



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Moral Ignorance and Blameworthiness

**Citation for published version:**

Mason, E 2015, 'Moral Ignorance and Blameworthiness', *Philosophical Studies*.  
<https://doi.org/10.1007/s11098-015-0456-7>

**Digital Object Identifier (DOI):**

[10.1007/s11098-015-0456-7](https://doi.org/10.1007/s11098-015-0456-7)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Philosophical Studies

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Moral Ignorance and Blameworthiness

Elinor Mason

## *Abstract*

In this paper I discuss various hard cases that an account of moral ignorance should be able to deal with: ancient slave holders, Susan Wolf's JoJo, psychopaths such as Robert Harris, and finally, moral outliers (people who, despite a normal background, behave in odious ways). All these agents are ignorant, but it is not at all clear that they are blameless on account of their ignorance.

I argue that the discussion of this issue in recent literature has missed the complexities of these cases by focusing on the question of epistemic fault. It is not clear that all blameworthy morally ignorant agents have committed an epistemic fault. There are other important issues that pull us in various directions: moral capacity, bad will, and formative circumstances. I argue that bad will is what is crucial, and moral ignorance itself can be a form of bad will. I argue that we should distinguish between two sorts of bad will, and correspondingly, two sorts of blameworthiness. Ordinary blameworthiness, requires moral knowledge, and is based on akratic action. The other kind of blameworthiness, objective blameworthiness, applies when the agent is morally ignorant, and when this indicates bad will. Objective blameworthiness can be undermined by unfortunate formative circumstances.

## 1. Introduction

Recent debate about moral ignorance has focused on historical cases, such as slaveholding in ancient times, and sexism in the 1950s.<sup>1</sup> On the one hand, it seems as though ignorance should exculpate, as it can in non-moral cases. Non-culpable ignorance of fact is usually a straightforward excuse. If I didn't know that the lever was connected to a puppy killing device, and there is no way I could have known or suspected that it was, then I am not blameworthy for pulling the lever. On the other hand, when the ignorance in question is *moral*, if, for example, I didn't know that killing puppies was bad, we are less inclined to straightforward exculpation.

In this paper I will use other examples of ignorance to argue that moral ignorance is not a simple exculpatory factor. What should we say, for example, about Susan Wolf's JoJo? JoJo is the son of Jo, an evil dictator. JoJo has been raised to accept and emulate his father's despotic rule, and JoJo does not question these values: he is cruel and ruthless. Yet JoJo has had such a limited moral education, it is hard to see how he could be fully responsible. (Wolf, 1987). Gary Watson (1987), also writing in the context of views about moral responsibility, discusses a case that is similar in some ways. Robert Harris murdered two people in cold blood and showed no remorse. Harris is not a paradigm case of moral ignorance: he claimed to know right from wrong. Yet, because of the horrendous childhood he had suffered, we are inclined to think that he is morally defective in some important way.<sup>2</sup>

---

<sup>1</sup> The cases are from Rosen (2003). See also Slote (1982), Zimmerman (1997, 2008), Levy (2011). The idea that one can be morally ignorant of course implies the converse, that there is moral knowledge. Throughout I make this assumption, even going so far as to refer to 'the moral facts'. This should not be taken to imply any sort of ambitious realism. A modest objectivism will do.

<sup>2</sup> I am not thinking of Harris as a psychopath, in the sense of a psychopath as someone whose moral deviance is innate, from birth. Rather, I am thinking of Harris as someone who probably would have been normal had his upbringing been normal. Obviously, these are stipulations for the sake of the argument.

These people seem at least somewhat blameworthy despite their moral ignorance. If we are inclined to let them off the hook at all it is not because of their ignorance, but because of their unfortunate formative circumstances.

