

MIT Open Access Articles

Saints, heroes, sages, and villains

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

Citation: Markovits, Julia. "Saints, Heroes, Sages, and Villains." *Philosophical Studies* 158.2 (2012): 289–311. Web.

As Published: <http://dx.doi.org/10.1007/s11098-012-9883-x>

Publisher: Springer Science + Business Media B.V.

Persistent URL: <http://hdl.handle.net/1721.1/76301>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike 3.0



Saints, Heroes, Sages, and Villains¹

This essay will explore the question of how to be good. But it will be more of a meander than a march: I won't be presenting any single, fully developed argument. Instead, I hope to explore a few different aspects of goodness that will, I believe, resist capture under any one definition. In short, there are several, importantly different, ways to be good.

My starting point will be a thesis about moral worth that I defended in an earlier paper.² My goal (or one of my goals) in that paper was to examine the thought that morally worthy actions weren't just right actions, but rather were right actions performed for the *right reasons*. I argued, very roughly, that the *right reasons* were the *right-making reasons*: that an action was morally worthy if and only if the reasons motivating it coincided with the reasons morally-justifying it. My focus in this paper is a bit broader: here, I'm interested in *morally admirable persons*, as well as in *morally admirable actions*. But an account of what makes for a morally admirable person can, I think, be quite readily extracted from my earlier account of morally admirable actions. While I think that account gets at one important sense of moral goodness, I will argue that it (unfortunately!) fails to capture several other important ways of being worthy of admiration on moral grounds. Moral goodness is, I've come to believe, more multi-faceted than my earlier account suggested. (This should perhaps not be surprising. It's a fairly familiar philosopher's-instinct to try to come up with a neat formula to explain some central feature of the world. But it's an instinct that as often as not leads us astray.)

My title is intended to capture some of that multi-facetedness: in the course of this essay, I want to think a bit about what *saintliness* (in a secular understanding of the term), *heroism*, and *sagacity* come to. To some extent, I use these as terms of art: our ordinary usage of these words is considerably more complicated, more ambiguous, and in some cases more specific than the way I will be using them in what follows. But I haven't just picked these labels to be cute – the variety of our common-sense moral ideals underscores the inadequacy of any one account of good character, and I think my discussion will help illuminate the distinct roles these ideals play in our everyday understanding of goodness.

At the end of this essay, I will say something of a very exploratory and provisional nature about the flip side of these ideals: about villainy.

¹ I owe thanks to Vanessa Carbonell, Terence Cuneo, Louis DeRosset, Tyler Doggett, Tom Dougherty, Joshua Gert, Alex Guerrero, Elizabeth Harman, Richard Holton, Stephen Kearns, Rebecca Markovits, Richard Markovits, Graham Oddie, Miriam Schoenfield, Paulina Sliwa, and Christina Van Dyke, as well as to audiences at MIT, the University of Vermont, the 2011 Bellingham Summer Philosophy Conference, and the 2011 Rocky Mountain Ethics Congress for very helpful comments on this essay.

² Markovits (2010).

1 The Coincident Reasons Thesis

I'll begin with my old account of moral worth. In my earlier paper, I defended the following thesis:

The Coincident Reasons Thesis: An action is morally worthy if and only if – and to the degree that – the noninstrumental reasons motivating the action coincide with the noninstrumental reasons that morally justify its performance.³

Some unpacking is (inevitably) in order.

The normative *reasons that morally justify actions* are facts: they are considerations that count in favor of performing the actions. For example, the fact that jumping into the water will allow me to save a child's life is a normative reason that morally justifies my jumping in. But the fact that jumping in will allow me to save a life can morally justify my jumping in only if it's a fact to which I have epistemic access – if there's no way for me to know that the child needs my help, then it's not the case that I morally-ought to jump into the water. The normative reasons that morally justify actions – the moral-obligation-grounding reasons, in other words – are *subjective*: we are morally required to do only what we have sufficient epistemic reason to believe it would be best to do, not what it would (in fact) be best to do. This is why, for example, doctors are morally obligated to prescribe the course of treatment their evidence tells them is most likely to cure their patients, not the treatment that (against all evidence) happens to be best; if all the evidence suggests that I need penicillin, and my doctor has no evidence that I'm allergic, she fails to fulfill her obligations if she refuses me penicillin, even if it turns out I *am* allergic. Cases like these illustrate that the reasons morally justifying our actions depend on our evidence. Because the reasons relevant to moral-ought claims are subjective – they depend on what an agent ought to believe about her situation – our normative reasons for acting *can't* be given by facts (like my unpredictable allergy) of which we're blamelessly ignorant.

An agent's *motivating reason*, as I'm using the term, is also a fact: one that supplies a certain kind of *explanation* of the agent's action. Motivating reasons are the kinds of facts we are after when we ask about an agent, "what were *her reasons* for acting as she did?" – those that appear in what have been called "rationalizing explanations." Other kinds of facts can also, of course, explain our actions, such as biochemical facts about our brains or facts about how much sleep we've been getting or how much coffee we've drunk; and these might also be called 'reasons.'⁴ But they could not be described as "*our reasons* for acting." The reason I snap at you might be that I haven't gotten enough sleep lately, but this can't be

³ Markovits (2010, pp. 230 and 238). Joshua Gert points out that the expression "morally justified action" is ambiguous: it could refer to an action we ought morally to perform, or it could refer to an action that is (merely) morally permissible. When I refer to reasons that morally justify an action, I have the former, stronger sense of morally justified action in mind.

⁴ As is often the case with explanations, these kinds of explanations need not crowd out rationalizing explanations – both may be true at the same time. For a useful discussion of the related idea of "interpretation" of an agent's actions (and of the contrast between "interpretive" and "mechanical" explanations of behavior) see Walden (2012, forthcoming).

my reason for snapping at you – it can't be my motivating reason.⁵ My motivating reason will always be some fact on the basis of which I chose to snap.⁶ If people were always perfectly self-aware and sincere, their account of what prompted them to choose to act as they did would always provide us with their motivating reasons. But I don't assume that people always recognize their own motivating reasons – that they always know on the basis of which facts they make their choices – or that it will always be a straightforward matter for us to identify the motivating reasons of others.⁷

The Coincident Reasons Thesis appeals to an agent's *noninstrumental* justifying and motivating reasons. What do I mean by "noninstrumental reasons"? Noninstrumental reasons are those reasons to act that are provided by the ends of our actions that are worth pursuing for their own sake (in the case of justifying reasons) or that we pursue for their own sake (in the case of motivating reasons). (Thus, as I'm using the expression, we have noninstrumental reasons to pursue even purely instrumentally valuable actions – they're provided by the end, valued for its own sake, to which we perform those actions.) Instrumental reasons, by contrast, are reasons provided by features of our actions (or their outcomes) that have instrumental value (or that we value instrumentally). So, for example, the fact that flossing removes plaque is a merely instrumental reason to floss, since plaque-removal is merely instrumentally valuable, not worth pursuing for its own sake.

Our most fundamental reasons for acting, normative and motivating, are always noninstrumental – our reasons cannot be instrumental all the way down. But it's not *only* our most fundamental reasons that are noninstrumental: less than fully fundamental reasons may also be noninstrumental. This is because there are plenty of things that we value and pursue for their own sakes – and that provide us with non-instrumental reasons – even if those things do not, in some sense, provide the *most fundamental justification* for our actions.

