DOING WITHOUT DESERT

BY

ERIN KELLY

Abstract: The idea of 'moral responsibility' is typically linked with praise and blame, and with the notion of 'the voluntary'. It is often thought that if we are free, in the relevant sense, we may "deserve" praise or blame; otherwise, we do not. But when we look at whether and why we need the notions of praise and blame, we find that they are not as intimately connected with desert as many philosophers have thought. In particular, this paper challenges the idea that forms of evaluation and behavior tied to our "reactive attitudes" (especially resentment) best further morality's aims, properly understood.

Justice – who knows what it is? Do I know? Does your honor know? Can your honor tell me what I deserve? Do you think you can cure the hatreds and maladjustments of the world by hanging them? If you hang these boys, you turn back to the past. I'm pleading for the future.

<div align="right">Orson Welles, Compulsion[1]</div>

The idea of 'moral responsibility' is typically linked with the notions of praise and blame, that is, with certain forms of moral evaluation. It is also usually associated with the notion of freedom, or 'the voluntary.' These ideas are drawn together by the following question: In what sense must we be free, or have acted voluntarily, in order to be appropriately subject to moral praise and blame? It is thought that if we are free (or our actions voluntary), in the relevant sense, we may "deserve" praise or blame; otherwise, we do not. While some philosophers find strict voluntariness to be too stringent a requirement, most concede that if an agent is to be thought deserving of praise and blame, it is important that he or she had or could have had some measure of control over his or her actions. Otherwise put, it is important that the agent had an opportunity to have acted otherwise. Moreover, it is thought that the extent of the agent's blame or praiseworthiness is related to her degree of control.

This brief depiction of the logic and metaphysical underpinnings of the concept of desert seems to me accurate. I will argue, however, that it is a mistake to attempt to rest the justification of praise and blame on that concept. Instead we should look to the normative role we think praise and blame should play in moral life. When we look at whether and why we need the notions of praise and blame, we will find that the forms of evaluation most important to morality are not as intimately connected with the notions of freedom and voluntariness as many philosophers have thought.

In asking whether and why we need the notions of praise and blame, I mean to be asking about the utility of these notions and whether their employment furthers the practical aims of morality.[2] The many powerful criticisms of utilitarianism notwithstanding, we should not lose sight of the important sense in which morality aims fundamentally to bring about a good result. J. J. C. Smart once proposed that instead of asking, "Who is responsible?" we should be asking, "Whom would it be useful to blame?"[3] This proposal comes off as outrageous, but I submit that we should not dismiss it altogether. We should investigate the value of engaging in blaming behavior, and where doing so generally is not morally beneficial, we should refrain from it. Our recoil at Smart's proposal may largely be the result of our rejection of the utilitarian goals of Smart's conception of morality, rather than its practical concern with consequences. The idea that morality's aim is to maximize happiness overall (or to maximize anything else, for that matter) may not be tenable, but although utilitarianism appears to get the practical aims of morality wrong, its teleological orientation with respect to the reasons we have for dealing out blame and various forms of punishment is a virtue.

In what follows, I will explore a certain conception of the nature and practical aims of morality. On this conception, the content of morality is determined by persons who have some concern for one another's basic needs and would be motivated to reason together as equals to reach an agreement on moral principles to regulate their social interactions.[4] In other words, I will assume that the content of morality is (metaphysically speaking) up to us, when we think of ourselves as engaged in a certain sort of common enterprise, and that its normative force is conferred by our common will, or what would be our common will, were we to reason together as equals to further that common enterprise. The moral motivation and basic interests of participants to moral reasoning act as normative constraints on the content of a reasonable agreement. They can be used to assess the relevance of considerations persons introduce; they are, in that sense, procedural constraints. Beyond this, however, there are no prior or independent moral facts that could be used to determine the agreement moral agents should make. Thus there is no room on this view for the position that reasonable persons must agree to treat each other in

a certain way because that is what they deserve. What persons deserve, on this view, can be determined only by how we agree to distribute praise and blame; it does not provide a normative basis for determining what the content of the agreement should be.[5]

An advantage of this philosophical conception of morality is that it avoids relying on the vague and metaphysically contested moral concept of desert. It seems that persons do not share an understanding of what persons deserve, at least not one detailed enough to serve as a guide to blaming behavior. Morality need not assume otherwise; it can serve as a framework of thought within which persons could work out a standard for distributing praise and blame that serves shared interests. This conception of morality is more feasible for a pluralistic and secular society than a morality that relies on the notion of desert. Because it is focused on the practical task of identifying and securing shared aims, I will refer to this conception of morality as a practical conception. Within the framework of a practical conception of morality, we can develop an account of the grounds of moral judgment and the reasons we have for engaging in moral criticism and other blaming behavior.

Praise and blame enter into a practical conception of morality in the following ways. First of all, morality will articulate a set of obligations and prohibitions to guide our behavior. These standards constitute the standards of *moral judgment*; our behavior can be evaluated relative to our moral obligations and prohibitions. But on what I am calling a practical conception of morality, we do not judge persons simply for the sake of judging. Moral judgment does not have a life of its own. It has an important role, but one that must serve morality's practical aims. The second section of this paper explores this role.

In addition to standards for moral judgment, morality also needs mutually acceptable criteria for determining how we are to treat people in view of moral evaluations, especially negative ones. Call these the standards of *blaming behavior*. The third section of this paper argues that the question that should guide our formulation of these standards is the following: what sort of expressions of our moral evaluations do we have reasons to endorse, when we reason together as equals and respect the basic interests of all? Appropriate notions of praise and blame should bring about a good result and not be unfair. I maintain that the best principles to utilize as guides to blaming behavior will protect people's rights but will not be much concerned with free will. They will make claims about the moral relevance of an agent's reasons for acting and her causal responsibility for harms. This information, however, falls short of what is needed to determine whether an agent deserves blame and punishment.

Some will say that it is not enough that our notions of praise and blame promote good results and do so in ways that do not violate people's rights. In addition, it will be claimed, our notions must allow for the

(public) expression and validation of certain moral sentiments; what counts as a mutually acceptable and reasonable way to promote good results cannot neglect our moral sentiments and must be involved with their expression. I agree that there are moral sentiments to which a practical conception of morality should be responsive, but I mean to raise serious questions about whether the forms of evaluation that track our "reactive attitudes" best further morality's practical aims, as they are properly understood. Reactive attitudes, and in particular, resentment, are commonly associated with a desire for retribution – with the thought that it is a good thing in itself that the offender pay a price. The social costs of collectively expressing such attitudes can be very high.

Many of these costs are incurred in the criminal justice system. The United States, for instance, now incarcerates two million persons, two-thirds of whom are non-violent offenders. It has been estimated by some that this is the highest incarceration rate in the world, having recently surpassed that of Russia.[6] The cost is about $40 billion a year. The death penalty, which enjoys widespread popular support, imposes additional costs.[7] In the state of Texas, a single death penalty case costs taxpayers an average of $2.3 million – three times the cost of life imprisonment of one inmate for forty years. In Florida, it costs six times more to execute than to imprison for life without parole. It is estimated that the state of California could save $90 million a year by abolishing capital punishment and that New York could save $118 million. The extra costs of capital punishment would appear to rely for their justification on the value of expressing retributive sentiments, since its deterrent value is in serious doubt.

Of course, the expression of our reactive attitudes, such as resentment and indignation, need not take the form of a legal sanction. We may, more simply, withdraw our goodwill toward a wrongdoer, break off a relationship, or engage in public criticism and censure. But while we may sometimes have good reasons for these responses, it is not clear that they need always, or normally, be regarded as morally desirable.[8] On a practical conception of morality, whether we have reasons to express ourselves in these ways depends on an assessment of the associated benefits.