Finally, there are the less easily explicable cases of ignorant wrongdoers: apparently ordinary people who knowingly cause suffering in pursuit of power or profit. These ‘moral outliers’, as I will call them, might think it is permissible to aggressively market infant formula in the third world, or to protect oil interests in an unstable region by funding the military oppression of protest. More extreme moral outliers amass enough power to start wars and commit genocide. Moral outliers (by hypothesis) did not have a terrible background, and were not raised in isolation from a normal moral education. Yet they are often, apparently sincerely, morally ignorant, and it is plausible that (at least in some cases) the ignorance was not acquired through a deliberate benighting act.<sup>3</sup>

## 2. The Akrasia Requirement

Both Michael Zimmerman (1997, 2008 ch. 4) and Gideon Rosen (2003, 2004) argue that unless there is some culpable act that directly caused the agent’s current ignorance, the agent is not culpable for acts done out of that ignorance. As Rosen puts it, this must be an act of “clear eyed akrasia” (Rosen, pp.306-308). In order to be culpable, the agent must knowingly be acting against her own best judgment about what she morally ought to do. There is rarely such an act in cases of moral ignorance, and so on both Zimmerman and Rosen’s view, agents are not usually blameworthy for what they do in moral ignorance.

Zimmerman and Rosen both claim that you cannot be *directly* responsible for acts done in ignorance. Zimmerman’s argument for this point is fairly subtle. He is not making a general claim that you are culpable for what you do in mental state M only if you are culpable for M itself. Zimmerman thinks that you can be culpable for acts done in anger, even though the anger itself is involuntary. Rather, there is something special about ignorance. As Zimmerman puts it, “one can act angrily while being aware that one ought not to perform the act in question, whereas one of course cannot act in or from ignorance of the fact that one ought not to do something while being aware that one ought not to do it.” (Zimmerman 2008, p. 177). So according to this view, being aware that you are doing wrong (having the correct self assessment), is crucial.

Rosen’s argument is very similar. Rosen, like Zimmerman, thinks it is clear that an act done in ignorance is not one for which we are directly responsible. Rosen argues that there is a general principle that we are not culpable when we are passive, and that we are passive when we are ignorant: Rosen compares it to being asleep at the wheel of a car. Again, knowing what you are doing is necessary for culpability (2004, p.300).<sup>8</sup> In the end, this is the

---

<sup>3</sup> That is Holly Smith’s phrase for the self-aware act that causes culpable ignorance (Smith 1983). Zimmerman’s and Rosen’s discussions both owe something to Smith’s account of culpable ignorance. (Smith 1983). However, Smith does not commit herself to the claim that acts done in ignorance are only indirectly blameworthy. As she frames the issue, blameworthiness depends on an act being done out of bad motives. In cases of culpable ignorance there is a benighting act, which certainly has bad motives. Her point is that it is hard to say whether earlier bad motives carry forward and become part of the later (ignorant) act. If bad motives are really *part of* the later act, then the agent is directly culpable for what she does later.

<sup>8</sup> By avoiding the claim that we are responsible for acting in a state M only if we are responsible for that state, Zimmerman and Rosen avoid the danger of a deterministic regress. They both leave the

claim that I shall dispute. Before getting to that I will examine Rosen's and Zimmerman's position, and some responses that the Akrasia Requirement in more detail.

The line that Zimmerman and Rosen take about culpable ignorance takes what the agent herself thinks she is doing as crucial to culpability. This lets a lot of people off the hook. It lets the ancient slave holders and the mid-century sexists off the hook. It also lets the moral outliers off the hook.<sup>9</sup> It lets JoJo off the hook. It probably lets Robert Harris off the hook.<sup>10</sup> The problem is, as Rosen points out, that almost all wrong action is done in ignorance in one way or another. Rosen himself admits that his view undermines much of our ordinary practice of blame assignment, and his own description of his view is 'skepticism about moral responsibility'.

