

FAIRNESS, SANCTION, AND CONDEMNATION

—draft—

Pamela Hieronymi

hieronymi@ucla.edu

July 22, 2020

I have long been puzzled about what philosophers have in mind when they talk about “basic desert,” “true moral responsibility,” or the “condemnatory force” of moral criticism.¹ I have been particularly puzzled by its presumed relation to some strong requirement of freedom. The presumption is that, if we are not “free,” in some very strong sense, then we are not truly morally responsible and so do not deserve condemnation.

I believe I have made some progress on this puzzlement, in an unexpected way. Although it is often noted that the fairness of a sanction requires an adequate opportunity to avoid it,² it is rarely asked why—why does the fairness of a sanction require an opportunity to avoid? By pressing, hard, on this question, I believe I have come to better understand why someone might think that “freely” choosing to do wrong will, itself, have an extraordinarily strong and immediate effect on the way in which the interests of the wrongdoer should matter to the rest of us. That is, I believe I have come to understand the relation between a strong form of freedom and a status that might be called “condemnation” or a kind of desert that might be called “basic.” I here present what I take myself to have learned.³

¹ Galen Strawson argues that “true moral responsibility,” the sort that would “make sense” of heaven and hell, would require self-causation (Strawson 1994). Derk Pereboom has argued that “basic desert” requires the falsity of determinism, while saying little about what “basic desert” is; if I capture it here, then doing without it will require less revision than Pereboom imagines (Pereboom 2014). Jay Wallace has criticized T. M. Scanlon’s conception of blame as lacking in “opprobrium” (Wallace 2011), Susan Wolf often talks of a special “force” of moral blame, connecting it to freedom (Wolf 1990); I am less confident about what Wallace or Wolf have in mind, but I suspect it is at least partially infected with what I here call “condemnation.”

² A classic statement occurs at (Watson 1996, 237).

³ Although every paper I have written has benefited tremendously from feedback, never has a paper been so transformed as this one. It has been extremely gratifying. I acknowledge many debts in the final footnote. I here thank Stephen White, T. M. Scanlon, Dana Nelkin, and Manuel Vargas for helping me to better locate my target and Sarah Stroud for asking for greater clarity.

Importantly, in presenting this understanding of “condemnation” or of “basic desert,” I am presenting an idea I do not endorse—in fact, one to which I am opposed. I present it hoping that, if what I say accurately captures what people have in mind, it will also show how condemnation or desert, in this sense, can be left behind.

SANCTIONS

To begin, let us examine more carefully the idea of sanction.⁴ To sanction someone is to impose something unpleasant or unwanted upon that person in response to his or her violation of some norm, demand, or expectation. That is to say, sanctions are consequences that are *created* and *attached* to certain failures. Parents sanction 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.

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.⁵ Sanctions are not simply negative consequences no one else is required to bear for you. They are negative consequences that have been created and attached, by some person or body, to certain violations or shortcomings. As such, they raise concerns of fairness.

TWO QUESTIONS

Now, two questions: First, *why* do sanctions, so understood, raise concerns of fairness? Second, why should the *fairness* of a sanction require an *opportunity to avoid* it?

Some would answer both questions with a quick appeal to desert: Sanctions require opportunity to avoid because they are otherwise *undeserved*, and, if undeserved, unfair.

⁴ “To sanction” can mean “to penalize” or “to endorse.” Throughout, I mean “penalize.” (I prefer “sanction,” for its distance from “punishment.”)

⁵ The mess is what Scanlon would call a “substantive responsibility” (Scanlon 1998, ch 6). I am, here, distinguishing sanctions from other (mere) substantive responsibilities, in a way that, to my knowledge, Scanlon does not.

I find this appeal to desert unhelpful. To see why, consider that it can be interpreted in two ways, which I will label “deflationary” and “substantive.”

On the first, deflationary, reading, the question of whether you “deserve” this or that reduces to, or is transparent to, the question of whether giving it to you or imposing it upon you would be apt, justified, or fair. Thus, saying that, without adequate opportunity to avoid, sanctions are undeserved is just a different way of saying that, without such opportunity, they are inapt, unjustified, or unfair. Thus, the deflationary interpretation leaves our questions untouched.

Even so, “undeserved” may be a useful label to introduce: It may pick out the specific form of justification at work in cases of penalties, prizes, punishments, and the like. I have no objection to this. However, we must bear in mind that, until we provide that underlying justification, our questions remain untouched—saying that fair sanctions require adequate opportunity to avoid because, without it, they are undeserved is like saying the sleeping pill works because of its dormative virtue.

Alternatively, one might think facts about desert can themselves provide *independent* reason to think sanctions are fair. Call this an appeal to “substantive desert.” On this view, we answer our original questions by appeal to desert, only immediately to face the further questions of why undeserved sanctions are unfair and why, without adequate opportunity to avoid, sanctions are undeserved. These further questions are difficult to answer.

One might take the difficulty of answering these further 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 opportunity to avoid is like asking why pain is bad: the claim is not empty (as on the deflationary reading), but it is not a question that needs an answer.

I find the notion of substantive desert opaque, and therefore I find the appeal to it as basic or bedrock unsatisfying. It seems we can reasonably 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 that our answers (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*) fair. (In contrast, answers to “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 answer, we no longer need the notion of substantive desert—we have returned to the deflationary interpretation.⁶

Those who disagree might simply consider what follows as alternative answers given by someone who is not content to take desert as bedrock.

Sanctions and Fairness

Here, then, 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 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.”⁷ (I take this to be about as bedrock a moral principle as you will find; I am content to treat it as primitive.)

Importantly, an appeal to “do unto others...” is not necessarily an appeal to fairness. On its face, it is an appeal to empathy. As noted by Kant, it can be used to ask for special treatment. In his example, a convicted criminal appeals to it to ask for mercy—for special treatment—from the sentencing judge.⁸ To understand “do unto others...” as counseling *fairness*, it must be understood as generating principles that apply uniformly across the population. So understood, it directs us to

⁶ One might yet disagree, thinking that desert is something like a “thick” ethical concept, something like courage or kindness (Williams 1985).

⁷ Or, at least, “Do not do to others what you would not have done to you.” In contrast, non-voluntary attitudes, such as distrust, can burden others without raising the same concerns of fairness. See (Hieronymi 2004, in progress).

An anonymous reviewer notes that unforeseen consequences and the results of negligence can also trigger “do unto others...”, though they are not voluntary. True, but this is so, I would argue, because one can reasonably be expected to avoid them through voluntary actions (unlike *unforeseeable* consequences). Nonetheless, my claim is only that voluntariness is sufficient to trigger concerns of fairness, not that it is necessary. See (Hieronymi 2004).

⁸ (Kant 1998, footnote at 4:430).

treat each person as best we can, consistent with treating others in the same way. That is, asks 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 that treat each symmetrically while doing adequately well by all. I will refer to these as principles of *fairness*.⁹

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 whether a given policy or principle treats the interests of each symmetrically and adequately well.

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. To propose that such interests be accommodated is to propose treating some asymmetrically poorly. Such a proposal is incompatible with treating others symmetrically and adequately well. And, therefore, these interests need not—in fact, *cannot*—be counted among those that must be symmetrically balanced and adequately served, in a determination of fairness. Such interests are, we might say, *silenced*.¹⁰ This idea of silencing will be important later.

In answer, then, to the first question: the institution or imposition of a sanction raises concerns of fairness because it is a voluntary action that burdens others. A sanction, or a system of sanctions, will be fair if but only if it treats the interests of each symmetrically while doing adequately well by all. I will take this much as read and turn to the second question.

