https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

SOME TENSIONS BETWEEN AUTONOMY AND SELF-GOVERNANCE

By Jonathan Jacobs

I. Introduction

The notions of autonomy and self-governance each capture something crucial about the moral dimensions of agents and actions. These notions are central to the ways in which we conceptualize ourselves and others. The concept of autonomy is especially crucial to understanding the distinct status of moral agents. For its part, self-governance has a significant relation to the evaluation of agents as individuals with *particular* characters, leading *particular* sorts of lives, and performing *particular* actions. Neither notion—autonomy nor self-governance—fully assimilates or dominates the other. Moreover, there are some important strains between them. There are certain forms of regard that the autonomy of an agent *demands* that are at odds with what an agent's exercise of self-governance *merits*. In this essay I plan to show this, and offer a diagnosis of why this is the case.

The discussion will proceed as follows. In Sections II and III of this essay, I shall suggest general ways in which considerations concerning autonomy and self-governance are distinct and also very important to moral theorizing. In Sections IV and V, I shall look at ways in which these considerations bear upon responsibility for character and action. In Section VI, I shall consider how the foregoing discussion raises important questions concerning the justification of punishment. Then, in Section VII and the brief concluding section, I shall look at why there are certain ways in which considerations of moral psychology having to do with self-governance put pressure on normative considerations having to do with autonomy. I shall also examine why this results in some important moral perplexities. Punishment is not the only context in which these perplexities occur, but it is an especially effective one for identifying and diagnosing them.

I shall refer to Kant rather extensively, especially in regard to autonomy. This is because of the way in which elements of Kantian moral philosophy have become fixtures of a great deal of moral thinking. In the discussion of self-governance, I shall likewise make several references to Aristotle, because of the importance of his arguments and insights. This is not a textual study, and there are, of course, many different notions of both autonomy and self-governance, but tethering the discussion to Kant and to Aristotle will lend it a valuable manageability and focus.

© 2003 Social Philosophy & Policy Foundation. Printed in the USA.

II. Autonomy and Moral Status

Autonomy is widely held to be the basis for the distinct status of rational agents. This status is not conferred, cannot be withdrawn, and locates each person in the moral world because each is a rational agent. A key element of the Kantian version of this concept is that respect for persons as participants in a moral order is neither discretionary nor dependent upon attaining a contingent status. It does not depend upon the sentiments of others, the achievement of some end, or the realization of one or another perfection. It is owed to each agent simply as a rational agent.

In Kant's view, as articulated in his *Foundations of the Metaphysics of Morals*, autonomy determines the distinct status of rational agents because "[a]utonomy of the will is that property of it by which it is a law to itself independently of any property of objects of volition."[1] Rational nature is the capacity to be normatively self-legislating, the capacity for participation in moral life. Self-legislation is crucial to moral status because through it, persons are ends in themselves: "Now morality is the condition under which alone a rational being can be an end in itself, because only through it is it possible to be a legislative member in the realm of ends. Thus morality and humanity, so far as it is capable of morality, alone have dignity."[2] Kant continues, "Autonomy is thus the basis of the dignity of both human nature and every rational nature."[3] The most familiar and synoptic rendering of this view is that persons are ends in themselves, which is a notion that has considerable currency even apart from explicit commitment to specific Kantian theses.

Kant asks, "What else, then, can the freedom of the will be but autonomy, i.e., the property of the will to be a law to itself?"[4] And he concludes, "Therefore a free will and a will under moral laws are identical."[5] The same capacity that formulates the moral law also provides conclusive reasons to fulfill moral requirements, and enables us to do so. Morality is essentially grounded in our rational nature, and rationality is sufficient to ascertain what is morally necessary. Thus, it is also part of Kant's view, as Stephen Darwall writes, that "the very considerations that make an action morally obligatory must also be conclusive reasons for acting—reasons on the basis of which we can *act*."[6] In acting morally, we are autonomous because we are obligated by reasons of which we are the authors, and only in acting morally are we fully free, rational beings.

[1] Immanuel Kant, *Foundations of the Metaphysics of Morals*, ed. and trans. Lewis White Beck (Indianapolis, IN: Bobbs-Merrill, 1959), 59.
[2] Ibid., 53.
[3] Ibid., 54.
[4] Ibid., 65.
[5] Ibid.
[6] Stephen Darwall, *The British Moralists and the Internal 'Ought', 1640–1740* (New York: Cambridge University Press, 1995), 322–23.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

A key feature of this conception of morality and autonomy as mutually entailing is that the demands of the moral law are not beyond our capacity—the moral *ought* implies *can*. With regard to the moral law, says Kant, "We ought to conform to it; consequently we must *be able* to do so."[7] In an earlier passage, he states, "For when the moral law commands that we *ought* now to be better men, it follows inevitably that we must *be able* to be better men."[8] Even one's own history of immorality does not fully alienate the agent from the moral law or defeat his capacity to act rightly:

> However evil a man has been up to the very moment of an impending free act (so that evil has actually become custom or second nature) it was not only his duty to have been better [in the past], it is now still his duty to better himself.[9]

Insofar as we are rational, we cannot alienate ourselves from the appeal and the authority of the moral law:

> When we present examples of honesty of purpose, of steadfastness in following good maxims, and of sympathy and general benevolence even with great sacrifice of advantages and comfort, there is no man, not even the most malicious villain (provided he is otherwise accustomed to using his reason), who does not wish that he also might have these qualities.[10]

No agent who is still in possession of his reason is morally irretrievable. Reason is the source of morality, and where it is not expunged it can re-engage the agent to moral considerations.