### 3. Epistemic fault

Responses to Rosen have argued that the agents Rosen discusses are blameworthy after all, because they either should have (Moody-Adams (1994)<sup>12</sup>, Guerrero (2007), Fitzpatrick (2008),) or could have (Harman (2011) known better. These responses are all, in various ways, trying to show that blameworthiness depends on epistemic fault, though it is not as simple as the fault that Rosen is looking for.<sup>13</sup>

We should separate two questions, first, whether an epistemic mistake has been made, and second, whether it is blameworthy. Rosen sometimes presents his argument as if he is arguing that if there is no epistemic error there is no blameworthiness. But that is not his argument. Rosen's argument is that if there is no *blameworthy* epistemic error, there is no blameworthiness. Rosen says that we have procedural epistemic obligations. These are "to take certain steps to inform yourself about matters that might bear on the permissibility of your conduct. You are obliged to keep your eyes on the road while driving, to seek advice before launching a war and to think seriously about the advice you're given; to see to it that dangerous substances are clearly labelled, and so on." (2004, p.301). On Rosen's view, we are blameworthy for failing to fulfil our procedural epistemic obligations *only if we do it knowingly*. If we fail to fulfil our epistemic obligations in ignorance, then we are not blameworthy. His point is that when we fail to fulfil our epistemic obligations, it is usually because we didn't know what they were, or failed to understand them in some way.<sup>15</sup>

---

vital space to say that we are responsible for what we do when we know that we are doing it, although of course that too is determined.

<sup>9</sup> Rosen discusses a "ruthless capitalist" who is akin to my moral outliers in 2004, p.305. Of course, if the moral outliers know that what they are doing is wrong, then they are not off the hook.

<sup>10</sup> Rosen discusses a case of what he calls "bare ignorance of the rational force of moral considerations". That is a fair characterization of Harris's ignorance. (2004, p.305)

<sup>12</sup> Moody-Adams' argument predates both Zimmerman and Rosen's articles, and her discussion involves much more than a critique of the position they espouse.

<sup>13</sup> George Sher (2009) calls the Zimmerman/Rosen view the 'Searchlight View', and like me, argues that we need a more character based account of responsibility to make sense of hard cases. Sher discusses the 'should have known better' responses in chapter 5.

<sup>15</sup> At points Rosen talks as if what can make an agent blameworthy is "recklessness or negligence" in belief formation. Sometimes this is modified by the word "culpable", but sometimes not. For example on p. 304 Rosen says that culpable recklessness or negligence can make an agent culpable for ignorance. However, on the next page, the word 'culpable' is dropped. Rosen says, "I contend that if you are careful to bear in mind the stipulation that in reaching his conclusion our capitalist has not been reckless or negligent in the management of his moral opinion, you will find it plausible that his moral ignorance is not his fault." (2004, p. 305). That implies that *unknowing* recklessness might be

One could try to argue that epistemic error is blameworthy whether it is knowing or not. But that would be unconvincing. There are clearly cases where I can make a mistake that in no way reflects on me. The mistake might be entirely due to misleading evidence. And even when a mistake is due to a fault in me, the fault may just be the fact that I am not ideally rational. We cannot blame agents for not being ideally rational. The larger point here is that we cannot blame agents simply for failing to get things right. There is an important difference between failing to get things right, and failing to get them right in a blameworthy way. Responses to the Rosen/Zimmerman line must respect that.

One good objection to the akrasia requirement is that akratic action is not always clear eyed, and perhaps bleary eyed akrasia is sufficient for culpability. Both Fitzpatrick and Moody-Adams make suggestions along these lines. Moody-Adams claims that the epistemic errors that have been made are blameworthy because the agents knowingly (at some level) caused their own ignorance. Her argument is that much moral ignorance is affected ignorance, and that affected ignorance is more widespread than we are willing to admit. Part of Moody-Adams's point is a point about our own attitudes to these cases. She speculates that part of our reason for wanting to let the ancient slave holders off the hook is that we think of them as 'just like us'. We identify with them, and so we realize that in their position, we would have believed, and done, the same. This gives us an illicit reason to find views that let them off the hook attractive. On Moody-Adams' view, the slaveholders are culpable because they are in a complex state of self-deception about the wrongness of their acts.

