

On the Limits of Reflection: A Theory of Evil*

Edward S. Hinchman

University of Wisconsin-Milwaukee

Since 9/11 we've heard much talk of evil – I mean, of the thing. And we've heard just as much talk of 'evil' – I mean, of the word. Some people can bring themselves to talk about evil *only* by talking about 'evil.' Hence the references to such curious entities as so-called evil, quote-unquote evil, and the species of evil to which one refers by rolling one's eyes as one aspirates the word with a sigh. People who talk like this think they can talk only about what others *believe* to be evil, and then only in terms that discredit the believers as superstitious or confused.

In this paper I aim to show that such beliefs need not be superstitious or confused, because 'evil' can be given a straightforward and purely secular definition that captures the point of much of its current usage. It will emerge that, so defined, 'evil' expresses an immensely interesting concept whose application to the world has important implications for moral psychology. An apparently moral-psychological concept, evil in fact lies outside the limits of what we can understand as deliberative reflection.

Understanding 'evil'

In this brief presentation I lack space to survey alternative theories, so I'll offer mine

* Version presented at the Central APA in Chicago, April 2004, with Michael Weber (Yale University) commenting.

undialectically.¹ Begin by observing that in traditional usage there are two interrelated species of evil: what we might call *passive* evil and *active* evil. Passive evil is simply undeserved suffering.² This is not, it seems, the species of evil nowadays under discussion. When Mr. Bush speaks of ‘evildoers’ he seems to have in mind *active* evil. He doesn’t seem to mean merely ‘producers of undeserved suffering.’ He seems to have in mind a quality of persons *as such*: of persons not merely insofar as they possess mental states or productive capacities but insofar as they’re also responsible agents. Though I can’t speak for others’ usage, and I emphatically do not aim to endorse the president’s or any political position, as a philosopher I can’t help wondering what, more specifically, this might amount to. I offer my proposal with exegetical humility and where interpretation fails in the spirit of conceptual reform.

Active evil, I propose, is the state of being systematically disposed to the production of passive evil through a willed failure to be receptive to considerations requiring empathy with the perspective of those to whom this passive evil is done.

That’s too much for one sentence, so let me spell out the clauses. S is actively evil, I propose, if and only if S:

- (i) in decision-making systematically fails to be receptive to considerations requiring empathy with the perspective of members of some group G,³
- (ii) wills that failure, and

¹ Recent work includes Claudia Card, *The Atrocity Paradigm: A Theory of Evil* (Oxford: Oxford University Press, 2002) and the April 2002 issue of the *Monist*.

² Thus the ‘Problem of Evil’ in philosophical theology is the problem of reconciling God’s nature as all-powerful, all-knowing, and all-loving with the existence of undeserved suffering.

³ The group may of course have only one member, though one-off failures of receptivity would typically fail to be systematic.

- (iii) actually subjects members of G to passive evil (i.e. undeserved suffering) through decisions so made.

At the core of active evil, I propose, lies a willed refusal to acknowledge the fundamental moral relationship in which you stand to another person when you're weighing whether to perform an action that will cause that person to suffer.⁴ You are not actively evil if the failure of acknowledgment does not result in suffering, or if the suffering it results in is somehow deserved – although we can say that you are *potentially* actively evil, since you satisfy the internal conditions (i) and (ii). Nor would you be actively evil if the failure of acknowledgment is not willed – if, say, it's the product of some cognitive or affective deficit or of sheer empathetic exhaustion.

Now I believe that this is in fact what many people mean when they frame their response to terrorism in terms of evil. They mean that we are confronting in the terrorists a willed refusal to acknowledge the most fundamental moral relationship in which they stand to their victims. Those who flew planes into the World Trade Center did not, it seems, regard themselves as owing any consideration to their victims. Nor do those who detonate nail bombs in crowded marketplaces. These are deliberated acts performed by otherwise conscientious people. What seems inexplicable about the acts is how the attitude with which such people perform them is compatible with the general conscientiousness. We use 'evil' to mark not the callousness but the background of conscientiousness.

