violate the rights of others; they only forfeit the standing to appeal to certain rights when they are the targets of appropriate expressions of blame. Even if he could justifiably be punished, Jones has a valid objection to being denied a good to which he has a right when that deprivation is not bound up with an appropriate expression of blame.

The moral standing account is also able to avoid Goldman's criticism of forfeiture because it is able to explain why the state is permitted to treat criminals in ways that

deprive them of different and greater goods than they have deprived of others in their crimes, without implying that any and all treatment of them is permitted. On the one hand, appropriate expressions of indignation do not always do to their targets what the object of indignation has done to his victim, and actions which express a proportionate degree of indignation toward wrongdoers sometimes deny them a greater good than they have denied their victims. Those who are fairly convicted of armed robbery cannot object to imprisonment that expresses the indignation of the community, provided that imprisonment expresses an appropriate degree of the community's indignation, even if it denies them a greater good than they have denied anyone in the community. On the other hand, criminals do not forfeit all their rights when they commit crimes on the reactive account. They can object being denied goods on the basis of their rights for various reasons, many of which I have enumerated above. For example, someone who commits a minor crime does not thereby lose his standing to object to being tortured on the reactive account, because torture is never an appropriate way to express the reactive sentiments.

I will conclude this section by comparing the reactive theory to the traditional alternatives with respect to the *no valid objection from rights desideratum*. In Chapter 1, I argued that consequentialist theories cannot make sense of the idea that people have rights in a meaningful way. I noted that the most famous criticism of consequentialist theories is they cannot explain why it is impermissible to punish the innocent when doing so would achieve the best consequences. The reactive account is not susceptible to this criticism. Those who are not blameworthy retain the standing to appeal to their rights in validly objecting to punishment. A related worry about consequentialist theories is that even when they justify state punishment of the guilty they do so based on the good social consequences of such treatment, even though the state cannot in general treat people in ways that would normally violate their rights on that basis. Alternatively, reactive punishment does not treat criminals merely as tools for producing social welfare but, rather, addresses them as responsible agents. Just as the reactive sentiments respect the agency of their objects, expressions of the reactive sentiments respect their targets. Proportionate expressions of blame sometimes treat their targets in ways that would otherwise fail to show someone the respect that all persons are owed. When this treatment is part of an appropriate expression of blame, however, it pays him a sort of respect. When Susie rescinds Jim's party invitation, for instance, she is respecting him as having the capacities to do better. Likewise, when criminals who are blamed by punishment that appropriately expresses the justified indignation of the community, those deprivations partially constitute treatment that respects them as moral agents. Though deterrent consequences play a role in why we punish on the reactive theory, we are not permitted to punish to an extent that would express a higher degree of indignation than is appropriate. Punishing past this point on deterrence grounds would treat the punished as a mere means to the security of society.⁹

⁹ I believe that the way that the deterrent function of punishment is bound up with expressions of indignation which constitutively treat the blamed as responsible agents gives the reactive theory a particularly robust response posed by the Kantian challenge that deterrent punishment treats criminals as means, compared to a theory that separates the expressive and deterrent elements of punishment. For comparison consider a justification of punishment suggested by T. M. Scanlon on which these elements

The idea that justified reactive punishment respects the punished suggests a possible explanation of *why* the targets of appropriate expressions of blame do not have the standing to reasonably object to those expressions. While people always have standing to object to treatment that fails to respect them, when one has committed a serious wrongdoing, treatment that would otherwise fail to respect one shows respect for their moral agency when it expresses appropriate blame. With respect to punishment, while people can usually complain about being deprived of their liberty or property without their consent on the grounds that such deprivations fail to show them the respect they are owed, the targets of reactive punishment are respected by those deprivations.¹⁰

This suggested deeper explanation behind the moral standing account of the permissibility of punishment relies on a questionable premise, however: that the only valid grounds that someone could have to object to any treatment could be understood as a failure to respect his agency. The idea that reactive punishment respects the punished explains why those who are deprived goods by it cannot reasonably object to it on the grounds that those deprivations fail to respect their agency. But it is not obvious that one cannot reasonably object to treatment that respects his agency on other grounds. Imagine I scream “I respect your moral agency,” in your ear. What I say respects you in a sense, but it seems clear that you can reasonably object to the way that I say it. While the criminal cannot reasonably object to his punishment on the grounds that it fails to respect his agency, the fact that it respects his agency might not fully explain why he has no valid objection to it. One reason to think this is the case is because not all valid objections to punishment directly appeal to the agency of the punished. I retain the standing to object to being punished for a private wrong I have committed, or to some forms of cruel punishment,¹¹ even though this punishment would respect my moral agency. I believe that the notion of the moral standing to object to blame is best understood as fundamental and not analyzable in other terms.

