

The Force and Fairness of Blame

Pamela Hieronymi
hieronymi@ucla.edu
July 23, 2004

I wish to consider the fairness of blaming a wrongdoer. More exactly, I wish to consider the claim that blaming a wrongdoer can be unfair because blame has a characteristic force, a force which is not fairly imposed upon the wrongdoer unless certain conditions are met—unless, e.g., the wrongdoer could have done otherwise, or is able to control her behavior by the light of moral reasons, or played a certain role in becoming the kind of person she is. Call this “the target charge of unfairness.”

While agreeing that blame has a characteristic force, I am skeptical of the target charge of unfairness. My skepticism concerns itself less with the particular conditions of fairness proposed than with the idea that blame can be rendered unfair by its force. If to blame a person were simply to perform certain disagreeable, burdensome, or harmful actions, then, as we will see, blame could be rendered unfair by its force. But to blame a person is not just to act in certain ways. To blame a person is, at least in large part, to adopt certain attitudes or make certain judgments. However, it is unclear how these attitudes or judgments carry “force” and also unclear whether they can be rendered unfair by their force. Examining these issues, I will suggest that much of the special force of blame is found in a set of judgments—most centrally, the judgment that a person failed to show proper regard for others. But, I will argue, once it is granted that such judgments are true, their characteristic force cannot render them unfair.

By the end, I hope to arrive at a conclusion similar to the one drawn by P. F. Strawson in his seminal “Freedom and Resentment”¹—though I hope to arrive at it by a different route. Strawson claimed that the conditions of moral responsibility, understood

as the conditions that render certain interpersonal attitudes appropriate, neither need nor admit of an “external” (metaphysical or utilitarian) justification. He made his argument by characterizing our interpersonal relations as a “practice.” I will argue for much the same conclusion, not by appeal to the idea of a practice, but rather by locating the force of blame in certain judgments and arguing that judgments, as such, neither need nor admit of an “external” justification.

THE TARGET CHARGE OF UNFAIRNESS

It is important, first, to isolate the target charge of unfairness from others with which it might be confused. Blame can be rendered inappropriate or unfair in many ways. For example, it may be that unless certain conditions are met, no wrong was actually done or no ill will shown.² If no wrong was done, then, of course, the person in question does not deserve blame. Blame is thus rendered inappropriate. Likewise, it may be that, unless certain conditions are met, the action or attitude in question cannot be rightly attributed to the person—it is not something with which the person can be “identified,” or it does not issue from that person’s “real self.”³ If the action or attitude is not fully the person’s own, then, again, *that person* showed no disregard, and so does not deserve blame. But the claim that blame is undeserved because the person showed no disregard contrasts with the target charge of unfairness. The target charge *grants* that the person showed disregard but claims that blaming the wrongdoer is nonetheless unfair, because imposing the characteristic force of blame is unfair, under the circumstances (because the person couldn’t have done otherwise, etc.).

The charge of unfairness is especially tempting in a certain kind of case. In such a case, someone has been especially callous or has in some way deliberately caused

unjustified suffering, while in full possession of her mental capacities and while knowing full well what she is doing. The action issues from the person's character, convictions, or values, without struggle or regret, and so seems attributable to her "real self." Thus, she has shown disregard for or ill will towards others, and seems to be blameworthy.

However, as the story unfolds, one learns a bit about this person's unfortunate formative years, and the history serves to make intelligible—indeed, all too intelligible—her action. In fact, once one understands the relevant history, it seems entirely predictable that this person would show such disregard or ill will. It begins to seem incredible that she would have instead behaved respectfully. The action in question, or one very much like it in all morally relevant respects, begins to look inevitable. There's a robust enough sense in which this person's "real self" lacks (at least some degree of) ability to have done otherwise—for reasons over which she has had little or no control. And so it seems unfair to blame her (to the same degree) for the wrong she does. It seems that the history, rather than the person, should be (at least partially) to blame.⁴

Why should blaming such a wrongdoer seem unfair? It is generally agreed that blame would not be unfair if blame were "mere description" or "grading"—if, as R. E. Hobart suggests, to blame someone is "simply to say that he is a bad act-maker."⁵ After all, if there are moral standards, and if one's actions or attitudes fall short of them, then it hardly seems unfair to note this fact. But blame, it is thought, goes beyond simple description or mere grading. As Susan Wolf has noted,

it is just not true that to blame someone for an action is simply to say that he is a bad act-maker. . . . Earthquakes, defective tires, and broken machines are, at least sometimes, bad act-makers, as are dogs and children and adults with various physical and mental handicaps. But we do not blame them, at least not in the way we sometimes blame normal adult human beings. More generally, inanimate objects, natural phenomena, lower animals, and normal adults in a wide range of contexts exhibit all sorts of behavior that can be seen to have all sorts of positive, negative, or neutral value. We can recognize that these individuals behave in these ways, and recognize that

their behaviors have the value they do, and yet find it completely inappropriate to blame . . . these individuals for the behaviors in question . . .⁶

Wolf goes on to distinguish what she calls “deep” blame from “superficial” blame.

Persons, she says, seem to “deserve a distinctive and more serious kind of blame for being deceitful or petty than pigs deserve for being sloppy or books for being frayed.”⁷

Likewise, being morally blamed involves a more serious sort of criticism than being told your vocal performance was flat, your cooking bland, your conversation dull, or your sentences opaque. Blame, unlike mere description, carries a characteristic depth, force, or sting.⁸ It is this characteristic force of blame that can render it unfair.

What is this force? How is it, exactly, that blame goes beyond mere description? This is a question with which we will be much concerned. Many philosophers believe the special force of blame is captured only by including in one’s account of blame something beyond a descriptive judgment. On such views, blaming a person “involves more than holding a particular belief about him; it entails a willingness to adopt certain attitudes towards that person and to behave towards him in certain ways.”⁹ On any such account, the one who is blamed is not only negatively graded, but also, on account of her poor performance, may be made the object of negative emotions such as resentment and indignation, and, further, may be subject to adverse treatment such as “avoidance, reproach, scolding, denunciation, remonstrance, and (at the limit) punishment.”¹⁰ These further blaming actions and attitudes are disagreeable, burdensome, or harmful. They have serious negative repercussions for the wrongdoer’s relations with herself and others. Additionally, they seem to issue a kind of demand or expectation to the one blamed: if you are so blamed, you must either justify your action (show it not wrong, after all) or acknowledge the wrongdoing and reform your ways.¹¹ So, on many views, the force of

blame is located in the harms, burdens, demands, and expectations contained in or resulting from those blaming actions and attitudes that go beyond the simple belief that ill will or disregard was shown.¹²

The target charge, then, grants that a person showed ill will or disregard, but takes the characteristic force of blame to saddle blame with special burdens of justification. Due to blame's characteristic force, blaming that person will not be fair unless some further condition is met—unless, e.g., she could have avoided the blame, or is someone capable of doing right, or can respond to the demands appropriately, or played a certain role in becoming the kind of person she is. According to the target charge of unfairness, the characteristic force of blame implies can.¹³

THE FORCE AND FAIRNESS OF BLAMING ACTIONS

The target charge would make ready sense if the force of blame were found entirely in ordinary intentional actions: confronting the person with her wrongdoing, expressing your resentment or indignation to her, issuing demands to her, making your judgment known to others, shunning, sanctioning, or punishing her. Such actions both carry a characteristic force—they are disagreeable or burdensome or harmful—and can be subject to questions of justification in light of that force. Issuing a demand, for example, may be inappropriate if the person in question is unable to meet the demand, or if the person in question is incapable of understanding it and responding to it. Issuing a demand to someone unable to meet it seems not only pointless but also unfair. An action that imposes harms on someone may not be justified unless it produces some good or serves some purpose, or unless the wrongdoer played a certain role in becoming the kind of person who deliberately harms others. (The question of whether the fact of

wrongdoing, itself, justifies the imposition of harms is a familiar question about retributive justice.) The force of an intentional action can render it unfair, for many reasons.¹⁴ Though assessing the *correctness* of the charge might prove difficult, the charge itself makes ready sense—it is a familiar sort of claim about the justification of an intentional action, in light of its significance or effects.

Notice that the target charge makes ready sense when applied to intentional actions because such actions are, in a certain sense, “voluntary:” a person decides to perform an intentional action by answering for herself (however implicitly) the question of whether so to act—a question that includes in it a representation of that action, under the relevant description.¹⁵ Because the crucial question includes a representation of the action, the agent can be expected, in answering it, to consider any particular feature or foreseeable effect of the action. Thus one can, in principle, appeal to any feature of the action or its effects in calling the justification of the action into question. And so one can readily claim, of a sanctioning or punishing action, that it carries a certain characteristic force, a force that renders it inappropriate if certain conditions are unmet.

So, insofar as the force of blame is located in the negative significance or adverse effects of ordinary intentional actions, the target charge is in good order.

THE FORCE OF BLAMING ATTITUDES AND JUDGMENTS

However, blaming someone is not simply a matter of acting in certain ways. You could, after all, perfectly well *act as if* you blame someone, without actually blaming her. Your actions would then be an insincere expression of blame. Likewise, it seems possible to blame a person without expressing that blame at all—without acting in any particular way. One could blame a person while behaving just as before, or without interacting

with that person at all. Even absent action, such blame would, it seems, carry enough of its characteristic force to raise concerns about fairness. Finally, certain forms of treatment (such as avoiding) will carry the force of blame (and so count as shunning) if and only if they (purport to) reveal certain underlying judgments or attitudes. The characteristic force of blame, then, is not imposed exclusively through ordinary intentional actions.

Where, then, are we to locate the remaining force of blame? Since the publication of P. F. Strawson's "Freedom and Resentment," considerable attention has been paid to what he called "reactive attitudes," attitudes such as resentment and indignation. Such attitudes are, no doubt, disagreeable and have an adverse effect on one's relations. Appealing to such attitudes in an account of blame is attractive, in part, because it seems to allow blame to be more than description—to carry a certain force—but still to be less than action. Reasoning in this way, R. Jay Wallace suggests that to find someone blameworthy is to find that person to be a fit object of such attitudes.¹⁶

If these attitudes carry the characteristic force of blame, and if that force is fairly imposed only if certain conditions are met, then, it seems, these attitudes will be fairly adopted only if those conditions are met.¹⁷ If the person could not have avoided wrongdoing (or would have had much greater difficulty in avoiding it or lacks the power to govern her behavior by the light of moral reasons), then it seems unfair that she should be made the object of such attitudes (to the same degree). The target charge of unfairness can thus seem to apply to these attitudes in the same ready way it applies to intentional action.

While I think these attitudes do capture much of the force of blame, I am skeptical that the target charge can be applied to them. Once we consider exactly how these attitudes carry their characteristic force and how they differ from intentional actions, we might doubt whether it makes sense to think that they can be rendered unfair by their force.

