

FAIRNESS, SANCTIONS, AND CONDEMNATION

— preliminary draft, not for citation or circulation —

Pamela Hieronymi

hieronymi@ucla.edu

May 19, 2019

Preface:

This text appears in the middle of a manuscript about free will and moral responsibility, provisionally titled *Minds that Matter*. The first half of the manuscript identifies a problem about understanding our own agency: once we grant that our thought and behavior are a product or function of our history and environment, it can seem that we are not really in control of ourselves or of what we do. I propose a solution to that problem, by appeal to an idea about control that I have been advancing in other work. However, even if we accept that solution and thereby resolve the difficulty about agency, a second problem remains: It seems as though there is some incompatibility between being or holding others morally responsible and understanding our thought and behavior as a product or function of our history or environment. This is a distinctively *ethical* problem, rather than a metaphysical one: even apart from worries about agency, morally *itself* seems to be incompatible with what I will call *determination*. This is the problem addressed in the second half of the manuscript.

In this text, I consider one aspect of the ethical problem, viz., the somewhat complicated thought that moral responsibility is incompatible with determination because, first, being morally responsible involves being vulnerable to demands and sanctions; second, demands and sanctions require, for their fairness, an adequate opportunity to avoid them; and, finally, an adequate opportunity to avoid is ruled out by the fact that our thought and behavior are a product or function of our history and environment. This complicated thought was put forward by Gary Watson (in his landmark “Two Faces of Responsibility”) as a diagnosis of our concerns about alternative possibilities. It has the considerable virtue of making plain the ethical nature of the problem: it fits well with the intuition that holding people responsible, if determinism is true, is not just inapt or confused, but is specifically *unjust*.

I would like to press an often overlooked question: why does the fairness of a sanction require an adequate opportunity to avoid it? By pressing this question, I believe we can better understand what people might have in mind when they talk about the “condemnatory force” of moral blame, or perhaps even “basic

desert.” If the understanding I propose in fact captures what people have in mind, it should also show how condemnation or desert, in this sense, can be left behind.

I am confident that I am onto something important, but I am also confident that I have not yet nailed it down. I am therefore eager for your feedback.

To begin, let us examine more carefully the ideas of demand and sanction.

What is it to make a demand? We sometimes say I make a demand, rather than a mere request, when I back that request with some negative consequence, i.e., when I make clear that I stand ready to take some action, should you disappoint me.¹ In such a case, one might say, I am *holding you to my* request. That is to say, it seems a request becomes a demand when it is enforced with a kind of one-off penalty, sanction, or threat. Demands of this sort can be understood in the same way that I will suggest, below, that we understand sanctions. I will postpone until the next chapter consideration of a more subtle sort of demand, one that is not created by the specter of some voluntarily imposed negative consequence.²

Turning, then, to sanction: To sanction someone is, as Watson agrees, to impose something unpleasant or unwanted upon that person in response to his or her violation of some norm, demand, or expectation. In holding accountable, he says, we “lay it down” that a certain consequence will follow if the standard isn’t met. (236–7) This seems true of sanctions. Sanctions are consequences that are *created* and *attached* to certain failures in certain contexts. Parents sanction or penalize children for misbehaving; society penalizes its members for violating civil law; nations impose sanctions on other nations for pursuing weapons programs; teachers penalize students for turning in work late. We even create and impose sanctions and penalties in games and sports.³

¹ CHECK recent paper about requests in PI.

² The more subtle sort: any criticism of your choices seems to call for some response—some justification or amendment. This “demand” does not rely to any threat of a negative consequence.

³ Note that, so understood, sanctions need not have any distinctively moral significance. I will take up the thought that blame is a sanction imposed within the social institution called “morality” later.

In contrast, a hangover is not a sanction for drinking too much. It is just a negative consequence. Similarly, the big mess in your kitchen is a negative consequence of last night's dinner party. The fact that you have to clean up that mess is not a sanction for having the party. It is just your job, and no one else's.⁴ Sanctions are not simply negative consequences that no one else is required to bear for you. They are negative consequences that have been created or instituted and attached, by some person or body, to certain violations or shortcomings.⁵ As such, they raise concerns of fairness. As Watson puts it, "It is unfair to impose sanctions upon people unless they have a reasonable opportunity to avoid incurring them." (237)

But: Why does the imposition of sanctions, understood as negative consequences that have been created and attached to certain failures in certain contexts, raise concerns of *fairness*? And why should the *fairness* of a sanction require an *opportunity to avoid* it?

We must carefully note and then diligently set aside one popular answer to this pair of questions, according to which sanctions require a reasonable opportunity to avoid because, without an adequate opportunity to avoid, sanctions are *undeserved*, and, if they are undeserved, they are, therefore, unfair.

This appeal to desert is unhelpful. To see why, consider that it can be interpreted in two ways, which I will label *deflationary* and *substantive*.

On the first, deflationary way of understanding desert, the question of whether you "deserve" this or that simply reduces to (or is transparent to) the question of whether giving it to you or imposing it upon you would be apt, justified, warranted, or fair. In this case, the claim that, absent an adequate opportunity to avoid, sanctions are undeserved is just a different way of saying that

⁴ It is what Scanlon would call a "substantive responsibility"—a question of where obligations and burdens fall. I am, here, distinguishing sanctions from other (or mere) substantive responsibilities, in a way that, to my knowledge, Scanlon does not.

⁵ One might worry that this definition leaves out things like "giving the cold shoulder" or dressing down. I will later place these under the label "guilt-tripping" and will understand them as attempts (perhaps successful attempts), by an individual, to institute and enforce a sanction. Thanks to Angela M. Smith for raising the question.

sanctions require an opportunity to avoid, or that, without such an opportunity, they are inapt, unjustified, unwarranted, or unfair. Thus, the deflationary appeal to desert leaves our questions untouched.

Even so, “undeserved” may be a useful label to introduce: Perhaps it will pick out the *specific form* of justification that justifies, and thereby shows fair, penalties, prizes, punishments, and the like. I have no objection to this usage—in fact, I will later suggest a specific form of justification to which the label might thus be applied. However, we must bear in mind that, until we provide that underlying justification, our questions remain untouched—saying that fair sanctions require an adequate opportunity to avoid because, without such an opportunity, they are undeserved is like saying that the sleeping pill works because it has a dormative virtue.

Alternatively, one might think that facts about desert *themselves* provide *independent* reason by which to answer the question of whether or under what conditions giving you or imposing upon you this or that would be fair. Call this an appeal to *substantive desert*. On this view, we can answer our original questions, by appeal to desert, but only immediately to face the further questions of why undeserved sanctions are unfair and why, without an adequate opportunity to avoid, sanctions are undeserved. These further questions are difficult to answer.