⁹ This is a broadly contractualist interpretation of the everyday notion of fairness. I believe such an interpretation is natural—the success of my diagnosis depends on it being so.

¹⁰ These interests are “unreasonable” in Scanlon’s terminology: incompatible with a commitment to finding reasonable principles. I draw the term “silenced” from (McDowell 1979).

Fairness and Opportunity

Our second question is, why would fairness require that sanctions be imposed only if there was adequate opportunity to avoid them? What is the connection between *treating each symmetrically and adequately well* and *ensuring adequate opportunity to avoid sanction*?¹¹

To start, we can note a general presumption against imposing burdens on others. That presumption must be overcome if a sanction is to be justified. It is typically overcome by some important purpose served by the sanction: typically, a sanction provides an incentive or deterrent—that is, it influences peoples’ choices—in order to achieve certain goals or to protect certain interests.¹² For example, instructors institute late penalties to provide students an incentive to submit their work by 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. (Importantly, the fact that a system of sanction serves a purpose does not the justification of sanctioning, in a particular case, “forward looking” or “consequentialist.” Once instituted, a system of sanctions is fair only if it is uniformly enforced. Enforcement is always “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 that is faced by each individual. It would be unfair, e.g., to completely impoverish those who turn in work late, even if that sanction were symmetrically enforced, because the risk, to each, of incurring such a sanction (though, say, bad luck or false accusation) far outweighs its benefits. It would treat each symmetrically, but not adequately well.

¹¹ Scanlon explores the broader question of why we have for wanting what happens to depend on our choices in (1998, chp 6). I here focus, narrowly, on sanction.

¹² In other case, it serves to manage risk—these cases of “strict liability” are insensitive to opportunities to avoid.

Notice—it will be important later—that the disproportionately severe sanction is ruled out, not simply because it is, in some simply brute or intuitive way, “disproportionate,” but specifically because, given the severity of the sanction, the risk of incurring it outweighs its benefit, and therefore it fails to treat each adequately well.

Of course, in the absence of any opportunity to avoid a sanction, imposing it could not provide the intended incentive or deterrent—without an opportunity to avoid it, there is nothing for the sanction to influence. Thus, if we were to institute a system of sanctions that remains insensitive to one’s opportunity to avoid it, we would put people at risk of incurring pointless harm. It would be better if our system made exceptions for cases in which there was no opportunity to avoid. In fact, such exceptions seem required by fairness.

Here, then, is a suggested answer to the second question: Fair sanctions require adequate opportunity to avoid, because, without such opportunity, imposing the sanction will burden interests for no purpose. Doing so fails to treat the interests of each symmetrically and adequately well.

A THIRD QUESTION: THE ADEQUACY OF OPPORTUNITY

We now face a third question, one that will ultimately bring us to the issue of freedom. It is this:

What makes a person’s opportunity to avoid sanction *adequate* for fairness?

Here is an initial, natural hypothesis: A student has 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.

This natural hypothesis can be supported in two very different ways. The first repeats the above: if the student did not faces choices by which they could avoid the penalty, there was nothing for the penalty to incentivize. Supported in this way, choice is *necessary* for fairness. The alternative appeals to control and to something like desert, reasoning that, if the student faced a choice, it was *up to the student* whether they incurred the penalty or not.¹³ Because the penalty was, in this way, in the

¹³ Thanks to Michael McKenna for highlighting the connection between desert and control.

student's control, they they cannot complain the burden is unfair. They made their bed, so they can lie in it. On this way of thinking, choice is *sufficient* for fair penalty.

The natural hypothesis is incorrect: The fact that someone faces a choice by which they would avoid the sanction is neither necessary nor sufficient to ensure that they had 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 may knowingly choose to incur the penalty, and could have chosen, instead, to complete the assignment and avoid the penalty (in however strong a sense of “could have” one likes), the penalty may nonetheless be unfair—and, unfair because the student did not have adequate opportunity to avoid it. Whether the student had *adequate* opportunity to avoid turns, in part, on the *quality* of the opportunity, so to speak—i.e., on the other burdens that one incurs in order to avoid it. Those other burdens may be unfair. If so, the fact that you choose which burden you incur does not render the sanction fair. Choice is not sufficient.

More surprisingly, choice is also not necessary for adequate opportunity to avoid. Consider a student who simply forgot about the assignment or who overslept and missed the deadline. There was no choice the student faced by which they would avoid the sanction. Yet, depending on the circumstances, the student may have had adequate opportunity to avoid it.

This surprising conclusion requires investigation. Let us focus on the student who overslept. They may try to advance the above argument: Enforcing the late penalty on oversleepers does not provide an effective incentive—those who oversleep cannot turn back time to correct their error. Thus, the sanction burdens them for no purpose; it is therefore unfair.

This line of reasoning can be challenged in two ways.¹⁴ First, one might argue that, though no one can turn back time, enforcing the late penalty on the oversleeper nonetheless provides an incentive: it creates an incentive for students to take steps to ensure they do not oversleep or to turn their work in a bit in advance, to guard against unforeseen difficulties.

In effect, this first challenge 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.

One might object: although oversleeping can *sometimes* be avoided through better planning, sometimes it cannot. Even the most conscientious of students, with excellent time-management habits, a reliable alarm clock, and a plan to submit work in advance, can find everything goes wrong and, really, there was nothing they could have done to avoid missing the deadline. Even the most conscientious sometimes fail, through sheer bad luck. Enforcing the penalty in such cases, one might object, is unfair.

In reply, notice, first, that when we say that, really, there was nothing the fully conscientious student could have done, in contrast to student who could have better managed their time, we are not marking a difference in the metaphysical freedom of these two students. When we say the fully conscientious student “could not have avoided” the penalty, we are not ruling out all counter-factual possibility of avoiding it. Strictly speaking, there were things the fully conscientious student might have done: they might have set a second or third alarm clock, for example. In fact, we could stipulate that determinism is false, that the student enjoys libertarian freedom, and that the student owns three working alarm clocks—even so, the fully conscientious oversleeper could still complain that there was nothing they could have done to avoid the penalty. In saying so, they are saying they *did everything that could reasonably be expected of them* and yet failed through bad luck. To say they failed

¹⁴ A third challenge: One might think late penalties function, not only as incentives, but also to level the playing field. Students who submit work late likely benefit from the extra time, and so their grade should be handicapped, so to speak. I set this aside and focus only on the late penalty as an incentive.

through “sheer bad luck” is to say they did everything that could be reasonably expected of them and yet, surprisingly, they did not thereby avoid the failure.¹⁵

In such a case, one might think, the late penalty serves no purpose and is, therefore, unfair. We should carve out an exception for the *conscientious* oversleeper—though perhaps not for the typical oversleeper.

We now encounter the second challenge: Sometimes making exceptions will generate costs greater than the cost of the sanction. If so, then choosing instead to enforce the sanction may yet treat each symmetrically and adequately well—it may be fair—even though the one on whom it is enforced did all that could reasonably be expected and failed through sheer bad luck. We might say that enforcing, in such cases, serves the *secondary* purpose of avoiding the costs of the exception.