I think that many cases of moral ignorance are blameworthy for the sorts of reason that Moody-Adams gives. Sometimes, epistemic error is blameworthy because it is knowing at some level. Clear-eyed akrasia may be rare, but bleary eyed akrasia seems reasonably common. Take the moral outliers I mentioned earlier: it seems plausible that much of the time, the CEO who says, 'bring me the profits, don't tell me how you got them' has an inkling that his profits are ill-gotten gains. The newspaper editor who says, 'get me the story, don't tell me if you are hacking phones', knows, of course, that phones are being hacked. In these cases we have a clear and sensible application of the claim that the agents 'should have known better'. There are harder cases: often ignorance is multi-layered, with self deception caused by complex motivations. It might be hard to say what is going on these cases, it is hard to say when epistemic errors are self-aware at some level. But being able to tell what is going on in particular cases is not crucial. Real life is messy.

However, Moody-Adams' picture does not apply in all cases. It is not clear that there is always epistemic fault when there is moral ignorance.<sup>18</sup> Just as with factual ignorance, there

---

blameworthiness rendering. The fact that recklessness usually is unknowing lends credence to that interpretation. However, Rosen's official view is clear: that blameworthiness attaches only to action done in full knowledge that the action is wrong.

<sup>18</sup> There is disagreement about this. Rosen has a firm intuition that it would have taken a moral hero to see that slavery was wrong (he is careful to use non-racially based slavery in his example, because it is more plausible that the evidence against that is unclear. I remain neutral on the question of whether the ancient slaveholders had misleading evidence). Arpaly also thinks that it is clearly possible that one could have a false moral belief as a result of misleading evidence. She gives an example of a young boy raised in an entirely sexist and segregated society (2003, p.104). On the other side, Harman thinks that we usually do have evidence for the moral truth (though this is not essential to her argument): "It is a hard question what constitutes evidence for moral claims, and in what circumstances a person's evidence is such that the right response to the evidence – the epistemically responsible response – involves believing the moral truth about a certain matter. But I claim that

might be cases where the evidence is misleading. I will return to that sort of case in the next section. In the remainder of this section, I shall discuss cases where ignorance is caused by a general flaw in the agent. Some moral outliers may fall into Moody-Adams' category of affected ignorance, but some may be, inexplicably, morally blind. Sometimes, the problem is not so much with what is known, but with how that is understood. Robert Harris claimed to know right from wrong, but he certainly lacked something, perhaps an ability to understand the practical force of moral reasons.<sup>19</sup> In these cases there may be no culpable error. I accept that there is some pull to exculpate on the grounds Rosen gives, but it still seems as though we are letting too many people off the hook. We have a sense that 'being bad' is important, even when that is not traceable to any previous culpable act. The agents in the hard cases above act badly, but more than that, they are bad.

'Being bad', in this sense, is not just equivalent to being blameworthy. I am using 'badness' here and in what follows to refer to a property that belongs to agents, that may or may not be blameworthy. Quite what the property of badness consists in is unclear of course. Part of what I aim to explore is how that concept works. The crucial point is that a bad agent is different from a bad inanimate object. Unlike a non-agent that is bad for us, like a flood, or a virus, an agent is bad in virtue of their intentions. This may or may not entail blameworthiness. What I mean when I say that it seems that badness is important is that it seems that bad intentions are important in themselves. This is the claim I aim to defend in more detail in what follows.

Like me, Fitzpatrick thinks it important that the agents who act out of ignorance are bad. However, there are two problems with his account. First, he is only interested in what I call moral outliers, agents who are bad without relevant extenuating circumstances in their background. Fitzpatrick argues, like Moody Adams, that many cases of moral ignorance are actually cases of affected ignorance. But then he goes on to argue that in cases where the ignorance genuinely goes all the way down, we might still be able to blame the agent on the grounds that they should have known better. Fitzpatrick cashes that out as follows: first, there are no relevant limitations on the agent's social context (note that this rules out blaming JoJo, or Robert Harris, and probably rules out blaming the slaveholders), second, the ignorance is the result of "the voluntary exercise of vices" such as overconfidence, incuriosity and so on, and thus, third, the agent could reasonable have been expected to do better. (Fitzpatrick 2008, p. 605).