One might take the difficulty of answering these questions as evidence that the claims are primitive, bedrock, or basic. One might think that asking why undeserved sanctions are unfair or why deserved sanctions require an opportunity to avoid is like asking why pain is bad. The claim is not simply a tautology (as it would be on the deflationary reading), but, one might think, it is not a question that needs or even could be given an answer.

I find the notion of substantive desert opaque and therefore find the appeal to it as basic or bedrock unsatisfying. It seems to me we can legitimately ask “*Why* does *S* deserve *x* on account of *y*?” or even, “*What is it* for *S* to deserve *x* on account of *y*?” and, further, that our answers to these questions—which cannot rely on the idea of substantive desert—should, themselves, show why

giving x to S or imposing x upon S because of y would be *pro tanto* both justified and fair. (In contrast, answers to the question “Why is S in pain?” or “What is it for S to be in pain?” do not show why being in pain is bad.) However, once we have an account that explains why giving x to S or imposing x upon S would be *pro tanto* justified and fair, an account which does not rely on the substantive notion of desert, then, it seems, we no longer need the notion of substantive desert. The work is being done by the account, for which “desert” serves as a label or placeholder—we have returned to the deflationary interpretation.

One might yet disagree, thinking that desert is something like a “thick” ethical concept,⁶ something like courage or kindness: We can answer specific questions about why someone or something is courageous or kind, but we cannot eliminate these concepts by replacing them with our answers to such questions. The role the concepts cannot be deflated or reduced.

For a moment, let us grant that “desert” is such a concept. Will our questions (Why is undeserved sanction unfair, and why is a sanction undeserved if one’s thought or behavior is a function or product of one’s history or environment?) then receive an answer? If so, then I invite the objector to provide answers, which we can then compare the answer to the one I will suggest below. If not, then “desert” has again been treated as a primitive. Since I find the appeal to desert as a primitive baffling, I will continue my inquiry. But I invite the objector to continue along and to consider the rest of what I have to say as an alternative answer given by someone who is not content to take desert as bedrock.

To continue the inquiry, then, we need to set aside the possibility of taking, as bedrock, the claim that “desert” requires the falsity of determination.⁷

Here is an alternative answer to the question of why sanctions raise concerns of fairness: Unlike a hangover or the mess in your kitchen, sanctions are created, attached to certain failures, and imposed

⁶ CITE Williams, Rippon.

⁷ CITE literature that depends on “desert”

in given instances, by some person or body, through some voluntary action. They are thus subject to the moral concerns raised by *any* voluntary action that impacts the interests of others: any such voluntary action comes under the general ethical principle of “do unto others as you would have them do unto you”—or, at least, “do not do to others what you would not have done to you.” (I take this to be about as bedrock a moral principle as you will find, and I am content to treat it as primitive.)⁸

Importantly, though, an appeal to “do unto others...” is not necessarily an appeal to fairness. It is, on its face, an appeal to empathy. As noted by Kant, it might therefore be used to ask for special treatment. In his example, a convicted criminal appeals to it in asking for mercy from the sentencing judge.⁹ To understand “do unto others....” as counseling *fairness*, it must be generalized: it must be understood as generating principles or directives that will apply uniformly across the population. So understood, it would counsel us to adopt principles or directives that treat each person as best we can, consistent with treating others in the same way.¹⁰ That is, would ask us to adopt principles that treat the needs and interests of each symmetrically while doing adequately well by all. I am, again, content to treat as bedrock the idea that we ought, in our voluntary actions, to comply with principles or rules that treat the needs and interests of each symmetrically while doing adequately well by all, and I will refer to these as principles of *fairness*.¹¹

In coming to conclusions about fairness, what counts as “symmetric” or as “adequate” will, of course, be hotly contested—but, if we are contesting fairness, then, I suggest, we are contesting

⁸ In contrast, non-voluntary attitudes, such as distrust, can burden others without raising the same concerns of fairness, as will be detailed in the next chapter (also, CITE Force and Fairness).

⁹ CITE. COMMENT about mercy.

¹⁰ NTS: the negative is hard to state. Current formulation is too strong, b/c “as best.”

¹¹ NOTE shift from “as best” to “adequately.” That’s substantive. ALSO “wrongness” “justice” and “fairness” problem. ADD cite to White, Preston-Roedder about the need for prerogatives. SAY SOMETHING about the interjection of principles, rather than direct appeal to outcomes. Not just rule utilitarianism, but a principle about the rules: they must be fair (rather than maximizing).

whether a given policy or principle treats the interests of each symmetrically and adequately well. John Rawls' thought experiment provides a useful heuristic: We can think about what would treat the interests of each symmetrically and adequately well by thinking about what policies or principles would be chosen from behind a "veil of ignorance," i.e., if you did not know which position you occupy, while motivated by self-interest and mutual disinterest. The hope is that, under those assumptions, you would want to do as well as you can for each one—that is, do as well as you can for any particular one, consistent with doing at least that well for others.¹² The policies or principles so adopted will be ones whose violations are unfair.

An important caveat: what if some people have a special interest in—take a special delight in or psychologically benefit from—belittling others? Or what if some people have a special interest in, or take special delight in, being given special treatment or enjoying special status—being treated asymmetrically well, relative to others? Must these "interests" be treated symmetrically with other interests, in determining principles of fairness? Plausibly, no. Advancing such interests would require treating someone asymmetrically. Thus, advancing such interests is incompatible with fairness¹³ and, therefore, these interests need not be counted among those that must be

¹² As was shown by Hasanyi, the hope may not be fulfilled. As argued by Scanlon in "Contractualism and Utilitarianism," the appeal to a veil of ignorance is simply a heuristic, and there are others: imagining that you will be assigned a place in society by your enemy; asking whether a society run by these rules would be a stable; engaging in the "trading places" thought experiment suggested by the Golden Rule. In each case, we are attempting to determine whether a given policy or principle treats each symmetrically and adequately well—whether it is, in Scanlon's terms, one that could not be reasonably rejected as a basis for informed, unforced general agreement by those motivated to find such a basis. For Scanlon, the symmetry and adequacy required by fairness are captured by providing each with a veto to any particular system, with the restriction that the veto can only be used when it is used consistently with two facts: (a) others have the same power of veto and (b) each of us are committed, first and foremost, to arriving at an agreement about how to live together cooperatively. CITE Rawls, Hasanyi, TMS, TMS paper.

It will be noted that I am providing a broadly contractualist interpretation of the everyday notion of fairness. I believe such an interpretation is the natural one—the success of my diagnosis of "condemnation" and "basic desert" depends on it being so.

¹³ incompatible with treating others symmetrically and adequately well

symmetrically balanced and adequately served, in a determination of fairness. Such interests are, we might say, *silenced*.¹⁴ This idea of silencing will be important later.

In sum, on the current suggestion, the institution or imposition of a sanction triggers concerns of fairness because it is a voluntary action that burdens others. A sanction will be fair if but only if it is part of a system that treats the interests of each symmetrically while doing adequately well by all. I will take this much as read and now turn to the questions of why a sanction is fair only if one had an adequate *opportunity to avoid* it.