Part of the practical task of morality is to limit the expression of our reactive attitudes, to call attention to the social and moral costs the expression of such attitudes may impose, and to construct social structures that may better serve our shared needs. Insofar as it is natural for us to seek channels for the public expression of our reactive attitudes, morality may need to serve a corrective function. The public expression of our reactive attitudes, I will argue, should itself be a practical aim of morality in only a limited way. I turn now to a fuller account of such attitudes.

1. *The reactive notion of responsibility*

In his article "Freedom and Resentment," P. F. Strawson illuminates what seems to be an important connection between the idea of moral responsibility and the reactive attitudes, in particular, resentment, love, gratitude, forgiveness, hurt feelings, and the like.[9] I will refer to this notion of responsibility as the reactive notion. Strawson analyzes the moral psychology of praise and blame in terms of the conditions that make possible our experience of the reactive attitudes he identifies. Our reactive attitudes indicate, he claims, that a personal interaction has fallen short of or exceeded our normal and ordinary expectations with regard to the attitudes of other persons toward us, and they represent our personal reactions to this shortcoming or special merit. In particular, our reactive attitudes indicate the importance within personal life of the attitudes of good or ill will (or indifference) that others display towards us in our interactions with them.[10] Resentment is a reaction to injury or indifference, a reaction we have when excusing conditions that would undermine the connection between the harm and ill will do not apply.[11] Strawson thus thinks our reactive attitudes make sense against the background of people's engagement in personal interaction in which ordinary standards of due regard for one another are normally satisfied. His view is that the expectation of this sort of personal regard is simply part of what it is to be involved or to participate in a normal human relationship.

Now our moral notions of condemnation and disapproval, Strawson thinks, can plausibly be thought of as the vicarious or impersonal analogues of these personal reactions: they are reactions to the quality of others' wills, not only towards ourselves, but also towards others.[12] (We may also feel guilt at the vicarious appreciation of others' reactions to our ill will.) These moral attitudes are at the core of our concept of moral responsibility, according to Strawson. In fact, he claims the making of moral demands *just is* the proneness to such attitudes.[13] The demands of our collective sense of morality carry with them the threat of sanction by our negative moral reactions for noncompliance.

Strawson describes our moral attitudes generally as having two notable features, both of which link them with what I will call blaming behavior. First of all, the experiences of moral indignation, disapprobation and condemnation, he claims, "tend to promote an at least partial and temporary withdrawal of goodwill; they do so in proportion as they are strong; and their strength is in general proportioned to what is felt to be the magnitude of the injury and to the degree to which the agent's will is identified with, or indifferent to, it."[14] Thus the first step in our moral reaction is to withdraw, to an extent we deem proportional to the degree of intentional harm, our goodwill. The second is that our attitudes give rise in us to a further interest in punishing the offender. Strawson says,

"[T]he preparedness to acquiesce in that infliction of suffering on the offender which is an essential part of punishment is all of a piece with this whole range of attitudes of which I have been speaking."[15] In these two ways, Strawson asserts an essential connection between our moral attitudes and blaming behavior. If we think of moral judgments as expressing our moral attitudes, then his view seems to be that making a (negative) moral judgment conveys our preparedness to engage in blaming behavior. I will argue that this identification of moral judgment with (potentially) blaming behavior is a mistake.

Strawson thinks we can, however, avoid launching into the varieties of blaming behavior, by deliberately suspending our moral attitudes. We may shift to a more "objective" point of view. This is a point of view from which we regard other persons as suitable subjects for therapeutic treatment or social policy, as persons "to be managed or handled or cured or trained; perhaps simply to be avoided. . . ."[16] From this point of view, we are no longer full participants in human relationships with them. We may observe our interactions with interest and strategize accordingly for the sake of constructive results, but the fact that we are not vulnerable to the full (normal) range of moral reactions means that we are not personally or morally engaged. We have suspended our moral attitudes at the cost of no longer viewing those we interact with as fully capable moral agents, and thus we have lost a moral dynamic that is normally central to interpersonal relations and interactions, or so Strawson claims. As we shall see, however, there is an important sense in which a practical conception of morality narrows the gap between the objective and participant points of view.[17]

Strawson's reactive account has generated considerable attention, for it presents itself as able to avoid metaphysical questions about whether we have free will by showing that the most illuminating account of moral judgment will be given by an analysis of when it is that our reactive attitudes tend to occur and how deeply those attitudes are implicated in our relationships with other people. As we have seen, Strawson's view is that to hold someone responsible is *just* to be vulnerable to these reactive emotions in one's dealings with that person (or vicariously). Moreover, he maintains that the conditions that give rise to the reactive emotions would not be undermined by our acceptance of the truth of causal determinism, because the excusing conditions that tend to diffuse our reactive attitudes, or assessments of incompetence that trigger our shift to the objective point of view, are not examples of how we would be forced to view all human behavior if we were to come to acknowledge the truth of causal determinism. But it is not clear that the reactive account can succeed in substituting moral psychology for a metaphysics of the will, and I believe there is an alternative approach that clarifies the significance of our moral practices while more successfully avoiding the sticky metaphysical issues.

One of the distinctive features of Strawson's account is that he does not attempt directly to justify our reactive attitudes. Rather, he is interested in describing their importance to us in a way that illuminates the essential role they play in our social lives. In that sense, his approach is not strongly normative. Yet it is interesting that discussions of the moral psychology of resentment often lead in a more avowedly normative direction. R. Jay Wallace asks, when *is it fair* to subject someone to our negative moral attitudes? The answer, he claims, refers us to the general rational powers of an agent, what he calls "the powers of reflective self-control." These powers include the general ability of persons "to grasp and apply moral reasons and to regulate their behavior by the light of such reasons."[18] Wallace argues that it is reflective self-control, rather than deep metaphysical freedom, that is presupposed by the psychology of blame and resentment. (When people lack reflective self-control, Strawson confirms, we tend to suspend our reactive attitudes.)

The Strawsonian approach, even as its emphasis is shifted by Wallace and others, retains a deflationary character: it aims to deflate the importance and urgency of sorting out the metaphysical underpinnings of our concept of moral responsibility by stressing that however those philosophical investigations might turn out, we cannot and would not want to give up the reactive attitudes that express our concept of moral responsibility. In other words, the fact that the experience and expression of the reactive attitudes is so highly valued and practically unavoidable within our normal interpersonal relations makes the truth of causal determinism irrelevant to moral philosophy, or so it is claimed. But despite the deflationary ambitions of the reactive account, it remains vulnerable to worries about whether the reflective self-control we possess is thoroughgoing enough to "justify" resentment. This appears to point to an important feature of the moral psychology of resentment: the concept of responsibility it presupposes is the concept of *justified* liability to sanctions (i.e., the sanction of our reactive attitudes and behavior). We seek a justification for resentment and find it (on Wallace's analysis) in the idea that a blameworthy (or praiseworthy) agent possesses a sufficient measure of reflective self-control. Resentment is justified by a defect in the agent's moral reasoning: it is directed at a person who understands moral reasons, and is able to guide her behavior accordingly, but fails to do so. We think such an agent *deserves* the sanction of our reactive attitudes and behavior and, generally, we think that it is a good thing if people get what they deserve.