To explain: The institution and administration of a sanction, in practice, will implicate interests beyond those most directly served by it. These will be highly context-dependent. For example, interests of both the instructor and students require that enforcing the sanction not involve a significant expenditure of the instructor’s time nor a significant invasion of student privacy. The interests of students require that the policy cannot be too easily “gamed.” These further interests will have different implications in a very large class 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 the less-than-fully-conscientious student is neither easy nor costless. Unless the instructor both is and is known to be, not only 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 a student is fully conscientious, or honest, by appeal to their “sense” of the students, without risking serious unfairness. There is, on

¹⁵ Thanks to Julie Tannenbaum for pointing out the possibility of a third alarm clock. For related discussion see (Hieronymi 2007).

the one hand, the worry that the instructor will show favoritism to those with whom they feel some affinity, and, on the other hand, the worry that students will be able to deceive the otherwise the judicious instructor. One might guard against these risks by requiring some sort of investigation into either the general habits or the fateful morning of the oversleeping student—but that will likely require 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 for other kinds of bad luck, 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, we *might* conclude that making an exception for the fully conscientious will create costs that are higher than the costs of late penalty. We might then conclude that the system that serves each students’ interests most fully and fairly will enforce the late penalty even in certain cases in which, it is admitted, the student made no poor choices and simply suffers from bad luck. The bad luck, we might think, will be acceptably distributed, and the costs of rooting it out of the system high, and so, we 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. Enforcing the penalty is fair, because it treats each symmetrically and adequately well.¹⁶

Summarizing the case of the conscientious oversleeper: The sanction is *triggered* by failing to meet the deadline, but it is possible to trigger the sanction even though one made the choices the sanction was designed to incentivize.¹⁷ The trigger is not the failed choices, but rather a typical consequence of them. In some of the cases in which the trigger was tripped through bad luck, an

¹⁶ Cf. Scanlon’s reasoning in (Scanlon 1998, ch 6) about the toxic waste project.

¹⁷ Thanks to John F. Horty for this helpful way of summarizing the position.

exception can be fairly made (as, for example, when someone is hospitalized and therefore fails to turn in their paper). Fairness then requires it. But sometimes making an exception introduces too many other costs, as with oversleeping. In such cases, the sanction may be fairly enforced even though no poor choices were made. Surprisingly, choice is not necessary for fairness.

This conclusion is controversial. Before considering the controversy, let us return to our third question: What makes a person's opportunity to avoid sanction *adequate* for fairness? We now face what turns out to be a purely notational choice about how to use "adequate:" One might want an "adequate" opportunity to ensure that, so long as you avoid poor choices, you also avoid the penalty. That is, one might want bad luck to render one's opportunity inadequate. If so, cases like the conscientious oversleeper show that an adequate opportunity is not required for fairness. Alternatively, one might insist that fairness requires adequate opportunity—but then an adequate opportunity will not rule out bad luck.

The first option (adequate opportunity rules out bad luck) fits nicely with intuitions about desert. But it carries two high costs. First, it will be extremely difficult, likely impossible, to determine when someone who tripped the trigger had an adequate opportunity to avoid doing so. We will have to determine whether the failure was due to poor choices; that will require specifying what could be reasonably expected at each decision-point that led to the failure. While some things will be obvious (do not stay up late without setting an alarm), there will be significant, yet reasonable, disagreement about many details (should you set a second alarm the night before every important deadline? how important?)—we may reasonably disagree about these details even while agreeing the deadline is reasonable. Worse, a person might fail in the details (fail to set the second alarm, e.g.) due to bad luck. And then, to determine whether they had an adequate opportunity to turn in the paper, we will need to determine whether they had an adequate opportunity to set the second alarm. The

problem will repeat, at that point. Thus, even if we were able to agree on the details, the task is still unworkably large.

A second cost might offset the first: though it may be impossible to determine whether someone had adequate opportunity, there may be no need to do so—because, if we insist that bad luck renders opportunity inadequate, then adequate opportunity is not required for fairness.

I prefer the second option: Insist that fairness requires an adequate opportunity, but allow that an adequate opportunity is no guarantee against bad luck. We will then determine whether an opportunity was adequate by determining whether enforcing the sanction would be unfair. If enforcing the sanction would be unfair because of bad luck (as when hospitalized or in a car accident), then the opportunity was inadequate.¹⁸ However, if bad luck does not undermine the fairness of the penalty, then it also does not undermine the adequacy of the opportunity (as in the case of the conscientious oversleeper).

A benefit of the second option: Not only can we grant, without arguing the details, that the conscientious oversleeper did everything we could reasonably expect to avoid missing the deadline, but we can also now add “short of meeting the deadline.”¹⁹ That is, we can maintain the deadline sets a reasonable expectation while nonetheless granting that it is sometimes possible to miss it despite making all the right choices.²⁰ The deadline, itself, remains reasonable for the same complex set of reasons that show the penalty remains fair.²¹

¹⁸ Note that *adequacy* of opportunity includes what I earlier called the “quality” of your opportunity. And, the quality of your opportunity may be extremely poor, though bad luck, and yet 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.

¹⁹ Thanks to Randolph Clarke for pressing on this point

²⁰ We can be confident that such cases will be rare: if bad luck were common, the risk of incurring sanction will outweigh the benefit, and the sanction will be unfair. See footnote XX.

²¹ The ethics drives 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. (Note the similarity to Strawson’s optimist: we are defining something that sounds like freedom by appeal to something more pragmatic. The optimist appeals to efficacy. I have, instead, appealed to fairness.)

This use of “adequate” will be controversial, but, again, it makes only a notational difference to the argument below. The real source of controversy is the claim that a sanction may be fairly enforced when the trigger was tripped due to bad luck. We consider objections in the next sections.

Two final observations: First, in order for the system of sanctions to treat each, including those who may incur it through bad luck, symmetrically and adequately well, the possibility of incurring it through bad luck must be minimized. The risk of incurring the sanction through bad luck must be weighed against the interests served by it, resulting in (what we might call) the minimally effective sanction.²² Therefore, the interests of the unlucky 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.

Second, notice that these same conclusions about fairness might be reached when the inability to avoid the sanction is due to constitutional weakness in the individual. Consider, for example, Professor Procrastinate. Like most professors, Procrastinate regularly takes books from the library. But, due to his inveterate procrastination, Procrastinate is sure to miss the deadline for returning them. Let us stipulate that Procrastinate lacks the ability to avoid the late fine. Is enforcing the late fine therefore unfair? Plausibly, no. Plausibly (details must be filled in²³), Procrastinate would instead, due to his constitutional weakness, fairly lose his lending privileges. (He may need to work in the library.)

We have now answered our three questions: Sanctions trigger concerns of fairness because they are created and imposed through voluntary actions that burden the interests of others. They require adequate opportunity to avoid because, without such opportunity, they burden interests for no

²² 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.

²³ Details include the importance and relative abundance of the good (books) being shared; the cost of providing that good, free and clear, to anyone who needs it; the prevalence of Procrastinate’s condition; social stigma or disadvantage associated with that condition, etc.

purpose and thereby fail to treat each symmetrically and adequately well. However, in *some* cases in which a person trips the trigger of a sanction despite doing all that could be reasonably expected to avoid it, enforcing the sanction may still be fair—it may still treat each symmetrically and adequately well (because it serves some secondary purpose). Thus, if adequate opportunity to avoid sanction is required for fairness, it is possible to have had adequate opportunity to avoid it even if you did all that could reasonably be expected to avoid it, as with the oversleeping student, and, it seems, even if you suffer from some constitutional weakness that prevents you from avoiding it, as with Procrastinate. Adequate opportunity does not rule out bad luck. (Or, if it does, it is not required for fairness.)