Take *happiness*, on a Kantian view, for example. The fact that some action will promote happiness is not, on a Kantian view, the most fundamental explanation of why we should perform it – a more fundamental explanation might appeal to the requirement to respect people's rational capacity to set ends, by helping them attain those ends, together with the claim that all (or almost all) rational people set happiness as (one of) their ends. Nonetheless, happiness is, on the Kantian view, something we value and pursue for its own sake: we don't value it (or rather, don't *only or primarily* value it) instrumentally – as a *means* to something else. So although Kantians and utilitarians may disagree about the

⁵ See also T.M. Scanlon's discussion of the distinction between merely explanatory reasons and intentional agent's reasons in *What We Owe To Each Other*: while there may be a reason (in the sense of an explanation) why a volcano erupts, we would not speak of *the volcano's reason for erupting* (Scanlon 1998, p. 18).

⁶ See Markovits (2010, p. 221). To say this is not to say that all motivating reasons are facts we take to be normative reasons, and certainly not to say that they are facts we take to be sufficient to justify our actions. We might sincerely offer a "rationalizing" explanation even when we believe our reasons for acting did not justify our action. I might offer as my reason for snapping at you the fact that your voice is rather shrill for this time in the morning, even if I know that the tone of your voice is no justification for snapping at you at all.

⁷ For further discussion of why we might be mistaken about or ignorant of our motivating reasons, see Markovits (2010, pp. 222-223).

fundamental justifications of action – our most fundamental normative reasons – they’ll likely find more common ground about our non-instrumental reasons.⁸

This distinction between fundamental and non-instrumental reasons is very important to making the Coincident Reasons Thesis plausible. This is because the thesis owes us some account of the *level* at which the reasons justifying and motivating an action must coincide for the action to be morally worthy. I don’t act worthily if the non-instrumental reason *justifying* my action is a merely instrumental *motivating* reason for me: if, say, I jump into the water to save a drowning child, but only because I desire to get my name in the paper. And I don’t act worthily if my non-instrumental motivating reason for doing the right thing is in fact just an instrumental justifying reason (say, I’m justified in saving money now to help pay, later, for my child’s education, but like a miser, I’m in fact motivated to save money for its own sake, not for the sake of my child’s education). So in the case of morally worthy actions, the overlap between motivating and justifying reasons will run deeper than mere instrumental reasons.

But it can’t be required to run too deep. Because what the right fundamental moral theory is is such a difficult, and very much open, question, our account of moral worth should make it possible for someone to act worthily even if she’s mistaken about the right fundamental moral theory, provided she cares non-instrumentally about non-instrumentally valuable things. (Even if Kantianism is right, for example, it should still be possible for people motivated on utilitarian grounds to act worthily in many, many cases.)⁹ Here’s where the distinction between non-instrumental and fundamental reasons becomes helpful: we needn’t, the Coincident Reasons Thesis says, be motivated by what most fundamentally justifies our actions to count as acting worthily, so long as we’re motivated by non-instrumentally right-making reasons.

So it seems to me that overlap of noninstrumental reasons represents the “level” of required overlap on which the Coincident Reasons Thesis is most plausible. So much, then, for the qualifier “noninstrumental.”

Now what does it look like when an agent’s motivating and justifying reasons coincide? One thing becomes immediately clear: we’re often motivated by several contributing reasons, and our actions are often justified by a combination of reasons. We might do the right thing for some of the right reasons. Or right-making reasons might be only part of the explanation of our acting as we should. So overlap between motivating and justifying reasons can and often will be a matter of degree.¹⁰ Correspondingly, the moral worth of our actions will also be a matter of degree – we act worthily *to the extent that we’re*

⁸ For more on a closely related distinction between intrinsic and non-instrumental value, see Korsgaard (1997).

⁹ In Markovits (2010), I argue that if the Coincident Reasons Thesis captures a necessary condition of moral worth, then the truth of (at least simple versions of) utilitarianism would entail that many intuitively worthy right actions performed from Kantian or common sense moral motives would not count as worthy, because they would be non-instrumentally motivated by merely instrumentally right-making reasons. I argue that this is grounds for doubting the truth of (simple versions of) utilitarianism. (See especially §4.)

¹⁰ I set aside here a number of thorny questions having to do with normatively or motivationally overdetermined actions. For some initial discussion, see Markovits (p. 224 and n. 66 (p. 238)).

motivated by right-making reasons. And because even wrong actions often have *something* that counts in favor of them, our motivating reasons can partly overlap with normative reasons even when we act wrongly. So even wrong actions can have a degree of moral worth on this view. This strikes me as an intuitively appealing feature of the Coincident Reasons Thesis.

The Coincident Reasons Thesis, in any case, has much to be said for it. I defended it initially as an alternative to a thesis about moral worth that is often attributed to Kant, according to which morally worthy actions must be performed “from the motive of duty,” or *because they are right*. That Kantian thesis captures the powerfully intuitive thought that in case of morally worthy actions, it’s no *accident* that the agent acts rightly. But it does so at the cost of making it impossible for someone who cares about and is motivated by what really matters but has uncertain or false moral beliefs to act morally. (The case of Twain’s Huck Finn, who helps protect his runaway slave friend Jim, despite believing himself to be acting very wrongly in doing so, is a popular counterexample in the literature.) Critics have also suggested that the Kantian thesis misrepresents the ideal moral agent’s motives: the ideal moral agent, they claim, is motivated not by thoughts about the moral status of her actions, but directly by facts about the needs and claims of those she helps.

The Coincident Reasons Thesis captures the non-accidental connection between an agent’s motives and the rightness of her actions while allowing for the possibility that we act worthily even when (like Huck Finn) we have false beliefs about what morality requires of us. Furthermore, it allows for the possibility of our acting worthily when we’re blamelessly ignorant of some relevant facts and when we’re mistaken or agnostic about the right fundamental ethical theory – the right account of what, ultimately, makes actions right. I believe it does a good job of capturing our most firmly held intuitions about morally worthy motives.¹¹ And it gives us some insight into why or how some acts, including wrong acts, are partially morally worthy: why they are worthier than they might have been, despite being less worthy than they might be.

The Coincident Reasons Thesis gives an account of what it is for *actions* to be morally worthy. But it leads readily to a closely-related view of virtue, or good *character*: while it is certainly possible, as the thesis suggests, to perform worthy actions *out of character*, morally worthy actions are the building blocks of virtue – a pattern of performing them makes up the life of a good person.

2 Saints and Heroes

So much for the Coincident Reasons Thesis. I was pretty pleased with it. But it has a weird implication, at least if we take it to provide the whole story about kinds and degrees of moral worth. And thinking more about this implication makes me think the thesis doesn’t tell the whole story: it fails to capture some significant aspects of moral admirableness.

¹¹ I haven’t really defended this broader claim here, but arguing for it was a central task of my earlier paper.

Here's the weird implication: the Thesis entails that any completely morally worthy action – any action that is entirely motivated by all the right-making reasons¹² – is just as worthy as any other. In other words, it entails that the predicate “morally worthy” is like the predicate “full” – any completely full glass is just as full as any other completely full glass, regardless of the volume of the glass. But intuitively, it seems like some completely morally worthy actions are – in some sense – worthier even than other completely morally worthy actions – “heroic” actions come to mind. Some have, we might say, a greater *volume* of moral worth.

In virtue of what might one morally worthy action be worthier than another? One perhaps natural initial thought is that an action's moral worth might also depend on the moral significance of the act: on the weight of the reasons morally justifying it. (The volume metaphor might suggest this: the bigger, so-to-speak, the relevant reasons, the “more” overlap there can be between motivating and justifying reasons.) This would explain why it seems much more impressive to give away half your income to an important charitable cause than to buy cookies from the girl-scout at the door or why saving someone's life might seem worthier than remembering to feed the cat on time. But thinking about a wider range of examples should quickly lead us to conclude that moral significance is at best a *necessary* condition for heroic acts: many people's lives depend on my decision to drive the right way down the highway; but my driving correctly is surely no *worthier* than many other decisions – such as a decision to help a struggling student with his essay – on which much less is riding.¹³

Nomy Arpaly's account of moral worth suggests a different answer. She defends a thesis like the Coincident Reasons Thesis as a *necessary* condition for the moral worth of actions.¹⁴ But, she argues, a further factor influences the *degree* of moral worth of actions – a factor she calls an agent's “degree of moral concern.”¹⁵ To determine the moral worth of an agent's actions – the amount of praise the agent merits, on moral grounds – we can't just look at what she does and what motivates her to do it, Arpaly claims. How morally worthy an act is, she says, depends not only on whether the agent acts for right-making reasons but also on how easily (or not) she might have failed to act for those reasons.