Compare resentment with distrust. Distrust also seems to carry a special, characteristic, and burdensome force. Notice, it might seem similarly unfair that someone who has become a generally unreliable person, due to formative circumstances outside her control, should be systematically subject to the burdens of being constantly distrusted. Yet it is not clear that those who interact with her can be charged with unfairness in distrusting her. In fact, if a person will always let you down, that seems to be solid grounds for distrust rather than a condition under which distrust would be unfair. The distrust, it seems, simply marks the fact that a person's unreliability or untrustworthiness is known. The force of the distrust is found in the importance of being, and being known to be, unreliable or untrustworthy. If a person is unreliable or untrustworthy, that force seems rightly in place; it is not unfair to distrust such a person (although perhaps there is something unfair in the fact that she became such an untrustworthy person due to factors beyond her control, and so is now rightly subject to distrust). Something similar, I will suggest, is true of the reactive attitudes of resentment and indignation. (Or, to avoid verbal dispute, something similar is true of a familiar and identifiable set of attitudes one might adopt towards a wrongdoer, which I will call resentment and indignation.)

Let's start by examining these attitudes a bit more closely. In an important insight, Strawson identified these attitudes as "reactions to the quality of others' wills towards us"

or towards others.¹⁸ Reactive attitudes, presumably, are reactions to one's *perception of* or *judgments about* the quality of another person's will towards oneself or others.¹⁹ Of course, a judgment about the quality of another person's will is a descriptive judgment. If the force of blame must lie in something beyond mere description, it can seem that force must lie in the resentment or indignation itself, as something added to the judgment of ill will (as a "judgment-sensitive attitude," perhaps).²⁰

What sort of addition are these reactive attitudes? What kind of "reaction" are they? And how could they carry this force? It seems clear enough these attitudes are not action-like. That is, they are not voluntary responses to a judgment in the way an intentional action might be: upon judging that ill will was shown, one does not, and cannot, then resent or be indignant simply by representing to oneself that attitude—resentment or indignation, with its characteristic features and possible effects—and deciding that there is something to be said for or against doing *that*—for or against adopting the attitude of resentment or indignation, in light of its features or effects. Being subject to such discretion, they would lose their characteristic significance.²¹ Rather, resentment and indignation are in this respect like judgments or beliefs: one resents or is indignant in response to (a representation of) certain features of the *object* of one's resentment or indignation.

If these attitudes are not, in this sense, voluntary responses to (one's judgment about) the quality of another's will, it is hard to see what sort of further "response" to that judgment they might be. An affective response, some might say.

Suppose, for a moment, that the reactive attitudes are only this—that they are simply the affect experienced in response to (a judgment or evaluation of) the quality of another

person's will. If so, then to have this particular blaming response—to resent another, say—will be simply to judge that that person showed you disregard and to experience this affect. To be indignant would be to judge that she showed disregard for another and to experience a somewhat different accompanying affect.

This is, to be sure, an overly simplistic account of these reactive attitudes. I will examine its shortcomings in due course. Still, the simplistic account serves to bring out two problems with any account that would rely on *affect* to explain the force of blame.

First, it is unclear how the affective accompaniment of a judgment could, itself, carry the characteristic force of blame. An affective accompaniment of a judgment would be a certain unpleasant emotional disturbance, occasioned by the judgment. But, the force of blame seems deeper, more serious or weighty than simply being the object of certain unpleasant emotional disturbance. The affect, itself, seems insufficiently robust.

Second, and relatedly, while it seems clear enough that blame can be fair or unfair, justified or unjustified, it is unclear how an affect, itself, could be rendered unfair. An affect is passive in a way that contrasts with both action and judgment. While a judgment is not voluntary, both actions and judgments are rational responses, and so either can be unjustified and either (I will argue shortly) can be unfair. But it is quite unclear how we would gain similar critical purchase on an affect. You could, of course, be a poor *manager* of your affects—just as you could be a poor manager of your allergies. But then it is the decisions you make about affect-management that are unjustified or unfair. It is unclear how an affect, itself, could be either.

There are a number of positions one might take to enrich the idea of affect.²² Perhaps the affect entails or signals certain dispositions, dispositions of future judgment and

action. It is unclear how this helps; we appealed to these attitudes precisely to avoid relying on either judgment or action. Appealing to disposition of thought and action, then, seems not to solve the puzzle.

Perhaps, instead, the affect serves to show the great importance of the judgment to the person—if I resent you, I make clear that I find your ill will of importance. This seems plausible (and close to the account I will propose), but it would mean that the force of blame lies not in the affect or even the attitude but rather in what it marks: the importance of your ill will to this person.

Alternatively, and sophisticatedly, one might appeal to the idea of adopting a “stance” from which these attitudes are (typically) experienced. Following Wallace, we can call this the stance of “holding a person responsible.” One adopts this stance when one thinks the attitudes in question would be appropriate, should a person do wrong. By adopting that stance, one somehow makes it the case that one (might) experience the corresponding affect upon judging that a person has shown ill will. Adopting this stance is an action (or at least is action-like): it is something you can decide to do by considering whether to do it. Adopting the stance thus could be fair or unfair. The affect could carry the force it does because it reveals this prior decision—the decision to hold the person responsible.²³

While the introduction of the decision to hold responsible allows these attitudes to be forceful and unfair, I find this solution unsatisfying. Thinking of the force of the affect as lying in the prior decision to hold the person responsible suggests that the affect, itself, is simply an instrument of sanction or punishment, employed when one deems it appropriate. To hold you responsible, on such a picture, would be to put you under threat

of affective response, or perhaps to deem you worthy of it, should you show ill will. But I suspect the force of the reactive attitudes is nothing of the sort.²⁴

Finally, one might abandon the attempt to capture the force of blame in an affect and instead claim that the reactive attitudes are sensitive to or manifestations of more than one judgment. One might, for example, claim that they are sensitive not only to the judgment that ill will has been shown, but also, crucially, to the judgment that, because of this ill will, one is entitled to treat the wrongdoer in certain ways. That is to say, one might add to the “purely descriptive” judgment of ill will a prescriptive (or practical or normative) judgment—a judgment about how others are entitled to act toward the wrongdoer. This prescriptive judgment, then, might both carry force, even absent any actual action, and yet be sensitive to all the conditions required to justify the actions. This account would allow blame both to carry a certain force and to be rendered unfair by whichever considerations render such treatment unfair.²⁵

While I will eventually propose that we understand the force of the reactive attitudes by appeal to more than one judgment, I think it too quick to locate the force of these attitudes in the judgment that certain forms of treatment are deserved. There is more to the force of blame or the force of the reactive attitudes than either affect or the threat of or the entitlement to certain forms of treatment.

I suggest, then, that we reconsider the other half of the overly simplistic picture: the judgment that a person has shown disregard or ill will. We looked to affect because it seemed that a merely descriptive judgment could not carry the force of blame. But, I will now argue, this is a mistake. This judgment, given its content, can carry much of that force, quite apart from either affect or (entitlement to) further negative treatment.

How can a judgment, a mere assessment of what is true, carry force? Doesn't such a judgment simply describe what is so?

Faced with the idea of “mere description,” one might become confused and start to worry about how a mere statement of descriptive fact can have any special force. One might think that, to account for the force of blame, we must look further, to (threat of or entitlement to) some unpleasant emotion or burdensome action. But looking further is quite unnecessary. Many truths are important to us. It matters to us whether our loved ones are healthy, and, surely, it also matters to us whether our sentences are opaque, our cooking bland, or our conversation dull. It matters to us if these things are true, and—importantly—it often also matters to us whether others *think* them true.

Quite standardly, then, a judgment gains force from the importance of its content and the importance of the opinions of others on that topic. Whether a judgment of disregard or ill will can carry the characteristic force of blame will depend on whether that force is well-captured by the importance of the quality of one's will towards others and their opinions about it.

Some philosophers identify the characteristic force of blame with the fact that one's will, or character, or moral worth is negatively assessed. That is to say, some philosophers think the force of blame lies, not in some further response or reaction to a descriptive judgment, but rather in the special depth or importance of the *object* of such a judgment—one's will or autonomy or character or moral value. Wolf seems to characterize what she calls the “depth” of blame in this way:

When we hold an individual morally responsible for some event, we are doing more than identifying her particularly crucial role in the causal series . . . We are regarding her as a fit subject for credit or discredit on the basis of the role she plays. . . . we are not merely judging the moral quality of the event with which the individual is so intimately associated; we are judging the moral quality of the individual herself. (*Freedom within Reason*, 41)

Thomas Nagel expresses a similar thought when he notes that, in judging a person, morally, “We are judging *him*, rather than his existence or characteristics.”²⁶ Taking the thought of “credit” quite seriously, Jonathan Glover claims that “our present practice of blame [involves] a kind of moral accounting, where a person’s actions are recorded in an informal balance sheet, with the object of assessing his moral worth.”²⁷ On any such account, the characteristic force of moral blame may be contained in a descriptive judgment—a judgment about one’s will, character, worth, or value—but such a judgment is not thought to be *merely* descriptive. It is rather a particularly deep form of grading or a particularly important sort of record-keeping.

I find this kind of account of the force of blame, taken strictly, quite beside the point. The judgment that one has shown ill will or disregard certainly has depth; it has depth simply in virtue of being an assessment of one’s will—an assessment of a choice or evaluation. An assessment of will goes deeper than an assessment of the frayed edges of a book. The ability to will (or to reason or to evaluate the worth of this or that) marks us out from mere things, and, it seems, helps to give us a sense of “self.” Our willings, choices, or evaluations, are, arguably, that with which we are essentially identified—that from which we cannot gain reflective distance, at least at the moment of making them. Thus an assessment of your will is indeed an assessment of you, in a rather deep way.

But focusing, in this way, on the depth of blame misses what I take to be the more important point. The characteristic force of a negative moral appraisal is not to be explained in terms of what deep thing in you is being negatively judged—your “agency,” or capacity to evaluate and set ends, rather than your height or hair-color. After all, one’s agency is on display in one’s chess game, one’s cooking, or one’s conversation, and yet

failing in these activities does not have the same significance of moral failure, nor do criticisms of these activities carry the force of blame. I suggest, then, that we distinguish between blame's depth and its force. To see why a judgment of ill will is more forceful than criticism of one's cooking or story-telling, we need to contrast moral assessments with assessments of other sorts, and so contrast moral failure with other sorts of failures.