FAIRNESS AND OPPORTUNITY

Why would fairness require that sanctions be imposed only if there was an adequate opportunity to avoid them? What is the connection between *treating each symmetrically and adequately well* and *ensuring adequate opportunity to avoid sanction*?¹⁵

To start, we can assume a general presumption against imposing burdens on others, a presumption grounded in the interests that would be thereby burdened. That presumption must be overcome if a sanction is to be justified. In the case of sanctions, it is typically overcome by the fact that instituting a sanction can serve some important *purpose*: typically, it can provide an incentive or deterrent—that is, it can influence people’s choices—in order to achieve certain goals or to protect

¹⁴ These caveats show up in the “motivational assumption” made by Rawls: that parties to the contract are self-interested and mutually disinterested. CITE. They are “unreasonable” in Scanlon’s terminology: incompatible with a commitment to finding reasonable principles.

¹⁵ CITE TMS who explores, more generally, reasons we have for wanting what happens to depend on our choices.

certain interests.¹⁶ For example, instructors institute a late penalty to provide students with an incentive to submit their work on or before the deadline, thus ensuring both that students have roughly the same amount of time to complete the assignment and that the instructor has adequate time to mark and return it, enabling learning during the term.

(An important side note: The fact that a sanction serves a purpose does not render the justification of the sanction, in an individual case, “forward looking” or “consequentialist.” Once instituted, a system of sanctions is fair only if it is uniformly enforced. Enforcement will be “backwards looking.”)

Because a system of sanctions is fair only if it treats each person’s interests symmetrically and adequately well, the importance of serving the purposes of the sanction (providing an incentive, getting papers turned in) must be sufficient to overcome the risk of incurring the burden faced by each individual open to the sanction. It will be unfair, e.g., to cut off fingers for turning in work late, even if that sanction were symmetrically enforced. Importantly, cutting off fingers would also be inhumane, cruel, and generally horrid—those are, in fact, the most vivid problems with the appalling suggestion. However, in addition to being inhumane and cruel, it would be *unfair*, because the risk, to each, of incurring such a sanction through false accusation, bad luck, or mistake, far outweighs its benefits. It would treat each symmetrically, but not adequately well.

¹⁶ [[WORK ON THIS other reasons for sanctions: Sanctions are sometimes instituted as a way to fairly distribute risk—and this might seem to result in sanctions that seem not to require an adequate opportunity to avoid, i.e., “strict liability.” For example, in some places those who keep wild animals are liable for any damage caused by the animal, regardless of whether there was any negligence on the part of the owner (where negligence would signal an opportunity to avoid)—they are strictly liable. Here, the system of liability allows a person to engage in risky behavior (keeping wild animals) by insisting that they take the risk created upon themselves. Notice that, in such a case, the animal keeper does not incur the penalty through simple bad luck (and so it is not the case that each of us runs the risk of incurring this penalty, as victims of outrageous fortune). Rather, we each have a reasonable opportunity to avoid the sanction by opting against keeping wild animals. Given the interests of all, this might seem fair. [NTS: do some research about strict liability. One might wonder about whether the cost of denying the right to keep wild animals, wholesale (a kind of liberty cost, I take it) weighs fully against the damage or loss that might be caused by the animals that cannot be adequately compensated by the animal keeper (e.g., a death)]. In contrast, no-fault insurance: the costs and complication of creating the disincentive are not worth the goods of the disincentive. Better to have a more streamlined system that treats it all as bad luck or as covered by a tax. Also: expressive reasons.]]

Importantly, though, without an opportunity to avoid the sanction, the sanction will not provide an incentive or deterrent, because, without an opportunity to avoid, there is no relevant choice for the sanction to influence. Thus it seems that, in cases in which there is not an adequate opportunity to avoid the sanction, the sanction will harm those who run afoul of it for no purpose. Thus, if we were to institute a system of sanctions that remains insensitive to an individual's opportunity to avoid the sanction, we put some people (perhaps each of us) at risk of incurring pointless harm. It would be better if our system of sanctions made exceptions for those cases in which there was not an opportunity to avoid. In fact, such exceptions seem required by fairness.

Here, then, is a suggestion: Fair sanctions require an adequate opportunity to avoid, because, without such an opportunity, imposing the sanction would burden certain people's interests for no purpose. It would thereby fail to treat the interests of each symmetrically and adequately well.

Next: What makes a person's opportunity to avoid a sanction *adequate* for fairness?

Here is an initial, natural hypothesis: A student has an adequate opportunity to avoid the late penalty just in case the student faces a choice (or, series of choices) by which they could successfully avoid the penalty.

There are two different, natural ways to support this initial hypothesis. The first appeals to something like desert, reasoning that, if the student faced a choice, then it was *up to the student* whether they incurred the penalty or not, and thus they cannot claim that the burden is unfair or that the system of sanctions treats them poorly. They made their bed, so they can lie in it. On this way of thinking, choice is sufficient for fair penalty. The second way to support the initial hypothesis appeals again to the purpose of the penalty: If the student did not face choices by which they could successfully avoid the penalty, then there was nothing for the penalty to incentivize. On this line of thought, choice is necessary for fairness.

This natural hypothesis is incorrect: The fact that someone faces a (series of) choices by which they would avoid the sanction is neither sufficient nor necessary to ensure that they had an adequate opportunity to avoid it.

First, choice is not sufficient. Suppose that the assignment was given on short notice during a period in which the student faced a series of time-consuming events of familial, religious, or cultural importance. It is plausible that, even though the student faces choices about whether to prioritize the assignment over the events, and even though the student may knowingly choose to incur the penalty—and, perhaps, could have chosen, instead, to complete the assignment and so avoid the penalty, in however strong a sense of “could have” one likes—the penalty may nonetheless be unfair—and it may be unfair because the student did not have had an adequate opportunity to avoid it. Whether the student had an *adequate* opportunity to avoid turns, in part, on the *quality* of the opportunity, so to speak—i.e., on the other burdens that one would have to incur in order to avoid it. Thus, the fact that one made a choice to incur the sanction, even freely and knowingly, is not, itself, sufficient to ensure that one had an adequate opportunity to avoid it.¹⁷

On the other hand, it may sometimes be the case that a student had an adequate opportunity to avoid a late penalty even though they *never* faced a choice by which they would avoid it. Consider a student who simply forgot about the assignment or who overslept and missed the deadline. There was no particular choice the student faced by which they would avoid the sanction. And yet, depending on the circumstances, the student may have had an adequate opportunity to avoid it. Surprisingly, choice is not necessary for an adequate opportunity to avoid.