Recall Kant's contention that morally worthy agents don't just *happen* to do the right thing as a mere matter of good fortune (as when his prudent grocer's profit motive fortunately leads him to charge a fair price for his goods). The Coincident Reasons Thesis (like Kant's motive of duty thesis) ensures that it's no accident that the motives of someone acting worthily lead her to act rightly. Arpaly's appeal to an agent's degree of moral concern – a measure of how easily she might have been deterred from acting well – seeks to guard

¹² Or at least, by normative reasons *sufficient* to justify the act. Again, for present purposes, I set aside complications having to do with actions whose performance or justification is overdetermined.

¹³ Thanks to Sin yee Chan pushing me to recognize that moral significance, while not sufficient for an act to be heroic, may well be necessary.

¹⁴ Arpaly (2003, p. 72). Arpaly calls her thesis “Praiseworthiness as Responsiveness to Moral Reasons.”

¹⁵ *Ibid.* (pp. 87-89).

against a further kind of accidentality: against the fear that it might be mere good fortune that a particular agent was well-motivated.

She introduces us to two characters, a “fair-weather” and a “foul-weather” or “die-hard” philanthropist: both perform the right (say, charitable) act, and for the right (i.e. right-making) reasons; but the former “would act benevolently even if severe depression came upon her and made it hard for her to pay attention to others,” whereas her “fair-weather friend ... acts benevolently as long as no serious problems cloud her mind, but [her] benevolent deeds would cease, the way some people drop their exercise programs, if there were a serious crisis in her marriage or job.”¹⁶ Arpaly concludes:

The first agent is more praiseworthy for her actions than the second agent, because to act benevolently for moral reasons while one is depressed takes more concern for those moral reasons than to do so in happy times.¹⁷

As I understand Arpaly, her claim is not just that Die Hard’s charitable actions are worthy *when she does them despite being depressed* but that they are worthier than Fair Weather’s *even when both agents act under the same circumstances*. They’re worthier because their goodness is less accidental: it depended less on the vicissitudes of circumstance.

Arpaly’s suggestion that an action performed for right-making reasons is worthier than another action performed for right-making reasons if the agent would have performed it under more difficult circumstances as well provides an alternative way to spell out the idea that some well-motivated actions are worthier than others. In this respect, it seems to offer a welcome amendment to the Coincident Reasons Thesis. But (for reasons I spelled out in my earlier paper, and will only summarize here) I think the appeal to counterfactuals on which Arpaly’s picture relies can often lead us astray. How are we to determine how easily someone can be deterred from doing her duty – how fragile her good will is? It can, I think, be unhelpful to ask of someone, “would she still have done that if...?” We can easily imagine that someone who willingly risks her life to save a stranger’s might nonetheless have been deterred from helping others in mortal danger by some comparatively minor concern – say, some threat to her dog, if she’s a doting enough owner. Is such a person fair-weather or die-hard?

The worries raised by the example aren’t just epistemic. There may be no fact of the matter about whether my imagined dog-lover is fair-weather or die-hard. But in any case, the proposed counterfactuals don’t do a good job of tracking our intuitions about moral worth. We do not think a relatively low-cost right action is made worthier by the fact that the agent who performs it (for the right-making reasons) would have done so even had the cost to him been higher. If saving the group of strangers requires my dog-lover to do no more than toss them a life-saver from the edge of the pier, we would not call her act heroic just because she would have risked her life for them had the emergency required it – her actual act required no particular heroism. Nor, as the case of the selfless dog-lover shows, do we think an action *less* heroic because the agent who performs it – still for right-making

¹⁶ Ibid. (p. 87).

¹⁷ Ibid. (p. 87).

reasons – might *not* have done so had the cost been different – might have been deterred by some consideration that strikes us as comparatively insignificant. So appeals, like Arpaly’s, to how an agent might have acted counterfactually won’t, it seems to me, help fill the gap in the Coincident Reasons Thesis: won’t allow us to account for our intuitions that some morally worthy actions, like heroic ones, may be worthier than others.¹⁸

So what does heroism come to? If a hero is not (necessarily) someone whose moral motivations are particularly resilient to changes in her circumstances, how are we to understand heroism? In virtue of what do we judge some actions – heroic ones – to be particularly worthy of moral admiration?

Thinking about our initial failed suggestions, and the range of examples that inspired and undermined them, suggests a more promising proposal: perhaps how morally worthy a morally worthy action can be is a factor, not (as Arpaly’s discussion suggests) of whether the agent *would* have performed it in more difficult circumstances, but of how difficult the circumstances were in which the agent *actually did* perform it. That is, perhaps we should ask not whether the agent would have performed the same actions had her circumstances been different but whether other agents would have performed the same action in the circumstances in which our agent actually found herself. We admire heroes in part because we think we would not have had it in us to act as they did had we been faced with the same decision.¹⁹

So let’s consider the hypothesis that a heroic action is a right action (of some moral significance) that most of us, judging the action, would not have had the moral strength to perform, had we been in the hero’s place.²⁰

Here are some interesting consequences of this way of understanding heroism:

First, it makes what counts as a heroic action (or as a hero) a *relative* matter – relative, I think, to the moral qualities of the speaker’s community. What stands out as a case of extremely unusual moral fortitude will depend, of course, on what counts as *normal*. This introduces a source of ambiguity into our ascriptions of heroism that is, I think, reflected in our linguistic practice. Identifying the relevant comparison class against which to determine whether a particular action is heroic won’t always be a straightforward matter.

¹⁸ Arpaly calls the “die-hard”-ness of an agent’s moral motivations one of several features “*associated*, other things being equal, with strength of concern” (p. 85). She doesn’t say die-hard motivation is constitutive of such stronger concern. For all she says, she may take counterfactuals like the ones concerning her two imagined philanthropists to play a merely *evidential* role in determining the moral worth of their acts. As I’ve argued, I don’t think they play even this evidential role.

¹⁹ This conception of heroism, which I think captures our everyday understanding of the term “hero” quite well, may be, in two respects, broader than ordinary usage: first, it sets aside, as an arbitrary and unnecessary restriction, the implication – common to many dictionary definitions (including the *OED*) – that heroes are male; second, it sets aside the requirement, common but not ubiquitous, that heroic actions exhibit *courage* – especially physical courage or valor. But it reflects the classical roots of the term as a label for men of near super-human ability – an idea with obvious echoes in the modern idea of a superhero.

²⁰ “Moral strength” in this analysis serves as a kind of placeholder for whatever it is about the agent’s character – courage, self-control, perceptiveness, the right desires – in virtue of which she can act well where most would act badly.