But why shouldn't we worry, then, about what caused the agent's defect (or success), and what would make it sufficiently willful? The concern to justify our reactive attitudes may lead us to look at the psychological history and social context of the agent's defective choice, and then to be skeptical about whether we can pin down a notion of reflective

self-control that is suitable robust to support judgments of desert. Agents lack control over many factors that influence the formation of their characters, dispositions, and ends. How much control do we really have? We might, for that matter, be skeptical about whether we can articulate a suitably circumscribed notion of the agent. Does Ulysses deserve praise for overcoming his weakness by having himself tied to the mast? He does succeed in overcoming, but not without the assistance of others. Think also about the importance to us that conferring with our friends about important decisions may have. The help this may provide us in coming out all right should lead us to check our reactive attitudes to persons who fail to get help and support under difficult circumstances. Yet we may not really be able to question their capacity for reflective self-control without questioning our own. We might thus come to worry about whether Strawson's catalogue of excusing conditions is thorough enough, or whether his distinction between the objective and participant points of view really holds up.

The reason the reactive account of responsibility is pressed to isolate a person's sphere of reflective self-control is that the account ties moral assessment and blaming behavior closely together. When negative moral assessment is understood to carry with it the withdrawal of goodwill and an interest in inflicting some further form of punishment, and since these are meant to be proportional to the amount of reflective self-control possessed by the agent, the justification of moral assessment depends on a fairly precise measure of the agent's freedom. This is difficult, perhaps even impossible to achieve. But I will argue that we need not accept this close association of moral assessment with blaming behavior and, therefore, we need not worry about the metaphysics of free will. The moral basis for assessing other people's conduct can be separated from our reactions to such assessments, including the reactive interest we may have in engaging in blaming behavior. We can then raise the question of whether something other than our reactive attitudes should guide blaming behavior as well.[19] In fact, we can raise the possibility of developing an account of moral assessment and behavior that bears no essential connection to the reactive attitudes; their role is a crutch we can do without. Rejecting it may have some real advantages.

2.   *Moral judgment*

I propose that the conditions under which moral qualities can be attributed to an agent's character or behavior can be examined and justified apart from our reactions to such attributions. The question whether and when we ought to engage in blaming (or praising) behavior can be treated separately. My proposal is similar in certain respects to one made by

Gary Watson.[20] Watson argues that judgments of an agent's faults need not imply that she deserves adverse treatment or "negative attitudes" in response to her faulty conduct, but may instead concern whether the agent's activities and ways of life are choiceworthy.[21] He claims we are interested in what is choiceworthy in order to develop "models and ideals of human possibility."[22] These consist in examples of virtue and vice that we aspire to emulate and avoid, respectively. Watson considers someone who betrays her ideals by choosing a dull but secure occupation instead of a riskier but potentially more enriching one, or endangers something of deep importance to her life for trivial ends (by sleeping too little and drinking too much before important performances, for example), and he concludes that we might judge that she has acted badly – cowardly, self indulgently, or unwisely.[23] But from this it does not follow that we withdraw our goodwill toward her or are inclined to attempt in some other way to punish her. When it comes to that, we might think that her behavior is none of our business.[24] Rather, our interest in appraising her may be to further our understanding or underscore our conviction about which considerations ought to guide future choices, our own and other people's.

What Watson says seems right thus far, and it emphasizes what I take to be the most fundamental reason why judgment of present and past behavior is important to morality: we rely on such judgments to guide present and future action. I also agree that when we appraise someone's choices or dispositions, we see her as someone who adopts ends for reasons, that is, as someone with evaluative capacities. We think her choices or dispositions can meaningfully be appraised morally insofar as they express evaluative commitments. Watson's point is that the basis of this appraisal does not deliver the justification sought by and for our reactive attitudes.

Recall that the reactive notion of responsibility seems to depend on whether we can isolate and defend the sphere of an agent's reflective self-control. This is important because our reactive attitudes reach beyond what an agent has done; they are directed toward the agent herself. They are involved with a judgment of what she is like, and they presuppose, importantly, that the agent is responsible for what she is like.[25] Our appraisals of an agent's evaluative commitments, however, need not depend on the truth of this presupposition, and thus they may fall short of justifying our reactive attitudes.

Now Watson appears to limit his analysis of "responsibility as attributibility" to assessments of virtue and vice that do not directly entail judgments about whether an agent has violated a moral duty. The latter sort of judgment, he thinks, is tied to the question of when moral sanctions are appropriate. But I do not see why this need be so. While Watson does not commit himself to any particular view of the source of morality's

claims upon us, I have proposed that our moral reasons are determined with reference to principles that could be the object of agreement among persons who care for one another's basic needs and acknowledge one another as equals. These principles can be relied upon to guide our judgments about where our duties and obligations lie and whether or not an agent has satisfied her obligations. But the suitability of these moral principles to serve as a guide to judgment may not entail that our reactive attitudes and blaming behavior are an appropriate response to an agent's failures. This is because an agent may not be sufficiently accountable for her failures.

The basis for our evaluations of what an agent is like would seem potentially to fall short of what is required to justify our reactive attitudes toward that agent even if we attempt to separate the reactive attitudes from blaming behavior, as T. M. Scanlon proposes. Scanlon denies that the justification of our reactive attitudes takes us into the territory of desert. Instead of separating the basis of moral judgment from the reactive attitudes, as Watson does in appraising a person's virtues and vices, Scanlon denies that the notion of responsibility that underlies our reactive attitudes need be understood as a desert-entailing notion.[26] He holds that we may justifiably experience resentment without experiencing a further propensity to engage in punitive behavior or to think that it would be a good thing if the offending party were to suffer. Thus we may speak about what justifies resentment without implying that blaming behavior is thereby warranted. A person is justified in feeling resentment when that person has been wronged.[27] The justification of blaming behavior should be based on further considerations.

Surely it is possible to separate the experience of certain sentiments from the conditions under which one ought to express those sentiments. In that sense, we may agree with Scanlon that our reactive attitudes need not be understood as guided by a desert-entailing notion of responsibility. Even so, it is not clear that the basis for our moral judgment of an agent's choice, action, or character suffices to justify our reactive attitudes toward that agent. I have argued that assessing whether an agent has violated a moral principle that represents a mutual agreement between persons who are morally motivated need not depend on an overall assessment of the agent's rational capabilities, and in particular, it need not lead us deeply into an assessment of the metaphysics of her self-control.[28] Suppose that someone who disregards an accident victim in order to be on time for a business meeting has acted wrongly. We can make this appraisal without knowing what has made her care so much about business meetings or so little about accident victims or whether we can rightly say how fully her values are under her control.[29]

In other words, moral assessment of an action or attitude need not require us to determine the extent to which various causal influences on

an agent's adoption of (good or bad) reasons threaten her reflective self-control.[30] We judge what she is now like only to the extent needed in order to isolate the reasons for her behavior. An agent's evaluative commitment may remain strong even when reflective self-control is uncertain, as in the case of a willing addict. The agent's causal history and the extent of her reflective self-control seem important when we wish to hold her responsible for what she is like. But when our aim is to understand what behavior to model ourselves after or avoid, the agent's attitude, character traits, deliberation and choice stand as model-types to emulate or avoid, and the particularities of her history and abilities become much less important.[31] This means that our moral evaluations may fall short of justifying our reactive attitudes.[32] Certain reactions toward an agent, such as resentment or indignation, may well be understandable, given what she has done or what she is like. But this is not to say that we are giving them a deeper sense of justification in terms of the agent's accountability for her actions and character.