Moreover, for Kant autonomy is the basis for a "systematic union of rational beings through common objective laws."[11] Kant holds that autonomy not only locates agents in a common moral world, but also *connects* them together in it. He argues:

> The concept of each rational being as a being that must regard itself as giving universal law through all the maxims of its will, so that it may judge itself and its actions from this standpoint, leads to a very fruitful concept, namely, that of a *realm of ends*.[12]

[7] Immanuel Kant, *Religion Within the Limits of Reason Alone*, trans. Theodore M. Greene and Hoyt H. Hudson (New York: Harper & Row, 1960), 55.
[8] Ibid., 46.
[9] Ibid., 36.
[10] Kant, *Foundations of the Metaphysics of Morals*, 73.
[11] Ibid., 52.
[12] Ibid., 51.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

We can recognize each person as an end in itself, and as such, as sovereign and subject in a rational moral order. This is Kant's basic idea of an "ethical commonwealth."[13] He explains: "A union of men under merely moral laws, patterned on the above idea ['the sovereignty of the good principle'] may be called an *ethical*, and as far as these laws are public, an *ethico-civil* (in contrast to a *juridico-civil*) society or an ethical commonwealth."[14]

Commenting on Kant's view of the realm of ends, David Wiggins remarks:

> This is the solidarity of all beings that partake in the noumenal, the solidarity of all rational beings. Even for creatures such as us, creatures who are not completely rational, the kingdom of ends, the systematic union of rational beings under common self-legislated rational laws, is an unforgettable rational ideal. This is the kingdom to which—in so far as we are rational—we cannot help but aspire to belong.[15]

We are participants in a common moral world because of reason, and we are able to construct shareable moral ideals because of reason. We can be cut off from each other by vice, inattention, selfishness, and in countless other ways. But it is rational autonomy that enables us both to be self-legislating and to share an ideal of moral community.

Overall, this conception of autonomy makes for a twofold moral democracy. Each individual has equal status as a participant in the moral order, and each is equally capable of sustaining that status, because *ought* implies *can*. We do not all act in ways that are morally worthy, but each of us is a locus of equal moral worth and the dignity that is founded on autonomy.

While many elements of Kant's conception of rational personality and its nonempirical nature are not widely embraced, key elements of his conception of morality as grounded in rational autonomy figure centrally in a great deal of recent moral theorizing. For example, the claim that the capacity for rational agency determines a distinct and equal status for individuals as participants in the moral world is a fixture of much moral theorizing. For example, John Rawls repeatedly refers to persons as "free and equal" and says the following about his view:

> The essential idea is that such procedures [by which first principles are selected] must be suitably founded on practical reason, or, more

[13] Kant, *Religion Within the Limits of Reason Alone*, 86.
[14] Ibid.
[15] David Wiggins, "Categorical Requirements," in Rosalind Hursthouse, Gavin Lawrence, and Warren Quinn, eds., *Virtues and Reasons: Philippa Foot and Moral Theory: Essays in Honour of Philippa Foot* (Oxford: Clarendon Press, 1998), 326.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

> exactly, on notions which characterize persons as reasonable and rational and which are incorporated into the way in which, as such persons, they represent to themselves their free and equal moral personality.[16]

This is Kantian in the sense that the principles that are properly constructed by rational agents must reflect the ways in which they are free and equal. Likewise reflecting the influence of Kant on recent moral theory, Alan Donagan writes:

> Since they are negatively free, the actions of a rational being have a causality higher than those of a brute animal; and it is because of that higher kind of causality that rational beings are ends in themselves—ends which are not producible but which exist independently of the actions done for their sake. It is as ends in themselves that rational beings find in their own natures a ground for the law they lay down for themselves.[17]

Frances Kamm, another philosopher who emphasizes rational agency, uses the idiom of "the high inviolability of persons."[18] She says, "If people are morally inviolable in a certain way, then, I believe, they have a *higher*—and not merely a different status."[19] This status is grounded in their being rational agents.

Focusing on a different aspect of autonomy, Christine Korsgaard argues, "If complete normative skepticism is to be avoided—if there is such a thing as a reason for action—then humanity, as the source of all reasons and values, must be valued for its own sake."[20] Normativity and moral obligation have their source in "[the] capacity for self-conscious reflection about our own actions. . . ."[21] This is the same as the basis for having reasons to act *at all*.

> If we do not treat our humanity as a normative identity, none of our other identities can be normative, and then we can have no reasons to act at all. Moral identity is therefore inescapable. Second, and for that reason, moral identity exerts a kind of governing role over the

[16] John Rawls, "Kantian Constructivism in Moral Theory," in Stephen Darwall, Allan Gibbard, and Peter Railton, eds., *Moral Discourse & Practice: Some Philosophical Approaches* (New York: Oxford University Press, 1997), 256.

[17] Alan Donagan, *The Theory of Morality* (Chicago, IL: University of Chicago Press, 1977), 233.

[18] Frances Kamm, "Nonconsequentialism," in Hugh LaFollette, ed., *The Blackwell Guide to Ethical Theory* (Oxford: Blackwell, 2000), 217.

[19] Ibid.

[20] Christine M. Korsgaard, with G. A. Cohen et al., *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 122.

[21] Ibid., 19.

> other kinds. Practical conceptions of your identity which are fundamentally inconsistent with the value of humanity must be given up.[22]

In Korsgaard's view, our capacity for reflection on action-guiding considerations gives us a kind of authority over our actions, grounded in autonomy. This is the basis of the moral identity of rational agents, and each rational agent has moral standing on account of it.

Part of the appeal of the Kantian notion of rational agency as determining the distinct and equal status of agents in a moral democracy is that this notion frees moral standing from determination by tradition, ancestry, membership in an estate, wealth, affiliation, or even birth order (all of which have much to do with luck or chance). The Kantian framework not only assigns a crucial role to reason—after all, so did Plato's and Aristotle's ethics, and Aquinas's, too, for that matter—but also takes reason to be autonomous in fashioning morality. Kant's view differs from these in the way in which the operation of reason in determining principles is autonomous in a particularly strong sense. Any determination of rational principles by considerations that are grounded in anything external to reason itself would, for Kant, constitute heteronomy. As Rawls has put it:

> Heteronomy obtains not only when first principles are fixed by the special psychological constitution of human nature, as in Hume, but also when they are fixed by an order of universals or concepts grasped by rational intuition, as in Plato's realm of forms or in Leibniz's hierarchy of perfections.[23]

Reason has normative authority in its own right, and its principles are accessible to any agent who has not lost his reason. The general contours of this approach to moral theorizing have been widely endorsed and variously developed, even when other elements of Kant's view have been abandoned. Whatever specific aspect of autonomy they might emphasize or take to be basic, it is plain that many philosophers hold the view that autonomy determines a distinct moral status for rational agents.

III. Self-Governance and the Voluntariness of Character

There is another fundamental dimension of moral evaluation and significance. It concerns features of agents that are not status-determining but concern what they are *like* and what they *do*. These features are objects of appraisal *within* the world of valuation of actions and agents. These

[22] Ibid., 129–30.
[23] John Rawls, "Kantian Constructivism in Moral Theory," 256.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

features reflect modes of self-governance, the activity through which a rational agent fashions a mode of substantive, practical engagement with the world. We are self-governing in that we *lead* lives—we fashion ends, we act upon values, we pursue interests and concerns that we conceptualize, and with regard to which we exercise deliberative rationality. This is not a metaphysical claim presupposing a metaphysical notion of free will. It is a point about the distinctive character of human lives and modes of self-conception, and how we conceptualize each other. Human beings give their lives valuative shape and content through specific modes of self-governance.