The second problem with Fitzpatrick's account is that it is not clear what he means by the *voluntary* exercise of vice. Discussing the case of Potter, the ruthless businessman from *It's a Wonderful Life*, Fitzpatrick says, "It is enough for culpability if Potter made his epistemically debilitating choices—cavalierly dismissing opposing arguments, insulating himself from open, critical discussion or relevant sources of information, and so on—out of indulgence of vices, in a context where he could reasonably have been expected to know better and to do a better job of informing himself morally, given his capabilities and culturally available opportunities." (2008, p. 606). As Fitzpatrick himself notices, there is a danger that if Potter himself is not aware at any level that his acts are wrong (and this is the possibility Fitzpatrick is considering here), then Rosen's point stands. What is the sense in which it is reasonable to expect Potter to do better? There is obviously a non-blaming sense,

---

ordinary people who know the non-moral facts of what they are doing, when they do wrong things, often do have sufficient evidence that their actions are wrong." (2011, pp. 461-462)

<sup>19</sup> Rosen explicitly includes this sort of practical ignorance in his discussion of moral ignorance.

in which we expect people to behave morally (that is just equivalent to saying that their acts are wrong when they are wrong). But that is not what Fitzpatrick intends.

We should distinguish between a subjective and an objective reading of the claim that an agent should have known better. On the objective reading, the claim means that a *reasonable* agent would have known better. I think that Guerrero's view must be interpreted in this way (Guerrero 2007). Guerrero says that the standards for belief formation are high when the moral values at stake are serious (he calls this 'moral epistemic contextualism'). But this is ambiguous between a subjective reading and an objective one. If standards are higher when the agent *believes that* the moral values at stake are serious, then the principle seems true (it seems true that we are culpable if we do not do extra investigation when we know that serious values are at stake), but this is just Rosen's view. Culpability depends on self-aware fault.

On the other hand, if moral epistemic contextualism is read as saying that standards for belief formation are higher when the values at stake *really are* serious, then there is no clear link to culpability. The agent can be non-culpably ignorant of the fact that the standards are higher in a particular case. Guerrero objects that the slave holder fails to see things that are *obvious*. But they are not obvious to the slave holder, clearly. Guerrero's argument amounts to saying that we can blame agents for not seeing things that are obvious to us. But if these things are not obvious to the agent we are blaming, that they are obvious to us is not relevant. In order for an agent to be blameworthy, there must be something that we can pin to the agent, something that the agent herself, as she is, did wrong. In other words, we must interpret the claim that she should have known better in a fairly subjective way.<sup>20</sup>

So Fitzpatrick's view cannot rely on the claim that Potter should have known better. Either that just amounts to saying that he acted wrongly, or it must depend on some sort of akrasia after all. However, there is no reason that Fitzpatrick's argument shouldn't simply rest on the fact that Potter acts viciously. We don't need the further claim that he should have known better than to act viciously. It is enough that Potter is bad. When agents are bad, and I think that the slaveholders, Jojo, Robert Harris and moral outliers such as Potter are all bad, they are blameworthy for their bad will. That is the view that I shall go onto to defend.

#### 4. Blameworthiness without epistemic fault

So what should we say about cases where moral ignorance is caused by external circumstances? Elizabeth Harman argues that you can be blameworthy for having a belief, even though you have made no epistemic error in coming to that belief (Harman 2011, p.459). Harman's view seems to me to be correct: epistemic error is not necessary for blameworthiness. However, the argument Harman gives for the view is not the right sort of argument.

Drawing on Arpaly's work, Harman suggests that some failures to believe the moral truth involve insufficient moral concern (Harman 2011, p. 460).<sup>21</sup> This is the category that

---

<sup>20</sup> George Sher makes the same point about the mistake in appealing to the idea of a 'reasonable person'. (2009, p. 27). See also Levy's much more detailed response to Fitzpatrick, which makes a similar point (Levy 2009).