Thus moral appraisal of an agent's action or attitude is less metaphysically complicated than moral appraisal of an agent's responsibility for her choice, character and dispositions. I have maintained that our reactive attitudes are directed toward the latter and that they are bound up with an assessment of the agent's self-control. Of course, some measure of control is involved in adopting and acting on reasons, and Watson concurs with this. He says, "absence of control indicates that the action was not attributable to the agent."[33] His point is that the basis for attributing an action to an agent points to an aspect of our concept of responsibility that is separate from, and no less important than, what is analyzed by the reactive account. On the core notion of "responsibility as attributability," Watson claims, an agent is responsible for her ends in the sense that she has adopted those ends for reasons that express her practical identity as an agent. But here we should be wary. It is misleading to employ the language of responsibility to describe the conditions of attributability because doing so seems to lead naturally to questions about the causal history of and possibilities for the formation of the agent's practical identity – what was the agent's role in shaping his or her own identity? How much control do we have over our characters? Are we responsible for our flaws? More to the point is the idea that an analysis of how actions (or reasons) are morally attributable to agents can bypass the notion of responsibility altogether. Establishing that the conditions of attributability are in place is a matter quite different from investigating the agent's mind and history as required by the moral psychology of blame and desert. The conditions of attributability should thus not be identified with the concept of responsibility.

Moral judgment, as I shall understand it, appraises an agent's reasons for action relative to the practical aims of morality; it assesses the extent

to which those reasons help to realize a mutually agreeable form of social life and are consistent with its principles. We situate actions in order to understand their significance as models of behavior relative to this ideal. Should we emulate them? Should we encourage others to adopt them? Simply attending to consequences, which could be radically contingent, will not enable us to answer these questions. We want to know what the agent intended to do and why. That is, we want to know what justification the agent thought (or could have thought) she had for what she aimed to do. Thus the paradigm case of the object of moral assessment is an agent's action situated in the context of her reasons for acting.[34] (According to Kant, this is captured in an agent's maxim.) But, in particular, we want to know what reasons the agent had for thinking that certain kinds of claims on other persons are warranted – claims that bear upon the rights, liberties, and basic needs that would be acknowledged by moral principles we can share.[35] Moral appraisal remains, in this way, quite focused.[36]

We can think of people's reasons as organized, normatively, in a hierarchy or in chains. Suppose I quickly take my change from the countertop and put it in my pocket so that the clerk will not notice the $5 bill she has mistaken for a $1 bill. There are several reasons here that might figure into an explanation of what I have done. I took myself to have reason (1) to pick up the change quickly and (2) to put it in my pocket. The reasons are instrumental to the end of (3) hiding the clerk's mistake. But so far we have not made much progress in understanding the justification I take myself to have for my action. Suppose that my behavior can further be described as self-interested. We have now ascended into the normative hierarchy of my reasons. A potentially morally neutral action (the quick pocketing of change) serves a morally relevant end (promoting my self-interest at certain costs to others). My opportunistic pursuit of self-interest represents a point in my cluster of reasons that is relevant to moral appraisal and, presumably, the appraisal would be negative.

Situating actions relevantly in their rational context, however, does not necessitate an evaluation of the causal history of an agent's adoption of her reasons for action, or of how an agent became someone who fails, within deliberation and self-regulation, to take certain (moral) reasons seriously. We need not, for instance, understand how I came to be someone who would act in such an instinctively self-serving manner, or whether I could come to alter my character. In that sense, we need not take up the issue of the extent or purity of an agent's reflective self-control. In fact, we need not even be particularly concerned with the agent's practical identity. We are interested in the agent's rationality in a more limited sense. It is enough that the agent has reasons that figure informatively into a certain kind of explanation of what she does. The explanation we seek is one that examines the reasons for an agent's actions in view of

how those reasons and the actions that serve them bear on the moral interests of other persons. Once we have identified a piece (or lack) of reasoning we can appraise morally, we can carry out our appraisal without exploring the causal conditions under which the agent acquired a disposition to reason as she has.

I have said that we are interested in assessing an agent's reasons because we are interested in what kind of behavior to model and emulate. Thus our interest in explaining an agent's actions in accordance with her reasons is connected with an investment we have in taking up a deliberative point of view in order to decide how to act or how to advise others. Nevertheless, the explanatory and deliberative points of view have different presuppositions.[37] When we deliberate about what to do, we aim to decide what to do based upon our assessment of the merits of the alternatives. Making a decision for what we regard as good reasons is fundamentally different from making a predictive claim based upon a survey of the causal factors impinging upon us.[38] The former excludes at the same time doing the latter. As Stuart Hampshire puts it,

It is essential to human action . . . that it should be impossible that while I am deciding what to do, and while I am therefore working towards the state of being certain what I shall do, I should at the same time ask myself the question – 'Knowing myself as I do, what am I in fact likely to do?'[39]

Deliberation proceeds under the presupposition that we are free to act on what we take to be the best reasons, otherwise it would be pointless. This does not mean that we are in fact free or even that we must believe that we are.[40] It only means that when we deliberate we think not about what is likely to happen, but about what we ought to do, in view of our survey of possible reasons for action. We imagine what we would do, if it were the case that nothing but the best reasons to act determine our judgment and will. By contrast, when we explain someone's actions by reference to reasons they have endorsed, we need not assume or imagine that they were free to have deliberated more soundly or to have acted on better reasons. In fact, there is some tension between a presupposition of the agent's freedom and our explanatory ambitions. In explaining human behavior we attempt to show that actions are the necessary result of antecedent conditions. We need not project onto the agent the libertarian presuppositions of the deliberative point of view.

I have argued that in moral appraisal, we aim to identify how a person's actions or dispositions bear on the morally relevant interests of other persons. We then attempt to determine why the action or disposition in question seems reasonable from the agent's point of view, and whether or not we can endorse that judgment. If the agent's action or disposition lacks normative structure because the agent appears to have no

evaluative commitments that bear on what she has done – for instance, what happened was an accident – then moral appraisal comes to a dead-end. It comes to a dead-end, that is, unless the agent was accident-prone for morally interesting reasons. It is useful here to draw an analogy with the "principle of quiescence" in chess. The principle states: always look a few moves beyond any flurry of exchanges to see what the board looks like when it quiets down.[41] In ethical assessment, we examine the network of reasons that explain what an agent does or is disposed to do, searching for "hot spots" where the agent's reasons (or lack thereof) have a bearing on the moral interests of other persons. These hot spots become the focus of moral appraisal, and what lies beyond them in the agent's norm-ative network of reasons is of diminishing import unless examining them calls attention to other hot spots.

The point is that we can do this without taking up the question of whether or not the agent has the capacity to adopt and live up to better reasons (or, for that matter, whether the agent took herself to have the reasons it makes sense to ascribe to her). We need not determine whether the agent freely or autonomously adopted her reasons or, if she acted badly, whether she could have known better or cared more about acting well. Thus the question of whether the agent "deserves" the moral assess-ment we confer upon her, i.e., how responsible the agent is for having turned out as she did, is drained of its urgency. The moral judgment of agency (i.e., the production of actions in accordance with reasons) should be less concerned with the causal history of an agent's adoption of rea-sons for action than many who have worried about moral responsibility have thought.