To a large extent, self-governance is a matter of how agents exercise their capacities for voluntary activity. A person does not become a voluntary agent at some particular point. The agent exercises voluntariness all along, although the character of it changes as the agent matures and grows in rational capacity and experience. Responsibility is calibrated accordingly. The behavior of a young child in reaching for something is voluntary, even though there may be no conscious, deliberative rationality involved, and the behavior of a sixteen-year-old in deciding what clique to join is typically voluntary, even if peer pressure and other influences are at work. The responsibility that we find appropriate to attribute to a thirty-year-old is more extensive because of the way in which his action expresses mature character. Voluntariness has a wide range of expression, and it can be exercised more or less consciously or deliberately. In general, however, it is on account of voluntariness that we find it appropriate to attribute responsibility to agents for their characteristics and their actions.

A great deal of early habituation is a training in self-governance, a matter of encouraging attachment to certain kinds of objects, and shaping modes of attention, perception, and receptivity. Habituation is not a matter of imposing dispositions on someone. In abusive or coercive forms it may be that, but more often it is a process of educating a person's capacities for voluntariness and self-governance. This is why it is appropriate to attribute responsibility to mature agents. Their characters reflect commitments, concerns, policies of action, and modes of valuing that, in turn, reflect who they are as voluntary agents. Self-governance is the form that voluntariness takes in a rational being. Or put another way, rational agents are capable of voluntariness that is also self-governance. Aristotle, for example, notes: "For children and the other animals share in what is voluntary, but not in decision; and the actions we do on the spur of the moment are said to be voluntary, but not to express decision."[24] Animals never become self-governing in the ways that human beings can, and habituation of the young is so crucial because it is training (for better or worse) in self-governance.

[24] Aristotle, *Nicomachean Ethics*, trans. Terence Irwin (Indianapolis, IN: Hackett Publishing Co., 1985), III, 2, 1111b9–10.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

Much of the habituation that we undergo is in the service of acquiring *rational* habits, habits of seeing things in certain ways, recognizing various considerations as reasons for action, and strategies of deliberation. Moreover, as we mature in our rationality we increasingly habituate *ourselves*. The process by which we come to have stable dispositions to respond, to choose, and to act is increasingly self-governed. Aristotle remarks: "We deliberate about what is up to us, i.e. about the actions we can do . . . [f]or causes seem to include nature, necessity and fortune, but besides them mind and everything [operating] through human agency."[25] The Aristotelian view, endorsed here, is that we are voluntary originators of action, both before our rational capacities are developed and once they are. In maturing, our self-motion is more and more deliberate, even if we reason poorly or inattentively.

An individual, in maturity, typically settles into more or less fixed dispositions. Indeed, a crucial dimension of what it is to *have* a character is that the agent has stable, action-guiding dispositions, and dispositions of affect, motivation, and receptivity. We do not regard a person's attributes as simply welcome or unwelcome, desirable or undesirable. We see them as features of a self-determining individual who does not just happen to be that way entirely on account of luck or necessity. While multiple influences are at work on each of us, what we make of these influences is partly determined by how we exercise self-governance. The character that we each acquire is, in substantial respects, voluntary.

Indeed, we tend to identify with many of our own persistent, characteristic dispositions. Sometimes we identify so strongly that even when we acknowledge their defects, the fact that they are our own, and are partially constitutive of us, poses an obstacle to changing them. Often, habit is pleasing just because it is "second nature." Although we are capable of adopting a critical stance toward some of our characteristics, we may withhold the effort to change them, not because we think change is impossible, but because there is comfort in the familiarity of "second nature."

In sum, to a substantial extent, we become who we are on account of the manner in which we exercise self-governance. Agents' states of character are more fully their responsibility as they are more fully able to decide to act on values and attitudes with which they identify, and which they recognize as reflecting their own commitments, concerns, and policies of action. As Bernard Williams puts it:

> [I]f one acknowledges responsibility for anything, one must acknowledge responsibility for decisions and action which are expressions of

[25] Ibid., 1112a31–33.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

character—to be an expression of character is perhaps the most substantial way in which an action can be one's own.[26]

Self-governance does not require that an agent be immune to luck or necessity. Nor does it require that a large "menu" of options be presented to the agent. The person who is raised in certain traditions, to the exclusion of seeing and doing things in other ways, and to whom it does not occur to think and act differently, can still exhibit a substantial degree of voluntariness and self-governance in leading his life. A duke may be acting exactly as the duke would act; namely, as the duke has acted for generations. He has been raised and educated to act that way. But the duke is no less a voluntary agent. In many cases, compliance, willingness, and endorsement of ends and practices are sufficient for the voluntariness of acts and states. The duke may give little thought to whether there is any other way for him to act. In fact, the thought that there might be any other way may strike him as absurd. Perhaps his endorsements are not thoughtful or critical. But they are the endorsements of a voluntary, rational agent, not an automaton.

We generally do not know what difference will be made to our characters by specific acts, and the fact that a state is voluntary does not imply that the agent had complete control over its initiation and establishment. However, as Aristotle remarks, we do know that, in general, "each type of activity produces the corresponding character."[27] We do not know how many forays into battle without cowering make for a courageous soldier, nor do we know how many lies cause one to be a liar—but it is facing risks and dangers without flinching, running, or equivocating that makes for courage, and it is lying that makes liars, and so forth.

IV. Character and Practical Necessity

A person may voluntarily acquire qualities of character that she did not intend to acquire, and which she is then unable to alter, and still be responsible for them and for the actions that flow from them. Although one may strive to establish a certain type of disposition, dispositions can be the result of voluntary activity even if what results is not something explicitly intended. In addition, character attributes determine, to a large extent, what is possible and what is necessary for an agent. In addition to the "have to" of law, of rule-following, of moral requirement, of social custom, and of biological imperative, is the "have to" of character and self-governance.