This final, surprising conclusion transgresses ordinary intuitions about the fairness of sanctions—intuitions that one might express using “desert.” One might, reluctantly, grant that a late penalty could be fairly enforced on the unlucky, for the reasons given above. But, one might feel that the unlucky do not really *deserve* the penalty—precisely because they did all that could be reasonably expected. We now turn to this objection.

OBJECTION: TRULY FAIR VS. FAIR IN PRACTICE

The objector grants that there is *some sense* in which a late penalty might be fairly enforced on the unlucky, but hopes to mark a distinction between what we might call the “truly fair” and the merely “fair in practice.” A late penalty is *truly* fair, the objector might say, only if the person has *not* done all that could be reasonably expected, short of tripping the trigger. If, instead, the person trips the trigger 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 or did not do all that could be reasonably expected—cases in which the sanction is or is not serving its original, rather than secondary, purpose. 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 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 bear on its fairness, so understood.

One might suggest that the unlucky cases arise only because of unfortunate practical limitations—because we cannot perfectly align the trigger with poor choice. In a perfect world, the trigger would be failure-due-to-poor-choice. Then, without a poor choice, the sanction would be unfair because, so to speak, *inaccurate*: the trigger was never tripped.

I see the temptation, here, but I doubt that, even in the ideal world, we would want to identify the trigger with poor choice—because I doubt that even the ideal world would include as much agreement about the details of planning and as much transparency of our thought and behavior to one another (or to ourselves) as that would require. There are good reasons to let the trigger be an outcome, simply.

One might instead idealize, not by imagining agreement or transparency amongst ourselves, but rather by imagining the sanction is somehow magically—or, perhaps better, divinely and omnisciently—enforced only when the person did not do all that could reasonably be expected.²⁴ Then the trigger would be failure-due-to-poor-choice, and in unlucky cases the sanction would be unfair because inaccurate. Thus, in the magically or divinely ordered world, one could always avoid fair sanction by making good choices.

This is, indeed, the ordinary intuition about desert—that truly deserved sanctions require poor choices. But we should again wonder: why should “true” fairness require omniscience or magic? That is, why should be a matter of *fairness* that the trigger for a sanction should be a poor choice, rather than some outcome—given that, short of agreement, transparency, omniscience, or magic, we are still able to treat each one symmetrically and adequately well?

One might answer: because we *ought*, as a matter of fairness, to be able to control whether bad things happen to us by making good choices.

²⁴ Thanks to Angela M. Smith for pressing this possibility on behalf of the objector.

This is a powerful intuition. I am very sorry to say, it seems, to me, a fantasy. It is an understandable fantasy, and, in certain ways, a noble one. But a fantasy nonetheless. Life is not, in this way, fair. Nor could we, in our interactions or social institutions, make it so—for the reasons given above (about agreement and transparency). That you could not have avoided a burden by choosing well does not *entail* that the burden is unfair—though, if the burden is a sanction, it makes it likely to be unfair, for the reasons above: sanctions should typically be avoidable.

Still, the intuition that an unlucky sanction is *always* undeserved and/or unfair remains strong. It needs further accounting.

OBJECTION: NOT A MORAL SANCTION

Someone might, at this point, raise the objection in a new guise: We have been misled by considering the wrong cases. Late penalties and library fines, understood as I have suggested, are not *moral* sanctions—as clearly shown by the two facts I highlighted: that the choices and the sanction can part company, as in the case of the oversleeper, and that one can incur the penalty due to constitutional weaknesses, as in the case of Procrastinate. Given these facts, no one would think incurring the penalty, *per se*, is a reason for guilt, condemnation, opprobrium, or moral judgment. The penalty has been stripped of any distinctively moral force or significance. That distinctive moral significance is what raises special concerns about fairness, luck, and opportunity.

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 describe them, are simply pricing systems, designed to achieve certain goals or solve certain problems. For the partying student, the late penalty is the price of partying. For the conscientious oversleeper, the penalty is bad luck—a kind of bad luck we created when we created the pricing system. The risk of bad luck, of this sort, it is price of the pricing system, so to speak. But *moral* sanctions are a different beast. Their

distinctive significance raises special requirements of fairness; their justification is not sensitive to secondary purposes in the pragmatic way outlined above.²⁵

What then, is a moral sanction? One might reply, flat-floodedly but accurately enough, it is a sanction for a moral failing. Further, moral sanctions differ from others only in their special significance. Thus, ordinary sanctions, such as monetary penalties, grade deductions, or restrictions of movement, can *become* moral sanctions, if incurred for a moral failing. They then, somehow, take on a distinctive significance. Likewise, certain forms of treatment (avoiding someone; explaining, in some detail, how they failed) might become moral sanction (condemnation, reproach, etc.) when incurred for moral failing. Moral sanctions do not differ from others “materially,” so to speak, but only in their special, moral significance.

Why, then, does this moral significance require a poor choice? As an initial answer, we can note that many simply *identify* moral failings with immoral choices.²⁶ It then follows that moral sanctions are unfair without immoral choice, because inaccurate: one did not trip the trigger.

By putting these thoughts together, we can account for the difference between “truly fair” sanctions and those that are only fair “in practice:” Without an immoral choice, the distinctive significance a *moral* sanction is unfair, because inaccurate. This may be so, even if the ordinary sanction (the fine or grade deduction) could still be fairly imposed, once stripped of the moral significance. Thus, sanctions carrying moral significance must be “truly fair”—incurred for immoral choices—while sanctions stripped of that significance may nonetheless be “fair in practice.”

This much seems both plausible and opaque.

²⁵ CITE Feinberg, thank Scanlon.

²⁶ Or, with actions-due-to-immoral-choices. This is a controversial assumption. Some think failures of character are moral failures, others think failures in action can be moral failures. See (Tannenbaum 2015). For present purposes, it is enough if the objector agrees that sanctionable moral failings are immoral choices. Thanks to Angela M. Smith.

Accepting it for the moment, we can ask two questions: First, *what is* this special moral significance? Second, why should *it*, whatever it is, be reserved for, or “deserved” only by, immoral choices? Call these “the remaining questions.”

I believe I have stumbled across an answer to the remaining questions, one that might make some sense of what people have in mind by “basic desert,” “condemnation,” or “true moral responsibility” and might connect those to an especially strong form of freedom. In what follows I present my answer by engaging in an admittedly odd exercise.

**ANSWERING THE REMAINING QUESTIONS:
SANCTIONS UNDER THE PRESUMPTION OF MORAL FREEDOM**

The odd exercise requires us first to assume what many ethicists, both incompatibilists and neo-Kantian compatibilists, have assumed: 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 will not prevent us from doing what is morally required. Call this *moral freedom* (short for “freedom to do what is morally required”). (Asserting moral freedom is similar to, but not the same as, denying determinism.²⁷)

Kant advocates moral freedom in the *Critique of Practical Reason*. He there asks his reader to imagine a man asked by his prince to give false testimony against an innocent person. A gallows is placed before the man, on which he will be hanged, should he refuse. Kant says of the man, “He would perhaps not venture to assert whether he would [tell the truth] or not, but he must admit without hesitation that it would be possible for him.” (5:30) Kant concludes that awareness of our obligation reveals our noumenal freedom (employing the rule that ought implies can).