And indeed, different comparison classes will appear relevant to different speakers. And so different speakers will, quite properly, disagree about what counts as heroic. A striking example of this, one that reinforces the account of heroism I am defending, is the frequency with which people held up as heroes demur, and resist the label. Consider the soldier – in one sense, a member of our community – who risks his life to disarm a bomb or pull a wounded comrade out of the crossfire. Such acts will strike those of us sitting on the sidelines as heroic: we could not imagine performing them ourselves. But they may well strike the soldier as normal: any one of his comrades would have done the same. We might both be right.²¹

I've suggested that heroic actions are right actions that most of us – belonging to the speaker's community – would not have performed were we in the agent's circumstances, and that the difficulty of defining the speaker's community provides one source of ambiguity in ascriptions of heroism. There's another, related source of ambiguity deriving from the difficulty of defining the agent's "circumstances": what are we to hold fixed when we try to imagine ourselves in the agent's shoes? Think again of the soldier. It is natural, in evaluating his act, to imagine ourselves, with our own actual histories and resultant dispositions, trying to do what the soldier does when faced with his choice. But actions that look heroic in that light might look less heroic if we imagine ourselves not just faced with the agent's choice but armed with his training, say, or experience. (This may be no easy imaginative feat.) The difficulty is again reflected in the kinds of demurrals that are common with ascriptions of heroism: he was just doing his job – what he was trained to do.

The problem of what to build into an agent's circumstances in evaluating her action is a version of one of the most difficult problems in ethics, and I don't pretend to have a solution to it. It's the same difficulty we encounter when we try to pinpoint what counts as a moral *excuse*, or what amounts to an *extenuating circumstance*. (Indeed, it's the same difficulty that underlies certain versions of the problem of free will and moral responsibility.) Is someone less blameworthy for some destructive action if he was coerced? What if his behavior can be traced to the effects of a brain tumor? What if he suffered abuse as a child? What if he is simply mean-spirited? We have a loose intuitive way of sorting such influences into external and internal forces – into forces that act on the agent from the outside and those that in some sense come *from* the agent – that reflect her character: coercion clearly absolves the agent of at least a certain amount of blame; and we don't excuse bad behavior that is motivated by mean-spiritedness. But the brain tumor and the bad upbringing cases are trickier. The tumor is in a mundane sense internal to the agent, but it strikes us as a kind of outside intrusion. We don't generally see tumor-caused

²¹ There's an interesting related question: are ascriptions of heroism across clearly different communities ever apt? Does it make sense to ascribe heroism to the acts of some person who inhabits a moral culture that is manifestly different from our own? Or is an ascription of heroism apt only when the speaker can reasonably see herself and the agent as occupying a shared community? I'll leave that question open, for the time being, because I don't know what I think about it. But my instinct is that we may be hitting up against what Bernard Williams saw as "the truth in relativism": certain forms of ethical appraisal are appropriate only when the gap between the appraiser's moral culture and the agent's is in some sense bridgeable. See Williams (1981).

actions as coming from the agent's character.²² The bad upbringing case is trickier still: it seems to be an external force that determines the internal forces – that *shapes* the agent's character.

Similarly, when we ask ourselves if we would have acted as another did had we been in her circumstances, we wonder what counts as the agent's circumstances, and what counts as part of her character. Clearly, factors completely external to the agent contribute to her circumstances: the burning house, the swirling tides, the grave danger. We build these in to our thought experiment. We do the same, I think, with *merely* physical characteristics and capacities of the agent: her brute strength, her height, her ability to swim, perhaps, or her sensitivity to smoke. (So, for example, what is heroic for a weak person might not be heroic for a strong person, if few of us would perform the action if weak, but many of us would perform it if strong. What is heroic for a non-swimmer might be quite ordinary for a swimmer.)

On the other hand, we clearly do not build purely moral characteristics of the agent into our thought experiment: when I evaluate the heroic actions of a very courageous person, I do not ask myself if I would have performed them had I been equally courageous.

But again, tricky, in-between cases are not hard to find. Phobias, for example, strike me the same way that the brain tumor does – as intuitively “external” to the agent's character in the relevant sense; a simple water rescue is, in my view, heroic if performed by an aquaphobe, because most of us could not have performed such a rescue if we suffered from aquaphobia. But what should we say of cases where external circumstances seem to shape character? The training of the soldier or the fire-fighter may fall into this category: it may be that we, with *their* training, would have shown similar courage; but aren't their actions heroic? It is surely relevant, in these cases, that soldiers and fire-fighters choose their training.

Relatedly, should we think the bar for acting heroically is set higher, in some sense, for a person with a good moral upbringing than for a person who was badly raised, because we (undoubtedly correctly) judge that we would have been less likely to act well if we had been raised badly than if we had been raised well? I think the suggestion has some plausibility. For example, it resonates with our attitudes towards early abolitionists. A seventeenth-century opponent of slavery naturally attracts greater moral admiration than an opponent of slavery today, though slavery was no less wrong in the seventeenth century than it is now. This reflects the fact that we recognize that had we been born into the seventeenth century instead of the twenty-first, the likelihood that we would have opposed slavery is small.

There's another interesting consequence of understanding heroism along the lines I have been suggesting: it represents heroic acts as falling on the far end of a continuum on which ordinary good conduct occupies the near end – a continuum defined by *how extraordinary* a

²² Still, the old skeptical worry rears its head: how does the tumor differ from more ordinary ways in which our biology might govern our natures?

particular good deed is.²³ Since I was looking at the case of heroic actions in order to discover a kind of *variation in degree* of moral worth that the Coincident Reasons Thesis seemed not to capture and since I pointed to heroic actions as *typifying*, rather than *exhausting*, the class of problem cases for the Thesis, this is a nice result.

Finally, it's worth thinking about what this account of heroism might teach us about a category of actions that is often taken to be closely related: that of *supererogatory* actions. J.O. Urmson reignited interest in supererogatory actions with his challenge, in his 1958 paper "Saints and Heroes," to the traditional "tri-partheid division" of actions into the moral categories of *prohibited*, *permitted* (or indifferent), and *obligatory*. He argued that heroic acts (among others) did not fall into any of these categories, being neither indifferent nor obligatory.²⁴ They are clearly praiseworthy, he argued, but cannot be obligatory, because obligatory actions are actions we have a right to "expect and demand of others." But we cannot "expect and demand of others" that they act heroically; heroism is rather something "we can merely hope for and receive with gratitude when we get it."²⁵

My account does not define heroic actions as non-obligatory. Indeed, because, on my view, whether an action counts as heroic (as opposed to *merely right*, let's say) is relative to the appraiser, the view sits uncomfortably with any view of supererogation according to which actions qualify as supererogatory in virtue of being heroic *and* supererogatory actions are by definition non-obligatory: this combination of views would entail an appraiser-relativism, not just about which actions are heroic but about which are obligatory. The very same action could be appropriately described as both required and not required (because heroic), depending on the speaker. I think we have good reason to resist such a strong form of moral relativism. But my account may provide the resources for an account of so-called "supererogatory" actions that explains how such actions could be obligatory despite its being inappropriate for us to "expect and demand" them of others. If we appropriately view an action as heroic – if it's heroic, relative to us – that means it's an action we know we and most of our peers would not be willing to perform, were we in our hero's shoes. In that case, "expecting and demanding" the action would indeed be inappropriate: it would show not just *chutzpah* but gross hypocrisy!²⁶ (This is not just true of heroic actions – we lack standing to demand compliance with quite ordinary moral obligations if we're not willing to comply with them ourselves. For example, if I always break my promises to you, I can't "expect and demand" that you keep yours to me.)

Let me pause at this point to deflect a natural worry about my account: perhaps how *unusual* an action is, in the sense I have been describing, only *indirectly* tracks its degree of heroism. Heroic actions, the worry says, are actions that are particularly difficult or costly

²³ An act may be statistically rare but not extraordinary, in this sense: for example, if the circumstances under which the agent acts are rare, but it would not be rare for an agent *in those circumstances* to act that way.

²⁴ Urmson (1958, pp. 198-216). Urmson actually distinguishes between "minor" heroism, and heroism "par excellence" – only the latter is non-obligatory on his view (p. 201).