[26] Bernard Williams, "Practical Necessity," in Williams, *Moral Luck: Philosophical Papers, 1973–1980* (Cambridge: Cambridge University Press, 1981), 130.
[27] Aristotle, *Nicomachean Ethics*, III, 5, 1114a7.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

Again, Aristotle is an especially valuable resource for developing this claim. He notes, for example, that the conditions for actions being virtuous acts are that the agent first "must know [that he is doing virtuous actions]; second, he must decide on them, and decide on them for themselves; and, third, he must also do them from a firm and unchanging state."[28] In Aristotle's view, and in the view that I am developing, the virtuous agent grasps and appreciates what is ethically salient, and also reasons well to what is required. He then proceeds to act in a way that is undistracted and free of motivational conflict. This is not thoughtless or mechanical, but is the result of employing an array of concepts and making certain discriminations, which inevitably leave many other, possibly action-relevant considerations out of the picture.

John McDowell uses the language of "silencing" to characterize the way in which a virtuous agent's manner of seeing things brings certain reason-making facts into view and excludes others. He says that for the virtuous agent:

> the dictates of virtue, if properly appreciated, are not weighed with other reasons at all, not even on a scale which always tips on their side. If a situation in which virtue imposes a requirement is genuinely conceived as such, according to this view, then considerations which, in the absence of the requirement, would have constituted reasons for acting otherwise are silenced altogether—not overridden—by the requirement.[29]

McDowell is talking about an ideal of virtue. It may be a rare individual whose character silences reasons for acting that conflict with what virtue requires. But the main point remains unaffected. It would be incorrect to say that the virtuous agent and the vicious agent apprehend the same action-guiding considerations but respond to them differently. Rather, these agents have different conceptions of what are the action-guiding considerations in the situation. As McDowell points out:

> It would be wrong to infer that the conceptions of situations which constitute the reasons are available to people who are not swayed by them, and weigh with those who are swayed only contingently upon their possession of an independent desire. . . . Their status as reasons is hypothetical only in this truistic sense: they sway only those who have them.[30]

[28] Ibid., II, 4, 1105a30–35.
[29] John McDowell, "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society* Supp. Vol. 52 (1978): 26.
[30] Ibid., 23.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

The vicious agent may be rational—he has conceptions of good and bad, and he may show care and resolve in enacting them. Vice is not clear evidence of a failure of rationality on the part of the vicious agent. Also, vice can silence as well as virtue can. In the seriously and persistently vicious agent, what is in accord with vice does not simply override what virtue requires. What virtue requires is silenced, and the considerations that have significance for the virtuous agent do not count as action-guiding reasons for the vicious agent even if they are brought to his attention. That is the difference, for example, between a cruelly vicious person who inflicts suffering because he really believes that his victim deserves to suffer, and the agent who succumbs to a vengeful passion, but is able to recognize it as a moral lapse. The cruelly or persistently vicious agent is doing what he thinks he ought to do, and he would stand by his actions even upon consideration. He is responding to what he takes to be rational imperatives. He is someone for whom what is morally necessary is not practically necessary, on account of his character. This agent might not even recognize what is morally required.

Much of what an agent does is practically necessary in the sense that it reflects firmly established states of character or perhaps the aspiration to exhibit or acquire a certain state of character. Those states or the determination to establish them can constitute real limitations on what an agent takes to be possible and what enters into her deliberation and intention. The exclusion of other possibilities is often not a matter of inattention, lack of information, or thoughtlessness (though it may sometimes be one of these). It may, instead, be a matter of the agent exercising a habit of judgment and action that particularly reflects what she takes into account and cares about (for better or worse). The acts of an uncoerced, aware, rational adult are voluntary acts even if they are of the sort of which we say, "she could not help it" because of the sort of person that she is. "Being in character" is not an alternative to being voluntary; "being in character" is a fundamental mode of acting voluntarily.

An upshot of this view of self-governance and character is that moral considerations are not equally accessible to all rational agents. These considerations may be *available*, in the sense that they are not epistemically out of reach to such an extent that agents cannot be reasonably expected to know them, but they still may not be *accessible* to all agents. (I shall say more about this distinction in Section V.) There are agents who cannot bring moral considerations into view and properly appreciate them. Given the wrongness or perversity of their values and motives, they do not acknowledge or respond to ethically sound considerations. They act for reasons that they endorse, and they weigh considerations, deploy normative concepts, and so forth. But their modes of self-governance preclude them from correctly appreciating and responding to ethical considerations, or at least some important kinds of them. Normatively sound

rationality is an achievement, and these agents progressively disable themselves for it, while remaining voluntary, responsible individuals.

Facts about self-governance, the role of voluntariness in the establishment of character, and the way in which states of character determine practical necessity and impossibility for individual agents push us toward the conclusion that, while moral requirements are categorically imperative, this does not imply that they are cognitively and motivationally accessible to all rational agents. There are agents who do not recognize their wrong actions as such. These agents do not appreciate their wrong actions as lapses from values and standards that they recognize; their actions are in accord with the values and standards that they endorse. Thus, there are some agents who cannot, in any practically effective way, appreciate the wrongness of their actions. In many cases, however, this does not defeat responsibility.

V. Character and Accountability

Gary Watson suggests a distinction between attributions of moral fault, on the one hand, and practices of moral accountability, on the other. This distinction will help us to explicate the tension between autonomy and self-governance. Judgments concerning attribution of moral faults are made from what he calls the "aretaic perspective,"[31] and we can think of this "as concerned with the question of what activities and ways of life are most choiceworthy."[32] In aretaic evaluation we are concerned with the degrees of excellence or the deficiencies of agents' characters and actions, that is, the ways in which they are admirable or not, or worthy of praise or not. We are not necessarily concerned with whether agents' actions are morally right or wrong. Judgments of *accountability*, however, concern "particular moral norms"[33] and the agent being "accountable to us or to others"[34] on the basis of "certain expectations or demands or requirements."[35]

Explaining what he means by accountability, Watson writes:

> To require or demand certain behavior of an agent is to lay it down that unless the agent so behaves she will be liable to certain adverse or unwelcome treatment. For convenience, I shall call the diverse forms of adverse treatment "sanctions." Holding accountable thus involves the idea of liability to sanctions. To be entitled to make demands, then, is to be entitled to impose conditions of liability.[36]

[31] Gary Watson, "Two Faces of Responsibility," *Philosophical Topics* 24, no. 2 (1996): 231.
[32] Ibid.
[33] Ibid.
[34] Ibid.
[35] Ibid., 235.
[36] Ibid., 236–37.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

In this way, "holding [someone] responsible goes beyond aretaic appraisal."[37] We sometimes find fault with agents without necessarily regarding those agents as being *at fault* for a specific wrong. They might be found selfish, shallow, cold, or perhaps even repugnant, but there may not be a specific wrong act for which they are liable. Similarly, we might admire and praise people for their composure, consideration, reliability, and the like, without also making specific judgments concerning their moral accountability. Watson says:

> Attributability has an importance to ethical life that is distinct from concerns about accountability. Responsibility is important to issues about what it is to lead a life, indeed about what it is to have a life in the biographical sense, and about the quality and character of that life. These issues reflect one face of responsibility (what I will call its aretaic face). Concerns about accountability reflect another.[38]

This distinction brings into relief the ways in which evaluations of agents answer to different kinds of considerations.