To assess a thing, generally speaking, is simply to determine how it fares against a standard; to criticize it is to say it fell short of some standard. But it seems that the reactive attitudes to which Strawson so helpfully directed our attention do not concern themselves with assessment of another's will against a standard *per se*, but rather concern themselves with what one person has communicated in his action or attitude about the worth or standing of another: I don't resent you because you failed to achieve some standard of human excellence or failed to live up to the norms internal to willing. Rather, I resent you because your action or attitude communicated disrespect or disregard for me. Likewise I am not grateful because you performed some particularly fine specimen of human action; I am grateful because you chose to benefit me in a way that you needn't have. I am more concerned with myself, or with the relation in which we stand, or with your concern for me, or with how I figure into your world, than I am concerned with your performance, as such—with how you fared against a standard of moral excellence or rectitude or goodness.²⁸ Likewise, if I am indignant about your treatment of another, I am less concerned with the fact that you have performed poorly than with the fact that you have disregarded the worth or standing of a fellow person. Talk of assessment against a standard can therefore mislead. The force of a judgment of disregard is not adequately captured simply by the thought that some especially important or deep

capacity of mine has failed, or even that I have failed in its exercise.²⁹ Rather, the significance of moral failure, I suggest, is to be found in importance of standing in certain sort of relationships with others: call them relationships of mutual regard.

It seems quite plausible to me that standing in relations of mutual regard is of considerable importance to creatures like us.³⁰ Thus the content of a judgment of ill will can carry a certain amount of force—despite being descriptive. If it is true, then you no longer stand in such a relationship. If you acknowledge its truth—if you make the judgment of yourself—that acknowledgement carries the corresponding sense of guilt.³¹ It also seems quite plausible that standing in relations in which the quality of one's will is *recognized*, both by oneself and by others, is of considerable importance. A change in what you or another person thinks about the quality of your will, in itself, changes your relations with them. Insofar as it is important to stand in relations in which goodwill is recognized, the judgment that you have shown ill will itself carries a certain force. That judgment—even if incorrect—makes it the case that you no longer stand in relations in which your good will is recognized on all sides. Thus the force of a judgment of ill will, I suggest, derives from the importance of standing in relations of recognition of mutual regard. The force of a judgment of ill will is found in and carried over from its content, even if the content is merely descriptive.

What of the demands of blame? Notice, first, that any criticism of one's will or reasoning or evaluation brings with it a certain demand: whenever some choice or willing or evaluation of yours is criticized or called into question—be it a strategic move in a chess game, a bit of mathematical reasoning, a career move, or a moral decision—you are called upon to either justify your choice or else acknowledge your error and revise your

ways so as to avoid repeating the error. A judgment of ill will carries these demands simply in virtue of being a negative assessment of one's choice or evaluations. However, calling your choices into question *morally* carries a further demand that questioning your choices in other ways does not. If you wrong me, and this is brought to your attention, then, just as with any other failure, you will be called upon to either justify your choice or to acknowledge your wrong. But, unlike an error in other performances, it will not do for you then simply to alter your future behavior so as to avoid repeating the mistake. Rather, your attention has been drawn to a fact about the relations in which you stand, a fact that needs addressing, perhaps by an apology or by some reparative gesture of goodwill.³² This feature is missing in a case of purely mathematical or strategic error. Again, the aspect of demand characteristic of moral criticism derives from the importance of standing in certain relations with others.³³

Thus the thought that blame involves more than mere grading is correct—in fact, it involves more than even an especially deep grading or an especially important record-keeping. But this is not because blame must involve more than a descriptive judgment—it is not because we must add to such a judgment some further action, attitude, or affect. The idea of grading or assessing against a standard fails to capture the characteristic force of blame, not because it involves only a descriptive judgment, but because it fails to mention the special significance of the moral failing it describes, viz., the importance of standing in relations in which mutual regard is recognized.

The effect of a judgment of ill will on one's relations with others seems to me to capture a central and essential part of the characteristic force of blame. But I do not claim that the relational import of a judgment of ill exhausts the force of blame. Surely blaming

actions are forceful in a further or different way, as is the belief that others are entitled to treat you in this or that way. We came to examine the judgment of ill will by considering the force of the reactive attitudes. But I do not claim that the force of the judgment of ill will exhausts the force of those attitudes, either. I will return to the reactive attitudes in due course. For now I have claimed only that the force of a judgment of ill will, as I have described it, captures a core aspect of the force of blame—an aspect central and essential to both the reactive attitudes and to blaming actions.

THE FAIRNESS OF BLAMING JUDGMENTS

Since the judgment of ill will carries much of the force of blame, I would like to examine, in some detail, the target charge of unfairness as applied to this judgment. Once we notice the way in which a judgment of ill will differs from an ordinary intentional action, we will see that, although this judgment carries much of the force of blame, making this judgment cannot be rendered unfair by that force. When applied to a judgment of ill will, the target charge of unfairness does not make sense.

The target charge of unfairness claims that, due to the characteristic force of blame, blaming the wrongdoer can be unfair unless certain conditions are met. Applied to the judgment of ill will, this charge would claim that judging that a person has shown ill will would be unfair, due to its characteristic force, unless certain conditions are met—unless the person could have done otherwise, or played a certain role in becoming the sort of person she is, or is capable of controlling her behavior by the light of moral reasons. I will now argue that although the judgment of ill will carries force, and although that judgment can be unfairly made, it cannot be rendered unfair by its force.

Let's return to a type of case that tempts one to make the target charge. Suppose that a coworker is both extremely competitive and extremely insensitive, due to untoward formative circumstances beyond her control. As a result, she is constantly putting others down and manipulating people and circumstances so as to gain the best of limited goods. As presently constituted, she is unable to respond to others in ways that show proper regard for them.

Suppose, then, you judge her to have shown disregard towards a colleague. You have thus made a judgment that carries a certain kind of force—it bears a certain importance for your relations with her. You no longer stand in a relation in which regard is recognized.³⁴ While I believe this judgment carries such force regardless of whether the person cares about it, let's stipulate that this person does care about the opinion of others. It pains her that others find her rude and selfish. But, given her history and her constitution, she is unable to translate that hurt into sensitivity. She is presently too insensitive and self-absorbed to treat the interests of others with proper respect.

The target charge, applied to your judgment, would claim it unfair: your judgment brings with it a certain characteristic force, and this force is appropriate only if the person has the ability to avoid the offensive behavior, an ability your coworker lacks, due to circumstances beyond her control. Given her helplessness to change, it is unfair that she should suffer your judgment.

Consider, first, what it would be to try to heed this charge of unfairness. To avoid imposing the force of the judgment, you would have to refrain from judging that your coworker showed ill will. However, the reasons for which you are to refrain from making the judgment—reasons about the fairness of imposing its force—do not bear on

whether the content of the judgment is correct. It is granted, by those who charge unfairness, that this content is correct—she has shown disregard. You may well have reasons that convincingly show this to be so. The reasons for which you are to refrain from the judgment show, not that your coworker did no wrong, but rather that judging her a wrongdoer carries a force which is unfair.

You are now in a bit of a bind. You have what are sometimes called epistemic reasons for judging that someone has done wrong and practical reasons to refrain from making that judgment.³⁵ In such a bind, you can certainly refrain from *expressing* any blaming judgment. You can, that is, avoid partaking in any blaming actions. But, in finding the reasons that bear on whether ill will was shown convincing, you have *already* done what you also judge unfair: you have made a judgment that carries characteristic force. You cannot now, for reasons that do not bear on the content of the judgment, refrain from making it. Nor can you, for such reasons, revise the judgment.

Or, rather, you cannot suspend or revise your blaming judgment *directly* for such reasons—you can't revise or suspend your judgment simply by making a new judgment on the same topic. However, if you happen to have at your disposal some means of self-deception or belief manipulation, you might utilize them to make it the case that you revise or forget the judgment in question. You can, for reasons that show a presumably accurate judgment to have unfair force, act on yourself in ways that will affect your judgment. (Even short of self-deception or self-induced amnesia, you might, if you know in advance that someone cannot fairly suffer the effects of blame, take steps to avoid considering the quality of her will.) Considerations that show the consequences of some judgment unfair can be reasons for engaging in various activities of judgment

management. But you can't, for such reasons, suspend or revise your judgment directly.³⁶

Faced with such a bind, it is tempting to argue that the charge of unfairness is itself unfair: you are charged with unfairness in your blaming judgment, but, in fact, you can't help but make that judgment, because you can't help but believe what is evidently true. If you can't help but make the judgment, then you can't be obliged to refrain from it. You could at most be obliged to partake in certain activities of judgment management to avoid making a blaming judgment or to forcibly eliminate it, if possible.

The temptation to put the argument in terms of fairness should be resisted. To think the charge of unfairness is itself unfair because you lack a power and so can't be obliged to exercise it is to mistake the kind of inability we here encounter. We have not encountered a lack of power but rather a fact about the activity of judging and the reasons for which it is done. The target charge, applied to a judgment of ill will, is not unfair; it is confused.

Judging is not action-like. One does not, and cannot, judge that *p* by representing to oneself the judgment, with its characteristic force and effects, and deciding whether or not so to judge. Rather, one judges by considering whether the content of the judgment is true. One judges that *p* by answering positively the question of whether *p*. The activity of judging just is the activity of answering a question. The reasons by which one answers a question will necessarily be the reasons one takes to bear on it. Thus, insofar as one judges that *p* on the basis of reasons, one will do so on the basis of reasons one takes to bear on the question of whether *p*—a question that makes no mention of your judgment. However, your judgment that *p* is not only your assessment *about* the world, it is also a

fact *in* the world—a fact about your psychology, an event or a state of mind, with its own characteristic psychological and social effects. The judgment of ill will, we have seen, carries a significance of this sort—it is important to others whether you make it. Thus you can have reasons which show that the judgment that *p* is itself, as a fact in the world, good or bad (useful or pernicious, beneficial or burdensome) which do not bear on whether *p*. They bear, instead, on a question *about* the judgment that *p*—whether that judgment is useful or pernicious. By finding such reasons convincing, you will arrive, not at the judgment that *p*, but rather at the *second-order* judgment that the judgment that *p* is useful or pernicious. You might then decide, for these reasons, to engage in some activity of judgment management. That is, you might decide to perform some (perhaps mental) action by which you try to make it the case that you avoid judging *p* or else bring it about that you judge that *p*. Reasons which bear only on whether a judgment is useful or pernicious are thus reasons for managing that judgment, not reasons by which you directly make or revise it.

Thus the bind in which you find yourself does not reveal any lack of power on your part. It rather reveals a fact about the activity of judging and the reasons for which it is done: one judges for reasons one takes to bear on the truth of the judgment. Other reasons, which only show the judgment useful or pernicious in some way, are reasons for which one might engage in the activity of judgment management.