²⁵ Urmson (p. 213).

²⁶ It might also be inappropriate to "expect" the action in the epistemic sense of that word.

for the agent, and such acts will, given human nature, naturally be rare as well; but it's their difficulty, or their costliness, and not their rarity, that makes them heroic.

This is indeed a natural thought. But it is, I think, mistaken. Not all heroic actions are *difficult* for the hero: rare good deeds may come easily to some very good people, and be no less heroic for that. And not all heroic actions are *costly* for the hero: the soldier who falls on the grenade to stifle the blast and save more distant comrades' lives may have died either way. Neither are all difficult and costly actions heroic, even when they're morally significant. Parenting provides a prime example of this: good parenting is, of course, no easy task; and parents regularly make significant sacrifices for their children. But though we're grateful for it, we do not see ordinary good parenting as heroic; indeed we *expect and demand* it. We *blame* bad parenting. By contrast, it seems to me, we don't blame a mere failure to act heroically.²⁷ This is, I have argued, because we lack the standing to blame people for actions we would not be willing to perform ourselves.

So: an *action* may meet the demands of the Coincident Reasons Thesis but lack a kind of moral worth – a kind that is characterized, in its most fully realized form, by the label “heroic.” What about a *person* who, as completely as possible, meets the requirements for *virtue* that the Coincident Reasons Thesis suggests? Is such a person also a *hero*? I'm going to call such a person – someone who always acts rightly, and always for all of the right-making reasons – a *saint*. (In this I loosely follow Susan Wolf, who defines a moral saint as “a person whose every action is as morally good as possible, a person, that is, who is as morally worthy as can be,”²⁸ and goes on to describe people she imagines to be always and only²⁹ motivated by what matters morally, to the extent that it matters morally.) So we can ask: are saints always heroes? The answer, it seems to me, is ‘no’. It's at least conceptually possible for us all to be saints: for us all to always do what we ought to do, for the reasons why we ought to do it. But if things were like that, the world would have no heroes: there

²⁷ Alex Guerrero has drawn my attention to the difficult case of Thomas Vander Woude: Vander Woude's son, Joseph, fell into a septic tank in his back yard. Vander Woude rushed to the tank, tried to help his son keep his head above the muck, and when it became clear that his son was not able to keep his head up, Vander Woude jumped in and pushed his son onto his shoulders, remaining submerged himself. He died holding his son above him.

What should we think of this case? Does Vander Woude act heroically in saving his son? Would many of us have done the same were it our child? These questions are hard to answer, but I can feel pull of “yes” answers to each. Still, I think such examples should give us pause: we greatly admire parents who make such sacrifices for their children. But we are also willing to condemn them if they *don't* make such sacrifices (consider how we would have felt about Vander Woude had he been unable to bring himself to save his son). We suspect, perhaps unfairly, there may be something deficient in such parents. This puts actions like these in a rather peculiar category: actions whose performance is seen as meriting special praise, but whose non-performance is seen as blameworthy. It seems to me there's something irrational about categorizing any act that way; and the shifting standards of evaluation we hold others to in cases like this reflect our uncertainty about what standards we would ourselves be capable of living up to.

²⁸ Wolf (1982, p. 419).

²⁹ The “only” is not, I believe, required by the Coincident Reasons Thesis – an action may be motivationally overdetermined without that detracting from its worth. (I should say that I don't share Wolf's views about the necessary humorlessness or blandness of moral saints – I would think that a good sense of humor, at the least, would be an essential asset to any saint.) See Susan Wolf (1982).

would be no one who acts rightly when doing so is, relative to her appraiser's community, extraordinary.^{30,31}

So saints may not always be heroes, though they often will be. Heroes, it is clear, won't always be saints. It takes just one heroic action to be a hero, whereas sainthood requires a pattern of reliably well-motivated actions across much of one's life. But I believe heroic actions *need* not even provide one piece of that pattern. That is, meeting the conditions set out by the Coincident Reasons Thesis is not only not sufficient but also not necessary for heroism. To see why, think back to the case of the soldier. Let's imagine our soldier is fighting in a just war; that fighting in it puts him in grave personal danger; that the goals of the war morally justify its costs; but that those costs – particular those borne by the soldiers fighting in it – are enough to deter most of us from enlisting. The military historian Richard Holmes, who interviewed veterans of many wars, describes the problem faced by such a soldier this way:

A soldier who constantly reflected upon the knee-smashing, widow-making characteristics of his weapon, or who always thought of the enemy as a man exactly like himself, doing much the same task and subjected to exactly the same stresses and strains, would find it difficult to operate effectively in battle. ... Without the creation of abstract images of the enemy, and without the depersonalization of the enemy during training, battle would become impossible to sustain. ... If ... men reflect too deeply upon their enemy's common humanity, then they risk being unable to proceed with a task whose aims may be eminently just and legitimate.³²

This might be so even if the "enemy's common humanity" underlies the justification for the war itself, and so provides a fundamental reason for fighting. That is, in a war fought on humanitarian grounds, soldiers may have reason to desensitize themselves to the common humanity of the inhabitants of an enemy state so that they can more effectively fight a war

³⁰ Thanks to Louis deRosset and Steve Yablo for helpful discussion on the right way to put this point. Even in the actual world, there could possibly be a moral saint (in my sense) who is not a hero: someone whose life simply affords her no opportunities for *extraordinary* moral fortitude – to act rightly when others would not. In practice, I believe most of us have opportunities to do good that the vast majority of us don't exercise – in that sense most of us pass up opportunities for heroism. But this may not be true of all of us.

³¹ Is it conceptually possible for us all to be heroes given then we aren't all saints? Or is a world full of heroes, like Garrison Keillor's imagined world full of above-average children, a sad impossibility? Fortunately, given our myriad imperfections, we might all be heroes, so long as we're all exceptional in different ways. Thanks to Mark Balaguer for pointing this out. A super-hero analogy might be helpful in illustrating the point: Superman was, of course, no super-hero on Krypton, his home planet, where everyone else shared his super-human strength, powers of flight, and x-ray vision. But the X-Men could all be super-heroes, even if there were no ordinary humans around to act as foils, because they all have different super-powers. (I suppose our children could, similarly, all be above average *in some respect*.)

³² Holmes (1985, p. 361). Soldiers in war also use the expression "tango down" to indicate that a hostile human "target" has been eliminated ("tango" represents the "t" in "target"), whereas they use the expression "man down" when one of their own fellow soldiers has been hit. See the online dictionary *Wiktionary*.

whose very justification is provided by that common humanity. If they have reason to fight in the war, and fight effectively, then they ought not be motivated to fight by that reason.³³

Pragmatic grounds not to be motivated by the reasons that apply to us are often generated when we are forced to act in emergency situations and against great odds, a fact that was strikingly demonstrated by post-crash interviews of Captain Chelsey Sullenberger, the US Airways pilot who miraculously succeeded in landing a commercial jetliner with no working engines on New York's Hudson River, improbably saving the lives of all 155 passengers and crew on board. Asked, in a *60 Minutes* interview by Katie Couric, whether he had been thinking about the passengers as his plane was descending rapidly towards the waters of the Hudson, Captain Sullenberger replied, "Not specifically. ... I mean, I knew I had to solve this problem. I knew I had to find a way out of this box I found myself in. ... My focus at that point was so intensely on the landing. ... I thought of nothing else."³⁴ It seems likely that years of training in emergency preparedness coached the Captain, with good reason, not to think about the ultimate reasons for successfully handling a crisis situation when faced with the need to do so.³⁵

Captain Sullenberger was widely hailed as a hero following his truly extraordinary feat of coolness under pressure. Soldiers' risking their lives fighting in just wars is often put forward as a paradigm case of heroism. Both cases of heroism seem not only compatible with but perhaps even to *require* not acting for right-making reasons.³⁶

³³ I don't want to suggest that such desensitization is easy to justify or usually justified. Indeed, the fact, described by Holmes, that soldiers often must be desensitized in this way to be effective soldiers is, I believe, one of the reasons why wars are hard to justify. Not only do wars require participants to "corrupt" themselves to be effective soldiers (a cost with immediate and long-term effects that should not be underestimated), but the need for such self-corruption also creates a significant risk that soldiers will prosecute a potentially justifiable war in a manner that makes it unjustified; as Holmes says, "if the abstract image [of the enemy, internalized by soldiers in training] is overdrawn or depersonalization is stretched into hatred, the restraints on human behavior in war are easily swept aside." (p. 361)

³⁴ Katie Couric interviewed Captain Sullenberger on *60 Minutes*, airdate February 8th, 2009.