Aretaic appraisal concerns matters that reflect the agent's mode of self-governance. "Because aretaic appraisals implicate one's practical identity," Watson writes, "they have ethical depth in an obvious sense." He further explains:

> This brings out the way in which aretaic appraisal involves an attribution of responsibility. To adopt an end, to commit oneself to a conception of value in this way, is a way of taking responsibility. To stand for something is to take a stand, to be ready to stand up for, to defend, to affirm, to answer for. Hence one notion of responsibility—*responsibility as attributability*—belongs to the very notion of practical identity.[39]

Accountability, in contrast to attributability, concerns what can be expected or demanded of an agent. It concerns liability to sanctions, not just susceptibility to appraisal. Accordingly, Watson notes that an important difference between accountability and attributability is that the former "raises issues of fairness that do not arise for aretaic appraisal."[40]

One way to see that autonomy and self-governance can be in tension is to recognize that agents can be accountable even when, in certain respects, it was not practically possible for them to do what they ought to have done, because aretaic features alienate them from what accountabil-

[37] Ibid., 238.
[38] Ibid., 229.
[39] Ibid., 234.
[40] Ibid., 235.

ity demands. There are cases in which what is morally required is *available, but is not accessible*, because of agents' habits and characters. The relevant moral understanding is abroad, but some agents are deeply alienated from it by their characters. Are *these* agents properly liable to sanction? Many of them are.

The agents of chief concern here are full participants in the moral order because they *could have* acquired correct understanding and morally sound motives, but these are now seemingly out of reach because of the states of character that the agents have settled into. They are not irrational, though they are firmly attached to wrong values, which they believe are normatively sound. Such agents reject attempts to get them to revise their beliefs and commitments, though this does not indicate incompetence or disability that would disqualify the agents for accountability.

Susan Wolf argues that "an individual is responsible if and only if she is able to form her actions on the basis of her values *and* she is able to form her values on the basis of what is True and Good."[41] She writes:

> Whatever the explanation that prevents the agent from being able to do the right thing for the right reasons, our intuitions seem to support the claim that the agent does not deserve blame. If an agent is incapable of doing the right thing for the right reasons, then it is not her fault that she stumbles into doing something wrong.[42]

When the deformation of an agent's values can be traced to factors that the agent could not control, we diminish her accountability. These factors, in Wolf's view, might be such things as the wrongness of prevailing norms, such as in the slaveholding American South in the nineteenth century or in Germany in the 1930s. (These are examples that Wolf employs, and I shall critique them below.) Wolf says, "[I]nsofar as we do regard social processes and norms as potential obstacles to sound moral judgments, we lessen the blame we would otherwise direct toward individuals who, surrounded by these obstacles, fail to reach these judgments."[43] The aretaic failures of nonresistant citizens in the slaveholding South or in Nazi Germany underwrite attributions of fault, but these failures do not, at least to the same extent, underwrite blame. Given the powerful influences operating on these agents, suggests Wolf, they are less than fully accountable because we can see why it was reasonable for them to have developed the values and commitments that they had. Their actions were voluntary, and they were acting on values that they endorsed, but they were prevented from seeing things in any other way. Thus, they are not fully accountable for their wrong actions.

[41] Susan Wolf, *Freedom Within Reason* (New York: Oxford University Press, 1990), 75.
[42] Ibid., 81.
[43] Ibid., 122.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

I have suggested a distinction between correct values being available and being accessible. The agents to whom Wolf refers are, I think, agents to whom correct values *were* available (even if, in some ways, not accessible), and that is a reason *not* to lessen blame. At least there is a question about whether to lessen it. These are agents who could have known better. First, they could have known better in the general sense that they were competent rational agents. We are considering agents who were not deranged, insane, or cognitively handicapped. Second, at crucial stages in their development, they did have the plasticity of character and capacity for learning that could have enabled them to acquire different valuative attachments, attitudes, and motivational policies. In the process of coming to have valuative commitments and firmly established states of character, these agents exercised voluntariness no less than agents who acquired virtues. For all of the influences at work on the agents whom Wolf cites in her examples, they were not victims of compulsion, coercion, or the destruction of their capacities for practical rationality—although they were encouraged in habits of mind that were wrong or perverse. (Of course, some agents are victims of destructive influences, and of course, such agents' responsibility is appropriately diminished. Being raised by abusive, violent people who keep you in confinement and deny you interaction with other people will destroy your capacity to acquire a character with normal emotionality and to exercise self-governance.) Third, a sounder moral understanding was not so remote from the agents in Wolf's examples that it is utterly unreasonable to expect any of them to have acquired it. There were, indeed, powerful forces militating against endorsement of values other than the prevailing ones. But, the prevailing ones were contested, and challenges to them were well known.

Let us (plausibly) suppose that Josef Goebbels, for example, reached a point where he could not consider a way of seeing things other than through Nazi ideology. Let us also (plausibly) suppose that there were powerful influences at work on him that were instrumental in motivating him to be a Nazi zealot. His surroundings shaped his character, and his character became fixed in a way that cut him off from "the True and the Good" (as Wolf puts it). Neither of these considerations automatically tips the balance against his accountability. The Nazi zealot and the defender of slavery live in social worlds that do not render them incapable of correct moral understanding, though they may be beyond the reach of right reason because of the habits of mind, motivation, and sensibility that have become second nature to them. This is to say, blameworthiness does not always require avoidability.