This surprising conclusion requires investigation. Let us focus, for a moment, on the student who overslept. The oversleeping student may try to advance the same argument by which we

¹⁷ Therefore the weaker claims, that you could have chosen to avoid the penalty, or that you face a choice that triggers the sanction, are also insufficient. (Note: This way of being unfair will be ruled out for sanctions for moral wrongdoing: in claiming that a choice is *wrong*, we have settled whether the quality of opportunity to avoid is adequate: one ought to have chosen against the wrongdoing.)

concluded that fair sanctions require an opportunity to avoid: Enforcing the late penalty on oversleepers, they might argue, does not provide an incentive to turn in work on time: a student who oversleeps cannot turn back time to correct their error. Therefore, they might argue, a system that does not make exceptions for oversleepers burdens them for no purpose. Therefore, a system that does not excuse oversleepers is unfair.

One could raise two doubts about this line of reasoning.¹⁸ First, one might argue that enforcing the late penalty, even in the case of the oversleeping student, does provide an incentive even though the oversleeping student cannot turn back time: it creates an incentive for students either to take steps to ensure that they do not oversleep or else to turn their work in a bit in advance, to guard against unforeseen difficulties. If oversleeping, forgetting, or otherwise accidentally missing a deadline are the sort of thing to which students are prone, then enforcing the late penalty in cases of oversleeping may provide a helpful incentive to take steps to guard against such accidents.¹⁹

In effect, this first reply simply insists that the oversleeping student *did* face some choice by which they would have avoided the penalty: perhaps a choice about whether to stay up very late, or whether to put off the work until the final hours available for the assignment.

However, one might object that, although oversleeping can sometimes be avoided through better planning, sometimes it cannot. Sometimes even the most conscientious of students, the one who has excellent time-management habits, a reliable alarm clock and a general plan to turn things in with a bit of time to spare, can find themselves in a situation in which everything goes wrong and,

¹⁸ A third doubt is this: One might think that a late penalty functions, not only to influence choices—that is, not only as a sanction—but also to level the playing field, so to speak—that is, also as a kind of handicap. Students who turn work in later than others are likely to have benefited from that extra time, and, therefore, the evaluation of the work should be handicapped, so to speak, relative to others. I will set aside this way of thinking about the late penalty and focus, instead, on the late penalty as an incentive or sanction—as a way of influencing choices. I will therefore consider only the remaining two doubts.

¹⁹ NTS: the turn-back-the-clock argument is looking at the consequences in the individual case. The reply is looking at it as a principle that applies uniformly across the population. Cf. Kant's criminal.

Alternatively, if oversleeping is granted as an exemption, then other foibles of time management might also need to be granted an exception, to ensure that the interests of all are treated symmetrically. But making too many such exceptions may become unwieldy, for the instructor.

really, there was nothing they could have done to avoid missing the deadline. Even the most conscientious student might sometimes fail to submit work on time, through sheer bad luck. In fact, all of us, as creatures of fortune, are liable to similar situations—it is, as we say, the kind of thing that happens to people. One might think that enforcing the penalty, in such a case, is unfair—after all, the student has made all of the choices the sanction was meant to incentivize.

Before considering this suggestion, it is important to notice that when we say that, really, there was nothing the fully conscientious student could have done, and contrast this student with the typical student who could have better managed their time, we are not marking a difference in the freedom of these two students, metaphysical or otherwise. When we say that the fully conscientious student “could not have avoided” the penalty, we are not ruling out all counter-factual possibility of avoiding it. After all, there were, strictly speaking, things that the fully conscientious student might have done: they could have set a second or third alarm clock, for example. In fact, we could stipulate that both determinism and determination are false, that the student enjoys libertarian freedom of whatever sort one likes, and that the student owns three working alarm clocks—even so, such a student might rightly complain that there was nothing they could have done to avoid the penalty. In saying this, they are simply saying that, as things happened, they *did everything that could reasonably have been expected of them*, to avoid it. To say that they failed through “sheer bad luck” is to say that they did everything that could be reasonably expected of them and yet they did not, thereby, avoid the failure.²⁰

If a student did all that one could reasonably expect to avoid the penalty, then, it may seem, there is no further work for the sanction, as incentive, to do (perhaps, one might say, it did its work). One might conclude that the sanction therefore serves no purpose and is, therefore, unfair. We

²⁰ CITE Julie on alarm clocks. COMMENT about Julie’s work. Say something about good will and character. AND blanking john: it was possible, but he could not avoid it.

should, again, carve out an exception for the *conscientious* oversleeper (though perhaps not for the typical oversleeper).

Before drawing this conclusion, though, we must notice that, even if a sanction does not provide an incentive (its original purpose), the making of exceptions must itself also be fair—it must also treat each symmetrically and adequately well—and sometimes making exceptions for certain cases will create costs greater than the cost of the sanction. If so, then failing to make the exception and instead enforcing the sanction will still treat each symmetrically and adequately well, and therefore enforcing the sanction will be fair.

To explain: The institution and administration of a sanction, in practice, will implicate interests beyond those most directly served by it. These further interests can be highly context-dependent. For example, interests of both the instructor and of the other students require that enforcing the sanction not require a significant expenditure of the instructor's time nor a significant invasion of student privacy. The interests of students require that the late policy cannot be too easily “gamed.” These further interests will likely have different implications in a very large class, in which the instructor does not know the students well, than in a very small one. They may also have different implications in an extremely competitive environment, in which gaming the system will be more tempting or more prevalent, than in a more cooperative culture, in which there is more trust.

Notice, next, that the task of discriminating between the fully conscientious and not-fully-conscientious student is neither easy nor costless. Unless the instructor both is and is known to be, not only very familiar with, but also a judicious and objective judge of, the character of all the students, the instructor will not be able to decide either whether the oversleeping student is fully conscientious, or whether the student is telling the truth, by appeal to their “sense” of the students, their “gut” or “good judgment,” without risking serious unfairness. There is, on the one hand, the worry that the instructor will show undue favoritism to those the instructor likes or feels some affinity with or is predisposed to think well of, and, on the other hand, the worry that students will

be able to game or to deceive the otherwise the judicious instructor. One might try to avoid these risks by requiring some sort of investigation into either the general habits of or the fateful morning of the oversleeping student—but that will require what is likely an unacceptable invasion of the student’s privacy. One might try to allow oversleeping, generally, to be a reason for exemption—but then one might need to grant a lot of exemptions, allowing other kinds of bad luck to exempt, in the interests of fairness. One might try to accommodate the bad luck by allowing, for everyone, a “one time” exception to the late policy. But this is an imperfect solution: it will reduce the effectiveness of the late policy, generally, and it does not address the fact that someone could encounter genuine bad luck twice in the term.