3.   *Blaming behavior*

My account separates our interest in and grounds for moral judgment from the reactive attitudes. I now take up the question when or whether we should engage in blaming behavior. By 'blaming behavior' I intend to refer to the expression of moral judgment and criticism, as well as to more intrusive forms of interference and physical restraint. In addressing the matter of blaming behavior, do our reactive attitudes surface again as essential guides to what is appropriate? On Watson's view they do, but I disagree. He is right that responsibility as attributability is not enough to tell us how to act toward an offender, but wrong to think that the react-ive account of responsibility is what must further direct us. In other words, I believe that the reactive account of responsibility can be separ-ated as well from the question of how morality should counsel us to act toward one another. What counts as fair and appropriate treatment, especially with regard to those persons who have acted badly, is not

determined by consulting our reactive attitudes. If this is true there are, of course, (further) deflationary implications for the reactive notion of responsibility – it would be left with a highly restricted domain.

I have argued that moral evaluation is concerned with the questions: Can we endorse the agent's aim in acting, or did the agent act for the wrong reasons? I will now defend the idea that if the agent's reasons are bad, we have grounds for expressing our negative moral judgments, because and insofar as doing so furthers aims that mutually concerned parties could endorse. I give only a bare sketch of how this might go, but enough to suggest a variety of reasons that need not rely on the notion of desert. Below are four ways in which the activity of moral criticism and censure can promote moral interests that we share:

(1) Moral criticism and censure can have deterrent force. They have this force for those who assign importance to their membership in the moral community, since for them moral failure is uncomfortable and to be avoided.[42] Criticism may lead to a person's loss of self-esteem as well as to a loss of her esteem in the eyes of others, by calling attention to her failure to live up to moral principles she accepts and takes herself to share with others. Deterrence is a familiar reason for moral censure and I won't say more about it.[43] The principle worry about forward-looking accounts of blaming behavior that invoke the value of deterrence is that such accounts are implausible when they rest entirely on that value.[44] So I turn to the further reasons.

(2) Moral criticism and censure can be instructive. They serve as an occasion to clarify on publicly accessible grounds the principles we expect others to accept if they are reasonable.[45] They do so in a context in which we address those reasons either directly or indirectly to someone who has acted for morally unacceptable reasons. This can provide the opportunity for further reflection on and argument about the nature of our moral principles. The grounds of moral justification are thus examined publicly and, if all goes well, they are strengthened.

(3) Moral criticism and censure serve the end of reconciliation. Understanding and acknowledging harms suffered is an important step in the process of recovery from them, and it is often especially important for this acknowledgment to come from people the harmed person respects and has confidence in, i.e., members of what she identifies as her moral community. These will be persons who share the moral principles on which the criticism is based.

(4) Finally, the activities of moral criticism and censure can facilitate a sense of solidarity in the moral community.[46] This point must be treated carefully. Solidarity in the relevant sense is created not through a sense of righteousness, but by underscoring a collective commitment to bringing about a better society. As I construe it, this is a commitment we make to helping one another to regulate our conduct by appropriate (shared)

moral principles, and to the idea that this commitment has a high priority in our overall scheme of values.[47]

None of the points I have gone over so far need involve significant interference with "blameworthy" persons, so clearly the account given thus far is incomplete. What about the justification of penalties and restraints? On this question, we should take seriously the large role played by (simple) causal responsibility, as opposed to moral desert, in justifying the regulation and restraint of persons who cause or imminently threaten harms to others.[48] The role within morality of what most would regard as a shallow notion of causal responsibility is under-appreciated by moral philosophers; judgments of causal responsibility contribute substantially to our judgment about proper treatment. If true, this means that the moral justification of our appropriate responses to and interference with the behavior of persons who cause harms should be less preoccupied with our assessment of the agent's will, much less its freedom, than many have thought.

The guiding principle is simple: we are usually justified in preventing harms by interfering with their causes.[49] An agent's causing harm to others provides us with grounds for restraining her. This idea receives considerable support from common sense and social policy. For example, we often restrain persons who have assaulted others and continue to pose a threat. We seem to feel justified in doing this pretty much regardless of the origins of the aggressor's anti-social behavior, provided that we do not violate his rights. On the approach I have endorsed, a person's rights are determined by reference to principles that equal and mutually concerned moral agents could agree upon. They protect our moral status and our freedom to lead a life compatible with a similar freedom for others. The limits of freedom should be negotiated with due attention to the equal moral status and basic interests of all persons. Restrictions would be unlikely to secure agreement unless they are necessary for the protection of shared and fundamental interests. When harms constitute a violation of such interests they provide grounds for restricting the harmdoer. If our understanding of the relevance of causal responsibility to the justification of penalties and restraints is constrained by this notion of people's rights, we can mostly avoid evaluations of the origins of harmful behavior. The origins of an aggressor's anti-social behavior may, however, enter into determining which restraint is appropriate. If the aggressor is not capable of being deterred by a prison sentence (much less by moral criticism), we may opt for some form of treatment instead. The deterrent value of punishment is not undermined by exempting some persons.[50] On the principle I am examining, the rationale for restraint is preventing harms, and more severe restraint than would be necessary to achieve this aim is not warranted.

The moral significance of the role of causal responsibility is highlighted by what we would concede to the interests of others from the first-person

perspective. We tend to believe it is not unreasonable for other persons to prevent us from harming them, even if we do not think that the threat we pose to them is our fault. For example, if I fall off a roof and am about to land on someone, I would not think it unreasonable of that person to move out of the way, even if that would mean more serious injuries for me. Or if a hypnotist causes me to aim a loaded gun at someone and instructs me to shoot, I would not think it would be unreasonable for the intended victim to preempt me, if that is the only way she could protect herself. Similarly, if I become infected with a serious and highly contagious disease, quarantine may be justified even supposing that I have done nothing wrong. According to the practical principle we are considering, however, the use of restraints must head off harms that would otherwise (imminently) occur, or it is not justified. Restraint is not here understood as a form of retaliation.

Consider another sort of case. We sometimes think it is justifiable to fight against and attempt to contain a society or group of persons who have unjustly aggressed or who threaten to aggress against another society or group (e.g., Serbian aggression in Kosovo), and that we may do this without investigating the source of the aggression in culture or history or political agendas of social control to see whether those who perpetuate it "deserve" the harms we may inflict upon them in response. Here it is important that the threat be imminent, if a transgression has not already occurred. If it has occurred, we must reasonably believe the threat continues to be a live one. In choosing our targets, we attempt (or ought to attempt) to incapacitate those who are causally responsible for the perpetuation of the unjust aggression, and to use minimum necessary force.

Also, we commonly think (or, I submit, we ought to think) that it is justified for a group to rebel against another group that oppresses it (e.g., the struggle against apartheid) without first investigating whether the individual members of the dominant group are "responsible" for their causal role in the oppression.[51] These acts of self-defense, as we might call them, are not about dealing out desert. They are motivated by concern to stop harms from occurring and to prevent them from reoccurring. This motivation is at the heart of a practical conception of morality.

These examples have focused on the moral importance of identifying and restraining individuals who have not only caused harms, but also continue to pose threats. Restraints are justified to stop harms from occurring and to deter further harms threatened by the agent who is subject to our blaming behavior, and only to the extent necessary to accomplish this. But I acknowledge as well that those who have caused harms yet do not pose further threats may rightly be penalized for the sake of the three additional reason-types outlined above (instruction, reconciliation, and solidarity). In these cases, of course, we reach beyond

the principle that allows us to prevent harms by interfering with their causes and return to the various reasons that justify moral criticism. In order to be justified, the penalties inflicted must advance these aims and not exceed what is required for doing so. This means restraints and other penalties can be justified only when it is reasonable to think that the agent (1) was causally responsible for harms or imminent threats, and (2) had morally faulty reasons for acting.[52]

4. *The moral sentiments*

I may seem to be coming dangerously close to the sort of flat, reductive view found in utilitarian accounts, one that does not appear to do justice to the role and complexity of our moral sentiments. Let me briefly turn to the question of how moral emotions bear upon or appear within a practical conception of morality.