In the antebellum South, many people, even among those who tolerated, condoned, or supported slavery, understood that it was wrong, but for a number of reasons failed to take steps to oppose or abolish it. These were agents who could perceive that slavery was immoral, even if they did not oppose it. There were other Southerners for whom correct values

were epistemically inaccessible. They were profoundly alienated from sound ethical understanding: they were resolute defenders of the institution of slavery and genuinely believed in its rightness. Once settled into their way of seeing things, perhaps they could not have known better, and they could not do what they ought to have done. What was right had become inaccessible to them because of their fixity of character and the way that it cut them off from correct moral understanding. They would not recognize accountability for perpetuating unjust practices, because they did not think that they were wrong. These persons too, I believe, are responsible agents, on account of the availability of correct ethical considerations and the fact that they voluntarily became the sorts of agents whom they were.[44]

That there are such agents is an important illustration of the tension between autonomy and self-governance. These agents merit respect as participants in the moral order. We owe it to them to address them as such, to regard them as moral agents, and to appeal to them as rational. Still, we might not have any reasonable expectation that they are capable of ethically appropriate responses because of the ways in which their characters shape what they take to be practically necessary.

VI. Perplexities Concerning Punishment

Even though some agents appear to have put sound ethical considerations out of their reach, there are both epistemic and moral reasons to be very cautious about concluding that such agents are incorrigible. We do not know with certainty, even in our own cases, if and when dispositions have become altogether fixed. Kant, for example, had few illusions about the human capacity for evil, but strongly cautioned against thinking that we are ever in a position to judge that an agent is so fixed in vice that we cannot interact with him as with a fully rational agent.[45] Respect for persons as rational authors of their actions requires a certain kind of humility. We are not to presume to know others so well that we can confidently regard them as irretrievably alienated from morality. Neither are we to regard an individual (except in cases where there clearly is coercion, compulsion, or some other unmistakably agency-defeating cause) as less than the author of his actions.

[44] See Philippa Foot's *Human Goodness* (Oxford: Clarendon Press, 2001), where she states a similar position. She says: "Ignorance itself may be voluntary, as when in present-day Britain, an arms dealer takes care not to enquire whether the weapons he is shipping to one country will not be shipped on to a repressive regime. Or if he 'enquires', his conclusion may be guided by self-interest rather than by the evidence at hand: he holds a convenient opinion that does not absolve him, because it is not held, as we say, in good faith" (70). In a footnote to this passage, she writes: "Many of the beliefs of slave owners, and of white people under South Africa's apartheid regime, must have been of this kind" (70).

[45] Kant makes this point in *Religion Within the Limits of Reason Alone*. See, for example, pp. 16 and 35.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

In addition to exercising this sort of humility, we also *morally* owe it to others not to abandon them, not to "write them off," so to speak, as incorrigible. To regard agents as less than fully accountable might appear to be a way of showing sympathetic concern, but it is actually a denial of their autonomy and a way of regarding them as less than full agents. R. A. Duff makes the moral point about respect and corrigibility as follows:

> We owe it to every moral agent to treat him as one who can be brought to reform and redeem himself—to keep trying however vainly, to reach the good that is in him, and to appeal to his capacity for moral understanding and concern. To talk thus of "the good that is in him" is not to make some psychological claim to the effect that he "really" cares for the values which he flouts: it is rather to combine the conceptual claim that every moral agent has the capacity or potential for moral development and reform, with the moral claim that we should never give up hope of bringing him to actualize that potential.[46]

If we cease to regard an agent as capable of correct moral acknowledgments and rationally motivated, ethical reform, then we have diminished the agent and damaged the unity of the moral world by making adjustments in that agent's status.

On the one hand, no rational agent is to be despised or treated with less than the respect owed to a participant in the moral order. This is because virtue is not a necessary credential for moral standing, and the autonomy of rational agents fixes that status. On the other hand, there are agents who not only act wrongly on a regular basis, but also seem to have characteristics that make it practically impossible for them to attain sound normative understanding and act upon it.

The tension between autonomy and self-governance is felt particularly strongly in respect to the issue of punishment. In punishing an offender, we deliberately impose harm as a strategy of appealing to him as a rational agent. Apart from consequentialist aims that punishment might have, it is a mode of address that includes the offender in the moral order. Duff expresses this view as follows:

> Punishment, like moral blame, respects and addresses the criminal as a rational moral agent: it seeks his understanding and his assent; it aims to bring him to repent his crime, and to reform himself, by communicating to him the reasons which justify our condemnation of his conduct.[47]

[46] R. A. Duff, "Expression, Penance, and Reform," in Jeffrie G. Murphy, ed., *Punishment and Rehabilitation*, 3rd ed. (Belmont, CA: Wadsworth Publishing Company, 1995), 198.
[47] Ibid., 174.

Furthermore, "punishment is part of that continuing dialogue with the criminal through which the law aims to guide his conduct by appealing to relevant reasons."[48] It is not merely a striking back, an institutionalized form of expressing outrage, or a form of control aimed at those who make mischief. Punishment takes seriously the agent's participation in the moral community and the importance of restoring the agent's relations to the rest of the moral community. Morally justified punishment is a strategy through which sound ethical understanding is intended to be made accessible to an agent who needs to connect or reconnect with it, through acknowledgment of having done wrong.

The significance of this cannot be undermined by pointing out that it is well known that punishment often fails to motivate moral self-correction, and that, indeed, it often worsens those who are punished. There are agents who *could* respond in a constructive way, but do not. This is indeed lamentable, but it is a distinct issue. Our present concern is the appropriateness of punishing agents who are not *able* to recognize the justice of sanctioning them, because of their firm attachment to wrong values. It is not merely that they do not respond by acknowledging their wrongdoing; rather, there are reasons to think that they *cannot*.

This issue is also distinct from certain other kinds of situations involving fundamental moral disagreements. Agents who disagree (say, on capital punishment, abortion, or physician-assisted suicide) may still recognize one another as holding their views in good faith, as willing to listen to reason, and as common participants in a shared moral world, even though their disagreements might become rancorous. On specific issues they might find that they can make no progress, and the dialectic grinds to a halt. But it is a *dialectic* that grinds to a halt, and there is, to a large extent, a common framework within which it takes place and halts. The situation is quite different if one party believes that it is perfectly all right to treat certain people in exclusionary ways, and to deny them basic rights, or to take no account of their welfare, *because* people "like that" are inferior and despicable, and they deserve no better. In this case, there is hardly room for meaningful dialectic if the advocate of this view, upon reflection, really believes it and dismisses considerations that confute it. There can be cases of deep disagreement that do not challenge or threaten the overall fabric of the moral community, and in which it is possible for agents to sustain respect for each other. Those are not the kinds of cases at issue here. The present focus is the character of the agents, not the contested nature of the issues.