Notice, then, the corresponding ways in which these different activities can be justified or lack justification. You are justified in judging that *p* insofar as you do so by answering for yourself the question of whether *p* on the basis of adequate reasons which bear on that question.³⁷ In contrast, you are justified in any particular activity of

judgment management if you have adequate reasons which bear on the question of whether to bring about or avoid making that judgment. Thus, if one has reasons showing a particular judgment to have bad effects or unfair force, one has reason both to make the second-order judgment that the judgment in question has bad effects or unfair force and to manage one's judgments so as to avoid making that one. If one makes the pernicious judgment nonetheless, and does so for adequate reasons bearing on its content, it is the activity of judgment management, rather than the judging itself, that may stand unjustified.³⁸

To make this stark, suppose that an evil demon has rigged the world so that someone will suffer terribly if you should arrive at a true judgment about some remote bit of mathematics, and suppose you know this. You are thus under an obligation to avoid learning that remote truth. If you nonetheless decide to engage in some sound mathematical reasoning and arrive at the forbidden judgment, you will have done something wrong. But the fault here lies not in drawing the conclusion that p , but rather in deciding to engage in the reasoning that led you to discover that p . Your judgment that p , itself, may well be entirely justified. For a more mundane case, perhaps it is simply no business of yours to know about my financial situation. If you nonetheless conduct an investigation into my monetary resources and conclude that I will default on my loan, you have been illicitly nosy. You may be entirely justified in drawing the conclusion that I will default but unjustified in being in a position to do so. You have again violated obligations of judgment management.

Contrast these failures of judgment management with cases in which you are unjustified in judging that I will default on my loan, and, in particular, unjustified in ways

that can be characterized as unfair. As we have said, you are unjustified in judging that *p* if you do so on the basis of inadequate reasons bearing on whether *p*. Your reasons might be inadequate in either quality or quantity. If your reasons are poor in certain ways, then your judging is unjustified in a way that can be called unfair. Suppose that you, as my loan officer, are entitled to know of my financial situation, but, as it turns out, you judge that I will default on my loan on the basis of prejudice or stereotypes. In such a case, your judgment is unjustified in a way that can be called unfair. Your judgment will also be unjustified and unfair if you judge that I will default without taking enough of the available facts into account before making up your mind—if you are hasty in jumping to your conclusion or inattentive to some range of evidence.³⁹

So one *can* judge that *p* unfairly, if one lacks good reasons showing that *p*. But the target charge of unfairness grants that *p*—it grants that ill will was shown. Further, it might as well also grant that one has good reasons showing that *p*. Thus the target charge of unfairness should grant that one might be justified in one's blaming judgment while nonetheless claiming that so judging is unfair because it is characteristically burdensome. But surely if you do something unfairly, you are, to that extent, unjustified in doing it, and so surely, if you unfairly judge that *p*, then you are, to that extent, unjustified in judging that *p*. Thus the particular charge of unfairness we have been considering, when applied to the judgment of ill will, does not make sense. The fact that making the judgment has its characteristic force—that it carries a certain interpersonal significance—does not bear on whether its content is true. Thus the force of the judgment cannot render the judgment unjustified, and so cannot render so judging unfair. At most this force, if indeed unfair, could generate obligations of judgment management.

FAIRNESS IN SUFFERING

I take this to be a significant conclusion: the characteristic force of a judgment of ill will cannot render one unfair in making that judgment; it could at most generate obligations of judgment management. Yet, one might grant this conclusion—one might grant that *imposing* the force, by making the judgment, cannot be rendered unfair—while still insisting that the force is nonetheless unfairly *suffered* and so generates such obligations (or, indeed, destabilizes our practices of blaming).

Sometimes we do call a judgment “unfair” even though no one did anything unjustified in either managing or making that judgment. Suppose, for example, that you innocently and justifiably yet inaccurately conclude that I will default on my loan. I might then claim that your innocent and justified belief depicts me in an inaccurately bad light and so might call your judgment “unfair.” I would not be claiming that you have been unfair in so judging; rather, I would be saying that your judgment is “unfair” because inaccurate.⁴⁰ Alternatively, if you quite innocently stumble across certain private information, it might seem unfair that you know what you are not entitled to know. Thus one can grant that making a judgment cannot be rendered unfair by its force while nonetheless insisting that the force of such a judgment is unfairly suffered.

We should grant that whether a judgment of ill will is fairly made does not settle whether its force is deserved.⁴¹ If a judgment of ill will is inaccurate (if one’s will was not ill) then it seems clear enough that its force is undeserved—even though the judgment may have been fairly made. If, on the other hand, the judgment is accurate, its force may seem deserved even if the judgment is *unfairly* made. Perhaps the judge jumps hastily to the conclusion that the wrongdoer has shown ill will, but is correct in her hasty judgment. Even though the judge is guilty of hastiness and so has been unfair in jumping to

conclusions, the force of the judgment may yet be deserved. So, whether the judgment is fairly made does not settle whether its force is deserved. In fact, it seems that once we allow that descriptive judgments can carry force, we should allow a gap between fairness and desert, to track the gap between justification and truth. It seems, at least at first, that the judgment is fair if justified but deserved if true.

How could we test this appearance, or establish when the force of a judgment is deserved? In asking whether a person deserves to suffer the force of a judgment, as opposed to asking whether the judgment is fairly made, we are considering the judgment not as an activity but rather as a part of a state of affairs—we are asking, “is it unfair *that* the judgment was made?” rather than “is it unfair *to* make the judgment?”⁴² How states of affairs can be unfair is a substantive question for moral theory; however, it seems clear that states of affairs can be, in some sense, unfair without anyone having done anything unjustified to bring them about. And it seems that accurate judgments are sometimes part of unfair states of affairs. Knowledge that violates another’s privacy, e.g., can seem unfair, even if the belief is justified and accurate. Likewise, the further effects of certain bits of knowledge—the effects of mathematical knowledge in our earlier case of demonic engineering or the effects of inside information in stock-market exchanges, e.g.—can seem unfair, even though the beliefs are accurate. So how can we be sure that the force of a judgment of ill will is deserved if the judgment is accurate?

Unfortunately, as I do not have on hand an account of what makes states of affairs unfair, I cannot address this question directly. Instead, we can notice that, if a state of affairs is genuinely unfair, it should support some individual or collective, other-things-being-equal obligation to mitigate the unfairness where possible. We have seen that, in

the case at hand, this would be an obligation of judgment management—an obligation to avoid learning certain truths. We can reflect on the plausibility of such obligations as a way to assess the plausibility of the claim that the force of an accurate judgment of ill will can be unfair.

Suppose you have arrived at the judgment that your coworker has shown ill will. You also know that, in some sense, she can't help it, and that this judgment of yours pains her. Suppose you also have at your disposal certain means of self-deception or forgetfulness—handy mind-altering pills, say, which would make it the case that you can selectively forget certain judgments without any harmful side effects. Would you then be under an obligation of fairness to bring yourself to revise or forget your judgment? If the force of your judgment really is unfair, it seems you should be so obligated. Notice, if you have knowledge that violates someone's privacy or that causes demonically engineered suffering (or that would qualify you as an inside-trader or disqualify you as a juror), it seems plausible that you might be under some obligation to take such pills. Yet it seems to me implausible to think that you would be under obligations of fairness to utilize the mind-altering pills in order to forget your accurate judgment of public ill will—even though that judgment will cause your coworker pain and will put you in an altered relation with her. I therefore suspect this concern for fairness is not rightly in place.⁴³

Obligations to avoid knowledge seem easily generated in cases in which certain bits of knowledge have contingent effects which are unfair (cases of demonic engineering or insider-trading, e.g.). But it seems that in cases in which the force of the judgment is found in the importance of its content and of other's opinions on that topic, constraints on

knowing are limited to special contexts of privacy. I suggest, then, the following principle for consideration: where the force of a judgment is found in the importance of its content and of others' opinion on that topic, then, absent special contexts of privacy, the force of the judgment is deserved where its content is accurate.

While accepting this principle, we can relocate the unfairness. We can still say that the *entire state of affairs* is unfair—it is, in some sense, unfair that your coworker suffered from the formative circumstances that she did, that she is therefore incorrigible in her disregard, and so is now *rightly* subject to the force of the judgment that she has been disrespectful. This global unfairness will, of course, generate obligations—we ought to remedy the unfairness, where possible. But, by accepting the principle, we insist that the global unfairness does not generate “local” unfairness. *Given* that your coworker is disrespectful, it is not unfair for others to judge her to be so, nor, given her disrespect, is it unfair that she should suffer the force of that accurate judgment.

Why should the unfairness of states of affairs fail to be transitive in this way? Answering this would, again, require a theory about the fairness of states of affairs. I won't attempt to offer one. I will only note that such a theory could allow that certain burdens—viz., the non-contingent force of descriptive judgments—are not generally subject to questions of fairness applied separately to them. Rather, absent special contexts of privacy, the fairness of suffering such burdens will turn on the accuracy of the judgment of which they are an immediate consequence.⁴⁴

So, once it is shown that the force of a descriptive judgment cannot render that judgment unfairly made but can only generate obligations of judgment-management, I think it

plausible to conclude that, absent special contexts, the force of such a judgment is deserved where the judgment is accurate.

If this much is correct, then we can conclude that the conditions for being rightly subject to the force of a judgment of ill will just are the conditions required to make the judgment true. I suspect that much the same conclusion can be drawn for the reactive attitudes. Let's now return to them.

THE FORCE AND FAIRNESS OF THE REACTIVE ATTITUDES

Thus far I have considered the force of blame and claimed that much of it lies in the interpersonal significance of a judgment of ill will. I have argued that the target charge of unfairness does not make sense when applied to this judgment; making such a judgment cannot be rendered unfair by its force. Judging is fair where justified. Further, I have suggested that the force of this judgment is generally deserved where the judgment is accurate.

However, earlier I granted that the force of a judgment of ill will does not itself exhaust the force of blame. Blaming actions, surely, carry further characteristic force. Because they are voluntary, they may be rendered unfair by that force. The reactive attitudes, as well, carry further force. But they are not voluntary. So we now return to the question of whether these attitudes can be rendered unfair by their force.

