1. Introduction

We do not employ the language of evil as freely as our forebears did. But call Hitler or the Holocaust evil and you are unlikely to arouse much disagreement. On the contrary: you will have better luck generating dissent if you refer to Hitler or the Holocaust merely as bad or wrong: “Hitler was a bad person, and what he did was wrong.” As is often noted, such tepid language seems terribly inadequate to the moral gravity of this subject matter. Prefix your adjectives with as many “verys” as you like; you still fall short. Only ‘evil’, it seems, will do.

It makes a difference whether we call someone evil, or merely (very, very . . . ) bad. The question is, what difference does it make? Does the concept of evil designate a significant moral category? Or is it merely a vehicle for shoddy thinking, serving no other purpose than to demonize whomever we find most distasteful? Philosophical theorizing about evil persons and actions has picked up in recent years, but for the most part theories (including my own) have been defended on grounds of intuitive plausibility, with little attention given to the motivation for distinguishing the evil from the merely bad. This leaves unanswered the question of why we ought to prefer one account to another: does one theory capture something of moral significance that the other does not? We wish to broaden our understanding of moral phenomena, not simply to describe folk concepts of dubious import.

In this paper I shall focus on the theory of evil character, arguing that most theories fail to support a robust bad/evil distinction. One intuitively plausible theory, however, does make such a distinction. I then argue that the concept of an evil person, as understood on that theory, defines one pole of a moral continuum that incorporates, at the opposite pole, the moral saint. This moral space differs substantially from the one in which some prominent theories place the evil person. Those views, I suggest,
confuse the evil person with the moral criminal, who is properly contrasted with the hero and not the saint. The concept of an evil person is, I conclude, morally significant, if perhaps less useful in practice than we might have hoped.

2. Two problems

A natural starting point is this: to be evil is to be the worst possible kind of person. Take your favored moral theory, consider its account of character, and then look to the negative end of the spectrum. This tactic might seem to provide all the motivation we need. Moreover, this approach respects the appearance that evil is an all-or-nothing affair: we tend not to speak of one person’s being “more” or “less” evil than another. We seem to add nothing by saying that someone is “very” evil.2

If correct, this approach would essentially obviate the theory of evil: it is just a special case of standard moral theory, an uninteresting extension of what we already believe.3 But this strategy cannot work. One worry about the proposal is that it appears to treat the concept of evil simply as a theoretical notion in ethics. Yet the concept has its home, not in moral theory, but in ordinary moral discourse. ‘Evil’ has an independent, if less than clear, meaning. Theorists of evil ought not to treat the term as if it were up for grabs, to be defined however it suits our moral theories. If we wish to take the moral phenomena seriously, we need to take the ordinary notion seriously. Otherwise we risk giving a theory of something that has little to do with evil as we know it. We may in the end decide to revise or even abandon the ordinary concept, but we can only make that decision once we know what we are talking about. It is possible that the notion of evil does in fact conform to this approach: it denotes whatever proves to be the worst possible character. But this would have to be shown by seeing how well the account conforms to the folk notion, not by simply annexing the notion into one’s favorite moral theory.

A far more serious difficulty is this: the proposal essentially defines evil out of existence, shrinking the range of possibilities to a mere point. On this sort of view the evil person would have, at every moment, to be doing the worst possible things, with the worst possible feelings and motives.4 This would make the theory and the concept of evil uninteresting: not even Satan, it seems, would qualify. (Wouldn’t Satan want a little time for himself?) A theory that requires a satanically depraved character
to qualify as evil seems too demanding. A theory that insists on an extra-
satanic dedication to malevolence is just silly.

We can improve the approach by toning it down: an evil person is
indeed the worst kind of person, but this kind of person comes in degrees.
One need not at every moment do, feel, and desire the worst possible
things; it suffices for being evil that one come close enough to this. And
the “close enough” mark may be far enough from the logical extreme that
the notion of evil actually has application, not just to the satanically fiendish,
but to mere mortals. Now we understand evil as a more or less broad
category that marks off the worst region on the scale of good and bad
character, not just the worst point. To my knowledge, this is how theorists
have actually approached the matter. But two further difficulties emerge.

Where on the scale does the merely bad give way to the evil? How
close to the extreme of utmost depravity is close enough? You might think
it doesn’t much matter: wherever we feel inclined to draw the line. And
this is likely to be rather vague, with a number of cases being neither
clearly evil nor clearly not. This appears to be a common sentiment amongst
commentators on the subject. Thus Laurence Thomas dubs people evil if
they are “often enough” prone to commit evil acts (1993, p. 82); and John
Kekes calls evil those who are “regular” sources of undeserved harm (1990,

Yet it does matter where we draw the line: for the distinction between
bad and evil is not merely one of degree. It is a qualitative difference. As
I noted above, we cannot get from bad to evil by adding any number of
“very.” Calling individuals evil places them in a different moral realm
from the merely bad. These do not appear to be arbitrary linguistic
artifacts. Moreover, to say that the dividing line doesn’t matter is just to
say that the distinction between evil and bad doesn’t matter. Which is to
say that the theory of evil, as such, doesn’t matter. We could just as well
call it the theory of the (very) bad. A theory of evil ought to accommodate
the idea that there is a significant moral boundary between the evil
and the merely bad—or, if not, explain the appearance that there is such a
boundary. We will see that the bad/evil distinction poses difficulties for a
number of accounts.

A second concern is that allowing for better and worse within the
category of evil appears to violate the sense that evil doesn’t come in
degrees. Perhaps this is just the price to be paid for ensuring that the
MORAL MONSTERS AND SAINTS

The notion can apply to an interesting range of cases. Moreover, the intuitions favoring the view that evil is an all-or-nothing affair are not overwhelmingly strong; some may not share them at all. The worry is not extreme. But it would be a plus if we could explain why it seems odd (to some of us, at least) to call a person or action “very” or “somewhat” evil.

We don’t just want a theory of evil that is substantively motivated somehow or other; we want one that makes sense of evil as a significant moral category. What are the options?