play distinct roles. Scanlon, T.M., 1999, “Punishment and the Rule of Law”; reprinted in his, 2003, *The Difficulty of Tolerance: Essays in Political Philosophy*, Cambridge: Cambridge University Press, p. 231: “The fact that an action is an unjustifiable intrusion... is a necessary condition for condemning it, and usually a sufficient condition for doing so. But it is not (on a nonretributive view) a sufficient condition for depriving the agent of liberty or inflicting other forms of harsh treatment on him. For such harsh treatment, some further justification is required beyond the desirability of expressing our judgments. This is where we must appeal to the utility of deterrence as a way of providing a kind of protection that we need and are entitled to, and the fact that everyone will have a fair opportunity to avoid liability to the penalties involved.”

¹⁰ In a recent article, Mitchell Berman recommends a similar line of thought to the retributivist. Berman, Mitchell, 2008, “*Punishment and Justification*,” *Ethics* 118, pp. 258–90, p. 279: “Because wrongdoers experience suffering as a bad, a usual way to respect them is to not cause them pain. But insofar as they have exercised their wills to violate legitimate interests of others, it is also plausible that causing them to suffer on account of their willing respects them too. Here, as is often the case (but often overlooked), our duties depend upon our motivations. We respect a person’s nature as sentient being by refraining from causing her pain. We respect a person’s nature as responsible agent by giving her what she deserves by virtue of the exercise of her will.”

¹¹ I claim in chapter 3 that a subset of forms of punishment that we consider cruel do not fail to preserve the moral capacities of the punished.

The reactive theory provides a more complete explanation of why punishment is permitted than retributive desert because desert claims alone do not generally override entitlement claims. Perhaps a retributive theory could incorporate moral standing in order to explain what is distinctive about retributive desert claims. On this understanding inflicting deserved harm or suffering on wrongdoers blames them.¹² Those who are subject to state punishment do not have the standing to reasonably complain that it violates their rights when those rights violations are required to appropriately blame them by giving them their deserved amount of harm or suffering.

Objections to retributivism that I have raised in previous chapters undermine the claim that giving wrongdoers their deserved suffering or blame is an appropriate way to blame them. I have argued that the retributive idea that wrongdoers deserve harm or suffering is ungrounded and that punishment can appropriately blame criminals without inflicting suffering on them or harming them in ways that retributivists often understand “harm.” In order to motivate skepticism about the retributive idea, I drew upon our blaming responses in the interpersonal context. In this context, expressions of third-personal blame with the end of the deserved harm or suffering of those who violate moral obligations are problematically vindictive. Consider the case where you forget to pick me up to drive me to the airport, and at a later date I say I will take you to the airport but fail to show up in order to deliver upon you your retributive matching penalty. It seems as though you can reasonably object to my vindictive treatment of you in this case. My argument that serious criminals lack the standing to object to punishment relies on an analogy with cases of interpersonal blame. The same argument is not available to a retributivist, insofar as retributive responses are inappropriate in the interpersonal context.

I believe that retributivism has at least one advantage over the reactive theory with respect to the permissibility of punishment: it has a clear explanation of why someone can reasonably object to being proportionally punished more than once for the same crime. If criminals deserve a fixed amount of harm or suffering in virtue of their crimes, and they receive it from their punishment, punishing them again would inflict undeserved harm or suffering on them. In the interpersonal context we think that someone can be blamed for the same wrong multiple times without reasonably objecting. This gives us further reason to think that our interpersonal blaming responses are not retributive. On the other hand, it might also give us reason to think that the practice of punishment is not continuous with our interpersonal practices, insofar as one has the standing to object to being punished more than once for the same crime. In modern states, however, the reactive theory understands state punishment to express the degree of indignation that members of the community have endorsed as appropriate in response to certain crimes. Only the state is in a position to punish and it is only permitted to punish a person once for each crime he commits. Were it to punish him twice, he could object that his the response to his crime is excessive on the grounds that the deprivation of

¹² Examples of retributive theories that emphasize punishment’s blaming role include Moore, Michael S., 1997, *Placing Blame: A Theory of Criminal Law*. Oxford: Oxford University Press, and Bennett, Christopher, 2008, *The Apology Ritual: A Philosophical Theory of Punishment*. Cambridge: Cambridge University Press.

combined goods expresses more indignation toward his crime than members of the community have endorsed as appropriate.¹³

This concludes my discussion of why punishment of those who commit serious crimes is morally permissible on the reactive theory. Combined with the argument from the earlier chapters, it provides a comprehensive justification of punishment. Those who commit serious crimes do not have the moral standing to appeal to punishment that expresses an appropriate degree of indignation toward their crime; and toward the most serious crimes the state has positive reason to punish all serious crimes in order to express a proportional degree of indignation in response to those crimes and thereby emphatically stand up for the victims of crime and show that all citizens are protected by certain moral obligations.