Taking into account all these difficulties and constraints, we *might* conclude that making an exception for the fully conscientious will create costs that are higher than the costs of late penalty. It may be that we cannot devise a system that can avoid the risk of harming the conscientious without introducing other, worse harms.²¹ We might then conclude that the system that serves each students’ overall interests most fully and fairly will enforce the late penalty even in certain cases in which, it is admitted, the student in fact did all that could reasonably be expected and simply suffers from bad luck. Doing so will be fair, despite the bad luck, so long as this policy symmetrically and adequately addresses the interests of each one. One might conclude it does. The bad luck, we might think, will be acceptably²² distributed, and the costs of rooting it out of the system high, and so, one might think, it is consistent with fairness to let the bad luck fall where it does. We might view the risk of incurring the late penalty through bad luck as a risk we can each accept, a cost worth incurring, given the gains to be had by implementing the late penalty. In that case, the penalty

²¹ Costs to *each* student, not aggregated costs. The costs to one might not be avoidable without introducing worst costs to another. If so, imposing the costs on the one may be fair.

²² (i.e., impersonally: what else? Beyond engineering against? “Random” won’t be true of Procrastinate?)

will be fair even if it does not, in every case, serve the purpose of providing an incentive. Again, it is fair so long as it treats each symmetrically and adequately well.²³

Summarizing the case of the oversleeper:: The sanction is triggered by failing to turn in one's paper by the deadline, but it is possible to *trigger* the sanction even if one has made all the choices that the sanction was meant to *incentivize*.²⁴ The trigger is not the failed choices, but rather some typical consequence of them. This will, I think, be a feature of any sanction. Sometimes, when the trigger is tripped without the failed choices, we are able to make an exception, withholding the sanction even though the trigger was tripped (as, for example, when someone is hospitalized and fails to turn in their paper). But sometimes making an exception in such a case introduces too many other costs, as with oversleeping. In that case, the sanction may be fair even though it could not have been reasonably avoided.

Returning to our question: what makes a person's opportunity to avoid a sanction adequate for fairness? An opportunity is inadequate if a general exception can be made, without introducing prohibitive costs, on the basis of the fact that you have done all that could reasonably be expected of you. It is otherwise adequate. If you lack an adequate opportunity, then fairness requires an exception.²⁵ If a penalty is fairly enforced, then it follows that you had an adequate opportunity to avoid it. It is possible to have had an adequate opportunity and yet incur the penalty through bad luck—in such cases, the opportunity did you no good, so to speak. In these cases, we will accept the perhaps surprising conclusion that you had an *adequate* opportunity to avoid the penalty even though,

²³ Cf. Scanlon's reasoning in WWO, chp 6, and in the Tanner lecture about the toxic waste project and putting each in a good enough position.

²⁴ Thanks to John F. Horty for this helpful way of summarizing the position.

²⁵ Note: the idea of *adequate* opportunity includes what I earlier called the "quality" of your opportunity. Further, it is possible for the quality of your opportunity to be extremely poor, though bad luck (you could choose to avoid the penalty, but doing so would impose unreasonable costs on you), and it may be adequate: it may still be fair to enforce the penalty, because making an exception would be more costly. Thanks to Angela M. Smith for asking about such cases.

as things happened, you did everything we could reasonably expect you to have done and did not avoid it. The opportunity was adequate because the penalty remains fair.²⁶

Of course, in order for the system of sanction to treat each, including those who may incur it through bad luck, symmetrically and adequately well, the possibility of incurring the penalty through bad luck should be minimized and, moreover, the severity of the sanction must be no more than is required for the interests served by it. The interests of the student who will incur the penalty despite having done all that could reasonably be expected²⁷ must be weighed against the interests served by instituting and enforcing the sanction, resulting in (what we might call) the *minimally effective* sanction.²⁸ Even if we conclude that the unlucky oversleeper is fairly penalized, their interests in avoiding the penalty are not silenced, in the determination of the penalty, in the way that an interest in belittling others or an interest in special treatment are silenced. To the contrary, their interests count just as much as anyone's, in constraining the sanction.

[[I skipping a section in which I apply the same line of thinking to Professor Procrastinate, who, like most professors, regularly takes books out of the library. But Procrastinate, due to his inveterate procrastination, is sure to miss the deadline for returning those books. Perhaps Procrastinate lacks the ability to avoid the late fine; in any case, it is not reasonable to predict he will avoid it. Yet, it is typically fair to enforce the late fine on Procrastinate; if it is not, he is likely to loose his lending privileges..]]

²⁶ Notice, here, that the ethics is driving the metaphysics: whether there was an adequate opportunity (whether you “could have” in the relevant sense) is determined by whether the penalty would be fair.

²⁷ NTS: this is true of anyone's interests

²⁸ The possibility of false accusation likewise requires the minimally effective sanction. In some cases, the risks of bad luck or false accusation may result in abandoning the sanction altogether.

OBJECTION: TRULY FAIR VS. FAIR IN PRACTICE

One might grant that the late penalty for a paper might be fairly enforced even on the unlucky oversleeper. Still, one might want to discriminate between what one might like to call “truly fair” penalties and penalties that are, instead, only “fair in practice.” A late penalty is *truly* fair, one might say, only if the person has *not* done all that could be reasonably expected. If, instead, the person has done all that could be reasonably expected and yet incurs the penalty through bad luck, the penalty is only “fair in practice.”

I have no objection to marking the difference between those cases in which a person did all that could be reasonably expected, in the circumstances, and those in which the person did not—that is I have no objection to marking a difference between cases in which the sanction is and those in which it is not serving its original purpose of influencing choices. But I doubt *this* difference is well labeled as a difference between the “truly fair” and something else. If fairness is a matter of treating each person’s interests symmetrically and adequately well, then, given the constraints of the situation, the penalty is fair, full stop, in both cases. The fact that, in certain cases, the sanction does not serve the original purpose of providing an incentive does not seem to bear on its *fairness*, so understood.

One might think that the unlucky cases arise only because of unfortunate practical limitations and constraints. That is, one might think that the difference between “truly fair” and “fair in practice” is marking a difference between something like “ideal” and “non-ideal” cases: The penalty is fairly incurred through bad luck only because we cannot figure out a way to fine-tune our system so as to avoid it—only because we cannot, so to speak, perfectly align the trigger and the choices. If we were not subject to unfortunate practical constraints of implementation, then sanctions would be triggered by poor choices, and so would be fair only in cases in which one did not do all that could be reasonably expected.

While it is true that a system of sanctions should come as close as practicably possible to serving its core purposes, I doubt that, even in an ideal world, we would match the trigger of the sanction to

the planning incentivized and thereby avoid bad luck—because I doubt that even the ideal world would include as much agreement about planning and as much transparency of our thought and behavior to one another (or even to ourselves) as that would require.²⁹

One could idealize, not as I was imagining, by imagining that we know one another's thoughts or that we have agreement on exactly which choices (vs. outcomes) are reasonable to make (vs. to expect that we avoid), but rather simply by imagining that the sanction is somehow, magically, only enforced in those cases in which the person did all that could reasonably be expected, in the circumstances (where agreement on what was reasonable to expect is not needed, for the magic to do its work). If we had access to such magic (or, one might think, such divinely made and executed judgment), then enforcing the sanction in the cases of bad luck *would* be unfair—but it would be unfair because, with access to that magic, enforcing the sanction would serve no purpose: making the exception for such cases no longer carries the prohibitive costs. The magic (or God) has taken care of that. It seems that this may be what people have in mind when they say that sanctioning the unlucky is not “truly fair.” However, if so, it again raises the question of why it should be a matter of *fairness* that the trigger for the sanction should be the choice, rather than the outcome—given that, even without without magic, we are able to treat each one symmetrically and adequately well.