3. The options

3.1 Harm-based accounts

Most discussions of evil either rely solely on intuitional evidence or treat the subject as a mere footnote to one of the popular varieties of moral theory. One noteworthy exception occurs in the writings of John Kekes. In his book Facing Evil and elsewhere,7 Kekes takes evil seriously as a philosophical subject while rooting his account in a broader view of morality’s purpose: viz., to promote human welfare, particularly by minimizing evil. With this in mind, Kekes argues that the evil action is an act that inflicts evil—understood as undeserved harm—on someone. His view of evil character derives from this: people are evil if they are regular sources of undeserved harm. We can set aside questions about how serious or frequent the harm-doing must be.

This account has three significant virtues. First, it has a principled basis in an attractive theory of the role of morality. It does not rely solely on pre-theoretical intuition. Second, it gives us a unified treatment of the various kinds of evil: a person is evil by virtue of being a frequent performer of evil actions; in turn, actions are evil by virtue of inflicting non-moral8 evils. Third, the theory offers an appealing explanation of our interest in evil: we care about it because it involves undeserved harms, and we want to minimize the suffering of such harms. The evil person, being a regular source of undeserved harm, is the most dangerous sort of person, the greatest threat to society. Isn’t this what we most care about in a person’s character?

Kekes does not endorse consequentialism, but his basic approach is congenial to a consequentialist outlook.9 For our purposes we can assimilate it to the consequentialist approach to evil: the consequences of our actions on each other are what ultimately matter. The motives, intentions,
or feelings of the agent count only insofar as they contribute to the suffering of undeserved harms. We don’t particularly care whether Hitler cried into his pillow every night.\textsuperscript{10} We care about the millions of lives he destroyed.

Be that as it may, harm-based accounts—as I shall call them—are not tenable. First, they are intuitively implausible: the concept of evil\textsuperscript{11} centrally concerns matters of motive and affect, and harm-based theories cannot account for this. For example, someone might qualify as evil despite never bringing about any harm. Take the vilest person you can imagine and make her a quadriplegic with no ability to communicate: living in silent spite, she wishes nothing more than the greatest suffering for her fellow creatures, and takes the greatest joy in witnessing others in agony. The handicap makes her less dangerous, but it scarcely makes her a better person. Nor does it make her anything better than evil.\textsuperscript{12} Even if harm-doing were necessary for evil, it is hard to believe that a person’s motives could fail to be relevant to the determination of whether one is evil. Those who intentionally harm others out of sheer malice are surely worse, and better candidates for evil, than those who do harm only as a means to achieving their ends.

The second problem with harm-based theories is their apparent inability to support a robust bad/evil distinction. How frequent and gross a harm-doer does one have to be to qualify as, not merely bad, but evil? To insist that evil people inflict undeserved harm whenever possible is absurdly demanding, and the view that evil persons usually do so is little better. Focusing on the magnitude of harm also seems inadequate: there appears to be no point at which, by committing still greater harms, we cross a moral Rubicon of the sort that might distinguish the evil from the bad. Besides, someone who is generally benevolent but would destroy the world whenever given the chance is not credibly regarded as evil, even if she is sane. Perhaps the evil person is one who never acts altruistically. But this is too undemanding: someone committed to a crudely egoistic moral theory might never act altruistically, but such a person may care enough about others not to qualify as evil. Kekes’s suggestion, that someone is evil who “regularly” inflicts undeserved harms, is more promising. But is there really such a great moral difference between someone who inflicts undeserved harm regularly and one who does so somewhat less often? It is not at all clear what the special significance of being, versus not quite being, a regular (undeserved) harm-doer could be. The worry is not that the distinction is vague, but rather that it lacks the kind of significance that the bad/evil distinction appears to have.
The evil character confounds harm-based theories, including standard consequentialist accounts of evil. Consider the case of a misanthrope who takes great pleasure in seeing people suffer at each other's hands. This pleasure is essentially voyeuristic: he must not harm anyone himself. He takes every opportunity to witness such suffering. This man feels nothing but malice for his fellow human beings. He is evil. Can a consequentialist say this? Evidently not: this man harms no one. At worst, he fails to help. Consequentialists could identify the evil person in terms of her motives: having those motives that tend to produce the worst consequences. But this person doesn't have such motives: his motivational structure tends not to produce especially bad consequences. Could a consequentialist say that merely having certain motives, regardless of whatever other motives one has or how they combine, suffices for being evil? It is hard to see on what grounds: if we are concerned to promote the good, then what matters is how a person's character disposes her to behave. This person behaves badly, but not that badly (from a consequentialist perspective). Call her evil, and the consequentialist will have to call quite a lot of people evil. My point is not simply a version of the familiar complaint against rule consequentialism that it fetishizes rules—or, in this case, traits—at the expense of abandoning its fundamental rationale. The point is more that we cannot judge a person's whole character simply on the basis of one or a few traits. We have to look at how it all fits together. Consequentialists might derive a credible list of virtues taken singly, but they cannot so easily account for the moral quality of people's characters taken as a whole.

It is a difficulty for consequentialism that, at least in its usual incarnations, this theory cannot readily accommodate the category of evil. It is more worrisome that consequentialists seem unable to say that some truly vile people are all that bad. For a consequentialist, a car thief likely counts as having a worse character than our depraved voyeur. This is implausible.

3.2 Motive-based accounts

The concept of an evil person is substantially inward-looking: it concerns not just the causing of undeserved harms—if it concerns that at all—but also the agent's motives. Perhaps evil persons are just those who are motivated in the worst possible way. We might say they have evil wills. This sort of view is broadly Kantian in tone, but one need not be a Kantian to endorse a motive-based view (as we may call it).
What is it to have an evil will? Two views are at least *prima facie* credible. First, we distinguish the very worst sorts of motives, the evil ones, and claim that someone is evil just in case evil motives govern her behavior often, or perhaps regularly. ("Usually" or "always" would be too strong, for now familiar reasons.) Laurence Thomas adopts something like this sort of view, basing it on his account of evil action. He claims that a person commits an evil act "if he delights in performing a harmful act that has a certain moral gravity to it . . . and if the person is not animated by understandable considerations" (1993, p. 77).16 Someone has an evil character if he is "often enough prone to do evil acts" (p. 82). This view is both motive-based and action-based, deriving an account of evil character from a theory of evil action, where the evil action is distinguished substantially by the agent’s motives. In particular, the evil agent does evil, not reluctantly or out of "understandable considerations" like rage at a severe provocation, but with delight, evincing a "deadening of moral sensibilities" (pp. 76–77). (This does not mean a failure to understand the relevant moral considerations, but rather a failure to be properly moved by them.) Thomas’s view is also partly harm-based, and thus looks both inward and outward. This strikes an appealing balance between our concern for the damage evil people do and our interest in the motives of the evil person.