⁴ If it makes sense to say that someone is unempathetic towards himself – and I think it does (though we're inclined to put it in terms of 'sympathy' rather than 'empathy' here) – then one may be actively evil in one's self-relations as well.

It is also what people seem to mean when they reserve ‘evil’ for the acts and attitudes of those who profess to fight evil. In the fall of 2001, the issue on everyone’s mind was the evil of terrorism. Today, the issue for many is the evil of those who use terrorism as a pretext for raw aggression. Those who believe that the Bush or Blair administration’s argument for war in Iraq embodied a strategy for gaining control of oil reserves in the region regard these politicians not merely as wrong and likely to do great harm. The belief commits one to regarding these politicians as unconcerned – or as not concerned in the right way – about the harms they are doing not only to Iraqis and coalition soldiers but to citizens and institutions in their home country. It is an extraordinary charge, since it portrays members of these administrations as at best differing only in degree from those with whom they profess to be at war. Since the politicians do not, on this interpretation, feel any need to argue for the policies that produce the harms from premises that they themselves believe justify those policies, thereby making a mockery of democratic political institutions, anyone who accepts the interpretation must view this indifference as not mere laxity but positively willful. The charge is, at heart, one of an indifference to the manifold harms so profound that it can only be willed. Yet nobody doubts that Bush or Blair is a caring parent, a loving friend, or a friendly and conscientious neighbor. So the question once again is how such broadly unmalicious people could adopt such perniciously callous attitudes. One needn’t share the scathing assessments of these particular politicians to be intrigued by the general form of the question.

Having acknowledged the unfortunate political landscape of 2004, let me cast back to a time of greater consensus. We can learn an important distinction from the history of terror-

ravaged Afghanistan.⁵ Perhaps we'll agree that active evil informed the Taliban's treatment of Shi'a Muslims in Mazar-e-Sharif in 1998, where thousands of Shi'ites were shot when they did not immediately renounce their faith. It also informed their ongoing treatment of women.⁶ Yet when the Northern Alliance had recaptured Mazar-e-Sharif in 1997, they'd treated the defeated Taliban so brutally that one can almost sympathize with the Taliban's equally vicious treatment of supporters of the Alliance the next year.⁷ The Northern Alliance moreover had a track record of massacring civilians in Kabul and elsewhere in Afghanistan, apparently from sheer ethnic hatred. This may or may not be evil, in my sense. It is evil if the ethnic hatred is itself *willed*. Of course the hatred in question has influenced the wills of those who feel it, and has led them to bring about much passive evil, i.e. undeserved suffering. But hatred is not actively evil unless what keeps it in place is an active rationalization. Clause (ii) is crucial.

By an 'active' rationalization I mean one that is not merely retrospective, not merely an effort to escape blame for the suffering. An active rationalization here would aim to *fuel* the hatred. It would figure in deliberation. There is an important moral distinction between sheer mindless hatred of those who are different from you and ideological hatred aimed at treating those differences as reasons why the suffering of the hated is deserved, reasons which then inform your practical deliberation. Callous indifference to suffering is a horrible attitude, but it

⁵ If you disagree with my take on these cases, it should be easy enough to substitute your own examples. Again, my view of evil does not depend on my assessment of any specific instances, and my argument in this paper is not a moral or political polemic.

⁶ The 9/11 terrorists were not Afghans, of course, but they were ideological soulmates of the Taliban both in their misogyny and in their willed incapacity to empathize with the suffering of those whom they would persecute on religious grounds. A striking illustration of this incapacity is bin Laden's complaint, on several occasions, about the UN's 'victimization' of his Muslim brothers in Indonesia when it finally intervened after a decades-long Indonesian rampage against the East Timorese. Bin Laden must have known what murderous things the Indonesians had been doing in East Timor, but the sufferings of the East Timorese, as non-Muslims, simply did not have standing for him.

⁷ The Taliban took Mazar from the Alliance in May of 1997, lost it several months later, and then regained it in July of 1998.

is not necessarily evil. It becomes evil only when it is informed by a view of the victims as unworthy of any other attitude.⁸

Understanding evil

Having defined the concept and illustrated how it may apply, let's turn to evil itself. How does active evil operate? What psychic mechanisms does it employ?