Kant’s use of the example is less an argument than a somewhat high-handed appeal for agreement—almost a dare. One can imagine a sincere, unselfconscious, but hard-bitten person

²⁷ Those who advocate libertarian freedom may think it ensures moral freedom; the reverse is not true.

refusing to admit that it would be possible for them to refuse the prince at the cost of their own life. Such a person might admit there is, indeed, a *question* that they are facing—whether to give false testimony—and therefore, in *that* sense, two possible answers in front of them. But they might insist it would be simply foolish to refuse to give the false testimony—as foolish as giving away all your money to the poor, or, for that matter, as foolish as concluding that the sum of two and three is one. As surely as arriving at either of the latter two answers is not, for them, a genuine possibility, so, too, they might claim, refusing the prince is not a genuine possibility.²⁸ In fact, even a reasonably sensitive person might conclude that, faced with the gallows, they will be in the unhappy position of Procrastinate: unable to do what they also believe they ought to do.²⁹ They need not be evading responsibility, in saying so.

Kant's picture of human freedom rules out both characters as either insincere or self-deceived, in some sort of bad faith. On his picture, when Reason becomes practical, it obeys its own law, unconstrained by empirical facts. Thus, anyone who is subject to the moral law can also satisfy it—so that, as St. Paul says, man is without excuse.

The assumption of moral freedom is extremely strong. If moral requirements require not only to making the right *choices*, but also bringing about certain *outcomes*, then we could not enjoy moral freedom. The world is often, obviously, uncooperative.³⁰ Thus, to make moral freedom plausible, we need to shrink our duties, so to speak: when things go wrong despite our best intentions, we have

²⁸ The Kantian will say that arriving at a mathematical error is not a genuine possibility for a rational creature, because it is against reason. (Of course, it is possible to make mathematical errors, though not, one might want to say, to do so knowingly.) The Kantian claims that acting contrary to duty is likewise against reason (though it seems possible to do so knowingly). The claim presently at issue is whether acting as (you know) duty requires is *always* possible, for those endowed with reason. This seems to me either a stipulation of what is meant by “endowed with reason” or else false. If the former, then it seems implausible to claim that only those “endowed with reason” are subject to moral requirements.

²⁹ The example is complicated by the fact that one could wonder whether giving false testimony to avoid certain death is in fact wrong. One might think the falsely accused is not wronged by a false accusation given in self-defense (so to speak). In fairness to Kant, alter the example until you feel truthful testimony is required but extremely costly.

³⁰ A fact explored by the literature on “moral luck.”

done no moral wrong.³¹ Moral failing is only a failing of willing.³² Let this be so, for the purposes of our diagnosis.

Even this restriction may not be enough. Humans are born unable to satisfy ordinary adult obligations, and moral development is messy and uncertain. Some seem to reach adulthood without the ability to make good choices—they are too insensitive, petty, frail of ego, etc. In the face of these apparent counter-examples, the advocate of moral freedom might shrink moral duties further: Those unfortunate souls who lack the ability to satisfy moral demands are, for that reason, not subject to them. Thus, some people are not actually failing, morally, when they seem to be—though it will be difficult to tell who, when. Let this be so, for the purposes of the diagnosis.

Opposite the presumption of moral freedom stands what we might call (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 motives, are always corrupt.³³ A weaker version claims that, though we are *sometimes* able to choose well, at other times we can find ourselves unable to choose as morality requires. It denies Kant at the gallows. Even the weaker (and I think very plausible) version of original sin is ruled out by moral freedom.

With the presumption of moral freedom in place, let us begin our odd exercise: Let us, as before, consider the fairness of sanctions (of any sort: emotional, monetary, corporeal), still understood as incentives, but now as incentives to make morally required choices.³⁴ Let us, further, stipulate that the trigger for the sanction is the immoral choice—so the sanction will be unfair, because inaccurate,

³¹ “Things going wrong” includes innocent mistakes.

³² See, again, (Tannenbaum 2015).

³³ I am appropriating “original sin.” Properly, it is the Christian doctrine according to which each person, from birth, carries the original sin of Adam and Eve and is therefore unable to achieve righteousness on their own. The doctrine of “total depravity” would be a better choice for this strongest claim, that our motives are *always* corrupt.

³⁴ This is odd for many reasons. One: if moral choice requires pure motives, it cannot be incentivized.

without an immoral choice. Now: given moral freedom, under what conditions would a sanction for an immoral choice be fair—i.e., treat the interests of each symmetrically and adequately well?

Notice, first, that, by stipulating that the choice is *immoral*, we have eliminated the concern about what we earlier called the *quality* of opportunity to avoid the sanction. Unlike the student who was given an assignment on short notice during a time of significant cultural or religious commitments, the quality of an opportunity to avoid wrongdoing is always adequate. To say that the choice is *wrong* is to say that costs of avoiding the choice—the burdens a potential wrongdoer will bear, to avoid wrongdoing—do not justify making it.³⁵

Notice, next, that by stipulating both that the trigger for the sanction is the choice itself and that everyone has the ability to make that choice regardless of their contingent history or circumstances, we have also eliminated the possibility of incurring the sanction even though you have done all that could be reasonably expected—we have eliminated bad luck. By stipulation, all that was expected was a choice, and, given moral freedom, that could be reasonably expected.

Unfortunately, we have not eliminated the possibility of false accusation.³⁶ Therefore we must continue to institute only the minimally effective sanction. Again, this requirement might appear under the label “proportionality:” the severity of the sanction must not exceed the importance of providing the incentive, lest it put the innocent at risk of needless harm.

So far so good. Now let us consider an even narrower, odder question. Suppose we could, somehow (with omniscience?) eliminate the possibility of false accusation. Given moral freedom, under what conditions would sanctions treat a *correctly* accused wrongdoer fairly? Or, to put the question differently, in establishing principles that govern a system of sanctions that treats each

³⁵ As required by moral freedom, I am presuming that in cases of “moral dilemmas,” either one horn is (tragic but) morally required or either horn is (tragic but) morally permissible (and, perhaps, leaving “moral remainders”).

³⁶ In fact, we may have made false accusation more likely by shrinking moral requirements in the face of inability, making it more difficult to tell who is exempted.

symmetrically and adequately well, how would we take into account the interests of those who (in fact) make immoral choices?

It seems to me the answer must be, we will not. Or, better, under the presumption of moral freedom, when considering the fairness of a sanction for an immoral choice, we *cannot* take into account the interests of one who makes an immoral choice. This is because, in presuming moral freedom, we have guaranteed that avoiding the burdens imposed by such a sanction is something any potential wrongdoer both *could* do (because of moral freedom) and *should* do (because doing otherwise is wrong), themselves, by choosing well. Thus, once we eliminate the possibility of false accusation, there is no risk of incurring the sanction without choosing wrongly—something that, under the presumption of moral freedom, everyone can always avoid. Thus, to seek to protect the interests that would be burdened by a sanction against choosing wrongly is, under the presumption of moral freedom, tantamount to seeking to protect an interest *in choosing to do wrong*—seeking to maintain the possibility of doing wrong without harming those interests (too much). But, an interest in doing wrong, or in choosing wrongly, is an interest in treating someone else's interests asymmetrically or inadequately well,³⁷ and an interest in that, like an interest in being treated asymmetrically well or an interest in belittling others, is incompatible with the aims of fairness. It cannot be coherently taken into account—it cannot coherently be added to those that must be treated symmetrically and adequately well. And so it must be silenced.³⁸

Objections will be raised to this line of reasoning; I will consider some below. But notice, first, how it answers the remaining questions: According to the above, by freely choosing wrongly, one *therein* silences one's own interests—one therein forfeits any claim of fairness that might be raised by

³⁷ Or, it is typically so. Arguments of this section may need to be restricted to actions that impact the interests of other people, what Scanlon calls the morality of right and wrong in (1998).