I have been criticizing accounts of moral responsibility (and, in particular, accounts of moral blame) that are centered around the reactive attitudes: resentment, indignation, and guilt. It is important that in rejecting the centrality of the reactive attitudes to the practical aims and orientation of morality, I am not rejecting altogether the role and importance of moral sentiments within morality. My positive account can and does maintain a sensitivity to moral emotions, although it presents a somewhat alternative take on which emotions are central to morality. Very briefly, I will make this case.

The view of morality I have been advocating focuses on our collective attempt to approximate an ideal, namely, a mutually acceptable form of social life that is suitably responsive to people's basic needs. The class of emotions most relevant to consider thus concerns our relationship to ideals: what it means to us to pursue an ideal, to fall short of it, or to embody it, however imperfectly. Pursuing an ideal may involve, for instance, the attitudes of respect and hope. The gains we make can give rise to pride and joy. Failing to attain an ideal may bring disappointment, sadness, anger, regret, and even shame.[53] Emotions such as these will be moral when the ideal to which they are responsive has moral content. I have proposed an abstract characterization of such an ideal. We may feel moral emotions, however, not simply with regard to abstract moral principles and the ideal they serve (although we may feel them in this way). We may also feel them toward other persons with whom we share a commitment (more or less) to realizing this ideal in human relationships. The ideal, after all, makes sense only because human beings and how they relate to one another matters. Our moral judgments are sensitive to what persons are like: how helpful and even adventurous they are in realizing our common sense of morality, and such assessments give rise,

importantly, to emotions such as the ones I have indicated. From the point of view of a practical conception of morality, we tend to feel respect and appreciation for someone's virtues. When persons fail, we feel disappointment in their shortcomings. Our disappointment is measured by our sense of the possibilities and opportunities lost, and is not, like resentment, fixed by a notion of what a person deserves.[54] Even our anger need not depend on our sense of how responsible persons are for having turned out as they did: what part exactly they played in their own failure. We may be sad, regretful, or angry that a person, for whatever reasons, did not turn out to be the person we hoped she would be, or thought she might have been. When human relationships work out well, on the other hand, we may celebrate the moral possibilities realized in and through them, and this can give rise to mutual love and admiration. The role each person has played in a relationship or community may not be sharply differentiated; and that is partly why the love and admiration that may emerge is not centered around feelings of gratitude. We may also feel gratitude toward persons who have done more than is expected, but this gratitude need not be thought of as the counterpart of resentment. It may not be important to distinguish the agent's "responsibility" for her accomplishment from the conditions that enabled it: we may simply be grateful for its possibility.

5.   *Beyond the psychology of blame*

I have argued that the prominence of the reactive attitudes in moral thought is not defensible on moral grounds. What, then, accounts for their persistence? Since a normative account is lacking, we must turn to psychology, history and sociology for an answer.

The best-known psychological account comes from Nietzsche.[55] Nietzsche thought that the notion of the will itself is an artifact of the psychology of *ressentiment*: when we become conscious of our powerlessness, and especially when others hurt us, we come to hate the power associated with their agency. In an attempt to give meaning to our suffering and to gain control over it when we are unable to act on it, we moralize this hatred and express it in the form of resentment and moral blame. In its purest form, blame attempts to isolate the power of agency in "the will." It positions the agent clearly behind and separate from the deed as its proper focus. It is thereby possible to posit that the harm need not have occurred – the agent could have prevented it by an act of will.[56] This frees us from an inevitable consignment to the status of victim.[57] Nietzsche maintained that, in this way, the psychology of blame reflects and serves a position of weakness: the perspective of the powerless. It serves this position by reactively elevating the moral status of the victim

above the "evil" of the agent who is supposed to have caused harm intentionally and voluntarily.[58] But this moral perspective is not liberating, he argued, for it is ultimately involved with the ascetic denial of desires and impulses, most notably our anger and aggression, that are implicated in the causal order of nature – a denial that is especially destructive and costly when it becomes internalized and thought to be good in itself.

Nietzsche's story seems to me plausible, but its details need not concern us here.[59] The point is that if the argument of this paper is valid, normative ethics will not be able to explain our attachment to the reactive attitudes. This means some other account must be found. If Nietzsche's story is not entirely correct, then something like it must be true.

The alternative to a psychology of blame I have proposed is to free our investment in the power of rational thought and control from the reactive attitudes and to connect it instead to the possibility of realizing the cooperative ideal of a just society. Justice, as I understand it, is not the meting out of resentment, but represents an alternative to and an overcoming of it.[60] While the Strawsonian picture rightly points to the fact of our investment in rationality, and to its importance to social relations, our investment in rationality need not be so tightly linked to the reactive attitudes. The reactive account of responsibility Strawson presents as an alternative to a metaphysics of desert does not succeed because the quest for a metaphysics of desert is driven by our reactive attitudes. It is because we seek a justification for our reactive attitudes and the blaming behavior that so often accompanies them that we attempt to stake out the conditions of desert. A practical conception of morality, on the other hand, can abandon the reactive account of responsibility and its metaphysics of free will and desert; it represents a collective response to our needs and conflicts that can replace the relief (perhaps misguidedly) sought in our experience of the reactive attitudes. The forward-looking, constructive quality of moral thought means that morality involves, in a sense, taking what Strawson refers to as an "objective" attitude in the collective construction and realization of moral ideals: it requires us to engage in practical reasoning about how to define and satisfy our aims as moral agents.

Recall the features of a practical conception of morality. Moral motivation involves each party taking some interest in the morally urgent needs of other people. The norms of morality are not the outcome of self-interested bargaining between persons. They presuppose that moral persons will take an interest in realizing a form of association that affirms the value of each person's basic interests, provided that others do so as well. This is expressed in persons' willingness to reason together with others as equals to arrive at a set of moral norms all who are similarly motivated could share. We might think of the "objective" point of view, then, as prescribing that we take one another's morally basic interests

and equal status into account in our social interactions. Morality would have us acknowledge the intersubjective validity of a certain set of needs and interests and the rights and liberties we ascribe to persons in view of these needs and interests.

A form of social life that is responsive to the morally urgent needs of each person must cultivate incentives for promoting shared aims.[61] Ideally, this involves enabling each person to realize his or her capacity and desire for moral agency. I have suggested that our moral attitudes are largely directed toward the relevant potentials and capacities in persons, and that our respect for persons as moral beings is expressed in our acknowledgment of these potentials and capacities.[62] When persons fail to realize their potential as moral agents, however, practical deliberation involves negotiating how acceptably to distribute this burden in the interest of most effectively promoting shared practical aims. This can lead us to important forward-looking rationales for identifying and responding to past wrongs. Expressing our disapproval of people's wrongdoing and penalizing them for it can play a crucial role in the process of halting abuses, compensating victims, returning stolen property, redistributing wealth, overcoming ethnic and racial conflict and, generally speaking, working toward greater justice. But this is a matter quite different from dealing out blame on the basis of desert. Our interest in the "stains" on people's records remains firmly linked to our interest in bettering the prospects for a tolerant and just society.

Thus we can see that on this view, morality is not the impersonal analogue of resentment, but an alternative to it. Morality need not attempt to isolate the self-generating power (or self-control) of the rational will. It depends not on a presumption of the autonomous forms of self-control that seem to underwrite the reactive attitudes but, rather, on the possibility of cultivating reasonable and just interrelations that generally offer sufficient incentives for compliance and safeguards for dealing with the failure of some persons to be motivated by morality.