<sup>21</sup> One way in which Harman's argument differs from Arpaly's is that Harman is arguing about responsibility of beliefs directly, whereas Arpaly thinks that we are primarily responsible for acts. The other difference, which I discuss above, is that for Arpaly, lack of moral concern applies only in cases where the belief is irrational. Harman thinks that an agent in possession of all the non-moral facts

the moral outliers fall into – they lack moral concern, and that is why they do not see what is obvious to everyone else, and on Arpaly’s view (as on mine) they are blameworthy because they are bad. However, Arpaly argues (using the example of a boy raised in an isolated sexist community) that it is possible that someone has false moral beliefs because of lack of evidence. Arpaly’s view is that if one’s false moral belief is *rational*, one is not blameworthy. An agent has a bad will only when the agent has ignored the evidence. If they did not have the evidence (as may be the case for the ancient slaveholders) then their will is not bad. (2002, p.104).

Harman parts company with Arpaly, in defending the view that even if false moral beliefs are epistemically justified, there can be blameworthiness. Harman’s thought is that the evidence can be obscured by the cultural circumstances, but that nonetheless the slave holder (for example) has an obligation to believe the truth. Harman argues that it is not impossible for the slave holder to believe the truth, and so nothing rules out that being his obligation. From there, she argues that one can be blameworthy for things not occurring to one, and so nothing rules out saying that the slave holders are blameworthy.

The problem with this argument is that there is a big gap between wrongdoing and blameworthiness that Harman does not pay enough attention to. She may be right that *obligation* can be based on the very bare notion of possibility that she uses. Lots of consequentialists have argued that we can be obligated to do what is actually best, even when that is something that we could not reliably predict. On this view, our options include all those that are possible for us, even if we could realize some of those options only by a fluke, or without knowing how they will turn out. Of course, many writers on the topic of objectivism and subjectivism about obligation have argued that this is too objective an account, even for obligation.<sup>23</sup> And most people take it for granted that this is too objective an account of blameworthiness. It is part of our concept of blameworthiness that blameworthiness is connected to the agent much more closely than obligation: it must have something to do with the agent’s will. So the claim that we might have an obligation to believe the moral truth is not very relevant to whether we might be blameworthy when we fail to do so.

This problem is illustrated by the fact that Harman’s examples are not convincing. Harman gives us several examples of people who are supposed to be blameworthy just because something did not occur to them. The first is a parent who forgets to pick her child up from school. The second is a doctor who forgets to check the patients chart for allergies. According to Harman, both the parent and the doctor have taken every reasonable precaution against forgetting (Harman 2011, p.463-464)<sup>24</sup>. Harman argues that they are blameworthy, even though they have not violated any of their procedural epistemic obligations. So what should we make of these cases?

I agree that the parent and the doctor seem blameworthy, but I suspect that this is because we do not in fact believe that they have taken every reasonable precaution. If you

---

usually has sufficient evidence for the moral facts. She thinks that in all the hard cases under discussion, the agent has sufficient evidence.

<sup>23</sup> See my 2013 for more on this.

<sup>24</sup> Harman has a third example: someone who smells a gas leak and then lights a match. This case is under described: most people know that you shouldn’t do this, but some have no idea. The ones who have no idea are surely blameless. For the ones who do know (at some level), the point I make about about the level of precautions that it is reasonable to expect applies. Because gas is so dangerous, we should be very cautious.



have special responsibilities, as parents and doctors do, you must take *extra* precautions. Forgetting is not an option. Imagine our reaction on finding out that someone has forgotten to collect her child from school: we would certainly not think that saying “I don’t know what happened, I just forgot” was an excuse. But that is because we expect parents to have internalised these things so deeply that they do not forget. And if someone has trouble with their memory, we expect them to set alarms and reminders and write notes to themselves and ask others to help them. If they still forget, having done all that, then we probably would not blame them. In fact, given that the precautions we expect parents to take are so stringent, we would think that something must be *wrong* with a parent, who, having taken all reasonable expectations, forgot to pick up their child.