One might at this point (feel frustrated and) make another appeal to desert: if you could not have avoided the sanction by making a more reasonable choice, then the sanction is not truly fair because it is *undeserved*.

But once again, we are at risk of simply restating the intuition. We could now restate our question by asking why a sanction is deserved only if you could have avoided it through making a more reasonable choice. If desert is more than a way of marking when a sanction is justified, we would, again, want to know more about desert.

²⁹ CITE Angie. CITE Rob on incentives even in a world of angels.

Still, the intuition that an unlucky sanction is undeserved and/or unfair is a strong one. It needs some accounting.

OBJECTION: NOT A MORAL SANCTION

Someone might, at this point, raise an important objection: We have been misled because we have been considering the wrong sort of cases. Late penalties and library fines, understood in the way I have suggested, are not *moral* sanctions. Even if one were tempted to think that turning in a paper late or failing to return a library book were moral wrongs, two facts clearly reveal that the penalty as I have described it is not a moral one: the fact that the choices the sanctions are meant to incentivize and that which triggers the sanction can part company, as in the case of the oversleeper, and the fact that one can incur the penalty due to constitutional weaknesses you are unable to overcome, as in the case of Procrastinate. Given these facts, no one would think that incurring the penalty, *per se*, is a reason for guilt, for example. Understood as I have suggested, the objector may continue, the penalty has been stripped of any distinctively moral force or significance. The penalty, *per se*, need not come with any condemnation, opprobrium, or moral judgment.

To make the objection more stark, we can introduce a distinction between a *price* and a *moral sanction*. The objector might claim that systems of sanctions, as I have described them, are simply pricing systems, designed to achieve certain goals or to solve certain problems. For the partying student who misses the deadline, the late penalty is the price of partying. For the oversleeping student, incurring the penalty is, as we have said, simply a matter of bad luck—a kind of bad luck that was created when we created the pricing system, and the risk of incurring this bad luck is price of the pricing system, so to speak (a price that, by stipulation, is worth paying). [add Procrastinate] But *moral* sanctions, the objector insists, are not mere prices. They are a different beast. They cannot rightly be incurred through bad luck, and whether one had an adequate opportunity to avoid them is not sensitive to the competing needs and interests in the pragmatic way outlined above.

Accepting this for the moment, we can ask, what is a *moral sanction*, or what makes a sanction a *moral* one? And, why are such sanctions fair only if one has made a poor choice that one could reasonably be expected to have avoided?

As an answer to the first question, one might reply, flat-footedly but I think accurately enough, that a moral sanction is a sanction for a moral failing. The moral failing is the trigger. And, moreover, moral failings, one might say, just are immoral choices.³⁰ Moral sanctions are thus unfair, or undeserved, or unjustified, if one has not made a poor choice, simply because they are, in such a case *inaccurate*: one has not failed morally. One did not trip the trigger. So, absent our divine or magical ability to enforce the sanction only in cases of failed choices, we will have to distinguish when a *moral* sanction is truly fair—because tripped the trigger by making an immoral choice—and when, instead, it is only fair in practice—because we cannot, given our limitations, tell that you did not, without incurring prohibitive costs.

But there is a further complication: “Moral” sanctions differ from others only in that they carry a specific kind of *significance* or *force*. That is to say, more ordinary sanctions, such as monetary penalties or grade deductions, can *become* moral sanctions if they are incurred for a moral failing. When such penalties are incurred for a moral failing, one might think, they will then take on the distinctive significance of a moral sanction. Likewise, certain forms of treatment (such as avoiding someone, failing to be friendly with them, failing to go out of one’s way to help them, or explaining to them, in some detail, the way in which they failed) might become a form of moral sanction (become, e.g., moral reproach, condemnation, scolding, etc) when incurred for moral failing. In fact, one might think that even a mere negative evaluation or criticism (“you failed against this or that

³⁰ This is a controversial assumption. Many would regard failures of character, e.g., as moral failures. However, I am at the moment attempting to diagnose the intuition that a sanction incurred through bad luck is unfair. Those who think that failures of character are moral failures and yet think that sanctions incurred through bad luck is unfair may be thinking, for some reason, that *sanctionable* moral failings are immoral choices. That will be enough for present purposes.

evaluative standard”) can become, not a sanction, but rather moral *judgment* or *condemnation* when the subject of evaluation is a moral failure.

By putting together this special significance with the thought that moral sanctions are to be triggered by immoral choices, we can account for the intuitive difference between sanctions that are “truly fair” and those that are only fair “in practice:” Without an immoral choice the distinctive force or significance a moral sanction is unfair, even if the ordinary sanction (the fine or grade deduction) could still be fairly imposed, once stripped of the moral significance. Thus sanctions that carry moral significance must be “truly fair”—they must be incurred or immoral choices—while those that do not may nonetheless be fair in practice.³¹

This much seems both plausible and opaque. We can ask, next, what is this distinctive significance that an evaluation, sanction, or other response might take on, when but only when it is a response to moral failing? Why should it be reserved for choices? And, finally, why should it raise concerns about “desert” and determination?

SANCTIONS UNDER THE PRESUMPTION OF MORAL FREEDOM

I believe we can arrive at a useful hypotheses to answer these questions if we begin by assuming what many ethicists, including both incompatibilists and neo-Kantian compatibilists, have assumed: that insofar as we are subject to moral requirements we are thereby *also* at each point able to avoid violating those requirements, *no matter* our contingent psychology, history, or environment. As it is sometimes put, ought implies can. We have this degree of freedom: our history and environment

³¹ Thus, if you were thinking of turning in a paper late as somehow a [perhaps slight] moral failing, then the penalty would carry a different significance in the case of the partying student and the case of the conscientious oversleeper. That extra significance would not be fair, as applied to the conscientious oversleeper.

NTS: trickiness, here. “morally poor choice” is not “moral failing”, b/c moral wrongdoing can part with blame. So you may want to give up on “moral sanction”, as sanction for choice, altogether. CF. Julie.

will not prevent us from doing what is morally required. Call this *moral freedom* (short for “freedom to do what is morally required”).³² [cf. Kant and the gallows, 2C]

The assumption of moral freedom is extremely strong. If moral requirements require us not only to make the right choices, but also to bring about certain outcomes—if satisfying the moral requirements requires the cooperation of the world, as it is sometimes put—then it is hard to see how we could enjoy moral freedom. The world is often uncooperative, and this fact is obvious.³³ Thus, to make the claim of moral freedom at all plausible, we may need to shrink our moral duties, so to speak: we may need to understand moral requirements to govern only the internal thing, the will or practical reason or what not.³⁴ That is, we need to assume that we are required to make certain choices, and to do what we can to follow through with them, but that, when things go wrong despite our best intentions, we have done no moral wrong. Moral failing, then, is only a failing of willing.³⁵ Let this be so, for the purposes of our diagnosis.