4) Conclusion

This concludes my argument that the reactive theory satisfies the three desiderata of a theory of punishment that I set out at the beginning of this dissertation. It satisfies the *positive reason desideratum* because we always have reason to express public blame in response toward serious crimes and punishment is accessible to all members of the community as a blaming response that takes wrongdoing seriously. It satisfies the *comparative proportionality desideratum* because a higher degree of indignation is called for in response to more serious crimes and more severe punishment expresses a higher degree of indignation from the community. Finally, it satisfies the *no valid objection from rights objection* because those who have committed serious crimes do not have the moral standing to object to appropriate expressions of the reactive sentiments.

While I have maintained that the reactive theory offers a distinct justification from consequentialism and retributivism, it embodies some of the features of our moral thinking that make those traditional theories plausible in the first place. The idea that punishment deters crime and thereby has good consequences is part of the reason that communities decide to spend their limited resources on justice systems and detention facilities. The deterrent function of punishment contributes to the explanation of why blame takes the form of punishment on the reactive account. Deterrence promotes those values that partially constitute the reactive sentiments expressed by blaming behavior, so deterrent punishments are constructive expressions of the community's indignation. Considerations of constructiveness inform our judgment about what expressions of blame are proportional. Retributivism, on the other hand, does not justify punishment based on its consequences, but instead based on the idea that it is a fitting response to wrongdoing. The understanding of punishment as a backward looking response to wrongdoing that acknowledges the agency of the wrongdoer seems essential to the way we think about it. The reactive theory shares this understanding of punishment with retributivism, without

¹³ I do not find this response completely satisfying. It leaves open the possibility that in the state of nature one could be subject to multiple punishments for the same wrongdoing without valid objection. I suspect that rather than a good argument for escaping the state of nature, this is an implausible account of what would be permissible in the state of nature.

relying on the controversial retributive idea that wrongdoers deserve to suffer or be harmed.¹⁴

Much of my motivation for the reactive justification has come from how it compares favorably to consequentialism and retributivism. Of course, there are other strategies for justifying punishment. I will conclude by briefly comparing the reactive theory to one of these strategies in a way that I hope illuminates some features of the theory. This third strategy justifies punishment based on its role in educating criminals about the wrongness of their crimes.¹⁵ Like the reactive theory, these moral education justifications employ the expressive function of punishment. Through punishment, the community communicates to the punished the judgment that his crime is wrong in a way that forces him to reflect on his wrongdoing and can lead him to improve his character and turn his life around. One of the unique features of the moral education approach is that the justifying aim of punishment is the *benefit of the punished*,¹⁶ as opposed to the benefit of society at large or the deserved harm or suffering of the punished. Typical metaphors employed by the theory invoke the parent-child relationship.¹⁷

I do not think a theory that has education as its justifying aim can straightforwardly satisfy the desiderata of a theory of punishment that takes commonsense morality seriously.¹⁸ First, it seems that moral education theories fail to satisfy the *positive reason desideratum* because intuitively punishment can be justified even if it does not educate. There are various reasons that educative punishment might fail: the punished might ignore the intended message of punishment, or fail to understand it, or understand it but reject it.¹⁹ Among those most likely to ignore the message are the most hardened criminals, those whose practical reasoning and behavior is least responsive to persuasion by the state or their fellow citizens. We think that imprisonment of unrepentant murderers is justified even if it does not reform them, and even if we have no legitimate expectation that it will. It does not seem that moral education theories

¹⁴ Though, see the conclusion to Chapter 3 for an argument that the reactive theory may be able to ground a nonstandard version of the retributive idea.

¹⁵ For a clear presentation of this view see Hampton, Jean, 1984, "The Moral Education Theory of Punishment," *Philosophy and Public Affairs* 13: 208–38. Hampton's considered view of the justification of punishment later changed to the retributive theory that I discussed in Chapter 1.

¹⁶ See Hampton, Jean, 1984, "The Moral Education Theory of Punishment," p. 214: "[The moral education theory] attempts to justify punishment as a way to benefit the person who will experience it. . . . [T]he moral good which punishment attempts to accomplish within the wrongdoer makes it something which is done *for* him, not *to* him."

¹⁷ Hampton, Jean, 1984, "The Moral Education Theory of Punishment," pp. 217–218, 226, 237.

¹⁸ Of course, a moral education theorist could claim, "so much the worse for commonsense morality." Hampton sometimes adopts this strategy. See, for example, Hampton, Jean, 1984, "The Moral Education Theory of Punishment," p. 242: "[O]ur society has been notoriously unsuccessful in coming up with punishments that are in any way morally educative. And I would argue that it speaks in favor of this theory that it rejects many forms of incarceration used today as legitimate punishments, insofar as they tend to make criminals morally worse rather than better."