A second approach is to look, not for extremity of motive, but consistency of motive: the evil person consistently—usually or always—governs her behavior by motives that are not morally good.17 (An insistence on having consistently bad motives would, again, be too demanding. Most behavior, even in moral monsters, is perfectly innocent.) I am unaware of any theorists who have explicitly taken this route, but S. I. Benn comes close (1985). Employing the language of wickedness, he suggests that "a person may be wicked because the maxims that order his life are, by and large, evil maxims, that is, maxims that no one ought to act on at all" (p. 796). If we read ‘evil maxims’ in an intuitive manner, this claim is too strong. "I shall steal a pack of gum whenever I wish to” may be an immoral maxim, but we would not regard it as evil. Benn apparently means nothing stronger than "immoral.” The evil person, then, is governed "by and large” by immorality. Even this is too demanding, but it is doubtful that Benn thinks it a necessary condition for being evil (thus he says "a person may be wicked . . . ”). Amongst the wicked he (with Kant) counts the merely selfish: those who do not actively seek immorality or the bad, but whose
Motive-based accounts of both types are inadequate. "Frequent evil-doing" views like Thomas's do seem capable of explaining the bad/evil distinction with respect to actions: to delight in immorally inflicting great harm on others does seem a special case of wrong-doing, worse in kind than any other. It does seem evil, and not merely wrong. But when it comes to having an evil character, such accounts fall short. For a frequent evil-doer could in most respects have a relatively normal character. Extremes of cruelty and sadism can surface, even with some regularity, in people who are normally dutiful, honorable, generous, even kind—not the sort of people whom, when fully informed about their characters, we would readily deem evil. This is one of the depressing lessons of the many mass atrocities of the last century. Thomas could insist on a frequency of evil-doing that rules out such complexity of character, but then his account will become implausibly demanding. More to the point, there does not seem to be any plausible degree of frequency or regularity of evil-doing that could provide a natural demarcation between the evil and the merely bad person—not even a vaguely specified one. The regular or frequent evil-doer seems different only in degree from the not-quite-regular or frequent evil-doer. The transition from one to the other does not seem particularly significant.

We should also wonder about the motivation for a motive-based account. If evil concerns us not simply for its effects, but also (or instead) for what it says about an agent's psychology, then why stop at the agent's motives? Why not also incorporate the agent's affective dispositions? Surely it makes a difference to one's character whether one takes pleasure in the suffering of others. Recall the voyeuristic misanthrope; this person is evil, not simply for his motives, for having an evil will, but because he enjoys seeing people in agony. Indeed, it seems that this more than anything is what drives us to call him evil. A decent person does not
rejoice at the spectacle of a child on the rack. We rightly think less of such an individual (to put it mildly). How we feel about the predicaments of our fellow creatures makes a difference in our characters, quite apart from how we are motivated to treat them. So why ignore such feelings in our account of the evil character, unless—with the harm-based theorist—we ultimately care only about the harm evil people do? If you include motives, you should include affect.

Someone might object that affect is more or less independent of volition—how we feel about things is substantially beyond our control—and thus we ought not to be morally assessed for it. But supposing that we should not be blamed for having certain feelings, it does not follow that we are no less morally estimable for having them. Our disgust with those who enjoy their confederates' travails does not depend on any assumption that those feelings are voluntary. I shall not argue the point at length, however, since I have addressed it elsewhere (1999).

I have challenged the grounds for including motives but not affect in our account of evil character. But we ought to include affect in our theory, however the account is motivated. Motive-based accounts are incomplete. In fact, while I used Thomas's view to illustrate the motive-based approach, his account appears to incorporate affect as such:20 to "delight" in performing an immoral and harmful act, and to exhibit "deadened moral sensibilities," is not merely to have certain motives. It is to have, and lack, certain feelings. Though ultimately unsatisfactory, his view points us in the right direction.

3.3 Affect-based accounts

Someone sufficiently impressed by the significance of affect for having an evil character might suppose that that is all there is to being evil: being disposed to take pleasure in the suffering of others, and perhaps also to suffer at the pleasure of others. Colin McGinn recently developed such an account (1997). While this sort of view is remarkably plausible considering how far it diverges from the traditional focus on action in moral theory, it too is incomplete: matters of motivation count as well. Governing one's life by evil or immoral maxims, or letting hate be one's guide, is clearly relevant to having an evil character. While we may dispute Kant's assessment of the unsympathetic but dutiful man, we
would be hard-pressed to regard a truly dutiful person as evil, however deranged her sympathies. A further worry is whether an affect-based theory can supply a firm bad/evil distinction. Once gain we would be asking too much to insist that evil persons must usually or always take pleasure in others’ suffering. Yet allowing a mere lack of sympathy or otherwise appropriate affect to suffice for being evil is too permissive. And placing the cutoff at “frequently” or “regularly” experiencing un-wholesome affect amounts to abandoning the idea that there is a qualitative difference between bad and evil.