In one respect, the issue of character might not be thought to be so important. It could be argued that what is important is that offenders be punished, and that we should only punish agents who are rational and competent, and not influenced by disease, coercion, or compulsion. We

[48] Ibid.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

can, the objection goes, make the relevant exculpatory and exempting discriminations, when appropriate, without complicating the issue with reference to the fact that there are rationally unreachable agents. Offenders need to be stopped and punished. Whether the criminal can ethically recompose himself is a separate matter, which does not determine the legitimacy of punishment. Perhaps the need to control those who engage in violent, wrongful acts is enough to justify punishing them. Watson, for example, notes that even if the agent suffered the misfortune of a very rotten background and various deprivations, this "does not shield victim-criminals from legal sanctions."[49] He adds, "We still protect ourselves against [criminals'] murderous assaults; we hunt them down, lock them up, shoot them. . . . Seeing the criminal as himself a victim will not prevent us from shutting the cage or pulling the trigger."[50] It is true that certain social concerns are met by punishing offenders, even when moral purposes are unfulfilled. Punishment is necessary to fulfill certain purposes of social regulation. However, it is important that Watson also notes that "these responses will then tend to seem regulative rather than retributive. In a disconcerting way, they lose their normal expressive function."[51] He finds this disconcerting because the expressive function ought to be tied to moral purpose, and not just to social control. The expressive function is the aspect of punishment through which it is a communicative engagement concerning values. I would add that, in a society in which the moral legitimacy of its institutions and practices is important as the basis of conformity to its norms, more than the regulative function of punishment is at issue.

Along these lines, Herbert Morris argues that among the conditions for morally justified punishment are that "the norms addressed to persons are generally just and that the society is to some substantial extent one in which those who are liable to punishment have roughly equal opportunities to conform to those just norms."[52] In addition, the theory "presupposes that there is a general commitment among persons to the values underlying them."[53] Suppose that a society has basically just norms and that its institutional practices are basically just. Still, it is almost certain that there will be members of the society who do not share a commitment to these norms, though they had the opportunity to acquire that commitment and to conform to them. The norms were available for adoption, although the combined effect of circumstances and voluntary activity made them inaccessible to some agents. Some of their actions are shame-

[49] Watson, "Two Faces of Responsibility," 240.
[50] Ibid.
[51] Ibid.
[52] Herbert Morris, "A Paternalistic Theory of Punishment," in Jeffrie G. Murphy, ed., *Punishment and Rehabilitation*, 165.
[53] Ibid.

less moral outrages, and the agents have no interest in revising their values and commitments. They may even scorn correct values.

These agents may still properly be subject to punishment because they possess the relevant sort of normative competence to be self-governing and responsible. *Normative competence* is not the same as virtue. To be normatively competent, an agent must use normative concepts in a coherent way and make various types of valuative distinctions and judgments. This is not a wanton or a psychopath who has no concern for what moves him to act. This is not someone who, out of desperation or rage, runs amuck. The agent in question may have articulate and elaborate rationales for his views and actions. He acts on the basis of valuations. However, he is so thoroughly alienated from the true and the good that society's norms are not *really* accessible to him, even though they were and are available.

VII. Forgiveness and Hatred

Considerations of moral psychology put pressure on the framework of morality as structured by autonomy in the way we have indicated. This pressure is especially evident in regard to punishment, but it also emerges in related questions concerning retributive hatred and forgiveness. By *retributive hatred* I do not mean malice, but an attitude by which we insist on the punitive suffering that an offender deserves and find it gratifying that the offender should suffer.[54] The morality of hatred in general involves many difficulties, but there is one that stands out in the present context: Can agents forfeit our respect for them as participants in the moral world and be properly regarded as loathsome, inferior creatures? Suppose that an agent is strongly attached to wrong values, regards those values as correct, and acts on them with a sense of justification. Given the agent's commitments and sincerity, we are moved to find him loathsome, as though he is "not one of us" and cannot be accommodated in the moral world along with other agents. Has he forfeited his distinct and equal status as a moral agent, and are we thus justified in our loathing?

According to Jean Hampton, while retributive punishment or hatred of a person's worst character traits may be morally well placed, we should not reach the point of regarding the wrongdoer "not merely as cloaked in evil, but as himself a bad thing."[55] In a passage a few pages earlier she writes:

> But I can remain emotionally opposed to someone's action, and still come to be supportive of, even reconciled to, *her*, if I am able to

[54] I borrow the term "retributive hatred" from Jeffrie G. Murphy, who introduces it in "Hatred: A Qualified Defense," in Jeffrie G. Murphy and Jean Hampton, *Forgiveness and Mercy* (New York: Cambridge University Press, 1988), 90.

[55] Jean Hampton, "The Retributive Idea," in Murphy and Hampton, *Forgiveness and Mercy*, 152.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

> dissociate her from the action and reapprove of her. This is precisely what I cannot do if I morally hate the wrongdoer.[56]

This is an attempt at retributivism without retributive hatred. It is also part of an argument for forgiveness following punishment. Hampton appeals to the considerations that I mentioned above, concerning our uncertainty about whether any agents are altogether incorrigible, as well as to the teachings of Christianity. She asks, "Aren't we obliged to have faith in a decent core within, even if it is a core which we are completely unable to see?"[57] Whatever evidence there is in the aretaic dimension, we may lack a sufficiency of evidence to determine that the person is "morally dead"[58] and thus reduced in moral standing. We cannot confidently claim that we are morally better. Perhaps we were lucky to have had better surroundings, better examples, and better opportunities. We can, and should, denounce the offender's acts and the character states that motivated them, but we should not hate the offender. We owe it to the offender's better self, and to our own better selves, to be both restrained in our hatred and generous in our willingness to reapprove of the agent. (This might be a way of recognizing serious aretaic defects without also allowing that recognition to disturb the respect owed to the individual as a rationally autonomous agent.)