But even this restriction to what is internal may not be enough. Given human weakness, given the fact that humans are not born able to satisfy their obligations and the fact that moral development and education is a messy and hazardous affair, it may be that some reach adulthood without the capacities required to make good choices. We may then have to shrink our duties even further: it may be that some of us are not in fact subject to the usual requirements, and so some of

³² I call this moral freedom, rather than the falsity of determination, because it is a different claim. It might, in principle, be true our thought and behavior is a product of our history and environment, and yet we each enjoy moral freedom—it could be the case that all histories and environments produce humans, or moral agents, who always choose to do the right thing (God might so arrange it, for example). In fact, moral freedom is consistent with the claim that we are all determined to do the right thing—because, as noted by Wolf, moral freedom is consistent with the claim that we cannot do wrong. But, since people in fact do wrong, we know that we are not determined to do the right thing. Thus it seems that determination is inconsistent with moral freedom, in the world as we know it. WORK ON THIS.

³³ CITE moral luck lit.

³⁴ The thing that, in the two-standpoints compatibilism, is in the standpoint that contrasts with explanation.

³⁵ CITE julie. [[One way to make the presumption of moral freedom more plausible is to allow the moral requirement to shrink to fit the abilities of the agent, so to speak.]] Also: some assumption that the inside is freer than the outside.

us not actually failing, morally, when we seem to violate them—our weakness and inability exempts some of us, some of the time, from the usual demands.

Opposite the presumption of moral freedom stands what we might call (somewhat sloppily) the presumption of *original sin*: the claim that we are unable to do what we are nonetheless morally required to do. At its strongest, the presumption of original sin claims that we are *never* able to avoid wrong choices—that our choices, or our motives, are always corrupt. A weaker version of the presumption would claim that, though we are *sometimes* able to choose well, at other times we can find ourselves simply unable to choose as morally requires. [denies Kant at the gallows]

Even the weaker (and I think very plausible) version of original sin is ruled out by the presumption of moral freedom. Again, on the presumption of moral freedom, each person is always able to do what is morally required, regardless of their history or environment. Ought implies can.

With the presumption of moral freedom in place, let us do something that is admittedly a little odd: Let us once again, as before, consider the fairness of sanctions (of any sort: emotional, monetary, corporeal), still understood as incentives, but this time as incentives to make morally required choices. Given the presumption of moral freedom, under what conditions would sanctions for moral violations be fair?

Note first that, under the presumption of moral freedom, although an individual is always able to choose against moral wrongdoing, it need not be easy.³⁶ Thus, a sanction can still serve as a useful incentive, helping to make the choice an easier one. By providing such an incentive, a moral sanction would help to protect the interests of potential victims. There is thus some reason to institute it.

³⁶ [why think a choice could be made difficult but nonetheless is always possible? Cf. “incline but not necessitate” literature.]

Notice, too, that, in stipulating that a choice is wrong, we have already determined that the costs, to the wrongdoer, of avoiding the choice—the burdens a potential wrongdoer will bear, in avoiding wrongdoing—do not justify making it. Burdensome though it may be, to avoid the wrongdoing, avoiding it is required. Thus, we already know that the “quality” of opportunity to avoid the wrong is adequate for an expectation that the wrong will be avoided. And thus we already know that a wrongdoer cannot appeal to the costs or difficulty of avoiding the wrong in order to claim that the prohibition or the sanction is unfair.

Finally, notice that, in presuming moral freedom, we have guaranteed that the interests that would be burdened by a sanction can always be protected by the wrongdoer, themselves, by choosing rightly. To put the point another way, under the presumption of moral freedom, no one can legitimately *expect* to do wrong. Correlatively, under that assumption, each of us can legitimately expect one another always to avoid wrongdoing. Thus, avoiding the harms of the sanction is something that any wrongdoer both *could* do (because of moral freedom) and *should* do (because doing otherwise is wrong). Of course, it may be *difficult* to do so—but that is why sanctions may be helpful. It helps to protect the legitimate interests of potential victims.

Because, under the presumption of moral freedom, none of us can, in good faith, expect to do wrong, the interests that might be burdened by a sanction against wrongdoing cannot legitimately enter the deliberations about the fairness of a sanction (as opposed to its decency or humaneness or proportionality). If someone were to try to enter those interests into deliberations about the fairness of a given sanction, they could rightly be rebuffed: given that each person always both can and should protect those interests, simply by avoiding wrongdoing, why should such interests be taken into account? Taking them into account seems incompatible with the assumption that fair rules or principles will treat each symmetrically and adequately well—to take them into account is to weigh the interests of the victim that would be protected by the sanction against interests that are at risk only if the wrongdoer chooses to do wrong—something that, by stipulation, the wrongdoer

both can and should avoid. Weighing those interests against one another seems incompatible with fairness—it would be, again, to take into account the interests of those who would like to do asymmetrically well, in comparison with others. But, as we saw above, such interests are incompatible with fairness and are therefore silenced. Thus it seems that the only interests that legitimately appear in deliberations about the fairness of a sanction against wrongdoing are those of the victim—those that might be protected by the sanction. Under the presumption of moral freedom, the interests of the wrongdoer are, so to speak, silenced by their freedom.³⁷

One might say that the presumption of moral freedom generates the “you made your bed” justification for sanctions: While the “you made your bed” justification would be rebutted by “I couldn’t have done otherwise” or “I couldn’t help it,” moral freedom ensures that you could have.

Importantly, in saying that the wrongdoer’s interests are *in this way* silenced—silenced in determinations of fairness, forfeited by the you-made-your-bed reply—we need not grant that they are silenced across the board, so to speak. Even given the assumption of moral freedom, we can nonetheless insist that the interests of the wrongdoer should still weigh with us: Simple human decency requires us to consider the interests, even of the wrongdoer. Certain sorts of sanction will be cruel or inhumane, even given moral freedom. Alternatively, we could advance some sort of principle of proportionality as a bedrock moral principle. Or, we might point out that especially severe sanctions are likely to lead to unrest or social strife. Perhaps most importantly, we must remember that we will make mistakes in our moral evaluations—we are likely to make false accusations. Given the possibility of false accusations, we should, again, look for the minimally effective sanction.³⁸ And, in practice, that will be the sanction that even the wrongdoer will face.

³⁷ Thus, the presumption of moral freedom would provides a contractualist rationale for what Scanlon calls “The Forfeiture View:” by choosing wrongdoing, a morally free individual forfeits their claims against treatment that burdens them.