3.4 An affective-motivational account

Harm-based views underestimate the importance of psychology for the theory of evil character, whereas motive- and affect-based theories take too narrow a view of the relevant psychology. Each type of theory risks failing to underwrite a robust bad/evil distinction. In two earlier papers I have outlined a pluralistic account that incorporates both affective and motivational components (Haybron 1999; forthcoming). Here I shall briefly summarize the theory and then argue that it does, like one’s version of Benn’s account, permit a robust bad/evil distinction. In the remainder of the paper I will argue that this view is well-motivated. We can think of being virtuous or vicious in terms of having a certain orientation to the good (where the good is understood in a broad manner that can include the right). Insofar as one’s disposition is to be appropriately moved and motivated by the good, we may say that one is aligned with the good. (I say “aligned with” because we need an expression that can encompass both conation and affect, and other terms like ‘loves’ have undesirable connotations.) One is virtuous insofar as one is at least adequately aligned with the good (some vices consist merely in being too-weakly aligned with the good). To be evil is, on my view, to be consistently vicious in the following sense: one is not aligned with the good to a morally significant extent. Evil persons are either wholly unaligned with the good, or they are moved and motivated by it so little that it makes no significant difference to the moral quality of their characters: morally speaking, they are not significantly better people than the wholly unaligned. They have no good side, but are consistently vicious. They do not show real compassion or conscience, among other things. Taking a cue from Thomas, we could
say that they possess profoundly deadened or perverse moral sensibilities. Yet, contra Thomas, they need never perform evil acts or do any harm at all (though that is most unlikely). Evil persons could even tend to conform their actions to duty, so long as this is not properly grounded in genuine respect or concern for others: perhaps they are merely being prudent, mechanically parroting rules they were taught, or pursuing aspirations of nobility. (A Hitler, for instance, might fulfill some of his duties purely out of the belief that not doing so would show him to be weak or ignoble.) Alignment or otherwise with the good is an internal, psychological matter: someone perfectly aligned with the good may inadvertently cause a disaster. This is not an Augustinian view, since evil need not be a mere privation; evil persons may, like Satan, be strenuously opposed to the good. My discussion will focus on less extreme cases, since those are the most controversial.

This view may seem to be both too demanding and too permissive. Too demanding, because most of those whom we tend to call evil—war criminals, torturers, mass murderers, etc.—would surely be excluded by the account. It is conceivable, if implausible, that even Hitler would not qualify as evil. Most evil-doers probably are significantly aligned with the good (though, of course, not enough): there is no reason to suppose that the typical Nazi war criminal, for instance, could not also have been a true friend and loving companion to some. Insofar as that is true, then those individuals were not evil; they truly had a good side. Our inclination to call them evil depends on the fact that that side of their characters is neither apparent nor particularly salient from our usual perspective. Were we fully appraised, not only of their worst acts, but also their best—if we knew about them what their loved ones know—it is unlikely that we would continue to feel comfortable in calling them evil. (This, I have argued, is what makes the lead character of the television series The Sopranos, Tony Soprano, so confounding [forthcoming]. Ordinarily we might see no difficulty in totally condemning such a mobster. But the show, which focuses on his home life, reveals too much of a complex character to permit us that comfort.)

My view may seem too permissive because it doesn’t appear to require any blatant extremity of vice to qualify as evil. One need not be extraordinarily cruel, for instance, to qualify as evil. Indeed, a law-abiding professional—a “good citizen”—might count as evil on my view, since
one can go through the motions of virtue without ever possessing it.\textsuperscript{27} Someone who perpetrated the horrors of a Mengele, by contrast, might not rate as evil. But complete disalignment from the good is more vicious than it sounds. It is not about being a mere ne’er-do-well. For among other things it requires a complete, or near-complete, lack of human concern for the well-being of others. What would we say of a sane man who allows, without compunction, a six-year-old to drink from a bottle of Drano, and impassively watches the child suffer an agonizing death? Or someone who would gladly make millions suffer if it suited her purposes? There is plenty of brimstone about the evil person; but sometimes we must look harder to find it.

Though this account of the evil person departs from ordinary ascriptive practice, I believe it coheres well with our considered intuitions, while enabling us to make sense of the bad/evil distinction: evil people are moved and motivated in ways that differ \textit{radically} from ourselves, and even from ordinary wrong-doers. Almost all of us are more or less on the same team, so to speak: we are, for the most part, moved and motivated in morally decent ways. Even violent criminals tend to share most of the moral sensibilities of decent citizens. There is nothing mysterious or alien about the psychological makeup of the average bad person. The evil person, by contrast, is not at all like us. He serves a very different master, and follows very different rules.

This account also explains the sense that evil doesn’t come in degrees: one either is, or is not, significantly aligned with the good. One either is or isn’t evil. Yet we need not sharply limit the range of evil characters: within the category of evil there is still morally better and worse.

4. Motivating the account

4.1 A broader characterology

The most serious objection to this account concerns its motivation: why should we conceive of evil \textit{this} way, however intuitively plausible the account may be, and however well it explains the bad/evil distinction? It is not obvious what purpose we serve by employing such a concept. This worry is sharpened by the fact that the theory seems to discount the importance of the harm bad people do. It also leaves unanswered the question of the relationship between evil character and evil action. Concerns about
motivation are most pressing in light of the *prima facie* appeal of frequent evil-doing accounts like those of Kekes and Thomas, so I shall aim my reply at them. Such views do seem to fit best with ordinary ascriptive practice, even if I am correct that the affective-motivational account better fits our considered intuitions. And, as we saw, those views offer a plausible explanation of why we should care about evil: evil persons are the most dangerous, the ones who commit the most heinous crimes.

I think the appeal of frequent evil-doing accounts depends on a confusion. Such theories do get at one kind of moral extreme; but this is not the kind of extreme that concerns us when we are deciding who has the worst sort of *character*. For in matters of character we are concerned, not with blame or credit, but with esteem and disesteem. And there is no obvious reason why a person’s moral estimability should hinge on what she does, *beyond what this reveals about her psychological makeup*. What made Hitler such a monstrous person is not that he happened to perpetrate the atrocities he committed. His actions were the *expression* of a depraved character; they did not constitute his character. Had he been thwarted, or a hopeless bungler, or were he too cowardly even to attempt what he wished to carry out, he would scarcely have been a better person. Ineptitude and cowardice cannot improve one’s character; they merely defang one’s other vices. Of course, it is only in extraordinary cases that depravity will fail ever to manifest itself in action. And in practice we can only judge people’s characters based on observed behavior. But this is an epistemic problem; it has no bearing on how people’s characters actually are.