One might call evil 'ineffable' meaning not that it's inexpressible or indefinable – claims I'd obviously reject – but that its role in causing suffering cannot be fully understood. We understand how someone could be callously indifferent to another's undeserved suffering because we know from the inside what it's like to have our capacity for empathy taxed to exhaustion, or constricted by selfishness, or disengaged by mere distraction. These are, alas, daily occurrences. But do we know from the inside what it's like to *will* this indifference?

I don't we think we *could* know this 'from the inside.' If I'm right, then, evil is in this respect ineffable. Key premise: to *will* indifference is to view the indifference as *justified*. If you view as justified the indifference that leads you to cause suffering, then you view the

⁸ Ethnic hatred is often kept in place by rationalization. It was thus no surprise when American neo-Nazi and other white supremacists celebrated the attacks of 9/11. White supremacist hatred clearly does satisfy my definition: it is kept in place by an elaborate theory of the races. This is homegrown American evil.

What then of the obvious candidates for systematic evil in U.S. history, white Americans' treatment of Native Americans and of African slaves? White Americans' attitudes toward these groups created much passive evil, but were they actively evil? Here we may wish to make a qualified claim: there was much rationalizing of the indifference to these groups' suffering, but not all of that rationalizing was active. Some of the rationalizing was merely retrospective, and some of the indifference was merely callous. Still, the indifference did receive support from hack theology, white supremacist racial theory, and the theory of 'manifest destiny.' It seems clear that U.S. history contains plenty of active evil.

I'll not speculate on what rationalizations may be at work in current U.S. politics.

suffering as to that extent *deserved*. So you could not will indifference to the plight of sufferers and at the same time view the suffering as undeserved. But an experience of your own active evil would require that you do both. Hence you cannot experience your own active evil – I mean, of course, experience it *as* actively evil.

To call someone evil is not obscurely to deny that he has grievances, goals or strategies.⁹ It is to characterize the attitude with which he pursues them. To note that the characterization prevents us from identifying with the perspective of this person is no more obscurantist than noting that one cannot imagine oneself occupying an unoccupied perspective. It is a logical point. In imagining oneself occupying a perspective, one imagines the perspective as occupied. And in identifying with a justification of indifference, one identifies with the conviction that the suffering it inflicts must be deserved. One cannot imagine simultaneously that the suffering is undeserved and that the indifference that produced it is justified.

Though we cannot for these reasons experience active evil ‘from the inside,’ most of us have experiences of active evil ‘from the outside.’ These come in two forms. Most of us have been observers or victims of active evil. But such an experience of active evil is indistinguishable from an experience of callous indifference. Active evil and callous indifference do not differ in their effects. Yet many of us have experienced the *temptation* to be actively evil ourselves, and this can help us to understand how evil works.

I pass the panhandler and think, ‘Get a job, you bum.’ I grow impatient with the unkempt

⁹ As Stanley Fish claimed in a post-9/11 editorial (*The New York Times*, October 15, 2001; revised and expanded in *The Responsive Community*, Volume 12, Issue 3 (Summer 2002); available online at <http://www.gwu.edu/~ccps/rcq/Fish.pdf>).

woman blocking my path and think, ‘Springer’s on, time to get back to your TV.’ However callous, this is not incipient active evil, because I do not endorse the thoughts.¹⁰ Still, many of us from time to time feel the temptation of such thoughts. We thus feel the incipient stirrings of a deliberative mechanism that in other educational or cultural circumstances might really lead us to our perdition.

The road to the spiritual condition known as hell is paved with such sweet-sounding thoughts. That we find ourselves stumbling by the side of this pavement from time to time is what enables us to imagine stepping onto it and walking more freely. A status as actively evil can come in degrees. The degree of one’s evil can be measured by the degree to which the ideology informing one’s willed indifference is entrenched and unshakable. Do you find it slipping when the sufferer is a friend from pre-revolutionary days? If so, you are not *as* evil as the monster who would rat on his spouse if the cause demanded it. If you keep altogether to the side of this road, you experience callous temptation but not what would give it direction.