I have argued that we should understand the moral value of the activities of praise and blame to be relative to the practical aims and ambitions of a morality. We should not confuse the psychology of blame and resentment with the moral interest we have in cultivating people's rationality (i.e., our capacity to recognize and act for reasons) for collectively acceptable aims and purposes. Our interest in fairness, furthermore, can be detached from our preoccupation with justifying resentment and extends more appropriately to the matter of how we are justified in treating others as we pursue our collective moral aims. I do not deny that people experience resentment and may sometimes have reason to seek acceptable channels for its expression. The psychology of resentment, the appeal of blame as punishment to inflict suffering because we think it is deserved, may well be a primitive psychology we may not be able fully to

divorce ourselves from.[63] The importance to people of expressing such attitudes is something that may need to be taken into account in our social policy on punishment. But we should attempt to minimize the importance of these expressions, for their social costs are high, they rest on dubious metaphysics, and they are not essentially linked to the practical aims of morality.

Department of Philosophy
Tufts University

NOTES

¹ *Compulsion*, dir. Richard Fleisher, with Orson Welles, Diane Varsi, Dean Stockwell, 20ᵗʰ Century Fox, 1959.

² Because my account justifies praise and blame directly by appeal to the aims and principles of morality, rather than via an analysis of the conditions under which agents act freely, it can be thought of as a version of what Stephen White calls "direct compatibalism." See his *Unity of the Self*. Cambridge, MA: MIT Press, 1991, p. 236.

³ Smart, J. J. C. (1973). "An Outline of a System of Utilitarian Ethics," in J. J. C. Smart and Bernard Williams (eds.) *Utilitarianism: For and Against*. Cambridge: Cambridge University Press, p. 54.

⁴ For a similar philosophical conception of morality, see Scanlon, T. M. (1998). *What We Owe To Each Other*. Cambridge, MA: Harvard University Press, Ch. 5; and Habermas, Jürgen (1990). "Discourse Ethics: Notes on a Program of Philosophical Justification," in *Moral Consciousness and Communicative Action*, tr. Christian Lenhardt and Shierry Weber Nicholsen. Cambridge, MA: The MIT Press.

⁵ Cf. Rawls's contrast between the idea of legitimate expectations within an institutional framework and a pre-institutional concept of desert. In Erin Kelly (ed.) (2001). *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press, pp. 73–79. For criticism of political liberalism's avoidance of the notion of desert, see Scheffler, Samuel (1992). "Responsibility, Reactive Attitudes, and Liberalism in Philosophy and Politics," *Philosophy and Public Affairs* 21, no. 4, pp. 299–323.

⁶ Lewis, Anthony (1999). "Abroad at Home; Punishing the Country," *The New York Times* 21 Dec., late ed.: A31.

⁷ The following statistics are from Richard C. Dieter (1997). "Millions Misspent: What Politicians Don't Say About the Hight Costs of the Death Penalty," in Hugo Adam Bedau (ed.) *The Death Penalty in America*. New York: Oxford University Press, pp. 402–3.

⁸ John Stuart Mill pointed to the power of the sanction of public opinion, regardless of legal enforcement. He was especially sensitive to the morally detrimental effects of expressions of the majority opinion. See (1859) *On Liberty*.

[9] Strawson, Peter (1982). "Freedom and Resentment," in Gary Watson (ed.) *Free Will*. Oxford: Oxford University Press, 1982, p. 62.

[10] Strawson, p. 71.

[11] Strawson, p. 70. Strawson groups excusing conditions into two categories. The first consists in conditions that undermine the connection between the injury and the agent's ill will (e.g., the agent was pushed). The second category concerns the existence of abnormal conditions that reveal that the ill will demonstrated is uncharacteristic of the agent (e.g., the agent was under great strain, or affected by posthypnotic suggestion, et cetera), or that the agent herself is psychologically abnormal or undeveloped. See p. 64ff.

[12] Strawson, p. 70.

[13] Strawson, p. 77.

[14] Strawson, p. 77.

[15] Strawson, p. 77.

[16] Strawson, p. 66.

[17] Psychiatric treatment itself may challenge the distinction between the objective and participatory stances by requiring the therapist to walk the line between the two. See, for instance, Dennett's interesting discussion of how successful therapeutic treatment of "multiple personality disorders" may require that the therapist demonstrate an attitude of respect for the alters, without actually endorsing or, one might say, fully embracing the participatory stance with regard to them. Dennett, Daniel (1998). *Brainchildren*. London: Penguin Books, pp. 57–8. The understanding of morality and agency I develop in this paper is intended to press us to think about how the objective and participatory stances intersect in our ordinary social and moral relations, and possibly also in the deliberative stance we assume with respect to our own action.

[18] Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, p. 155.

[19] My criticism here is similar to one made by T. M. Scanlon, although, as discussed below, the view I defend breaks the connection between moral judgment and the reactive attitudes more radically than his does. See Scanlon, T. M. (1988). "The Significance of Choice," *The Tanner Lectures on Human Values 1988*. Salt Lake City: University of Utah Press, pp. 160–166.

[20] See also Scanlon, T. M. (1988). *What We Owe To Each Other*, Ch. 6; Smart, J. J. C. (1961). "Free Will, Praise and Blame," *Mind* 70, pp. 291–306; Feinberg, Joel (1991). "Collective Responsibility," Larry May and Stacey Hoffman (eds.) *Collective Responsibility*. Savage, MD: Rowman and Littlefield, pp. 74–5; and Wallace, p. 52ff.

[21] Watson, Gary (1996). "Two Faces of Responsibility," *Philosophical Topics* 24, p. 231.

[22] Watson, p. 243.

[23] Watson, p. 231.

[24] Watson, p. 231.

[25] This concerns what Susan Wolf refers to as "deep responsibility." Wolf, S. (1990). *Freedom within Reason*. New York: Oxford University Press, p. 41.

[26] Scanlon, *What We Owe To Each Other*, pp. 274, 276–77.

[27] Scanlon, *What We Owe To Each Other*, p. 271.

[28] Scanlon would agree with this because he believes that the justification of our reactive attitudes is provided by the reasons we have for thinking that an agent has acted wrongly. I question his claim that the reactive attitudes do not presuppose that an agent is responsible for her actions and character. But perhaps our attitude of resentment, for instance, can take different forms. See footnote 54 below.

[29] Cf. Slote, Michael (1990). "Ethics Without Free Will," *Social Theory and Practice* 16, no. 3, pp. 369–383. Slote describes a non-utilitarian and non-egoistic virtue ethics that

avoids the notions of moral praise, blame and responsibility altogether. The account I offer does not eliminate reference to praise and blame in this way, but it nevertheless bears some resemblance to Slote's approach.

[30] Scanlon denies that moral assessment requires us to determine whether a wrongdoer is responsible for having become the kind of person she now is. (*What We Owe To Each Other*, p. 284) In that respect, his view is close to mine. Nevertheless, Scanlon is more concerned than I am with whether the agent has a coherent psychology that is stable over time. (pp. 278–79) I believe that this is because he is interested not simply in articulating the grounds for moral judgment, but also in the further task of justifying our reactive attitudes.

[31] This is so even though character traits cannot be understood apart from knowing the person. We may know a person well without being familiar with her developmental history.

[32] Scanlon holds that the justification of our reactive attitudes need not presuppose that the wrongdoer is responsible for what she is like. (*What We Owe To Each Other*, pp. 284–85) I believe this significantly revises our familiar concept of resentment. See n. 54.

[33] Watson, p. 234.