Notice that frequent evil-doing accounts are attractive also because they fulfill a seemingly important desideratum: they make it easy for us to say that the usual exemplars of depravity, the “successful” Hitlers and Bundys of the world, are evil. It is plausible that those who would stress this putative requirement would not be similarly concerned to account for the world’s frustrated—but no better constituted—Hitlers and Bundys. Kekes, for instance, explicitly requires an evil person to have established a pattern of evildoing (1990, pp. 48–49). This, I would suggest, is because frequent evil-doing accounts do not characterize what it is to have the worst sort of character. Their subject matter is, rather, the *moral criminal*, or an extreme variety thereof. To be a moral criminal is, roughly, to have performed seriously immoral acts that reflect major deficiencies of character. The qualifier ‘moral’ is meant to stress that this is not a legal but
a moral kind (through I will sometimes omit the qualifier). We often refer to individuals as criminals merely by virtue of their having done something bad, but there is more to it than that. For we would hesitate to denounce someone as a criminal if we believed his crimes to be completely out of character, aberrant bits of behavior that fail to reflect any well-embedded disposition. We might wish to add that criminals must be, on the whole, vicious, or at least not virtuous: they are not good people. Criminality is partly a matter of character and partly a matter of history. Those whom Kekes and Thomas—and lay ascriptive practice—would deem evil are better described as an extreme variety of criminal. And because criminality is linked mainly to the performance of certain sorts of actions, the notion can only encompass part of a person’s character. Much of moral significance is necessarily left out. The notion of criminality is too blame-oriented and too narrow for the purpose of specifying the worst sort of character. For that, we need to set aside questions of history and consider every morally important aspect of an individual’s character, not just those that pertain to certain sorts of action. We need a distinct, purely aretaic, notion.

Some proponents of frequent evil-doing accounts might object that their concern *is* with estimability and not with blameworthiness: what makes the frequent evil-doer evil is not the actual performance of evil but rather the propensity to attempt it. And it is worse to be disposed to perform frequent acts of evil than to be, say, a quietly amoral bureaucrat. To some extent this is correct: the frequent evil-doer does have some important vices that are worse than the corresponding vices of the quiet amoralist. Yet some evil-doers have important virtues that the amoralist lacks: they have a good side, often exhibiting genuine compassion, conscientiousness, and other moral virtues. And the depth of an amoralist’s depravity may become apparent only in unusual circumstances, as when gross immorality is clearly prudent, or when faced with opportunities for profound callousness. That the evil-doer is worse in some very important respects does not oblige us to call him a worse *person*. (We may, given that he actually does do evil, consider him worse on the whole in one way: he is more blameworthy, more deserving of censure and punishment. And consider whether he would really seem worse than the quiet amoralist if he somehow did not manage to perform evil acts. The evil person is not our only exemplar of immorality; the criminal is another.)
The distinction between evil and criminality mirrors a distinction we find at the other end of the moral spectrum: that between moral saints and moral heroes. Like being a criminal, being a hero is partly a matter of character and partly a matter of history: the hero has performed some great deed or deeds, and has done so not by accident or by acting out of character, but by acting from a virtuous disposition. The hero need not be perfectly virtuous, and may in fact be deeply flawed; indeed, some of the most interesting heroes are seriously flawed. We would not, however, regard as a hero someone whose character was generally poor. The hero is a good, or at least decent, person who has done great things.

The saint, by contrast, is the image of moral perfection: this person has no significant moral flaws or vices, but is perfectly virtuous, or nearly so. He is pure, the positive counterpart of the evil person. As with the evil person, the saint’s status depends partly on her motivational economy and partly on her affective dispositions: perhaps Kant’s unsympathetic but dutiful man qualifies as a good person, but he is no saint. In fact it is the saint’s profound compassion that most impresses us, not her commitment to duty. (Consider the popular image of such esteemed individuals as Mother Teresa or the Dalai Lama.) The saint must not only promote the welfare of others; she must care, in a humane and empathetic way, about the welfare of others. I would suggest that the saint is best characterized in parallel with the evil person: as (morally) perfectly, or near-perfectly, aligned with the good. (In keeping with our definition of the evil person, we could say that the saint is “not significantly unaligned” with the good. But this is awkward.)

Someone might object that saints differ in an important way from evil people as I have characterized them: whereas evil does not require the performance of seriously immoral actions, sainthood does require the doing of many good deeds. Sainthood consists substantially in performing the right kinds of actions, even if evil does not. We would not, for instance, call a completely inactive person a saint. Let’s assume that the worry does not simply reflect an epistemic limitation: in practice we would have to basis for evaluating the character of someone who displayed no observable behavior. I think we can explain any apparent difference in terms of general differences between virtues and vices. For example, virtues are more closely linked to action than vices. A truly kind person will act kindly when the occasion arises; whereas it is easier for a callous person
not to act callously: perhaps she fears retribution. In general, vices may be prevented from issuing in action by various factors that do not impact the saint, including other vices. Someone’s failure to act viciously when the occasion permits it need not reflect a less vicious character; whereas someone’s failure to act virtuously does reflect a less virtuous character. While claims about the unity of the virtues are overstated, they do reflect an important truth: virtues are less separable than vices. In general, virtue is more demanding than vice.

In fact it is not so clear that sainthood requires good woks: suppose a proven saint were badly crippled in an accident, so that she could no longer do good deeds. This would not entail a moral decline in her character, but would merely hinder the expression of her virtues. Indeed, we can readily imagine someone’s character improving through such a turn of events.