³⁵ There may be a difference between not (consciously) thinking about some consideration when one acts, and not being motivated by it. It might be suggested that Captain Sullenberger was motivated by the fact that landing the plane would save his passengers' lives despite the fact that he was not thinking of them at the time. I'm not sure how to understand this suggestion – I'd want to hear more about what makes it the case that Sullenberger was motivated by this reason. It can't just be that it was this reason that motivated him to coach himself not to think about these reasons in emergencies. Second-order motivations aren't transitive in this way – they are discharged when the primary motives they target are formed. They needn't linger on to co-motivate, so to speak, the actions motivated by those primary motives. In any case, even if Sullenberger can be characterized as sub-consciously acting for right-making reasons, the case of the desensitized soldier seems to me very hard to recharacterize along these lines. Thanks to Christina Van Dyke and Tom Dougherty for pushing me on this point.

³⁶ It is noteworthy, however, that though neither the Captain nor the desensitized soldier, as I've described them, acts for right-making reasons, in both cases right-making reasons plausibly play a higher-order motivating role in the back-story of their actions. In both cases, as I've imagined them, it's plausible that these agents were motivated by genuine normative reasons *not* to be motivated by such reasons when being motivated by them might undermine their ability to act as they have most reason to act. So even in these cases, right-making reasons seem to play a significant motivating role somewhere 'upstream' of the act. It may be that if they did not play this higher-order motivating role, we would not judge the 'downstream' actions to be heroic. (Thanks again to Christina Van Dyke for helpful discussion.)

So sainthood and heroism seem to me to represent the end-points of quite independent axes of admirableness. But both contribute to how worthy of admiration someone is on moral grounds. To see this, note that heroism – performing a truly exceptional right action – can *make up for* a certain deficiency in saintliness.

Consider the case of Oskar Schindler. By all accounts, Schindler did not live a life of virtue either before he risked his life to undertake an extraordinary and extraordinarily risky rescue of the Jewish workers in his factory. Nor can he be said to have experienced a conversion, of sorts, to a more general saintliness during the war. Take, for example, the stories of what happened to the ring that some of his workers forged for him from their own gold fillings as a token of their gratitude: in one version, he is said to have lost it gambling; in another he traded it for a bottle of Schnapps. Here is what Schindler's ex-wife, a sometime victim of his vices, both before and after the war, is reported to have said about him in a later interview:

Oskar had done nothing astounding before the war and had been unexceptional since. He was fortunate, therefore, that in that short fierce era between 1939 and 1945 he had met people who summoned forth his deeper talents.³⁷

Schindler deservedly earns our moral admiration despite his flaws because of his exceptional moral accomplishments. I think we rightly don't feel the same way about him as we would about someone with a more ordinary mix of virtues and vices.

3 Saints and Sages

I think there's at least one other, very important, kind of moral admirableness. To see why, I'll begin by explaining what appears at first to be a reason to think the Coincident Reasons Thesis identifies at least a necessary condition for morally worthy actions.

This apparent support for the Coincident Reasons thesis comes from the solution it seems to provide to a puzzle that has recently been getting a fair amount of attention from moral philosophers: the puzzle of moral deference. The puzzle is generated by what Sarah McGrath, in her exposition, calls "the datum":

Pure moral deference seems more problematic than deference in many other domains.³⁸

("Pure" moral deference is deference about moral matters that is not reducible to deference about non-moral matters.) The datum purports to highlight a *disanalogy* in our attitudes towards forming judgments on the basis of "pure deference" about what are otherwise *apparently analogous* matters of fact. It seems entirely unproblematic for me to defer to my accountant about how much tax I owe or to my doctor about which medicine to take or even to a stranger on the street about the best way to get to Boston Garden. But there's

³⁷ Reported by Thomas Keneally (1982, pp. 396-397).

³⁸ McGrath (2011, p. 115).

something intuitively *off* about someone who stops eating meat *simply* because her friend told her eating meat is wrong, without coming to appreciate, herself, *why* it is wrong. What explains this difference?

In two recent papers,³⁹ McGrath and Alison Hills have (independently) defended explanations of the datum that appeal to a condition for moral worth like the Coincident Reasons Thesis. The problem with moral deference, they argue, is that morally deferential agents can't act for *right-making reasons*, even when the testimony they rely on is that of a true moral *expert*, and even if the testimony *can* engender moral *knowledge*. And acting for right-making reasons is, as the Coincident Reasons Thesis claims, a condition of at least an important kind of moral worth.

Say I stop eating meat because my friend Sally, whom I know to be an excellent judge of such things, tells me that eating meat is wrong, and I want to act rightly. Say that (as Hills and McGrath both allow) I come to *know*, in this way, that eating meat is wrong. If I don't *understand why* eating meat is wrong and if I'm not motivated to stop eating meat by *the reasons why* it's wrong, there seems to be something lacking, in me or in my action. And the Coincident Reasons Thesis seems to explain what that is. Here's Hills, discussing a version of a similar example from Arpaly:

[Such an agent] has good motivations, he wants to do what is morally right and chooses in accordance with those desires, and he has moral knowledge too, but more is required for morally worthy action: you need to act for the reasons that make your action right.⁴⁰

And here's McGrath:

Notice that insofar as pure moral deference does not deliver understanding, pure moral deference does not put one in a position to do the right thing for the right reasons even in those cases in which it delivers genuine moral knowledge. If the child refrains from lying on a given occasion because he knows that lying is wrong but has no grasp on why it is, then his refraining from lying on that occasion is not based on the reasons that there are not to lie.⁴¹

And later:

[E]ven granting that such an agent might very well deserve moral praise, he or she still falls short of an important ideal associated with moral agency: that of doing the right thing for the reasons that make it right. Thus, the realist might seek to accommodate the datum by claiming that our relative aversion to moral deference is due to our recognizing that someone who relies on moral deference for her moral views will inevitably fall short of this ideal. By contrast, when a person defers about some empirical matter, her

³⁹ See McGrath (2011) and Hills (2009).

⁴⁰ Hills (2009, p. 117).

⁴¹ McGrath (2011, p. 133).

doing so does not similarly frustrate the achievement of any ideal associated with agency. Thus, deference about the weather, geography, or some esoteric scientific matter strikes us as unproblematic in a way that moral deference does not.⁴²

McGrath calls this proposal “the most promising of the strategies [for accommodating the datum within a realist framework] that [she has] considered.”⁴³ If there is indeed something *wrong* with moral deference and if the Coincident Reasons Thesis can explain what that is, so much the better for the thesis.