So where does this leave the morally ignorant agents in our cases above? If we grant, as Harman does, that the slave holder has made no epistemic mistake, then how could we hold the slave holder blameworthy? I agree with Harman that there can be blameworthiness without epistemic fault. Harman’s suggestion is that blameworthiness can be connected to the mere fact that the slave holder *could have* come to believe the truth. But this is not enough. Harman is right to think that we can sometimes be culpable for things that did not occur to us that could have occurred to us, but we need an account of the way in which ignorance in these cases is connected to the will of the agent. We need to know how an agent with subjectively rational and subjectively morally acceptable beliefs can nonetheless be showing lack of moral concern.

Perhaps it is not obvious that we think that the ancient slave holders have bad wills. However, I think that when we reflect about that case it is not difficult to see that we are, at the very least, ambivalent. Compare our reactions in two sorts of case. Imagine that we hear about a simple factual mistake: Jake waited at the bus stop in ignorance of the fact that the bus drivers were on strike. We would find it easy and inconsequential to say, ‘I would have done the same’. We find this considerably less easy in the case of the ancient slave holders. It is not easy, and certainly not something we can take lightly, to say: ‘If I had been born into the slaveholding class in a slaveholding society I would have owned slaves. I would have whipped them periodically, and let them sleep outside in the cold, and I would have sold their children.’

This is a complicated point. On the one hand, our discomfort with the thought that we might have been a slave holder indicates that we think that being a slave holder is blameworthy. On the other hand, the thought can get twisted in the other direction. This is the mistake that Moody-Adams rightly warns us to be cautious of (Moody-Adams 1994, p.302). Because we start off thinking that we are culturally similar, we want to exonerate them. Our thought is something like this: ‘it is sheer luck that *I* am not a slave holder, therefore it cannot be blameworthy.’ We should resist the change of direction in the argument. We could indeed have been slave holders, and if we were slave holders, we would have been bad. The thought in the end is just a thought about luck: it is sheer luck that the slave holders were bad, and we are not. But we should not let that distract us from the fact that they were, indeed, bad.

## 5. Quality of Will

In one of his earlier papers on this topic, Rosen briefly wonders whether Strawson’s account of moral responsibility might undermine his view. (Rosen 2003, p.72). Rosen revisits this in his third paper on the topic (Rosen 2008.) Rosen imagines a man called Kleinbart. Kleinbart

is trying to tell a funny story at a dinner party but unfortunately, the story is too personal, too other-revealing, and upsets his wife. However, Kleinbart is not aware, as he is talking, that this is the effect he is having, and his ignorance is not traceable to negligence.

“the only reason Kleinbart fails to notice his wife’s distress is that he simply does not care that much about her. He is dutiful enough. He would never knowingly offend her and he takes care not to do so inadvertently. We may even imagine that he overcompensates in this regard. Aware that indifference dulls the sympathetic faculties, he is somewhat more attentive to his wife than he would otherwise be. I want to stipulate that Kleinbart has compensated adequately in this regard. It was neither negligent nor reckless for him to launch into the story in the first place, and as the conversation unfolded he complied with the procedural requirements under which he labored. I want to stipulate, in other words, that his ignorance cannot be traced to prior negligence or recklessness of any sort.” (Rosen 2008, p. 607).

As Rosen points out, Kleinbart does not care enough about his wife, so he has an objectionable quality of will. If objectionable quality of will is what makes agents blameworthy, Kleinbart should be blameworthy. Rosen wants to resist this conclusion, although he admits that he has no decisive arguments. His strategy is to paint the view in an unappealing light. Rosen keeps pointing our attention back to the fact that poor old Kleinbart has done nothing deliberately wrong. In fact, all his deliberate actions are good. Kleinbart has done nothing to bring about his lack of regard for his wife, and he tries hard to compensate for it. Rosen insists that to blame Kleinbart seems unfair.