What, then, is the remaining force of the reactive attitudes? We earlier considered an account according to which these attitudes are simply a judgment of ill will accompanied by a certain affect. But this account is overly simplistic—and not just in its reliance on affect. The reactive attitudes must be sensitive to more than a judgment of ill will if we are to accommodate various moral commonplaces. Moral indignation, e.g., can be

rendered inappropriate, even if disregard is obvious, for reasons that have to do with the standing of the moral judge. If you showed the same disregard yesterday, you have little room to be indignant about it today. I have argued elsewhere that if we are to make sense of forgiveness we must be able to revise our resentment without revising the judgment that ill will was shown.⁴⁵ Additionally—and importantly—these attitudes are conditioned by one’s assessment of the importance or significance or standing of the will that was ill. If the wrongdoer is unimportant, or lacks a certain standing, then one won’t resent or be indignant about her wrongs. So contempt undermines resentment or indignation—one doesn’t dignify the wrongs of the contemptible with such reactions. Likewise, we don’t fully experience the reactive attitudes towards children, and paternalism, even towards an adult, will mitigate these attitudes.⁴⁶ If you should learn that you are not confronting a person, at all, but are instead confronting a brute or a thing, these attitudes are undermined entirely. Finally, in our imperfect world we sometimes simply tolerate certain forms of disregard or ill will. Perhaps a relative or colleague or student is especially insecure or operates with long-entrenched stereotypes. We might acknowledge that some range of her behavior is demeaning, and worthy of resentment, and yet we might simply discount it or overlook it, finding it tolerable, without condoning it, forgiving it, or objectifying its author.⁴⁷

Each of these further complexities—the possibility of failing to be resented, say, in each of these ways—reveals the complex interpersonal significance of these attitudes. If someone fails to resent your display of ill will toward her, it will matter to you why: have you been forgiven, held in contempt, treated as a child, objectified, tolerated, or just ignored? Likewise, when someone does resent your disregard, her resentment will

contain an implicit assumption of her own standing, an assumption that you are not child-like, an acknowledgment of both the importance of your will and the seriousness of the wrong, and a lack of forgiveness. The force of these attitudes, then, is multifaceted.⁴⁸

How it is that resentment and indignation manage to carry this complex, multifaceted force? Affect, we have seen, is not a rich enough vehicle. Nor are these attitudes action-like. I suggest we handle this multifaceted force by taking these attitudes to involve (to be sensitive to, or to entail, or to be the affective face of, or simply a label for) a complex set of judgments or evaluations—a set which necessarily includes, as a unifying feature, the judgment that disregard was shown, but which also includes judgments along a number of other dimensions. Each of these other dimensions is important, as are others' opinions regarding them. In this way these attitudes both carry multifaceted force and yet appear to be reactions to ill will: given the right circumstances and background assumptions, the judgment that someone has shown ill will may well be all that is needed to complete the set, and so all that is needed to resent or be indignant. Yet, the reactive attitude will in fact reveal a larger set of judgments about a number of features of one's relations, and so will carry a multifaceted force.

Determining the members of such a set of judgments or evaluations (or, indeed, whether there are determinate members) will be a difficult task. However, if the force of the reactive attitudes can be understood as the force of a set judgments, then the question of the fairness of imposing this force will be a question of the fairness of making this set of judgments. We have seen, in some detail, that the fairness of making a judgment can be called into question only by considerations that bear on its content. Thus adopting the reactive attitudes could be rendered unfair only by considerations that bear on the content

of the judgments they reveal.⁴⁹ These attitudes would be fair where the judgments they reveal are justified and (one might think) deserved where those judgments are true. The target charge of unfairness, applied to such attitudes, would not make sense.⁵⁰

Of course, if the reactive attitudes are or represent a complex *set* of judgments, it is possible that we will find, among the judgments implicit in them, judgments regarding just those the conditions employed in the target charge of unfairness. That is to say, we could find, in the complex set, the judgment that the person in question could have done otherwise, or is capable of controlling her behavior by the light of moral reasons, or played a certain role in becoming the kind of person who shows disregard for others. We could, then, satisfy the intuitions that fuel the target charge without having to employ it.

If failing to met such a condition were either to mitigate the significance of the disregard or to undercut the importance of the wrongdoer—if it were to show, e.g., that the disregard was minor, or that wrongdoer is no more than a child or isn't a person at all but is rather a beast or a thing or a mere instrument of some other will—then failing to met such a condition would undermine resentment or indignation in ways we have already considered. These attitudes, we have seen, reveal judgments about both the importance of the disregard and the standing of the wrongdoer. Anything that undermines those judgments will render the attitude inappropriate—not because unfair, but because inaccurate.

Suppose, though, that we grant that some significant form of disregard was shown by someone whose will, we want to insist, is both adult and important. Mightn't we still find, among the judgments these attitudes reveal, a judgment about whether the person

could have performed a better action or whether she played a certain role in becoming the kind of person she is?

While there may well be attitudes that include such judgments, and while these may be what many have in mind, I doubt that Strawson's reactive attitudes of resentment and indignation are among them.⁵¹

We considered the view that blame reflects a judgment of moral performance or a form of moral record-keeping. The conditions in question could certainly enter into an attitude that reflects such a grading of moral performance, so long as one thinks that, given the nature of moral performance, such an assessment should, in all fairness, be handicapped for hardship—so long as one is grading moral *effort*, so to speak. An attitude that reflected a grade of moral effort, an assessment of what one has done with the limited resources made available by circumstances, would contain, in it, a judgment about ability or contribution or hardship.⁵²

We have also considered the view that blame reflects the prescriptive judgment that some action—some degree of punishment or sanction or ill treatment—is in order. As we have seen, actions admit of a wide range of conditions. Any attitude that reflects the endorsement of an action will be conditioned by anything that conditions that action.⁵³

Perhaps certain interpersonal attitudes are essentially therapeutic or purpose-driven, adopted in order to bring about change in another. Such attitudes would be action-like, and conditioned, in familiar ways, by concerns of paternalism and of the other's ability to change.⁵⁴ Or perhaps certain attitudes are essentially communicative—that they attempt to relay to the other the importance one gives to relations of mutual goodwill. Such an attitude should be sensitive to the ability or willingness of the other to communicate.⁵⁵

While there may be such attitudes, I believe there is another sort, at least equally worthy of the names resentment and indignation, which Strawson had in mind. While we could, if we were interested, embark on the project of grading another's moral effort, or endorsing this or that form of treatment, or trying to change someone, or communicating with her, the reactive attitudes to which Strawson directed our attention seem to me to be something less than any of these. As he put it, he wanted to draw attention to "the non-detached attitudes and reactions of people directly involved in transactions with each other" and so asked us to consider "the kind of importance we attach to the attitudes and intentions towards us of those who stand in [various personal relationships] to us, and the kinds of *reactive* attitudes and feelings to which we are ourselves prone."⁵⁶ These attitudes and feelings, I suggest, are *simply* reactions to—that is, they simply mark or acknowledge—the importance of a display of ill will or disregard between persons standing in relation. They are thus sensitive to just those facts that make the wrongdoing interpersonally important—facts about one's own standing, the standing of the victim, the standing of the wrongdoer, and the seriousness of the disregard. They are not sensitive to facts that do not change, in some understandable way, the significance of the wrongdoing for one's interpersonal relations.⁵⁷ Thus, the considerations that fuel the target charge will show appear among the judgments contained in these attitudes only if those considerations can change, in some understandable way, the interpersonal significance of the wrongdoing. If such a consideration does change the importance of the wrongdoing, the attitude would be, not unfair, but inaccurate.

The reactive attitudes, so understood, are "non-detached" in two different ways.⁵⁸ First, they do not reflect a grading or an assessment of moral performance from the vantage of

some detached observer or judge, but rather reflect the significance of moral failure for the relations between persons embedded in relation or in community. One does adopt such attitudes by standing back and ranking moral performance, but rather by acknowledging the importance, for oneself and others, of the quality of a person's will towards oneself or those like oneself.

Second, the reactive attitudes, so understood, do not allow room for the sort of "detachment" characteristic of voluntary action. As we have seen, one acts intentionally by answering for oneself a question that includes in it a representation of that action. One acts intentionally by asking and answering for oneself the question of whether so to act. In asking this question, one stands at some reflective distance from the action, considering it. Thus one can consider whether so acting would be fair. However, on the picture of the reactive attitudes I am here advancing, if one has judged that certain things are true, then one has *already* adopted those attitudes. One cannot then stand back and consider whether these attitudes are fair. There is no room to ask that further question.

So understood, resentment and indignation are like distrust. Though distrust also carries a special, characteristic and burdensome force or significance, and though it may be unfair that a person has suffered from formative circumstances that render her generally unreliable and so leave her systematically subject to constant distrust, those who interact with her cannot be unjustified in distrusting her on these grounds. The distrust simply marks the fact that the untrustworthiness is known. Its force is inherited from the significance of untrustworthiness. Likewise, it may be unfair that a person is subject to formative circumstances that render her unable to show proper regard for others and so leave her systematically subject to other's resentment and indignation. But,

once it is granted that the relations in which she stands are of a certain quality, the attitudes which simply acknowledge those facts cannot be unfair.

Finally we can note that, if what it is to be morally responsible just is to be a suitable object of resentment and indignation should one show significant disregard for others, then the conditions that render one morally responsible just are the conditions that put one in characteristically interpersonal relations with others. The abilities required for moral responsibility are those required to stand in the kind of interpersonal relations whose violation carries the significance marked by these attitudes.

This is, of course, Strawson's suggestion of forty-some years ago. Strawson claimed that the conditions of moral responsibility are to be identified with the conditions that render such attitudes appropriate, and the appropriateness of these attitudes depends on the existence and importance of certain familiar interpersonal relations. Strawson made his argument by characterizing our interpersonal relations with others as "practices," and insisting that these practices neither need or admit of "external" (metaphysical or utilitarian) justification. Some have thought that Strawson's arguments left open the possibility of locating, within our moral practices, some condition on resentment or indignation that would apply to everyone if determinism is true, or some condition that shows our practices inconsistent.⁵⁹ One might think that the target charge of unfairness could be so located and then employed to either of these ends.

I have come to much the same conclusion by a different route. I have suggested that the characteristic force of attitudes such as resentment and indignation lies in the force of a set of descriptive judgments. The sort of "external" justification which these attitudes neither need nor admit is, I suggest, the equivalent of a pragmatic justification for a

belief. The inappropriateness such justifications can be traced to facts about the activity of judging and its justification, together with some reflection on the fairness of states of affairs. Thus the conditions on fair blame must be located, not simply “within moral practice,” but rather within the set of considerations that bear on the judgments that give blame its force—that is to say, among the facts render certain displays of disregard important for one’s relations.

CONCLUSION

I hope to have shown that the conditions on being rightly subject to blame are not those render certain burdens fair but rather those that allow for the characteristic interpersonal significance of moral failing. That is to say, the conditions on being rightly subject to blame just are those conditions that provide blame with its characteristic force.⁶⁰

¹ *Proceedings of the British Academy*, 48 (1962): 5–20. Reprinted in Gary Watson, ed. *Free Will* (Oxford: Oxford University Press, 1982): 59–80.

² I am here glossing over what might be an important difference between the thought that wrong was done and the thought that ill will was shown. In this discussion we are interested in ill will or disregard—that is, in those choices, evaluations, or pieces of reasoning that do not take the interests or importance of others rightly into consideration. (I will also gloss over important differences between disregard and ill will. I will use them interchangeably to refer to such choices, evaluations, or pieces of reasoning.) Showing disregard or ill will is, I take it, wrong. However, it may be that one can also do “wrong” innocently, so to speak—i.e., one can do something impermissible without displaying ill will (or without displaying as much ill will as such a violation typically marks). These would be ready cases of excuse. On the other hand, perhaps one can show ill will without doing anything impermissible. Still, it seems, the ill will, itself, would be wrong. The relation between impermissibility, wrongness, and illness of will is complex and of considerable interest, but it is not the topic of this paper. This paper focuses on those “wrongdoings” that involve ill will or disregard—that is, on those choices, evaluations, or pieces of reasoning that do not take the interests or importance of others rightly into consideration.