Our account of the evil character is not ad hoc or sui generis, but belongs to a broader characterology that is itself plausible. The evil character defines one pole of a moral continuum that has the saint at its other pole. Or rather, the evil person and the saint define the humanly possible ends of this space. For beyond the evil character we might posit, at least as logical possibilities, individuals whose capacities for hatred, malevolence, and the like outstrip those of mere mortals: demons or devils; at the extreme, Satan. Similarly, saints may be surpassed in goodness by angels, gods, or, at the summit, God.33 The relevant moral space is purely aretaic: from most vicious to most virtuous.34

Heroes and criminals occupy a different moral space, one that is not purely aretaic—does not just concern moral estimability—but largely concerns creditworthiness and blameworthiness: which individuals deserve the most credit for their actions, and which ones deserve the most blame or censure? Again, the hero and the criminal define the humanly possible ends of the spectrum, whereas God and Satan comprise the logical extremes. But here they occupy different roles: as exemplars not of perfect virtue and vice, but as the perfect “good-doer”—for want of a better word—and evil-doer.

The reader may grant the preceding account of moral saints and evil characters yet ask why heroes and criminals should be conceived in such a mixed manner. Wouldn’t it be more natural to eschew any references to character, construing them purely in terms of what they have done? Such an approach would be simpler, but intuitively less plausible. The reason is
likely practical: heroes and criminals would be less useful as exemplars of morality and immorality if their characters could diverge too sharply from the quality of their noteworthy actions. We would not wish to venerate basically rotten people as moral heroes, whatever they have done. Likewise, we should be reluctant to demonize people whom we believe to be basically good.

If all this is right, then we can see where frequent evil-doing accounts go wrong: they are concerned with a legitimate moral extreme; just not the right kind as accounts of the evil person. And the affective-motivational theory of evil character gains plausibility from its role in a broader moral scheme: it is not ad hoc, nor is it the ill-conceived progeny of a confused folk morality. It is both intuitively plausible and reasonably well motivated.

4.2 An objection: only action counts

Advocates of frequent evil-doing views may object that this sort of motivation is relatively weak: why should we care about the evil person—or for that matter the saint—except insofar as she does bad (or good) things? Action is all that really counts in the final analysis, and we should evaluate people's characters in relation to what they do. All that the foregoing arguments have shown is that we can situate a dubious theory of evil character in a larger, but no more compelling, characterology. Only frequent evil-doing accounts are adequately motivated. Or so the objection goes.

I shall not attempt a complete rebuttal of this objection, since the fundamental question is why, in general, we ought to care about people's characters, if not simply because we care about what they do. That question is too broad for this paper. (Though note, for starters, that such a fundamental of life as friendship would be impossible without genuine concern, quite apart from how such concern issues in action.) Why might we care specifically about whether someone is evil as I understand it, versus merely bad or an evil-doer?

One reason is that the evil character provides us with a moral anchor, an anti-ideal. Even if no real person qualified as evil, the notion would still be useful for illuminating our moral ideals and defining the moral space within which we situate less fiendish individuals. It is useful to see which traits are most abhorrent, and which individuals best approximate the moral nadir of the evil person. Knowing what it is to be evil tells us who we most want not to be.
A second reason is that the ascription of evil affords various simplifications. The evil person, in the consistency of her vice, presents us with none of the difficulties posed by more complicated, merely bad individuals—mere criminals. She has no good side. We need not worry that we shall lose our resolve to deal appropriately with her if we learn too much about her. And we can confidently predict how she will behave: never from decent motives. When she does do evil, we need seek no special explanation: lacking the moral qualms that inconvenience the rest of us, then of course she will do awful things.

Moreover, the evil person is beyond ordinary moral criticism and dialogue: he has no better nature to which we can appeal. Morality has no significant foothold in him. He is arguably beyond redemption through rational deliberation; nothing short of a conversion or reprogramming, it seems, could rehabilitate him. He understands morality, and may be perfectly capable of moral decency, but he rarely if ever exercises this capacity. Because of this, the evil person is also beyond society: a moral exile. No one can expect good from her, or engage in genuine friendship with or expect love from her. She may exhibit at best a kind of reptilian affection for others, devoid of genuine empathy, human concern, or respect. One reason why criminals and their families so often go out of their way to argue that they aren’t evil, that they are really just like everyone else, is precisely to avoid this implication of the label. How could even a mother love an evil person? And how could one regard as love the affection shown by a father or husband who merits the term ‘evil’? Seriously to regard someone as evil is to claim him ineligible for any human relationship. This is no small condemnation. (Consider the Jewish victims of the Holocaust, whose extermination was facilitated by the view that they were evil [Thomas 1993].)

The evil person is something of an alien, lying somewhere between the human and demonic. We call her, not coincidentally, a monster. The appellation ‘evil’ thus serves to distance its subjects from the rest of us, to emphasize the profound moral and psychological gulf between them and us. Interestingly, calling someone a saint has a similar effect: the saint occupies a point in moral space between the ordinary person and the angelic, and is almost superhuman in his moral purity. Unlike most heroes, the saint need not struggle to overcome temptation: uninfected by the impurities of a normal person, he is largely beyond temptation, for acting well comes naturally. We should not be surprised, then, that few
people would really want to be regarded as a saint. The saint seems sufficiently alien to be a less than ideal friend or relative (save perhaps as a parent). For one thing, she lacks the usual human weaknesses and thus, we may suppose, the ability fully to understand or empathize with the difficulties her compatriots must face. For another, her very perfection creates a deep moral asymmetry between herself and her friends and relatives: she is manifestly a better person—indeed, perfect. Who wants a friend like that? Relationships require a certain degree of equality. They flourish through the matching of strengths and weaknesses, and through the complementarity of vice as well as through the virtues. Otherwise we shall have too much, or too little, to offer each other. Like the notion of an evil person, the notion of a saint serves to distance its object from the rest of us. One is clearly better than the other, but no sensible person really wishes to be either.37

The fact that the notion of evil underwrites these kinds of simplification may seem to count against its utility: people are rarely so simple, and it is even more rarely that we can know enough about someone justifiably to regard them as evil. The term is thus liable to be abused more often than not. And its legitimate uses will prove highly restricted. This seems right, but it does not count against my view: the notion is still useful insofar as it designates a real moral extreme. Evil persons are probably neither common nor easily identified, but they do comprise a genuine and interesting class of moral characters. The utility of a notion is not simply a function of how often we may be warranted in applying it.