³⁸ Though, again, the trouble with false accusation is a problem with implementation. It would not come into play, if we were omniscient with respect to the choices that people make. [Say something about Kant’s empirical realism.]

Still, when considering the possibility of implementing sanctions for morally wrong choices, there is one extremely important thing one cannot say, under the presumption of moral freedom: One cannot appeal to the interests of the wrongdoer that would be harmed by the sanction in order to claim that the sanction is *unfair*—that a sanction fails to treat the wrongdoer adequately well. Under the presumption, the wrongdoers’ interests always can be and always should be protected by the wrongdoer, themselves, by avoiding the wrong.

This is a significant change. On the presumption of moral freedom, the wrongdoer, in doing wrong, thereby earns a *status* from which their interests, *qua wrongdoer*, no longer weigh in determinations of fairness. They have, in choosing to do wrong, alienated or banished themselves, so to speak, from the moral community or from the pool of moral concern. The wrongdoer, *qua wrongdoer*, is an outsider, and, as such, incurs a vulnerability to sanction that is unlimited by concerns of fairness grounded in their own interests.³⁹ Thus, insofar as you have, in fact, done wrong, you are, in a certain way, defenseless—you are without a justification, grounded in your own interests, by which to object against the sanction as unfair.⁴⁰ Even though we can, and should, continue to take your interests into account, as a matter of decency or humanity, *you* cannot insist on it or complain about it, as a matter of fairness. If you ask others to do unto you, they can fairly reply that you made your bed. You cannot ask for justice, but only mercy.

This status would, I think, provide a distinctive force or significance to sanctions imposed for moral wrongdoing, in particular. In fact, it seems to me plausible that this vulnerability, this alienation from the pool of moral concern, provides the force one might have in mind when talking about “moral condemnation.” I will, hereafter, use the term *condemnation* as a label to refer to this particular status: to be condemned is to have your interests silenced by your own free choices. In

³⁹ The claim is not that wrongdoing silences one’s interests across the board, but only those interests burdened by the sanction.

⁴⁰ CITE JDV on “normative vulnerability”

fact, one might understand the claim that something is “deserved,” in the sense of “basic desert” to be the claim that one’s claims against it have been silenced by one’s choices under moral freedom. This seems to me a satisfying interpretation of what people seem to have in mind.

This sting or force of moral condemnation would, of course, be unfair without the the presumption of moral freedom. Without that presumption, we cannot be confident that each person will always be able to protect their own interests against moral sanction. Rather, we may find ourselves in the position of Procrastinate, trapped by our own constitution, expecting ourselves to make poor choices. Without the assumption of moral freedom, it is therefore legitimate to plan to do wrong, so to speak (as we saw, in Chapter Two, that Procrastinate must do). To put the point differently, without moral freedom, acknowledging that one might, regrettably, do wrong is not incompatible with a commitment to fairness and so not incompatible with membership in the moral community. And thus follows that, without moral freedom, the fact that a person chose to unfairly burden the interests of another—the fact that someone chose wrongly—will not silence the wrongdoer’s interests against sanction. Rather, they will continue to weigh the same as anyone’s.

The thought is sometimes expressed in the phrase, “There but for the grace of God go I.” It is a thought made appropriate by (the possibility of) original sin. Under the assumption of original sin, the wrongdoer, like Procrastinate and like the unlucky oversleeper, may still be fairly sanctioned. But they may not be condemned. And thus, if *condemnation* is what provides the distinctive force or sting of specifically *moral* sanction, and if we do not we enjoy moral freedom, then the distinctive, condemnatory force of moral sanction is unfair.

There are, I think, two strong sources of resistance to giving up “basic desert” or condemnation. The first is an extremely deep rooted resistance to the idea that we cannot control the bad things that happen to us. The freewill defense to the problem of evil is, to many people, extremely attractive: it seems better, to many, to have my suffering be my fault than to have it be simply tragic

or brute. If my suffering is my fault—if I made this bed—then there is some hope that I control it or that others could avoid my fate. Thus, giving up on condemnation might seem to require accepting something like the problem of evil. I will not pursue this.

The second source of resistance is, I think, less deeply rooted. It is the thought that, if we give up condemnation, then we are giving up what is distinctive about responses to moral failures—we are moving to a mere pricing system. I disagree.

In the chapter that follows this one, I present to a different way to understand the distinctive significance of moral failure and responses thereto. The distinctive significance of moral failure is not that of condemnation, nor even that of meriting a sanction or penalty. It is, instead, the significance or importance of standing in certain relationships with other people, relationships of regard or respect. Matter to others, not meriting negative consequences, is what morality, and moral responsibility, is about.

CONCLUSION

To review: We pursued Watson's suggestion that the responses to which we are subject, when morally responsible—demands, sanctions, and the like—raise concerns of fairness, and that this concern for fairness is the source of the concern about ability, opportunity, or possibility.

We saw that, although the fairness of a sanction does require an adequate opportunity to avoid it, what counts as an adequate opportunity to avoid will depend on how best to adjudicate the competing interests in the circumstances. It may be that someone could fairly incur a sanction even if they did all they could be reasonably expected to do, to avoid it (as with the oversleeping student) or even if they could not help but incur it (as with Procrastinate).

We then considered an important objection: so understood, sanctions do not carry any distinctively moral force or significance, and it is this distinctive force or significance that requires the falsity of determination.

It is difficult to say exactly what this force or significance is. I offered a hypothesis: if you first presume that we enjoy what I called *moral freedom* (which, plausibly, we do not enjoy if determination is true) and then consider the fairness of moral sanctions, you will find that, under the presumption of moral freedom, the severity of a moral sanction will be unlimited by concerns about the interests of the wrongdoer. Limits will appear only due to concerns about false accusations or concerns other than fairness (e.g., decency). I suggested that to be vulnerable to sanctions unlimited by concern for one's own interests, to have one's own interests in this way silenced, is a form of *condemnation*. Further, sanctions could carry such condemnatory significance or force even if they are limited, for other reasons, in their severity.

If moral sanctions are those that carry this condemnatory force, then they will be unfair unless we enjoy moral freedom—and so, it seems, they will be unfair if determination is true.

But this does not show that fair sanctions for moral wrongs *must* be incompatible with determination—it only shows that such unlimited, or *condemnatory*, moral sanctions are. We could give up the condemnation.

Some people will worry that, by giving up condemnation, we are giving up what is distinctive about moral responsibility. In the next chapter I will suggest that this is not so; we can give up condemnation without giving up the distinctive significance or force of moral failing and responses thereto. We can locate a significance that is more than dispraise, less than condemnation, and different than pricing—different, even, than sanction. Then in Chapter Nine, we will consider, in detail, various attitudinal or emotional responses to moral failing, before turning, in the final chapters, to the question of moral standards.