[34] For a discussion of the complexities of moral interpretation, see Moody-Adams, Michele M. (1997). *Fieldwork in Familiar Places: Morality, Culture, and Philosophy*. Cambridge, MA: Harvard University Press, especially Chs. 4 and 5.

[35] We also appraise people's failure to assess the bearing of their actions on other persons' rights, liberties, and basic needs.

[36] Watson's criteria of choiceworthiness would seem to extend beyond what I am identifying as the province of moral appraisal.

[37] For a discussion of these two "standpoints," see Bok, Hillary (1998). *Freedom and Responsibility*. Princeton: Princeton University Press, Ch. 2.

[38] On the distinction between a prediction and a decision, see Hampshire, Stuart (1967). *Thought and Action*. New York: The Viking Press, pp. 107–112. See also my "Clarke's Unlikely Success Against Hume," forthcoming in *Archiv für Geschichte der Philosophie*; Korsgaard, Christine M. (1989). "Morality as Freedom," in Yirmiahu Yovel (ed.) *Kant's Practical Philosophy Reconsidered: Papers Presented at the Seventh Jerusalem Philosophical Encounter*. Dordrecht: Kluwer, pp. 26–27; Korsgaard, Christine M. (1996) in O'nora O'Neill (ed.). *The Sources of Normativity*. Cambridge, Cambridge University Press, especially pp. 94–97 of Ch. 3 entitled, "The Authority of Reflection"; Hill, Thomas E. (1992). "Kant's Argument for the Rationality of Moral Conduct," in *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca: Cornell University Press, pp. 116–122; and Hill, "Kant's Theory of Practical Reason," in *Dignity and Practical Reason in Kant's Moral Theory*, pp. 131–138.

[39] Hampshire, *Thought and Action*, p. 109.

[40] See Korsgaard, "Morality as Freedom," p. 26.

[41] Dennett, Daniel (1995). *Darwin's Dangerous Idea*. London: Penguin Books, p. 499.

[42] See Scanlon, *What We Owe To Each Other*, p. 276.

[43] For discussion of the justification of punishment on grounds of deterrence, see Farrell, Daniel M. (1985). "The Justification of General Deterrence," *The Philosophical Review* 94, pp. 367–394. Farrell stresses the difficulties of justifying punishment for the sake of general deterrence, that is, punishing an offender for the sake of deterring others from committing similar crimes. I agree that general deterrence lacks justification when we cannot establish a causal connection between the offender's behavior and the harms we aim to avert.

[44] See Wallace, pp. 54–59. See also Bennett, Jonathan (1980). "Accountability," in Zak van Straaten (ed.) *Philosophical Subjects: Essays Presented to P. F. Strawson*. Oxford: Oxford University Press, especially pp. 19–21.

[45] This reasoning is similar to that found in moral education theories of punishment. See, for example, Hampton, Jean (1984). "The Moral Education Theory of Punishment,"

Philosophy and Public Affairs 13, pp. 208–38; and Morris, Herbert (1994). "A Paternalistic Theory of Punishment," in Antony Duff and David Garland (eds.) *A Reader on Punishment*. Oxford: Oxford University Press. Moral education theories of punishment tend to focus on the aim of morally educating the offender. Morris writes, for example, "a principle justification for punishment is the potential and actual wrongdoer's good." (p. 97) I am stressing, however, the benefit to the wider community.

[46] Durkheim argues that this is the primary function of punishment. See Durkheim, Emile (1933). *The Division of Labor in Society*. New York: The MacMillan Co., pp. 108–110.

[47] Cf. Joel Feinberg (1970) on the expressive function of punishment, *Doing and Deserving*, Princeton: Princeton University Press, pp. 95–118. Feinberg stresses that the function of punishment is to express our reactive attitudes, in particular, to express a kind of vindictive resentment. (p. 100) I agree with Feinberg that punishment has an expressive and symbolic significance, but I am suggesting a different understanding of what it can and should convey.

[48] I am indebted to Lionel McPherson for this point.

[49] This is a principle of blaming behavior that lacks moral content and perhaps should not even be classified as a principle of blaming behavior.

[50] Daniel Dennett stresses that punishing those incapable of being deterred would undermine the rationale for punishment as deterrence. See his *Elbow Room* (1984). Cambridge, MA: The MIT Press, Ch. 7. This point depends, of course, on the idea that it is the agent who is punished that we aim to deter. It does not count against the prospects of deterring others. But general deterrence is, for other reasons, difficult to justify. See footnote 43.

[51] Cf. Calhoun, Cheshire (1982). "Responsibility and Reproach," in Cass R. Sunstein (ed.) *Feminism and Political Theory*. Chicago: University of Chicago Press. Calhoun argues that in certain abnormal contexts, we may be justified in using moral reproach as a tool for social improvement even when the individual being reproached is not blameworthy. The point I am making here, however, does not involve making use of moral reproach.

[52] Farrell argues plausibly for the relevance of the following principle of distributive justice: "a person's (informed) choices make her liable to suffering certain harms if, in the light of those choices, it is inevitable that someone be harmed and she is one of the individuals who can be harmed in order that someone else be saved." ("The Justification of General Deterrence," p. 373)

[53] It is interesting that people often feel shame about things that are not in their control, such as their heritage or height. A similar point can be made about regret: we may regret actions and consequences of actions that are beyond our control. Moreover, we tend to regret harms that we have caused, regardless of their relationship to our will. See Bernard Williams (1981). "Moral Luck," in his *Moral Luck*. Cambridge: Cambridge University Press. See also his (1993), *Shame and Necessity*. Berkeley: University of California Press, Chs. 3 and 4.

[54] As discussed above, Scanlon proposes an understanding of resentment that separates it from the notion of desert. See *What We Owe To Each Other*, Ch. 6. I believe his understanding of resentment moves resentment closer to disappointment.

[55] Nietzsche, Friedrich (1887). *The Genealogy of Morals*.

[56] Williams, Bernard (1995). "Nietzsche's Minimalist Moral Psychology," in his *Making Sense of Humanity*. Cambridge: Cambridge University Press, p. 73. Williams adds that blame actually involves trying to make its subject into such an agent (i.e., into someone who would have prevented the harm through an act of will). He writes that it involves the thought that, "I, now, might change the agent from one who did not acknowledge me to one who did." (p. 73) In other words, the psychology of blame involves the fantasy that the

expression of blame itself might retrospectively have prevented the injury. See also Bernard Williams (1995), "Internal Reasons and the Obscurity of Blame," in his *Making Sense of Humanity*. Cambridge: Cambridge University Press.

[57] See Nietzsche, *Genealogy of Morals*, Essay 1, section 13.

[58] Durkheim argues that the moralization of our reactive sentiments leads to a sense of alienation, for in avenging them we project them outside of ourselves, and attribute them to the idea of morality or a divinity, which we regard as superior to ourselves. See *The Division of Labor in Society*, pp. 100–101.

[59] For further analysis and discussion, see Scheler, Max (1961). *Ressentiment*, Lewis A. Coser. (ed.) New York: The Free Press.

[60] Nietzsche, *Genealogy of Morals*, Essay 1, section 11.

[61] A practical conception of morality must also solve coordination problems.

[62] Moral education as an aim of punishment is relevant here, as are the aims I have connected with the activity of moral criticism. See note 45 and Section 2.

[63] See, for example, Mackie, J. L. (1985). "Morality and the Retributive Emotions," in Joan Mackie and Penelope Mackie (eds.) *Persons and Values: Selected Papers, Volume II*. Oxford: Oxford University Press.