Alan Gibbard has interesting things to say about the relation between wrongdoing, blame, guilt, resentment, and the contours of morality. See, e.g., his *Wise Choices, Apt Feelings* (Cambridge: Harvard University Press, 1990) and “Moral Concepts: Substance and Sentiment” *Philosophical Perspectives* 6 (1992): 199–221. T. M. Scanlon has pursued some of these questions in “Blame” (presented as The Jack Smart Lecture, Australian National University, July 30, 2003), “Intention and Permissibility I,” *Proceedings of the Aristotelian Society* Suppl. 74 (2000): 301–17, and “Moral Assessment and the Agent’s Point of View” (unpublished).

³ Talk of “identification” was introduced by Harry Frankfurt in a series of articles reprinted in his *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988). Susan Wolf introduced the idea of a “real self.” See her *Freedom within Reason* (New York: Oxford University Press,

1990). (I am skeptical of views which restrict a person's "real self" to some sub-set of her psychology, but that is not the topic of this paper.)

⁴ This case, of course, raises many complicated issues and is badly under-described. Because nothing in my argument turns on the characterization of the particular, proposed conditions under which the force of blame supposedly renders it unfair, the details of the case and of the proposed conditions matter little. It is important only that there are cases in which it is tempting to grant that disregard or ill will was shown and yet to insist that imposing the force of blame is, for some reason, unfair. The reader can choose a case of his or her liking.

If one finds the target charge tempting in cases in which a person's history seems to dictate certain choices, it is worth noting the breadth of this loose description. The particular cases marshaled to illustrate it often involve horrific childhood abuse and deprivation. One is sometimes asked to consider accounts of people raised by seriously debilitated alcoholics and drug addicts, in circumstances of considerable poverty, who go on to cause considerable pain in the lives of intimates and strangers. (For a particularly insightful and influential treatment of such a case, see Gary Watson, "Responsibility and the Limits of Evil," in Ferdinand David Schoeman, ed. *Responsibility, Character and the Emotions* [Cambridge: Cambridge University Press, 1987]: 256–86. For a quite different treatment of such cases, see Sarah Buss, "Justified Wrongdoing," *Noûs* 31, no. 3 [1997]: 337–69.) These are complex cases of serious concern which raise many issues beyond the scope of this discussion. But the loose description fits a wide range of cases: it would capture the callous or brutal actions of the callous or brutal characters sometimes produced by an environment of wealth, power, and ambition; the racist or sexist or bigoted behavior of a member of a particular culture who was raised to think that members of a certain race or sex or ethnicity are less than moral equals; or, in fact, any case in which a person is unable to refrain from wrongful acts in the way you might find yourself unable to betray a friend or turn down a job offer. Just as you might say, "Yes, it will require a lot of sacrifice, but this is exactly the position I have been made for. I couldn't turn it down," someone else might say, "Yes, I see it will cost him his job, but I didn't know, when I promised, that I could double my money. I'm all about making money. I certainly can't turn that down."

For discussion of constraints of will that do not involve any psychological impairment, see Harry Frankfurt "The Importance of What We Care About" and "Rationality and the Unthinkable," both reprinted in *The Importance of What We Care About*, 80–94, and Bernard Williams, "Practical Necessity," reprinted in *Moral Luck* (Cambridge: Cambridge University Press, 1981): 124–31 and "Moral Incapacity," reprinted in *Making Sense of Humanity* (Cambridge: Cambridge University Press, 1995): 46–55.

Finally, it is worth emphasizing that I am *not* arguing that any of these cases are inexcusable. Rather, I am arguing that a particular, proposed form of excuse or exemption—the target charge—is unsuccessful. One could readily claim any of these cases excusable by claiming, e.g., that some feature of the case renders the person incapable of wrongdoing or renders that person's ill will of little importance. By the end of the paper we will be able to see the form I think particular excuses will have to take.

⁵ "Free Will as Involving Determination and Inconceivable Without It," *Mind* 43 (1934): 17–18.

⁶ *Freedom within Reason*, 40.

⁷ *Ibid*, 64.

⁸ On the depth, force, significance, or weight of blame, see, e.g., J. J. C. Smart, "Free-Will, Praise, and Blame" *Mind* 70 (July 1961): 291–306; Jonathan Glover, *Responsibility* (London: Routledge & Kegan Paul, 1970): 56–7, 62–84; Susan Wolf, *Freedom within Reason*, 35–45, 63–65, 77–78; R. Jay Wallace, *Responsibility and the Moral Sentiments* (Cambridge: Harvard University Press, 1996): 52–54, 62; Gary Watson, "Two Faces of Responsibility," *Philosophical Topics* 24 (1996): 227–48; and T. M. Scanlon, *What We Owe to Each Other* (Cambridge: Harvard University Press, 1998): 267–77, and, more recently, "Blame." (The account I will give of the force of blame owes a considerable debt to Scanlon's account in *What We Owe*. His account in "Blame" seems to me a departure from his earlier view.)

⁹ John Martin Fischer and Mark Ravizza, *Responsibility and Control* (Cambridge: Cambridge University Press, 1998): 1. Fischer and Ravizza say this of moral responsibility, rather than blame, but the point is widely taken to apply to blame as well.

¹⁰ Wallace, *Sentiments*, 54. Wallace characterizes blame as involving the reactive attitudes of resentment and indignation and refers to the quoted activities, which express these attitudes, as sanctions (*Sentiments*, 12).

¹¹ This aspect of demand is found in Scanlon's *What We Owe*, "moral criticism claims that an agent has governed herself in a way that [is wrong] . . . this charge calls for her to explain why this claim is mistaken or to acknowledge that it is valid and that her self-governance has been faulty" (268). The demand is discussed explicitly by Wallace in "Scanlon's Contractualism," *Ethics* 112 (April 2002): 441–49. Scanlon's reply is found in the same volume, in his "Reasons, Responsibility, and Reliance: Replies to Wallace, Dworkin, and Deigh," 510–13. Wallace, it seems, thinks of the demand as something in addition to the negative description of the action. Scanlon seems to think of the demand as implicit in that description. I will eventually side with Scanlon. For a nice discussion of Scanlon and Wallace, see Matthew Talbert, "Ultimacy, Normative Competence, and the Significance of Responsibility Claims" (unpublished).

¹² Smart holds that to blame a person is to grade him poorly and to "imply" that he is responsible for his fault, where this means that his fault can be (and presumably should be) influenced by threat, promise, or punishment ("Free Will, Praise, and Blame," 302–305).

For Glover, blaming "consists in adopting an attitude of . . . disapproval towards him on account of what he has done or failed to do" (*Responsibility*, 56).

Wallace and Watson both identify the force of blame as residing in something beyond a descriptive judgment, and they both take this something further to be what raises concerns of fairness. Wallace locates the force of blame in certain attitudes that, he claims, find their "natural expression" in sanctioning behavior: "Blame would be rendered superficial on [an account in which blame simply expresses the belief that certain obligations have been breached], reduced to a way of describing what an agent has done . . . True moral blame, by contrast, is a form of deep assessment, reflecting an attitude towards the agent that finds its natural expression in sanctioning behavior (avoidance, denunciation, reproach, censure, and the like). The reactive attitudes are needed to explain this attitudinal aspect of true moral blame, and to account for its natural connection with sanctioning behavior" (*Sentiments*, 78). The "natural" connection with sanctioning behavior is what, for Wallace, raises concerns of fairness: ". . . this behavioral side of the stance of holding people responsible opens the stance to assessment in light of our practical norms [of fairness]" (*Ibid*, 93–94).

Watson is particularly sensitive to this distinction between descriptive judgments and further responses of action and attitude. He therefore divides moral responsibility into two "faces," the "aretaic" face and the "accountability" face. The aretaic face simply attributes some act or attitude to a person and evaluates the person together with the act or attitude against a moral standard. This face is descriptive. While it is not, Watson believes, a superficial view of blame, he does not take it to raise any concerns of fairness. Rather, those concerns appear, he claims, only for the "accountability" face, or what he calls "holding responsible." This face adds to the aretaic appraisal what he calls "reactive entitlement;" it involves a kind of blame that "suggest[s] that some further response to the agent is (in principle) appropriate," where these responses include not only actions but also attitudes such as resentment and indignation ("Two Faces," 230). Only the accountability side of blame raises concerns of fairness: "Because [our accountability practices] involve the imposition of demands on people . . . they raise issues of fairness that do not arise for aretaic appraisal. It is these concerns about fairness that underlie the requirement of control (or avoidability) as a condition for moral accountability" (*Ibid*, 235). Again, "it is only insofar as and because blaming responses (at least potentially) affect the interests of their objects adversely that moral accountability raises the issues of avoidability that have been central to the traditional topic of moral responsibility" (*Ibid*, 239). While I agree with Watson that "aretaic appraisal" carries the depth of blame, I believe that many would take it to involve its force and fairness as well. I argue, here, that even merely aretaic appraisal carries both depth and force, but cannot be rendered unfair by that force. (Importantly, I will suggest that the reactive attitudes cannot be separated from "aretaic" appraisal.)

¹³ The target charge thus contrasts with views that take either obligation or membership in the moral community to require ability. Unlike the target charge, these alternative views would not admit the initial description of the action as "wrong" or as "falling short of the moral standard." Of course, given the tight

connections between being a member of the moral community, being subject to obligation, and being rightly blamed, one could move from the claim that one of these things requires ability to the conclusion that the others do, as well. One might claim that, because a person does not fall under obligations, he does neither wrong nor right, cannot be rightly blamed, and is not a member of our moral community. Or, one might think that because a person can't fairly suffer the force of blame, he therefore doesn't fall under obligation, does neither wrong nor right, and is not a member of our moral community. The proponent of the target charge would understand the order of explanation to run from the fairness of the force of blame to claims about membership or obligation.

In "Incompatibilism and the Avoidability of Blame" Michael Otsuka argues for what he calls the "Principle of Avoidable Blame" which states, "one is blameworthy for performing an act of a given type only if one could instead have behaved in a manner for which one would have been entirely blameless" (*Ethics* 108, no. 4 [1998]: 688). One way to make the target charge of unfairness appeals to the characteristic force of blame to justify employing this principle to exempt a wrongdoer from blame.

Wallace advances the following principle as placing a condition on holding a person to an obligation: "it is unreasonable to demand that people do something—in a way that potentially exposes them to the harms of moral sanction—if they lack the general power to grasp and comply with the reasons that support the demand" (161). Since, on Wallace's view, one blames only those that one holds to obligations, and since, for Wallace, the unreasonableness at issue is unfairness, this principle states a condition on the fairness of blame. The condition itself—possession of a general power to grasp and comply with moral reasons—is a little abstract. It is a bit unclear to me both what "moral reasons" are and exactly what lacking this power amounts to. But if it is granted that a person who lacks this general power is still capable of showing ill will or disregard (as seems plausible, given an example Wallace uses at 231–33), this principle amounts to the target charge of unfairness: it appeals to the force of blame—in particular, to the harms of moral sanction—to justify placing a condition on the fairness of blaming a wrongdoer.