Attributing active evil to another cannot, therefore, involve the detection in that other of a motivational state with which one is familiar from the inside. So the basis of the attribution cannot be an empathetic identification with that person. We cannot know ‘from the inside’ what makes an evil person tick. In this respect, to view him as evil is to view him as alien.

¹⁰ The thoughts are so manifestly callous, in fact, that no one *could* seriously endorse them. There is no prospect, thankfully, of anyone’s constructing an ideology according to which the poor must simply be lazy or the socially disenfranchised simply stuck in the muck of their gene pool. Such an ideology would be absurd on its face, much too obviously evil for anyone *we* know, at any rate, to adopt it.

Being evil

To view someone as evil is to view him as thereby excluded from moral community. In general terms, moral community is achieved among a group of people insofar as they are able to reason with each other about how to live in part by empathetically identifying with each other's deliberative predicaments. I have argued that one cannot empathetically identify with someone whom one regards as evil.

I think it's a good use of the word 'civilization' to equate civilization with moral community in this sense. We build civilization insofar as we realize moral community in specific social institutions. An evil person cannot be party to such community. To the extent that parties to would-be civilization are evil, the institutions that result are uncivilized. It is therefore appropriate to describe an evil action that does great harm to civilized people as an attack on civilization itself. It is an attack on the very possibility of moral community. Since I presuppose no specific conception of practical reasoning or of empathetic identification, in making this point I presuppose no specific conception of moral community. There is no *a priori* reason why communities should not embrace diverse conceptions of moral community. My point is that whatever conception a community adopts, its status as a community, so defined, is incompatible with its active members being actively evil. An actively evil person both fails to empathize with others, and because he wills this failure cannot be empathized with in turn. Whatever empathy is, exactly, and however it figures in practical reasoning, this two-sided failure is incompatible with the very idea of membership in a moral community.

The pathology goes deeper than we like to think. Christian theology and much philosophy mistakenly locate the human propensity for evil in what is just the brutally natural,

pre-moral fact that we are often callously indifferent. We have an obligation not to act on these indifferences, of course, but evil enters the story only through our more-than-brutely-natural – our second-natural – tendency to sidestep that obligation by *willing* the indifference through active rationalization. The acculturation we acquire as second nature can cope well enough, one might hope, with the brutality build into our first nature. It cannot cope with our active rationalizations of that brutality. Such active rationalizing uses acculturation – concepts, reasons, norms – in the service of its own debasement. It is therefore not *mere* brutality.

This puts the actively evil in an important respect beyond the reach of moral intervention. If you are actively evil, it will never seem to you that you are. Someone with a roster of grievances, goals and strategies in the service of evil is thus, we can say with some understatement, an *especially bad* thing. He is, more accurately, the worst thing a human person can be. Active evil is the purest perversion of our capacity for moral agency and personhood, since it directly undermines our efforts at well-informed practical self-reflection. It also renders us deaf to moral criticism. Here is a normative reason, then, to treat the actively evil as morally and culturally alien. Even if you could enter the mental life of the actively evil, you'd know in advance you could never get a word in.

Being good

I have argued both that evil is the worst thing we can be and that we can never directly experience ourselves as evil. (By 'direct' I mean, as I earlier put it, *from the inside*.) A single

argument yields both results, assuming only that a capacity for self-reflection lies at the core of our status as persons. I'll not here argue for that assumption.¹¹

If we cannot directly experience ourselves as evil, how do we guard against it? Though, switching metaphors, we may fall into the condition by degree, the fact remains that at each stage we'll be motivated to deny that that's where we're falling. Once fallen, we'll be incapable of gauging our predicament first-hand, and this will also be true *as* we fall to the precise extent that we've fallen. The harms we inflict will strike us as justified, and the justification itself will seem well grounded. Avoiding evil therefore seems a matter of luck.