³⁸ Again, it is not clear what would count as treating an interest in *disrespecting* others symmetrically and adequately well. One might say that the interest is not in disrespect, per se, but in, say, gaining the goods to be gained by the immoral choice (the money, e.g.). But, *that* interest (in the money) has already been taken into account, in determining that the choice was wrong—it cannot be double-counted.

appeal to them.³⁹ Thus, by freely choosing wrongly, one therein, immediately, incurs a kind of status, a kind of defenselessness or moral vulnerability to sanction. This silencing of one's interests, this moral vulnerability or defenselessness, might be the special significance of specifically *moral* sanction. I will hereafter call it “condemnation” (the label seems appropriate, if we hear, in it, “condemning to” death, prison, hell, or to some other kind of exile, banishment, or separateness from those of us who are not condemned).

Notice, too, that if the above line of reasoning is correct—or, rather, if it is even *available*—we not only have a way to understand the special force or significance of moral sanction, but we also have an explanation of why such significance is “deserved,” in a “basic” way—without further story—when one freely chooses to do wrong: a freely made immoral choice silences one's interests, in determinations of fairness. Thus, if the above line of reasoning is available, we have a way of understanding why someone might think that a freely made immoral choice will, itself, have an extraordinarily strong, immediate effect on the way in which the interests or well-being of the wrongdoer should matter to the rest of us.⁴⁰

OBJECTIONS

Turning, now, to objections. The lynchpin of the above argument is the claim that, given moral freedom, an interest in avoiding sanction for a wrong done is tantamount to an interest in doing wrong. One might object. In particular, one might point out that, even given moral freedom, we should still *expect* ourselves sometimes to do wrong. Humans are subject to temptation—and will remain so even if graced with moral freedom. Thus, even though, given moral freedom, everyone

³⁹ The above line of reasoning derives what Scanlon calls the “Forfeiture View” from contractualist reasoning, given the presumption of moral freedom. My appeal to “forfeit,” here, is meant to echo Scanlon's—given moral freedom, choice, itself, seems to carry a certain basic or fundamental moral power (1988, 1998). I do not mean “forfeit” to suggest a privilege that one might lose. Thanks to Tommie Shelby for helpful conversation.

⁴⁰ We have seen how a freely made immoral choice entails condemnation. We have not shown it is required for condemnation. But, it seems plausible that condemnation should be avoided unless required.

I have said nothing to explain why some should think it a positive good that wrongdoers suffer. That further stretch of retributivist thinking needs another story. Thanks to T. M. Scanlon for pointing this out.

both *could* and *should* choose rightly, we must acknowledge that choosing against wrongdoing is often extremely difficult. And so we should expect ourselves to sometimes do wrong, *even though* we always have the ability to avoid it. But, if we can expect ourselves sometimes to do wrong, then, one might argue, an interest in avoiding sanction for a wrong done is not tantamount to an interest in *doing* wrong—it is, rather, an interest in fair treatment after the (predictable and regrettable) fact of wrongdoing.

Importantly, we can indeed, consistent with moral freedom, *predict* that we will sometimes freely do wrong. (Kant imagined this, at the gallows.) Nonetheless, under the presumption of moral freedom, we cannot, consistently with the aims of fairness and in good faith, *plan* to do wrong. That is, moral freedom rules out the possibility that one might find oneself in the position of Procrastinate (or Ulysses approaching the *Sirenum scopuli*), needing to anticipate and then plan around one's own constitutional weakness—because, under the presumption of moral freedom, one has no such weakness. If one were to plan around one's own predicted wrongdoing, one would immediately, to so speak, confront one's victim: to plan to do wrong is to plan to treat someone else's interests asymmetrically or inadequately well. Thus, if one were to plan around one's own predicted wrongdoing, one's predicted victim could rightly protest: a better plan is simply to avoid the wrongdoing. In fact, fairness requires it. To plan “around” the wrong action is to plan to do it, which is to plan to treat them unfairly. Given moral freedom, one cannot, consistently with the aims of fairness and in good faith, make such plans.

In fact, a prediction of wrongdoing, together with the interests of potential victims, is what provided the justification for instituting sanctions understood as incentives. Although we cannot, in good faith, plan around unavoidable wrongdoing (because wrongdoing is not unavoidable), we *can*, in good faith, make plans to reduce temptation—to make choosing well less difficult.⁴¹ We thereby provide ourselves with a way to strengthen our resolve and protect the interests of potential victims.

⁴¹ My thinking, here, is indebted to Berislav Maruši'c's careful treatment of “difficult action” in his (2015).

Thus, if a potential wrongdoer predicts wrongdoing, then, once we remember the victim, it may seem the right response is not to *curtail* the sanction but rather to *strengthen* it—to make it more daunting, and thereby more effective. The stronger the sanction, the more we have reinforced the wrongdoer’s native ability to choose well.⁴²

If a potential wrongdoer objects, suggesting the sanction instead be curtailed, the potential victim could make the following argument: The victim is unable to protect themselves against the burdens imposed upon them by the wrong—they are, so to speak, at the mercy of the wrongdoer. But, given moral freedom, the wrongdoer can avoid both the burdens they impose on the victim and the burdens that will be imposed on them, in the sanction, simply by choosing well—which is, by stipulation, what fairness requires. For the wrongdoer to claim the sanction is unfair is to suggest that the victim accept a risk that the wrongdoer both could and should neutralize by incurring only those costs that have already been outweighed in determining that the action is wrong. To suggest that the victim should accept a risk that the wrongdoer should and could costlessly (so to speak) neutralize is to claim that the interests of the victim should be treated asymmetrically poorly. Such a claim is incompatible with deliberations of fairness, and will be silenced.⁴³

This silencing response is hard.⁴⁴ Can it be softened? Could we find some other way to respond to the prediction of wrongdoing by making it less difficult to avoid, without silencing the interests of the wrongdoer?⁴⁵

⁴² Could we make a sanction severe enough to *always* prevent wrongdoing? Probably not. But, if we enjoy moral freedom, then to choose to incur a maximal sanction (to incur the wrath of hell, say) would be to prefer self-destruction and the wronging of another over self-preservation without that wrong. Given moral freedom, it is unclear why *fairness* requires that we prevent someone from implementing that preference. (See all the caveats below.)

⁴³ I thank T. M. Scanlon for “costlessly.” Note, too, the reemergence of the fantasy: were we morally free, we could, collectively, control all non-natural evil, simply by choosing well. Fairness would require it.

⁴⁴ The spirit of the silencing response might be seen in an old sketch by Bob Newhart: <https://www.youtube.com/watch?v=Ow0lr63y4Mw>

⁴⁵ Thanks to Stephen White and Elizabeth Fricker for pressing for other options.

Let me note, again, it is sufficient for my diagnostic goal if the silencing response is *available* as way to understand the condemnatory significance of moral sanction and its connection to a strong sort of freedom. Even if silencing is not required by moral freedom, it may yet provide a way of understanding certain strongly held but rarely explained intuitions about desert and condemnation.

Nonetheless, I am (at the moment) convinced that silencing follows from moral freedom. But I will postpone consideration of softer responses in order to clarify the implications of silencing—because, although hard, they not *as* hard as might at first seem.

The silencing response claims that, under the presumption of moral freedom, a wrongdoer has forfeited any *complaint of fairness* they might have raised, *on their own behalf*, about burdens they suffer as a sanction for wrong choice. It claims that when one freely chooses wrong, one thereby silences one's own interest and so incurs a particular kind of status that might well be labeled “condemnation.”