¹⁴ These ordinary intentional actions, such as confrontation or reproach, are quite often unjustified or inappropriate for many different kinds of reasons, not all of which refer to the fairness of imposing burdens or demands. Confronting a person with her wrongdoing may be unjustified because, e.g., one has, oneself, committed the very same wrong. A confrontation would then be hypocritical. Alternatively, a confrontation may be unjustified because the person is, herself, already engaged in self-reproach and reform. Further reproach from you would be both unnecessary and callous. Expressing a judgment of blame may be uncalled for in certain cases in which the wrongdoer is undergoing certain kinds of suffering not because of concerns of fairness, but because considerations of kindness preclude it.

¹⁵ This is, of course, a substantive claim, but I do not think it a particularly contentious one. If you act intentionally, you act under a particular description. Further, if you act intentionally, you are then subject to certain questions and criticisms, namely, those questions and criticisms which would be answered by the same considerations that would answer the question of whether to have acted under that description (an insight I take from G. E. M. Anscombe, *Intention* [Oxford: Basil Blackwell, 1963]: 9). That you are subject to these questions and criticism shows, I take it, that you have at least implicitly or tacitly answered the question of whether so to act. (If one resists the thought of implicitly or tacitly answering a question, it is enough to grant that, in acting intentionally, one is *answerable* to a question that includes in it a description of that action. Granting this much will secure the possibility of criticizing the action on the basis of its features or effects.) (Note, this understanding of "intentional action" will include as actions certain "mental acts," such as supposing *p* for the sake of argument or imagining there's no heaven.)

¹⁶ *Sentiments*, 70, 78, 82–83.

¹⁷ As we have seen, Wallace's account has this result. One thinks these attitudes appropriate, according to Wallace, only if one "holds a person responsible" for violating an obligation, and it is fair to hold a person to an obligation only if certain conditions are met ("Responsibility," chap. 3 in *Sentiments*).

¹⁸ "Freedom and Resentment," 70.

¹⁹ Strawson talks of beliefs: "The central commonplace I want to insist on is the very great importance that we attach to the attitudes and intentions towards us of other human beings, and the great extent to which

our personal feelings and reactions depend upon, or involve, our *beliefs* about these attitudes and intentions” (*Ibid.*, 62, emphasis added).

²⁰ On “judgment-sensitive attitudes,” see Scanlon, *What We Owe*, 18–22.

²¹ It is important that certain attitudes not be subject to such control. See Harry Frankfurt, “The Importance of What We Care About” and Robert M. Adams, “Involuntary Sins,” *Philosophical Review* 94 (1985): 3–31. (Note, too, that if they were subject to such control, there would be little distinction between sincere and insincere expressions of the attitude—just as there is little room for a distinction between sincere and insincere imaginings.)

It may be worth noting, at this point, that there is an attitude, perhaps sometimes called resentment, which people sometimes adopt towards others as a way of punishing them or making them feel guilty for their wrong. Adopting an attitude of this sort seems to me action-like: goal directed and, in principle, voluntary. This is not, of course, the kind of attitude I have in mind. (A Kantian might note that adopting an attitude in this way involves a kind of contradiction: one pretends, for a purpose, to have an attitude that cannot be had for a purpose if that attitude is to have its significance.)

²² There is, by now, a rich literature on the emotions. For an argument that the appropriateness or inappropriateness of an emotion must take into account the emotion’s “fit” with the world, see Justin D’Arms and Daniel Jacobson, “The Moralistic Fallacy,” *Philosophy and Phenomenological Research* 61, no. 1 (2000): 65–90. Though I prefer to talk about judgments (as I find it cleaner), I believe any account of the emotions which allows that emotions are a reaction to the world that can have or lack “fit” with the way things are would be congenial to my conclusions.

²³ Wallace talks of a stance, but I am not confident that the position I describe is his.

²⁴ I would expect Wallace to deny that his account takes this form. He both argues at length against what he calls “the economy of threats account” and is at pains to distinguish his account from retributivism. The economy of threats account fails, according to Wallace, on three points: it treats blame behavioristically, it is “forward-looking” rather than “backward-looking,” and it fails to explain the apparent tension between responsibility and determinism (*Sentiments*, 54–59). Retributivism is vengeful in a way that blame need not be (*Ibid.*, 59–61). Wallace proposes to solve these problems by appeal to the reactive attitudes. He characterizes these attitudes as follows: “they are focused emotional responses to the violation of moral obligations that we accept. In expressing these emotions . . . we are demonstrating our commitment to certain moral standards, as regulative of social life. Once this point is grasped, blame and moral sanction can be seen to have a positive, perhaps irreplaceable contribution to make to the constitution and maintenance of moral communities: by giving voice to the reactive emotions, these responses help to articulate, and thereby to affirm and deepen, our commitment to a set of common moral obligations” (*Ibid.*, 69).

While emotions thus characterized avoid the difficulties Wallace cites for the economy of threats approach, it is unclear to me that, simply so described, they get us a great distance from the spirit of the views he wants to avoid. This description casts our emotions as expressive of “our commitment to . . . moral standards, as regulative of social life.” The decision to hold you responsible, then, seems to be the decision to admit you into the society of persons subject to moral expectation, where that seems to be the decision that you are a fit object of a certain negative reaction, should you fail to meet those expectations. We need to know why subjecting you to this reaction is neither vengeful nor purpose-driven. I believe the account of the reactive attitudes I will sketch in this paper would show why they are neither (though they may, of course, have regulative effects). But this account does so by understanding these attitudes in terms of a set of descriptive judgments, and so eliminating any place for adopting a “stance” from which they are experienced.

²⁵ It is worth noting that, on this account, it is action, rather than attitude, to which one takes oneself to be entitled. We are, at this point, hoping to give an account of the reactive attitudes, and considering the view that they reflect one’s endorsement of certain forms of action. This position is close to Watson’s view. At one point Watson says the reactive attitudes involve what he calls a “readiness” to treat people in certain ways (“Two Faces,” 239). But earlier he seems to say that holding a person accountable involves thinking oneself entitled both to certain forms of behavior *and* to the reactive attitudes (*Ibid.*, 231). If the reactive

attitudes are themselves among the things to which one counts oneself entitled (rather than the counting-onself-entitled, itself), it remains somewhat unclear what, exactly, they are. (I will consider the possibility that the reactive attitudes contain a self-referential judgment about their own appropriateness in a later note.)

²⁶ “Moral Luck” reprinted in Watson, *Free Will*, 184.

²⁷ *Responsibility*, 44. Nagel likewise makes crucial appeal to what he calls a person’s “moral record,” at 182.

In an interesting footnote, Fischer and Ravizza refer to views which “construe ascriptions of responsibility as primarily *judgments* (concerning an agent’s moral value)” as “ledger views,” and contrast them with the “Strawsonian” approach, which focuses on the reactive attitudes. They note that, on a ledger view, “the reactive attitudes and the associated practices of praising and blaming take on a secondary role, following these primary assessments of moral worth like practical consequences” (*Responsibility and Control*, 9–10n). They remark, “someone might argue that it is conceptually true that if one has a mark on one’s ledger, then one is *eo ipso* an appropriate candidate for the reactive attitudes” (*Ibid*, 10n). While I would strongly resist the “ledger” metaphor (as well as concern with “moral value”), I believe the view I will present would have this result.

²⁸ Of course, your action can communicate disrespect or thoughtfulness towards me only if we in fact have some standards about what counts as respectful or thoughtless behavior. And so talk of standards and assessments against them is not completely out of place. Nonetheless, my primary concern, if I resent what you’ve done, is that you have shown me disregard. (This connects with the interesting issues about wrongness and ill will referenced earlier.) I am grateful to Angela Smith for helpful conversation on this matter.

²⁹ Sometimes people prefer to talk of standards, thinking that it is somehow more respectful to simply inform another person of how they are doing with regard to some standard (which perhaps that person has adopted) than it is to call attention to oneself, or one’s injury. Drawing attention to the standard, rather than to oneself, can seem more loving or selfless. I think this appearance is badly misleading. But this is a separate topic.

³⁰ I wish I had more to say about why standing in such relations is important to creatures like us—self-conscious, self-governing, rational, communicative creatures. I suspect that investigation of this question will lead us back to our capacity for self-consciousness and for choice—just the capacities that the tradition has associated with moral agency.

It is also worth noting that not every breach of such a relationship will have the same significance. It matters, of course, whether the violation of the relationship was initiated by you or by the other. Further, the significance of the wrongdoing will vary with exactly *which* interests of the other were disregarded (likely with both the “objective” and the socially recognized importance of these interests) and to what *degree*. Finally, as we will see, the circumstances surrounding the wrong, as well the characteristics or capacities of the wrongdoer, can change the significance of the violation. Wrongdoing or ill will comes in myriad forms and flavors.

³¹ I say “the corresponding sense of guilt” because guilt takes many forms. Strawson identified it as among the “self-reactive attitudes associated with demands on oneself for others.” (“Freedom and Resentment, 71). Scanlon characterizes a form of guilt that he calls “moral self-reproach” or “guilt in the narrow sense” which is the “realization that one has acted contrary to reasons flowing from [the value of others] as rational creatures” (*What We Owe*, 269–72). J. David Velleman, in a very interesting paper, ties guilt to distrust. See his “Don’t Worry, Feel Guilty” *Philosophy* 52, Suppl. (2003): 235–48. For an excellent series of papers on guilt, see the work of Hebert Morris. (That work was the topic of a special issue of *Law and Philosophy*. See volume 18, no. 4 [1999].)

³² If I am your instructor or your friend, I may criticize your cooking or your chess game and then go on to offer you some suggestions for improving it. But if I tell you that you have offended me in some way, it would be odd indeed if, without looking for some reply from you, I simply went on in the mode of advice, counseling you about how to avoid such blunders in the future.

³³ Wallace, in his criticism of Scanlon, seems to miss the relational aspect of justification in Scanlon's view. He quotes Scanlon as locating the force of blame in the fact that a description of one's moral wrong "differs from mere unwelcome description because it calls for a particular kind of response, such as justification, explanation, and admission of fault" ("The Significance of Choice," in Sterling M. McMurrin, ed., *The Tanner Lectures on Human Values*, vol. 8 [Salt Lake City: University of Utah Press, 1988]: 171). Wallace points out that if we were simply concerned with justification, we would have to assimilate moral criticism to criticism of "peoples' political and aesthetic and scientific opinions," but, as he notes, "It is one thing to criticize a philosopher's views about causation and quite another thing to blame the philosopher for supporting racist or sexist hiring practices; the difference seems to consist in the fact that moral blame has a quality of opprobrium that is lacking in criticism of beliefs or opinions. Hence we cannot hope to account for this special force solely in terms of the connection of moral blame and responsibility with justification" (*Sentiments*, 80–81).