Not so fast. Yes, avoiding evil *entirely on your own* seems a matter of luck. But why think you're all on your own here? Given what's at stake, it may matter enormously to others that you remain good. Why not let them help? You could let them help by not insisting that their attempts to influence you pass through the possibly distorting medium of your deliberative rationalizations. Of course, you have to deliberate, and to that extent you have to rationalize. But if avoiding evil is something more than a matter of luck, it must be possible to let others influence you in a way that bypasses your deliberations but does not for that reason fail to count as your agency. I lack space to elaborate my understanding of this possibility here.¹² I merely note its role in ensuring that avoiding evil is more than a matter of luck. If you suspect that there

¹¹ Except to note that I don't aspire to *explain* personhood or agency by appeal to self-reflection. (For an attempt at such an explanation, see J. David Velleman, *Practical Reflection* (Princeton: Princeton University Press, 1989), esp. Chapters 1 and 2, and *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000), especially the Introduction. For a critique of this attempt, see Michael Bratman, "Cognitivism about Practical Reason," reprinted in his *Faces of Intention* (Cambridge: Cambridge University Press, 1999). I give different grounds for rejecting a core premise of Velleman's argument in my "Trust and Diachronic Agency," *Noûs* 37 (2003).)

¹² I argue for this possibility at length in "Telling as Inviting to Trust," forthcoming in *Philosophy and Phenomenological Research*, and "Advising as Inviting to Trust," forthcoming in *Canadian Journal of Philosophy*.

is no such intelligible notion of non-deliberative practical rationality, it seems you're stuck with the moral luck hypothesis.

Can the hypothesis even make sense here? The past quarter-century has seen the revival of Sophoclean conceptions of moral luck,¹³ but Sophocles was not operating with the present conception of evil. As Martha Nussbaum explicates it, “[w]hat happens to a person by luck will be just what does not happen through his or her agency, what just *happens* to him, as opposed to what he does or makes.”¹⁴ But the mechanism of active evil, as I’ve explained it, ensures that it is fully agential. Active evil does not befall you – nor do you, despite my earlier talk, fall into it. You *dig* your way down. Active evil – marked by *too much* reflection, not too little – is paradigmatically agential. If your making yourself into this sort of doer is just a matter of your bad moral luck, we can’t explain moral luck by contrasting it with what you do or make.

We tend to think of the unempathetic and antisocial as suffering a lack, and the question then becomes, ‘How can what they do as a result of this lack be their responsibility if the lack itself is not?’ This sort of lack seems merely to befall the lacking. It’s a failure to develop, at any rate a blind spot. But active evil is not a failure to develop, or any other lack.

Being good – at least, avoiding evil – therefore begins in the thought that you must sometimes give others’ perspectives on your normative standing as such; you mustn’t view them solely through the lens of your rationalizations. What we might call proto-evil lies in the conviction that you can avoid evil without ever thus being normatively receptive to the wills of

¹³ Some of this work – including seminal contributions by Bernard Williams, Thomas Nagel, and Martha Nussbaum – is collected in Daniel Statman (ed.), *Moral Luck* (Albany : State University of New York Press, 1993).

¹⁴ *The Fragility of Goodness* (Cambridge: Cambridge University Press, 1986), p. 3.

critical others. It is a most general refusal of empathy. As Hume saw, it is not enough to sympathize merely with potential victims; you must sympathize with potential *critics* – themselves perhaps in sympathy with the victims – to get a motive sufficient to forgo the tempting bad act. I’ve defined evil from the perspective of its victims, but an evil agent will prove just as incapable of resonating to the perspectives of critics. The problem is not that the criticism can’t reach him but that he won’t let it. He isn’t deaf but inwardly shouting it down. The insensitivities to victim and critic are two sides of a single deliberative disposition.

In sum

Evil must be confronted, but from what I have said here nothing follows about how we should do so. (Of course, it will depend on the case.) I have argued only that the concept of evil is perfectly intelligible within purely secular morality, although understanding how it works requires self-reflective sensitivity to one’s own temptations to it. Understanding not only the concept but the reality of evil is thus a prerequisite for full participation in anything recognizable as moral community. Such an understanding would express acknowledgment both of the natural human temptation to rationalize one’s own callous indifference to the suffering of others and of the consequent precious fragility of moral community, since it is premised on participants not giving in to the temptation. These are prerequisites for full participation in moral community, because they are signs that one is open to moral influence.