The silencing response does *not* claim that, under moral freedom, there are *no* limitations or restrictions on morally appropriate sanctions. There are. Because of the real-world possibility of false accusation, sanctions must be minimal, and, to be fair, they must be enforced uniformly. Thus, the minimally effective sanction must be the one even the genuine wrongdoer faces, in the real world. Thus, the silencing of the wrongdoer's interests will not have any effect on, so to speak, the *material* sanctions it is fair to institute. It will only have an effect on what we earlier called the *significance* of those sanctions: Those who are confident in the accusation and convinced of moral freedom will likewise be confident that the wrongdoer stands morally vulnerable, “condemned.” They will see those same material burdens as now carrying that condemnatory force.⁴⁶

Importantly, the silencing response also does not claim that, in doing wrong, the *wrongdoer* is silenced. Rather, in doing wrong, the wrongdoer has silenced their own *interests* for the purpose of

⁴⁶ The proponent of moral freedom may advise doubting any such confidence: Condemnation may be reserved for the omniscient. As I understand it, Kant's position is somewhere in this neighborhood.

determining the fairness of a sanction. But the wrongdoer is not, *themselves*, silenced: they may make other important complaints. Here are some:

Most broadly, simple human decency requires us to consider the interests of all creatures, even wrongdoers. Certain sanctions are cruel or inhumane, and those should be avoided, independently of their fairness. The wrongdoer can appeal to their own interests and ask for mercy or complain of indecency or a lack of generosity.⁴⁷ This possibility remains, even given moral freedom.

More narrowly, a genuine wrongdoer may even complain of the *unfairness* of the sanction, so long as they do so without reference to *their own* interests. If subjected to a more-than-minimal material sanction (one that is “disproportionate”), the wrongdoer could correctly complain that it is unfair. But, importantly, a genuine wrongdoer could do so only by invoking the interests of *others*—viz., the potentially falsely accused.⁴⁸ Similarly, the wrongdoer might point out that certain kinds of sanctions, or certain severity of sanction, are likely to lead to social strife, and therefore put the interests of the innocent unfairly at risk. The wrongdoer could, again, thereby complain about the fairness of the sanction, without appeal to their own interests.

The only thing ruled out by moral freedom is a complaint of *unfairness* by appeal to *their own* interests.⁴⁹ Still, this is a significant limitation. *Qua* wrongdoer, you stand, in one important way, defenseless—your interests do not constrain the burdens that may be fairly imposed upon you. Even though *we* can, and doubtless should, take them into account as a matter of decency or

⁴⁷ Thanks to Julie Tannenbaum for emphasizing generosity.

⁴⁸ Or, perhaps to themselves *qua* potentially falsely accused, but not to themselves *qua* wrongdoer.

I have identified concerns about proportionality as concerns about the minimally effective sanction, given the possibility of false accusation. One might, instead, think there is some bedrock principle of proportionality. If so, the principle of proportionality is playing the same kind of independent role as decency or generosity—it is something that we ought to attend to. It does not allow the wrongdoer to stand on their own interests and complain that they have not been treated symmetrically and adequately well, relative to those of others. Thanks to Marshall Thompson for pressing this point.

⁴⁹ One might note that the *unlucky* also cannot complain of unfairness by appeal to their own interests—the sanction is also fair, in their case, despite the burden on their interests. However, the interests of the unlucky are not silenced—their interests ensure the sanction is minimal. The silencing generates the “condemnatory force” of the sanction. Thanks to Rachel Achs.

humanity or proportionality or good (fair) policy, you cannot insist on it as a matter of *fairness to you*. If you ask others to do unto you, they can fairly reply that you made your bed. For yourself, you cannot ask for justice, but only mercy.

Now, can we find a softer response? We have seen that, given moral freedom, we cannot take wrongdoing as unavoidable and plan around it. We must, instead, plan to make avoiding the avoidable less difficult. Instituting sanctions is one way. Could we find another? Could we, e.g., arrange society so there are fewer opportunities to choose wrongly?

We certainly could and should—but, to be fair, we cannot ask potential victims to accept risks or costs that could instead be neutralized by freely avoiding immoral choice. This will be a significant limitation in making such arrangements.

One might think we have not yet made full use of the fact that we can predict our own wrongdoing—that occasional wrongdoing is ~~our common fate~~ common among us. One might point out that we may *all* sometimes ~~find ourselves too weak of will to avoid wrongdoing, here and now~~ correctly predict that we will freely choose wrongly, in the future. Given this recognition of our ~~common weakness~~ moral failing, can we find a way to take the interests of the wrongdoer into account, as a matter of fairness?

I do not yet see how, given moral freedom. Acknowledging that we each have packaged together, in one skin, across time, both victim and free wrongdoer, does not provide us an answer to a victim's complaint. Each victim can still rightly insist that a better plan is to avoid the wrong.⁵⁰

Suppose we deny moral freedom and allow, not just moral failing, but true moral weakness — suppose we deny Kant, at the gallows. We thereby allow it possible to find oneself in the position of either the hardbitten or the reasonably sensitive man: seeing two options, but finding one

⁵⁰ One might conclude that, in such a world, no one would avoid condemnation, and so we should adjust our expectations. We would thereby avoid condemnation by changing what counts as wrong.

impossible to take, not a genuine option, even though the alternative is wrong. Our wills are then weak in something like the straightforward sense in which our muscles are sometimes weak. If we anticipate such weakness, we then find ourselves in the position of Procrastinate (or Ulysses), needing to plan around predicted wrongdoing. But now, the victim cannot protest that a better plan is simply to avoid the wrongdoing—though avoiding wrong is, certainly, a better plan, it may not be available. Thus, once we acknowledge the possibility of moral weakness, we lose the lynchpin of the original argument: It is no longer the case that seeking to protect the interests burdened by a sanction against wrongdoing is tantamount to seeking to protect an interest in doing wrong. It may, instead, be to protect one's interests given the predictable, inexcusable, but nonetheless unavoidable fact of wrongdoing.⁵¹ Given moral weakness, we may plan around our own wrongdoing in as much good faith as we are capable of mustering.

Of course, to plan around wrongdoing is still to plan to do it, and so it is still to plan to treat potential victims unfairly—such plans are inconsistent with the aims of fairness. But the inconsistency is now in *us*, so to speak. It is no longer in the reasoning about the fairness of a sanction. When reasoning about the fairness of the sanction, we can take into account the interests burdened by unavoidable wrongdoing, just as we can take into account the interests of the unlucky. It can be sanctioned, but it will not be condemned.

One might wonder: though we *could* take into account the interests of morally weak, why would we? Why would we allow the morally weak—those who treat us badly—to be counted among those whose interests should be treated symmetrically and adequately well? Why not reserve fairness for the fair?

We might *now* appeal to our common humanity, to our original sin: none of us are morally free. Our ability to avoid wrong must be developed, and its development is fraught with contingency and

⁵¹ How can something unavoidable be inexcusable? In denying moral freedom, we deny that ought implies can. We can then acknowledge weakness without denying responsibility. The full argument will be presented in *Minds that Matter*.

luck. In that development, each of us has encountered our own weakness—or, if not, that would be, itself, a matter of tremendously good luck. And so, were we to exclude the weak, few if any would be left. More, those excluded could say, “Here but for the grace of God go you.”

Interestingly, they would then not be seeking mercy, but appealing to fairness.