I take the difference to consist, not in a quality of opprobrium, but rather in the importance of standing in certain relations with others. Scanlon takes relations of mutual regard to be relations in which each recognizes the other as one to whom justification is owed. So, for Scanlon, a connection with justification is a connection to the relations I have been highlighting (this fact comes out much more clearly in *What We Owe* than in the earlier lectures to which Wallace is responding).

³⁴ Perhaps you have not stood in *this* relation for some time. Again, I do not mean to imply that there is only two sorts of relation—recognition of mutual good will and its negation. There are, of course, a vast array of possibilities. Additionally, perhaps one can reach the point, with such a person, at which an additional acknowledgement of yet another display of ill will changes one's relations little. This could, of course, be accommodated in the theory I am proposing. In such a case, the additional judgment changes the relation little because earlier judgments have already done the work I have described.

³⁵ In fact, I think the labels "epistemic" and "practical" are misleading. The relevant distinction, here, is between reasons which bear on whether the content of the judgment is correct and reasons which do not bear on this question but nonetheless count in favor of the judgment by showing that it is, in some way, good or useful to make. See my "Controlling Attitudes" (unpublished manuscript).

³⁶ I argue, at length, that one cannot directly form or revise beliefs for reasons that do not bear on the truth of their content in my "Controlling Attitudes" (unpublished manuscript).

³⁷ This is an "internalist" conception of justification: whether one is justified in judging depends on the adequacy of one's reasons for doing so. Since we are ultimately concerned with a moral charge of unfairness, and since I take moral criticism to be directed at the quality of one's reasons, I will focus on this conception of justification. However, it seems the charge of unfairness would fare no better on an externalist conception of justification because even there whether one is warranted in one's judgment turns on whether the mechanism by which one arrives at that judgment reliably tracks the truth. The conditions which supposedly render the blaming judgment unfair do not bear on its truth.

³⁸ Here, perhaps, we encounter a place to employ ought implies can—not as a principle of moral fairness or reasonableness, nor as a principle about obligations and powers, but as a claim about justification: an activity can have or lack justification only insofar as one has or lacks the kind of reasons by which one directly partakes in that activity.

³⁹ Your judgment will be unfair, in such cases, even if accurate.

⁴⁰ I am not entirely happy to concede this point about "unfairness." If we grant that one did not make the judgment unfairly, then the judgment, as an activity, is not unfair. Of course, we can consider the judgment, not as an activity, but simply as part of a state of affairs. We might think this state of affairs somehow unfair—perhaps because the inaccurate judgment has unfair effects. In this case, it seems clearer to me to say that it is unfair *that* the judgment was made, rather than that the judgment is, itself, unfair.

⁴¹ I use "deserved" here simply to mark the fairness of suffering some burden or receiving some benefit—fairness in passivity, so to speak—as opposed to the fairness in the activity of judgment or action. I do not mean to pick up any assumptions about retributivism.

⁴² These can come apart precisely because the judgment is not voluntary—because one does not judge by representing one’s judging and considering whether so judging is appropriate.

⁴³ One way we can become confused about the fairness of blame is by thinking of the harms and demands involved in blame as burdens which are themselves to be allocated according to principles of justice, applied directly to them. As Wallace notes, we can think of “moral responsibility as a system for distributing the benefits and burdens of human favor,” and then “imagine distributive outcomes that we would not consider fair” (*Sentiments*, 201). Because our practices of blaming are social practices, it is very tempting to think of the burdens of blame as costs to be distributed according to principles that regulate the distribution of social resources and risks. Looked at in this way, the burdens do indeed seem unfair—the victim of a morally impoverished childhood is not at fault for the formative circumstances which leave her disrespectful of others, and yet she is blamed for the disrespectful behavior that results. Since the distribution of the negative repercussions of the blaming judgment follows the unfair distribution of these formative circumstances, these burdens can seem unfairly suffered. I will argue, to the contrary, that the unfairness of the entire state of affairs does not require the unfairness of the burdens of blame, taken separately.

⁴⁴ Though I can’t explore it here, this principle would be supported in part by the kind of reflection on agency and justification we have already seen. It may turn out to align with some of the familiar paradoxes surrounding certain “deontological” moral prohibitions.

⁴⁵ See my “Articulating an Uncompromising Forgiveness,” *Philosophy and Phenomenological Research* 62, no. 3 (2001): 529–55.

⁴⁶ For an account of why this is, see Tamar Schapiro, “What is a Child?” *Ethics* 109, no. 4 (1999): 715–38.

⁴⁷ Toleration, as I am here imagining it, contrasts with adopting what Strawson characterized as “objective” attitudes towards a person—opting out of interpersonal relations altogether and treating the person merely as a force to be reckoned. (It may be a matter of treating the person in something like this objective way, over some limited range of issues.) It also contrasts with declining to “hold a person responsible,” as Wallace paints it. If I fail to hold someone responsible, according to Wallace, I think the reactive attitudes inappropriate. But if I merely tolerate some wrong, then I acknowledge these attitudes would, in some measure, be appropriate—I would not criticize you for feeling them in some degree—though I, somehow, tolerate the wrong. (Toleration is quite difficult to conceptualize. See, e.g., the papers collected in David Heyd, ed., *Toleration: An Elusive Virtue* [Princeton: Princeton University Press, 1996].) I am grateful for conversation with Mark Johnson on this point.

⁴⁸ If one identifies the force of blame with the force of resentment or indignation, then, of course, the force of blame is also multifaceted—capable of various modifications and subtleties. Treating blame or its force as a single item, rendered appropriate or inappropriate by the same set of facts, seems to me a mistake. I hope, here, to have displayed a bit of its complexity.

⁴⁹ I am here assuming that these attitudes are adopted by (implicitly or tacitly) making judgments, and so can be fairly or unfairly adopted only in the ways that a judgment can be. However, the view that such attitudes can be rightly criticized only in the way a judgment can be criticized—by appeal to features of its object—can be made plausible even on a “non-judgmentalist” view of emotion. For this, see D’Arms and Jacobson, “The Moralistic Fallacy.”

⁵⁰ One might think that the attitudes could contain or be sensitive to a self-referential judgment about the appropriateness of imposing their characteristic force (e.g., a judgment to the effect that the force of the attitude is fairly suffered in this case). One might then think that the target charge would again have application. If the self-referential judgment does not specify particular conditions of appropriateness, it will result in an action-like attitude. If you can decide that adopting this or that forceful attitude is appropriate or inappropriate for whatever reason, then the attitude will be voluntary—and so could be unfair in just the way an action can be unfair. The target charge would have application. If, instead, the conditions are specified, then it is unimportant (for our discussion) whether the judgment is self-referential or whether the set of judgments simply contain the conditions in question without self-reference. In either case, if the conditions the judgments contain are unmet, the judgment will be undeserved because inaccurate—but it may not, for that, be unfairly made.

⁵¹ If I am right that these attitudes don't already contain such conditions, I don't know what it would mean for us to "revise" them, to bring it about that they do. Here perhaps we find Strawson's appeal to practice.

⁵² If one thinks the assessment in question is the basis for eternal punishment or reward distributed by a just and merciful God, then one might think that the assessment should be conditioned in exactly this way—there should be a handicap for hardship. The distribution should be corrected for luck and for providence. It should be based on only your own, isolated, contribution. But then one might, upon reflection, start to worry whether there will be any basis for distribution left, once the handicap has made its corrections. There but for the grace of God go I, one might think. (Or, "He hath mercy on whom he will have mercy, and whom he will he hardeneth. . . . Why doth he yet find fault?") And one might, then, start to worry about whether one makes any contribution to the world, at all. One gets into considerable tangles thinking of the significance of moral achievement or failure, or even of one's contribution to the world, along the lines of a ranked performance which earns reward or punishment.

⁵³ Wallace thinks of the reactive attitudes as essentially connected with ("find expression in") what he calls "sanctioning activity" (*Sentiments*, 93). Watson holds that they "involve a readiness to adverse treatment" ("Two Faces, 239). Perhaps, then, the attitudes that Wallace and Watson have in mind reflect the judgment that certain forms of treatment are in order, and so are rationally conditioned by whichever considerations would condition that treatment. (On such a reading, the reactive attitudes would be like intentions: not themselves voluntary, but conditioned by whatever conditions the voluntary action with which they are concerned.)

For the claim that resentment need not include any endorsement of or desire to punish, see my "Articulating an Uncompromising Forgiveness," especially pages 539–40 and 547–48.

⁵⁴ These attitudes would contrast with attitudes that merely happen to have certain therapeutic effects—effects which, perhaps, they would not have were they adopted for that end.

⁵⁵ Watson considers the importance of communication in his insightful "Responsibility and the Limits of Evil." I doubt that resentment or indignation are *essentially* communicative. Though they do, not doubt, communicate, and though we will, no doubt, be frustrated by a refusal or inability to "hear" on the part of the other, I don't think the attitudes, themselves, are revised by it. But my intuitions here are not strong.

⁵⁶ "Freedom and Resentment," 62, 63, emphasis in the original.

⁵⁷ I am grateful for conversation with Gary Watson on this point.

⁵⁸ Though these are not, I think, what Strawson had in mind in talking of "detachment," I do think these map onto the two opponents with which he was contending. The tendency to want to detach and judge moral effort (either with an eye to record keeping or with an eye to reward and punishment), and so to miss the inescapable interpersonal significance of moral failing, is characteristic of the pessimist. The tendency to think that we have before us simply a question of how to act—or, worse, of which attitude to adopt, where this is thought to be somehow subject to our discretion—is characteristic of the optimist.

⁵⁹ Wallace discusses Strawson's methodology and the possibility of adopting what he calls "the generalization strategy" in *Sentiments*, 95–103. Nagel's "Moral Luck" seems to me a fine example of an argument that operates within moral practice to generate skepticism about its coherence.

⁶⁰ Earlier versions of this material have benefited from comments and questions from many people, including William Bracken, Franklin Bruno, Tyler Burge, Christopher Ciocchetti, Justin D'Arms, Sharon Dolovich, Bennett Helm, Barbara Herman, Thomas Hill, David Jensen, Daniel Jacobson, Aaron James, Mark Johnson, David Kaplan, Sean Kelsey, Aimée Koeplin, Niko Kolodny, Michael McKenna, Richard Moran, Herbert Morris, Andy Sabl, T. M. Scanlon, Tamar Schapiro, Seana Shiffrin, Angela Smith, Michael Thompson, Manuel Vargas, Gideon Yaffe, members of the Southern California Law and Philosophy Discussion Group, an audience at the Pacific Division APA, an audience at UCSB, and members of the Riverside Autonomy Workshop, especially John Fischer and Gary Watson.

Work on this paper was generously supported by funding from the American Council of Learned Societies.