¹⁹ In presenting a different communicative view, Duff recognizes that rational communication necessarily leaves open possibilities such as these. Duff, R.A., 2000, *Punishment, Communication, and Community*. New York: Oxford University Press, p. 122.

explain the positive reason we have to punish in such cases.²⁰ Other counterexamples to moral education theories with respect to the *positive reason desideratum* include cases in which non-punitive modes of communication would educate criminals at least as effectively as punishment at a lower cost. These problematic cases also explain why moral education theories have difficulty satisfying the *comparative proportionality desideratum*. The severity of punishment we have reason to inflict on criminals is not a function of the seriousness of their crimes, but rather a function of what punishment will be most effective in morally educating criminals given their individual circumstances.

Moral education theories have a straightforward answer to the *no valid objection from rights desideratum*: criminals cannot reasonably object to being punished because it benefits them. As one might expect from a theory that utilizes the parent-child analogy, this answer is paternalistic. If citizens can object to the state violating their rights for their own good on paternalistic grounds, criminals have a valid objection to educative punishment. Even if one does not find this paternalism problematic,²¹ a related objection applies in some cases given that punitive communication is not always the most effective means of moral education. If my character would be more effectively improved by voluntarily attending my local weekly Bible study than by being imprisoned, it seems like I can reasonably object to the state imprisoning me, and making it impossible for me to attend that Bible study, for the sake of improving my character.

The inability of moral education theories to satisfy the desiderata brings out some of the appealing features of the reactive alternative. The reactive theory avoids the objections to educative punishment with respect to the *no valid objection from rights desideratum*. Our interpersonal blaming practices reveal that we think the reactive sentiments and their expression are appropriate responses for adults to have toward the wrongdoing of other adults, so are not paternalistic. Further, because the aim of reactive punishment is not to benefit the punished, criminals cannot object to their reactive punishment on the grounds that it does not benefit them. Again consider the interpersonal context. Jim cannot complain about the blaming behavior of Susie and her father on the grounds that other ways in which they might have blamed him would have more effectively improved his character. The idea that successful blame need not educate the blamed explains why the reactive theory is not susceptible to the same objections with

²⁰ Hampton recognizes this possibility but argues that respecting the agential capacities of criminals means always treating them as educable. Hampton, Jean, 1984, "The Moral Education Theory of Punishment," *Philosophy and Public Affairs* 13: 208–38, p. 231. I find this response unconvincing. There seems to be a practical irrationality in attempting to do what one is nearly certain will fail. If we see someone jumping up and down outside the Sears tower while saying, "I'm certain I can't do it but I'm trying to jump over the Sears tower," we take it as an unfunny joke rather than taking him at his word—he is acting as if attempting to do what he takes to be something he will never achieve. For a criticism in this spirit see Shafer-Landau, Russ, 1991, "Can Punishment Morally Educate?" *Law and Philosophy* 10: 189–219, p. 217.

²¹ Hampton maintains that the sense of paternalism that her theory relies upon is unproblematic. Hampton, Jean, 1984, "The Moral Education Theory of Punishment," p. 219: "[The moral education theory] views the state's punishment as interference in [the individual's] life plans without his consent for his own good. But why should paternalism in this sense be offensive? It would be strange indeed if philosophers insisted that the state should only try to prevent further harm to the community by actively intending to harm, or use, or at least be indifferent to, the people it punishes!"

respect to the *positive reason* and *comparative proportionality desiderata* as moral education theories. Unlike successful communication, a successful expression of blame does not require any action by its target. Punishment can successfully express the indignation of the community regardless of whether the punished accepts the blame and turns his life around. Hardened criminals are blamed by their punishment, even when they are not educated by it. So the reactive theory gives us positive reason to punish serious criminals to an extent that tracks the seriousness of their crimes, whether or not doing so will morally educate them or more effective ways of improving their characters are available to us.

In response to this last point, one may worry that the success of punishment on the reactive theory is too focused on the expresser of blame at the expense of the potential positive impact that punishment can have on the target of the expression. However, I maintained in Chapter 3 that considerations of constructiveness give us reason to punish in ways that make it possible that punishment will lead the punished to turn their lives around. Successful reactive punishment is certainly *consistent* with the idea that punishments have this effect, as successful blame sometimes does have the effect in the interpersonal context.

I conclude by emphasizing this point: not every feature of our system of punishment need itself be justified by what gives us positive reason to punish proportionally and makes punishment permissible. Those features need only be consistent with it and otherwise morally defensible. Society's other values ought to be reflected in systems of punishment. Yet from our complex network of social values, I believe that those that explain why punishment is justified are the values that are always bound up with justified blame.