The present worry also indicates that the theory of evil matters. We often make judgments about those whom we take to be highly vicious, and sometimes we refer to them as evil. This is an extreme form of condemnation, and we had best get it right. If I am correct about the simplifications involved in ascribing evil, then regarding individuals as evil amounts to treating them as moral write-offs, as monsters who are not fully human and certainly not fit for any kind of society.38 They are not fit even for friendship or familial relations. And if I am correct about the looseness of the connection between evil-doing and evil character, then we are probably mistaken in most of our ascriptions of evil: we regard the wrong individuals as moral write-offs. Most of those we call evil probably are not; and worse, we have failed to recognize the level of depravity that can reside even in seemingly respectable individuals. The very worst
members of the human race need not be criminals at all, much less murderers, rapists and the like; perhaps a few of them reside in the U.S. Congress, in the executive offices of major corporations, or in high church positions.39

Similar mistakes occur regarding those whom we venerate: we tend to overstate the virtues of do-gooders while understating the virtues of less conspicuous people. Only those who do great works gain widespread recognition, so only heroes are eligible to be hailed as moral saints. But real saints—or their nearest cousins among us—may lack the opportunity or inclination to be conspicuous doers of good works. Humility, a desire not to “show up” those around them, or simply a situation that does not call for (or allow) prominent acts of heroism, may cause a saint’s goodness to garner little attention.40 A saintly individual who finds herself destitute is unlikely to gain notoriety. Most heroes are not saints, and there are few people so well known by the public that their status as reputed saints is truly warranted by the evidence. Mother Teresa did wonderful things, but was she morally perfect, or very close? Not likely. Likewise for other popular candidates for moral sainthood such as Gandhi and Martin Luther King. Exemplary persons they were; but morally better ones may well reside, anonymously, in most communities.

If we ascribe evil incorrectly, particularly as we are wont to, then we shall fail to understand why evil-doing occurs: most evil actions are not the product of evil people. (This observation is hardly original, but we have seen that the connection between evil-doing and evil character is looser than most writers suppose.) They are, rather, performed by people much more like ourselves than we may care to admit. As a result we may blind ourselves to the risk that we, or those near to us, might participate in moral atrocities. (Consider the reluctance of many Americans to believe that the Oklahoma City bombing might have been done by a fellow countryman, despite clear indications to this effect.) And those who do evil may be less willing to come to terms with their actions: they may falsely consider the admission to having done evil tantamount to considering themselves evil. Moreover, the conflation of evil persons with evil-doers may lead us to punish the latter far out of proportion to the severity of their crimes or the turpitude of their characters. (Such confusion may help to explain some of the excesses of the American penal system.) At any rate, it is unjust to regard an individual without a hint of compassion or
conscience as a better person than someone who is largely decent but has some serious character flaws, simply because the latter alone has done something horrible. This individual may merit greater blame and punishment, but not less esteem.

In general, an understanding of the evil character illustrates the importance of clearly distinguishing matters of esteem and disesteem from matters of credit and blame. Frequent evil-doing views appear to conflate the two issues.

5. Conclusion

Loose ends remain. I will mention just one: the affective-motivational account does not clearly connect with any theory of other kinds of evil. Frequent evil-doing theories, by contrast, can provide a unitary account of the various sorts of evil: an action is evil if it causes the suffering of non-moral evils, and people are evil if they frequently perform evil actions. This is a significant advantage, and it would be odd if our theories of evil—particularly those of evil character and action—did not connect in any meaningful way. I would suggest that we reverse the usual order of explication and understand the evil action in terms of its relation to the evil character: e.g., perhaps, an action is evil if it manifests profoundly deadened or perverted moral sensibilities—the sensibilities characteristic of an evil person. (One need not be evil to do evil on such a view: one may have the usual sensibilities while not exercising them at all times.) Such an approach still fails to encompass non-moral evils, but it is far less clear that a single framework need incorporate those: we often refer to anything bad as an evil, simply because it is awkward to speak of things like pains as “bads.” ‘Evil’ is not a superlative in such contexts, but merely denotes the contrary of ‘good’. However, a full discussion of such questions will have to wait for another occasion.

Daniel M. Haybron

Department of Philosophy
University of Arizona

NOTES

1. I wish to thank Douglas Husak for helpful comments on an earlier version of this paper.
2. In an earlier paper I did refer to degrees of evil (1999). This now strikes me as unnatural; as I note later, it is preferable to speak of morally better and worse than more or less evil.
3. One related approach would not have this result: start with an account of evil as the worst possible, then build a complete moral theory from that. While the theory of evil may well shed light on general moral theory, this tactic surely goes too far.

4. Or some subset of these, depending on one’s theory of character. The evil person, thus understood, appears to be the negative counterpart of Susan Wolf’s saint, “whose every action is as morally good as possible . . . who is as morally worthy as can be” (Wolf 1982, p. 413). As a serious account of sainthood, this view suffers the same deficiencies as the present view of evil character. However, Wolf’s purposes in that paper do not require a full-blooded account of the conditions for sainthood, so this is not really a problem for her.

5. As Garrard notes (1998, p. 44). Her paper also addresses the significance of the bad/evil distinction.

6. As such it would still be interesting: extremes of moral depravity are worth studying whether or not we employ the concept of evil.


8. I use this term to denote the bad things, like pains, that we commonly refer to as evils. Though I refer to them as “non-moral,” they may (but need not) be implicated in immorality. We might call them natural evils, but this term is commonly used to denote only evils not caused by human agency.

9. Consequentialists unhappy with Kekes’s reliance on desert needn’t worry: that aspect of his account is not relevant to the discussion that follows.

10. As Frederic Schick put it to me.

11. There is actually a family of evil concepts, but for convenience I often refer to “the” concept of evil.

12. This example is essentially the same as one given by S. I. Benn (1985).

13. I originally described this case in Haybron (1999). The point that follows is new. This case differs from Benn’s handicapped person in that the voyeur lacks the motivation to harm others, whereas in the other case only the means are lacking.