Even if an individual has seemingly destroyed the basis for respecting him, it is not appropriate (Hampton suggests) for us to withdraw respect for this person as a moral agent. That respect is not contingent on virtue. Hampton notes that a presupposition of retributivism, as she understands it, is a "Kantian theory of human worth, which makes people intrinsically, objectively, and equally valuable."[59] We should always look to the agent's capacity to morally redeem himself, rather than withdraw respect because of his vices. This harks back to what we noted about autonomy at the outset of this essay. Autonomy determines the status of an individual as a moral agent, based on his rationality, which is the capacity to participate effectively in the moral order and to orient himself to what morality requires. Self-governance may disfigure this capacity through the way that self-governance is exercised, but it is the capacity for autonomy that determines the individual as a moral agent with special standing.

It is undeniable that retributive hatred can become unruly and minatory in ways that fail to respect the hated agent as a fellow rational being. Still, the willingness to forgive and to reapprove a moral offender can distract us from the genuine evil that he has done. It can be symptomatic

[56] Ibid., 148.
[57] Ibid., 151.
[58] Ibid., 153.
[59] Ibid., 124.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

of an unwillingness to see evil for what it is. Sometimes, forgiveness is motivated by the fact that the offender is no longer dangerous or fearsome simply because he is in custody and at our mercy. Adam Smith remarks that "[w]hen the guilty is about to suffer that just retaliation, . . . when he ceases to be an object of fear, with the generous and humane he begins to be an object of pity."[60] To counterbalance this urge to pardon the criminal, Smith recommends that the generous and humane remind themselves of

> the general interest of society. . . . They [ought to] reflect that mercy to the guilty is cruelty to the innocent, and oppose to the emotions of compassion which they feel for a particular person, a more enlarged compassion which they feel for mankind.[61]

It can be very difficult to reconcile ourselves to the fact that some agents may be as deeply or enduringly vicious as they seem to be. Perhaps some are hateful in ways that merit retributive hatred. There is a danger that forgiveness will be corrupted (whether or not we admit it) into a kind of narcissism, in which the concern is primarily for our own (perceived) virtue, rather than for the evil with which we must deal.

Granted, Hampton was not suggesting that we forgo punishment. We can punish *and* forgive. This would be a way of sustaining respect and avoiding retributive hatred. However, the moral purpose of forgiveness is fulfilled only when there is some realistic chance of uptake of the gift. Some agents are unreachable, unrepentant, and utterly devoted to wrong or perverse values. God can forgive them because God can infuse receptivity to grace. But *we* cannot do that. There certainly seem to be some agents who are unreachable by us. The fact that we might find gratification (which is not the same as joy) in their punishment and suffering does not disqualify retributive hatred as a proper expression of fidelity to correct values and norms. Epistemic and moral humility are in order here, but so, too, is resolve to see that correct values are upheld and that "the victim's value receives its proper defense."[62]

In most cases of punishment, an offender is capable of acknowledging his wrong and recognizing the justice of his punishment. The respect owed to the offender is not in doubt, and retributive hatred is either out of place, or it should be forgone without difficulty. But the agent in question in less typical cases is not like that. He gives every sign that there is no better self to reach through punishment or to respond to forgiveness. It may be asking too much to expect virtuous agents *not* to feel retributive hatred for such an unrepentant agent, or to forgive him, even after pun-

[60] Adam Smith, *The Theory of Moral Sentiments* (Indianapolis, IN: Liberty Fund, 1984), 88.
[61] Ibid., 88–89.
[62] Hampton, "The Retributive Idea," 131.

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press

ishing him. There are times (many of them) when it is wrong not to forgive. But there are also agents whose vice is sufficiently profound and so centrally a part of their practical identities that forgiving them only makes sense on the basis of a misrepresentation of their characters and their capacities to connect (or reconnect) to correct values. There seem to be agents whose viciousness is of such a nature that it threatens our respect for them as rational beings. At least it calls that respect into doubt, even though it is a type of fundamental regard that is supposed to be protected from judgments or decisions to confer or withdraw it. If there are agents who merit retributive hatred, then there is a source of serious strain on the normative framework for autonomy, to which respect is crucial.

This is not a concession to our baser selves. It is an acknowledgment of important features (involving sentiment and judgment) of the commitment to justice and to correct values. As Jeffrie Murphy puts the point, "we may say that retributive hatred is a strategy designed to see (and to let the victim see) that people get their just deserts; as such it is neither irrational nor immoral."[63] The hatred can reflect the strength of our attachment to certain values, rather than mere or sheer malice. Retributive hatred is not always wrong simply because it is hatred—and forgiveness is not always right, simply because it is forgiveness. A virtuous agent respects other agents, can tell the difference between vengeance and retributive hatred, and works at keeping the former from infecting the latter. However, a virtuous agent finds no justice or satisfaction when a wrongdoer does *not* suffer at all: Kant remarks at the beginning of the *Foundations of the Metaphysics of Morals*, "It need hardly be mentioned that the sight of a being adorned with no feature of a pure and good will, yet enjoying uninterrupted prosperity, can never give pleasure to a rational impartial observer."[64] We need to consider seriously whether a vicious and seemingly incorrigible agent can be found so repugnant that we find it not just acceptable but gratifying that he suffers (just punishment).

VIII. Conclusion

Rational agency *demands* respect, we might say, while the virtuous exercise of self-governance *merits* it. Respect is owed to rational agents as having a distinct and equal status. The agent whose self-governance is profoundly and persistently vicious may merit our loathing to such a degree that we experience doubts about his standing as a participant in the moral order. This is not a license to degrade, humiliate, or abuse that person, but, at the same time, there may be serious obstacles to forgiveness and restoration of respect. A shameless, dreadfully bad person who

[63] Murphy, "Hatred: A Qualified Defense," 95.
[64] Kant, *Foundations of the Metaphysics of Morals*, 9.

has no inclination to morally retrieve himself may exhaust the bases for the full measure of respect that rational agency otherwise seems to demand.

Part of the explanation for this forfeiture of respect is that aretaic features are not merely accessories to an agent's rational nature. They are the expression and realization of it. The degree to which the normative framework that requires respect for persons actually structures the moral world depends upon the extent to which agents are virtuous. The normative framework is merely ideal or notional *unless* the ways in which people exercise self-governance actualize and sustain it. Vicious self-governance can threaten the respect that is owed to an agent on the grounds of autonomy. What autonomy demands and what self-governance merits may pull in opposite directions. Thus, there remain difficult questions about whether one's standing in the moral order can be relinquished on account of how one exercises self-governance. This is one way in which the concepts that we use to map the moral landscape may fail to bring complete clarity.

Philosophy and Religion, Colgate University

https://doi.org/10.1017/S0265052503202090 Published online by Cambridge University Press