But: What about cases in which our wills are *not* weak? Even allowing for original sin, most of us, most of the time, are able to avoid this or that immoral choice. Wrongdoing is not always—or, even, very often—unavoidable.⁵² Yet, often enough, we choose wrongly. It seems, then, that most of us, most of the time, stand condemned when we make an immoral choice.⁵³

This conclusion is sometimes sought: by distinguishing the able from the weak, we could rule out condemnation in certain cases (perhaps, cases such as Robert Alton Harris or Jo Jo), while preserving it for ordinary, garden-variety wrongdoings.⁵⁴

This seems to me the wrong conclusion to draw from the fact of original sin—though I am, at the moment, unable to clearly state why.⁵⁵ One quick problem: the odd exercise in which we have been engaged (in our attempt to understand why unlucky sanction should be less than “truly” unfair) is misguided: moral choices cannot be incentivized, and so are not legitimate objects of sanction.⁵⁶ More, it seems to me we have no business sanctioning one another for choices rather than outcomes—fraught though they are with luck. Thus, “moral sanction” is, itself, misconceived.

Many will disagree. Some will feel that, if we give up moral sanctions and use only pricing systems, we are giving up all that is distinctive about responses to moral failure—that we are moving

⁵² In fact, I suspect that it must typically be avoidable. See (Hieronymi 2020).

⁵³ Though, perhaps, actual condemning requires omniscience, for its implementation.

⁵⁴ Those who seek to distinguish the able from the weak use “blame” or “opprobrium” where I use “condemnation,” and they would deny that the special significance of moral sanction entails the silencing of interests. See (Wolf 1990, Wallace 1996, 2011). Elsewhere I provide an alternative account of the significance of moral failure, one that is not a sanction (2004, 2019, in progress). My alternative does not mark this difference between the able and the weak.

⁵⁵ Thanks to Richard Moran for extremely helpful conversation.

⁵⁶ A *moral* choice is a choice for certain reasons and not others.

to a mere “utilitarian” system of social control. I strongly disagree. Rather than avoid the ugliness of social control by appeal to the distinctive significance of moral *sanction* (which seems, to me, trading one attempt at control for another), we should instead consider the distinctive significance of moral *failure*—which is not, I would argue, that of being liable to a sanction. It is, instead, the significance or importance of standing in certain sorts of relationships with other people, relationships of regard or respect.⁵⁷ Mattering to others, not meriting negative consequences, is what morality, and moral responsibility, is about.

CONCLUSION

To summarize: Sanctions raise concerns of fairness because they involve voluntary actions that burden others; *any* such action raises such concerns. The fairness of a sanction typically requires an opportunity to avoid it, because otherwise it burdens interests for no purpose. However, sometimes making an exception is, itself, burdensome, and so it is sometimes possible to fairly incur a sanction through bad luck.

So understood, sanctions do not carry moral significance; they are pricing systems. It is tempting to think, if we instead consider *moral* sanctions, which do carry a distinctive significance, then sanctioning the unlucky is always unfair.

What is this distinctive significance? And why does it rule out luck?

I offered a hypothesis: If we first presume moral freedom and then consider sanctions for immoral choices, we not only rule out bad luck, but we also find that the fairness of such sanctions is unlimited by the interests of the wrongdoer. To have one’s interests thus silenced by one’s freely made immoral choice is a kind of *condemnation*. This condemnation might be the special significance of moral sanction—it follows, immediately, from a freely made immoral choice. We would then

⁵⁷ It will be crucial that interpersonal responses to moral failure are not voluntary—and so not attempts at control. In Strawson’s words, they “do not merely exploit our nature, but express it.” Thanks to Christine Korsgaard and Tamar Shapiro for pressing for clarification. For further elaboration, see (Hieronymi 2019, 2020, in progress).

have a way to understand why “basic desert” or “true moral responsibility” might require an especially strong form of freedom.⁵⁸

For my own part, I do not think we enjoy moral freedom, nor do I think we should be in the business of instituting sanctions for choices. I therefore think we should give up on condemnation.

Some will worry that we would then give up what is distinctive about moral responsibility—we would move to mere social control. To the contrary, I believe we can give up condemnation without giving up what is distinctive of moral responsibility—but that is a topic for another day.⁵⁹

⁵⁸ See note XX (follow from vs. required by)

⁵⁹ This paper has benefitted tremendously from the input of many. Thanks are due, of course, to the audience at NOWAR, as well as to audiences at UCSD, Harvard, Notre Dame, Yale, the Oxford/Toronto/UCLA Workshop, *Varieties of Agency* at Stanford, University of Miami, FSU, and *Taking Responsibility and Holding Responsible* at the Center for the Study of Mind and Nature, Oslo. In addition, I would like to thank Brendon de Kenessey, Berislav Marusic, Michael McKenna, T. M. Scanlon, and Angela M. Smith for written comments, and David Brink, Andreas Brekke Carlsson, Randall Clarke, Jorah Dannenberg, Stephen Darwall, Elizabeth Fricker, Mark C. Johnson, Shelly Kagan, Martin Kavka, Sean Kelsey, Christine Korsgaard, Gavin Lawrence, Richard Moran, Dana Nelkin, Calvin Normore, Jonathan Schwenkler, Tamar Shapiro, Tommie Shelby, David Shoemaker, Sarah Stroud, Julie Tannenbaum, Marshall Thompson, Patrick Todd, Manuel Vargas, Stephen J. White, and Yuan Yuan for helpful questions or conversation.

BIBLIOGRAPHY

- Hieronymi, Pamela. 2004. "The Force and Fairness of Blame." *Philosophical Perspectives* 18 (1):115–48.
- Hieronymi, Pamela. 2007. "Rational Capacity as a Condition on Blame." *Philosophical Books* 48 (2):109–123. doi: 10.1111/j.1468-0149.2007.00435.x.
- Hieronymi, Pamela. 2019. "I'll bet you think this blame is about you." In *Oxford Studies in Agency and Responsibility: Essays on Themes from the Work of Gary Watson*, edited by Justin Coates and Neal Tognazzini, 60–87. Oxford: Oxford University Press.
- Hieronymi, Pamela. 2020. *Freedom, Resentment, and the Metaphysics of Morals*. Princeton: Princeton University Press.
- Hieronymi, Pamela. in progress. *Minds that Matter*.
- Kant, Immanuel. 1998. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor. Edited by Christine M. Korsgaard. Cambridge: Cambridge University Press. Original edition, 1785.
- Marusic, Berislav. 2015. *Evidence and Agency Norms of Belief for Promising and Resolving*. Oxford University Press.
- McDowell, John. 1979. "Virtue and reason." *The Monist* 62 (3):331-351.
- Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. Oxford U.P.
- Scanlon, T. M. 1988. "The Significance of Choice." In *The Tanner Lectures on Human Values (Vol. 8, pp. 149-216)*, edited by Sterling M. McMurrin. University of Utah Press.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Strawson, Galen. 1994. "The impossibility of moral responsibility." *Philosophical Studies* 75:5-24.
- Tannenbaum, Julie. 2015. "Mere moral failure." *Canadian Journal of Philosophy* 45 (1):58-84.
- Wallace, R. Jay. 1996. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.
- Wallace, R. Jay. 2011. "Dispassionate Opprobrium: On Blame and the Reactive Sentiments." In *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar and Samuel Freeman, 348–72. New York: Oxford University Press.
- Watson, Gary. 1996. "Two Faces of Responsibility." *Philosophical Topics* 24 (2):227-248.
- Williams, Bernard. 1985. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.
- Wolf, Susan. 1990. *Freedom within Reason*. New York: Oxford University Press.