15. However, one should never underestimate the resourcefulness of the consequentialist. For an argument defending the ability of consequentialists to account for evil, see Calder (forthcoming). Consequentialists who conceive of virtue, not in terms of producing good, but in terms of agents’ attitudes toward the good, may not be susceptible to the criticisms given here. Thomas Hurka, for instance, has recently developed such a theory (Hurka 1992; 2001). Hurka defends his conception of virtue by taking virtuous attitudes to be among the intrinsic goods. Still, one might wonder whether his consequentialism really allows him to divorce the virtues from their consequences in this way. He can call the voyeur evil, but as a consequentialist he can truly regard relatively harmless monsters like this as morally worse than the average thief, who may do much more harm? (That is, without placing an implausibly high intrinsic disvalue on vicious attitudes.)

16. Todd Calder advocates a variant of Thomas’s theory (forthcoming). Eve Garrard defends a roughly similar account of evil action (1998). Instead of requiring delight, her view takes the evil action to consist primarily in the “silencing” (see McDowell [1978]) of the reasons against committing a wrongful act. As Garrard notes, we can readily imagine extending this account to cover evil character.

17. And meets the conditions for moral agency.

18. Garrard’s account of evil acts naturally suggests a view like the one under consideration: the evil agent is one for whom the reasons against wrongdoing are consistently silenced. This theory might also seem insufficiently demanding. More worrisome, I think,
is that it seems not to allow for more virulent strains of wickedness: those for whom the undeserved suffering of others is not silenced, but is rather a source of attraction. This is a general problem for perceptual theories of moral motivation: it is implausible that evil is never more than a kind of moral blindness.

19. For a depressingly vivid account of this history, see Glover (1999).

20. Recall that it also departs from a purely motive-based approach by incorporating harm.

21. One’s orientation toward the bad also counts, but I will set aside this qualification in what follows. This account is influenced by Hurka’s theory of virtue and vice (1992; 2001). Ronald Milo also takes evil to consist in an affective-motivational defect of character (1984; 1998).

22. I.e., having the appropriate affective responses. Being “appropriately” moved and motivated by the good consists in being moved and motivated positively by the good (and not, say, repelled by it); being moved to the right degree (and not moved, say, to the point that one is overly squeamish about inflicting necessary pain); and being moved by the good for the right reasons (not, e.g., out of some robotic impulse).


24. I will usually shorten this to “one is not significantly aligned with the good.” We shall eventually need some explication of just what it means to be aligned with the good, not just simpliciter, but in a morally significant way. For instance, an evil person may be concerned about her own good or exhibit such executive virtues as resourcefulness and courage (if we may call it that), and a lover of art might count as evil. But we need not resolve such worries here. At any rate, the qualifier “morally” is probably less important than it seems: there appears to be a perfectly natural sense in which an evil person, however prudent or aesthetically enlightened, may be completely opposed to the good. Satan, for example, is often said to have many virtues. (As Loren Lomasky pointed out to me.) But these are not morally redeeming; nor do they give cause for regarding him as in any important way a friend of the good.

25. They need not lack the capacity for these, however. Perhaps, e.g., they are wholly devoted to some monstrous ideology. I argue that evil persons qualify as moral agents in (1999; forthcoming).

26. We should distinguish genuinely human love, caring, and respect from mere affection. Even an evil person can have great affection for others, but this may be no more morally significant than someone’s affection for his favorite possessions.

27. Consider the phenomenon of “industrial psychopaths”—individuals who successfully pursue professional careers despite meeting the criteria for psychopathy (Sherman 2000). Whether psychopaths qualify as moral agents or not, such individuals illustrate just how little it takes, morally speaking, to pose as a respectable member of society.

28. For a helpful discussion of the distinction between blame and estimability in relation to the emotions, see Oakley (1992).

29. We might wish to add: against others, or involving serious harm to others. Such niceties are not important here.

30. But recall that Kekes, at least, denies this.

31. Cf., among others, Blum (1988), Urmson (1958), and Wolf (1982). My view differs substantially from Urmson’s, though this may result largely from a difference in our purposes. The notion of a saint that concerns us is purely a moral, and not religious, notion. Notice also that saints are specifically moral exemplars; as Wolf argues, they need not be the best people all things considered.
MORAL MONSTERS AND SAINTS

32. Melville referred to Billy Budd's Claggart as "the direct reverse of a saint" (1924).
33. Two points. First, the notions of God, Satan, and other supernatural beings function here as logical placeholders; the present discussion makes no assumptions about the existence of such figures. Second, one might wonder how it is possible for the saint's perfect alignment with the good to be exceeded. This is an interesting question, but it should suffice for present purposes to say that the saint's perfection is relative to the type of being she is: she is aligned with the good as strongly as is humanly possible.
34. Psychologist Jonathan Haidt contends that human beings naturally conceptualize the social domain in terms of a moral space, one of whose dimensions is arguably this one (forthcoming). Seeing vice, he argues, triggers a moral emotion of disgust, whereas virtue elicits the opposite of disgust, an emotion he calls "elevation." He does not distinguish the monster-saint dimension from the criminal-hero dimension, but it is plausible that his claims refer to the former.
35. Blum stresses this point in his discussion of Oskar Schindler (1988).
37. Wolf quotes the following passage from a piece on Gandhi by George Orwell: "sainthood is . . . a thing that human beings must avoid. . . . it is probable that some who aspire to achieve sainthood have never felt much temptation to be human beings" (Orwell 1945, p. 176), cited in Wolf (1982).
38. To call them "moral write-offs" is not to say they are irredeemably evil. Something might prompt a conversion for the better. But attempts to reason with them morally seem futile.
39. Psychologist Robert Hare reports that psychopaths comprise "perhaps 1%" of the general population (1998), using criteria that are similar to those for an evil person (however, the measures do not assess whether these individuals count as moral agents; perhaps many of them do not). As I noted earlier, some psychopaths lead seemingly respectable lives as professionals; perhaps the reader works with one.
40. This connects with questions of moral luck. See, e.g., Nagel (1979), and Williams (1981).

REFERENCES