Is there something wrong with moral deference? As both Hills and McGrath point out, there is reason to approach the datum with some skepticism. For one thing, the examples often used in the literature on the topic as anti-moral-deference intuition pumps often fail to establish a disanalogy between the moral and non-moral cases. The moral cases (for example, the ubiquitous meat-eating example) often concern matters of some controversy, about which considerable disagreement exists, even among supposed “experts”; and deference about similarly unsettled questions in non-moral cases seems similarly problematic. For example, there seems to be something faulty about drawing conclusions about the relative dangers to the economy of too little stimulus and too large deficits by asking the first economist you come across.

But as others of McGrath’s examples make clear, the problems with *moral* deference seem to persist in non-controversial cases as well: there’s something “odd”, as she says, about someone’s basing his belief in the wrongness of slavery purely on the testimony of a reliable authority (despite being in possession of all the relevant non-moral facts). In cases like this, however, we might well question whether someone who doesn’t know on his own that slavery is wrong can be relied upon to identify a moral expert worth deferring to. As McGrath and Hills both point out, in the case of moral experts (unlike the case of medical or tax experts), it often takes one to know one.

It’s less clear to me than it seems to be to McGrath and Hills that there is a surviving “problem” of moral deference – that there is a remaining *disanalogy* between moral and non-moral deference left to explain. It seems to me that, in the moral case, we are more likely to find ourselves surrounded by moral *peers* than by relative moral *experts*. And just as I should, when I encounter disagreement about non-moral matters from an epistemic peer, reduce my credences in the proposition at issue, it strikes me as entirely appropriate to reduce confidence in my moral beliefs when I learn that someone I consider a moral epistemic peer disagrees with me, even when I don’t know the reasons for her disagreement. Furthermore, when we focus on the rare cases where we think someone *really is* in a position to recognize that someone else is a moral expert relative to her, deference seems quite unproblematic. We sometimes judge young children to be old enough know that they ought to defer to their parents’ judgments about moral matters they are too young to understand themselves.

⁴² Ibid. (p. 135).

⁴³ Ibid. (p. 135).

But the problem of moral deference seems to many people (Hills and McGrath among them) to survive these efforts to explain it away. And the appeal to the thought that morally worthy actions must be motivated by *right-making* reasons, and not just by the reason *that they are right*, seems to explain the problem *without* explaining it away.

As will by now be clear, I agree that an important dimension of moral worth can be attained only by performing an action for the reasons *why* it is right (rather than simply for the reason *that it is right*). But I *don't* agree that deference prevents an agent from acting on *right-making* reasons.

That is because, as I argued at the outset of this essay, the moral reasons for us to perform some action are *subjective* – we are morally required to do only what we have sufficient epistemic reason to believe it would be best to do, not what it would (in fact) be best to do. Certainly, it is *subjective* reasons that are relevant to the assessment of an action's (or an agent's) *moral worth*. If expert testimony gives us most reason to believe some act would be best, then that testimony *is the reason* we ought to perform that act. This is why (to recall my earlier example) if my generally reliable doctor tells me that my child needs penicillin then (absent any contrary evidence I ought to be aware of) I fail to fulfill my parental obligations if I refuse to give it to her, *even if* she turns out to be allergic. Cases like this illustrate that our normative reasons depend on our evidence.

The case of *moral* testimony is analogous: in those – perhaps rare – cases in which the advice of a (recognizable) moral expert provides our agent with sufficient evidence for the belief that a particular act would be best, our agent's acting so may be *made* right by the fact that the moral expert advises him to do it – indeed, this may be true even if the expert (for once) advises him badly. The agent's belief that the act is right, which motivates him to act, is based on the expert's advice. And his act is right *because* it follows the expert's advice: this is the reason *why* he ought to perform it. Think again about the young child who should defer to his parents as moral experts: sometimes “because I told you so” is a perfectly good reason to do something.⁴⁴ So the deferential agent may, it seems to me, act for right-making reasons after all.⁴⁵

⁴⁴ As Joseph Raz has pointed out, parents of children who do the right thing but against their explicit instructions feel torn in their praise for just this reason: the whole *point* of, and justification for, the instructions was that the children were more likely to act rightly by following them than by relying on their own judgment. Acting on the parents' instructions was, in this case, what the children had most subjective reason to do, even if it was fortunate that they disobeyed. See Raz (1999, p. 43).

⁴⁵ The fact that a recognizable moral expert tells our agent to perform some action is, I've said, a genuine right-making reason to perform it. But it is not a *non-instrumental* right-making reason to perform it. If the agent performs the act *just* because that would conform to the expert's advice, and not because he (legitimately) thought conforming to the expert's advice would lead him to do what is (objectively) best, his action still has no moral worth. Compare the parallel case of non-moral deference: if my reliable doctor tells me to give my daughter penicillin, and I do, not because I think deferring to my doctor will be best for my daughter, but rather out of some bizarre non-instrumental motivation to do as my doctor tells me, I'm non-instrumentally motivated by a merely instrumentally right-making reason, and my act has no moral worth. (This is why my claim about evidence doesn't entail that, e.g., right actions motivated by a desire for a reward are morally worthy, even if the reward is reliable evidence of the objective rightness of the act, and so, on my view, a right-making reason: it's at best an *instrumentally* right-making reason.) Thanks to Paulina Sliwa for bringing this worry to my attention.

So moral deference does not get in the way of attaining the kind of moral worth captured by the Coincident Reasons Thesis. To return, once again, to the language of the last section: there can be deferential *saints*. But this is not to say that there isn't something *morally admirable* about moral agents who don't need to defer – who are themselves moral experts. It's just to say that what is admirable about such people is not captured by the Coincident Reasons Thesis. It suggests, instead, yet another axis of moral admirableness. We admire people who act virtuously, but we also admire people whom we can turn to for moral guidance – who can not just tell us what we ought to do, *given our current epistemic boundaries*, but can effect what we ought to do, by *expanding those boundaries*. I'll call such people moral *sages*.

Like saints and heroes, moral sages occupy the far end of a continuum that ordinary moral agents can move along – in this case, by gaining in moral understanding: in the kind of moral knowledge that, as Hills notes, can be applied to new and relevantly similar situations and can be passed on to others. And as before, we seem to have identified an independent category of moral goodness: saints, I've said, may not be sages. Heroes certainly may not be. And sages may be neither saints nor heroes: indeed, they may often not be motivated to act rightly at all. (But a sage who doesn't practice what she preaches will be harder than ever to recognize as an expert.)

4 Villains?

I wanted to end by quickly considering *villainy*. There is, I've suggested, more than one way to be good. Is the complexity of moral goodness mirrored by a similar complexity of evil? I won't answer that question satisfactorily here, but some initial thoughts are suggested by the categories of virtue I've been considering.

A *saint*, I've said, is someone who is always motivated to act rightly by the reasons why her action is right. We can imagine a kind of villain who is motivated to act wrongly by the very reasons why his action is wrong. Nomy Arpaly calls such reasons "sinister reasons," and writes: "Other things being equal, a person is more blameworthy for a given wrong action if she acts out of ill will (for sinister reasons) than she would be if she were to act out of a lack of good will (for neutral reasons, while ignoring moral reasons to the contrary)."⁴⁶ (Such "indifference" to moral reasons Arpaly characterizes as a severe lack of *moral concern* – a kind of extreme fair-weather philanthropy. She makes clear that she thinks some very bad actions done out of such moral indifference may be more blameworthy than some merely mildly immoral acts done for sinister reasons.)

Is this right? I'm not sure. It certainly has some plausibility. But an example gives me pause:

In February 2004, Cameron Todd Willingham was executed in Texas for allegedly setting fire to his house in order to deliberately kill his three children. Willingham consistently denied starting the fire. The initial conviction rested almost entirely on the testimony of the local fire marshal,

⁴⁶ Arpaly, p. 80.

who claimed that the burn patterns in the house indicated arson. Shortly before the execution, a leading fire investigator completed his review of the original investigation and concluded, as other leading fire investigators would in the years to come, that there was no scientific basis to determine that Willingham had set the fire, and that the original fire investigators had based their theories on discredited methods. His report was sent to authorities asking to postpone the execution. Neither the Texas Board of Pardons and Paroles nor Texas Governor Rick Perry responded to these appeals. When, after Willingham's death, the Texas Forensic Science Commission agreed to investigate the case, they hired nationally recognized fire investigator Craig Beyler to provide an independent opinion on the evidence. Beyler confirmed earlier expert findings that the finding of arson had no evidential support. But shortly before the Commission was to hold hearings on his report, Gov. Perry replaced three members of the Commission, postponing the hearing indefinitely.⁴⁷

In deciding against granting clemency or a stay of execution to Willingham, Governor Rick Perry most likely did not act for "sinister" reasons, in Arpaly's sense, but for straightforwardly selfish ones. In Texas, where appearing "tough on crime" is a necessity for getting elected, granting a death row inmate clemency would come with significant political risks. When Willingham filed his petition, the State of Texas had approved only one application for clemency from a prisoner on death row since 1976. And any subsequent conclusion that the State executed an innocent man would of course have had similarly detrimental effects on Perry's political career. It seems most likely that Perry was motivated to act as he did by a desire to keep his job. He acted less out of ill will, in Arpaly's sense, than out of moral indifference.

But there seems to me to be something particularly abhorrent about such terribly wrong actions performed out of such total indifference – more abhorrent, in some ways, than the same actions performed out of ill will. Why is that? It may be, in part, because wrong actions that reveal such moral indifference are more likely to also *merely use* their victims, to *objectify* them, to treat them as *tools* rather than as persons. Someone who kills out of ill will, say, must at least recognize his victim to be a center of subjectivity – a source of aims and desires of his own. Not a saving grace. But its absence leaves a distinctive chill of its own.

A *hero*, I've said, is someone who performs right actions that most of us – judging the actions – would not have been able to perform had we been in the agent's circumstances. We have a special kind of admiration, I argued, for right actions that are in this sense *unusual*. I think we also have a special condemnation for wrong actions that reflect *unusual* sins. That is, our condemnation of wrong actions tracks not just the *strength* of the wrong-making reasons or whether the agent acted *for* those wrong-making reasons but also whether the agent acted wrongly in circumstances where most of us, doing the judging, would have acted rightly. Extraordinary sins draw more condemnation than ordinary ones, independent of the moral cost of the sin.

⁴⁷ See Grann (2009).

Here are two possible examples:

First, Peter Singer has famously drawn attention to the fact that, while someone who opted not to rescue a child drowning in a shallow pond because he didn't want to ruin his new leather shoes would draw near-universal condemnation, almost all of us breezily opt against giving money to famine relief – money sufficient to spare a child from starvation – so that we can instead use that money to buy such a pair of shoes.⁴⁸ The reasons to give to famine relief seem just as powerful as the reasons to save the child. But we clearly judge the two omissions quite differently. This is largely, I suggest, because the one sin is so much more ordinary than the other.

Second, I have vegetarian friends who firmly believe that eating meat will come to be viewed as this century's slavery – as the morally horrific practice whose horror the majority willfully ignored. And yet most of these vegetarian friends would much more readily befriend a meat eater than, say, an unusually malicious gossip.

As in the case of heroism, a disinclination to hypocrisy plays a role here: just as we feel we cannot demand actions of others that we would be unwilling to perform ourselves, we feel that we cannot *condemn* others for sins we are ourselves guilty of. The very ubiquity of ordinary sins provides them with some shelter from condemnation.

Finally: a *sage*, I said, is a moral expert – someone with an unusually high degree of moral understanding, who is therefore particularly well-placed to give advice and provide moral guidance. Contrast terms may again reflect a distinctive and important kind of villainy. Two recognizably dangerous figures come to mind: the first is the *demagogue* – who pretends to a moral expertise she doesn't have to further her own goals; the second is the misguided *fanatic*, who is driven by mistaken but sincerely held moral beliefs to lead people into wrongdoing.

Conclusion

I've argued that our intuitions about virtue, as well as about villainy, reflect a complexity to moral goodness that resists capture under any one account. To close: a comment on my methodology. I'm not sure if the work I've done in this paper should be seen as descriptive or normative. That is, to some extent I've been trying to systematically describe some of our intuitions about moral worth. Our practice of moral evaluation seems to me to reflect a veneration of heroes and sages as well as saints, and to differentiate between villains of many stripes. To the extent that such a description can help *make sense* of our practices, it constitutes a partial defense of them. But it constitutes *at best* a partial defense. Some explanations are debunking, not rationalizing. And in practice it's not always so easy to tell the difference.

I think that facts about how we are willing to act play a role in answering the (normative) question concerning what kinds of acts we have standing to praise and blame. But the question of what actions and agents we have standing to praise and blame *may* come apart from another, less well-defined, normative question, concerning what kind of praise and

⁴⁸ Singer (1972).

blame actions and agents *appraiser-neutrally* merit. In light of this possible distinction, I feel least confident that heroism, and especially the corresponding kind of ‘extraordinary’ villainy, merit the disproportionate levels of admiration and condemnation we allot them.

References

- Arpaly, N. (2003). *Unprincipled Virtue: An Inquiry into Moral Agency*. New York: Oxford University Press.
- Grann, D. (2009). A Sudden Dismissal. *The New Yorker: News Desk Blog*, October 1st, 2009. URL: <http://www.newyorker.com/online/blogs/newsdesk/2009/10/david-grann-a-sudden-dismissal.html>, accessed February 27th, 2012.
- Hills, A. (2009). Moral Testimony and Moral Epistemology. *Ethics*, Vol. 120, No. 1, pp. 94–127.
- Holmes, R. (1985). *Acts of War: The Behaviour of Men in Battle*. New York, NY: The Free Press (Simon & Schuster, Inc.).
- Keneally, T. (1982). *Schindler's Ark*. New York: Scribner.
- Korsgaard, C. (1997). Two Distinctions in Goodness. In *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, pp. 249-274.
- Markovits, J. (2010). Acting for the Right Reasons. *Philosophical Review*, Vol. 119, No. 2, pp. 201-242.
- McGrath, S. (2011). Skepticism about Moral Expertise as a Puzzle for Moral Realism. *Journal of Philosophy*, Vol. 108, No. 3, pp. 111-137.
- Raz, J. (1999). *Practical Reason and Norms*. Oxford: Oxford University Press.
- Scanlon, T.M. (1998). *What We Owe To Each Other*. Cambridge, MA: Harvard University Press (Belknap).
- Singer, P. (1972). Famine, Affluence, and Morality. *Philosophy & Public Affairs*, Vol. 1, No. 3, pp. 229-243.
- Sullenberger, C. (2009). Interview with Katie Couric. *60 Minutes*. Feb. 8th, 2009. CBS News.
- Tango down. Appendix: Glossary of military slang (under “T”). In *Wiktionary*. URL: http://en.wiktionary.org/wiki/Appendix:Military_slang#T, accessed August 18, 2010.
- Urmson, J.O. (1958). Saints and Heroes. In A. I. Melden (ed.), *Essays in Moral Philosophy*. University of Washington Press, pp. 198-216.
- Walden, K.E.D. (2012). Laws of Nature, Laws of Freedom, and the Social Construction of Normativity. *Oxford Studies in Metaethics*, Vol. 7 (forthcoming).
- Williams, B. (1981). The Truth in Relativism. In *Moral Luck*. Cambridge UK: Cambridge University Press, pp. 132–43.
- Wolf, S. (1982). Moral Saints. *Journal of Philosophy*, Vol. 79, No. 8, pp. 419-439.