There are two reasons we might be attracted to the quality of will view. The first as Rosen points out in his earlier piece, is that the quality of will account may be the best compatibilist view we have. On a compatibilist picture, moral responsibility is based on some naturalistic feature of the agent’s action. The fundamental compatibilist insight, which we find in Hume and more recently in Strawson and Frankfurt, is that when we blame people, we are not blaming them on the basis that they could have done otherwise. Rather, we blame them because the act manifests some important feature of the agent. As Strawson puts it, we are interested in the *quality* of the agent’s will. To put it another way, we are interested in the motivations of the agent, not where those motivations ultimately came from. All compatibilist accounts are based on this idea, though the details vary.

We have no reason to worry about Rosen’s account just because we are compatibilists. Compatibilist theories can be based, in one way or another, on choices an agent has made.<sup>25</sup> But, the second reason we might be attracted to a quality of will account is

---

<sup>25</sup> In his 2003 paper, Rosen appeals to R.J. Wallace’s compatibilist account of moral responsibility, according to which we should blame the agent when it is fair to do so (p.74). It is not fair to blame someone when they did not know what they were doing. Of course, as this is a compatibilist view, we need to know why it is still fair to blame someone when determinism is true. Wallace’s answer is that it is fair to blame someone when they have *made choices* that fail to live up to moral standards that we hold them to. These choices must be self-conscious and controlled. Thus we cannot hold people responsible for states of emotion or feeling (Wallace 1998, pp.131-132). Wallace’s view, like Rosen’s, implies that none of the agents in our hard cases above are blameworthy. They all lack the reasoning capacities that are necessary for attributions of responsibility to be warranted (Wallace 1998, pp.166-180). Wallace’s view is thus like Rosen’s in that it fails to vindicate our inclination to apportion at least some blameworthiness to these agents.

that it gives us a way of saying that agents are bad even when they have not made morally self-aware choices to be or act that way. As I said, we have a strong intuition that some of the agents in these cases are blameworthy. Perhaps, that *all* of them are at least partially blameworthy. They seem blameworthy because they seem bad. It does not matter that there is no akrasia, and no epistemic fault. Whipping slaves reflects badly on someone even if we grant that their epistemic environment was such that their ignorance about the moral status of slaveholding is not culpably acquired. As Rosen himself puts it at one point, “an objectionable attitude towards others is often constituted by a form of moral ignorance: ignorance about the consideration due to others. So in these cases there will typically be no basis for concluding, from the fact that the agent acted from blameless ignorance, that his action does not express one of the underlying attitudes that makes blame appropriate.” (Rosen 2003, pp. 73-74.)

We should thus look to attributability views about responsibility. According to attributability views, what matters for responsibility is the agent’s deep self, where that is not merely a causal notion.<sup>30</sup> An agent’s acts and attitudes are attributable to that agent if they are appropriately related to the agent’s deep commitments. On this view, slaveholders are callous, cruel and disrespectful. These attitudes are in line with their deep commitments, and thus attributable to them, and they are fully blameworthy. JoJo is deeply bad in this way, so is Robert Harris, so are the moral outliers.

However, the problem with the attributability view as applied to cases of ignorant wrongdoing is just the converse of the problem with Rosen’s view: *all* the agents come out as fully blameworthy. What both of these views about moral responsibility are missing is the complexity of our hard cases. We want to be able to distinguish between the different cases, and the different reasons that we might want to grant some clemency. In the slave-owner case, we are worried by the fact that the ignorance is so widespread. Similarly, in JoJo’s case we are worried by the fact that his epistemic circumstances have been so limited; in Robert Harris’s, that his emotional development was so stunted. The complex cases involve various factors: the capacities the agents have varies; the completeness of moral ignorance varies; and the formative circumstances vary. In what follows I shall give a brief sketch of the relevance of these factors.

## 6. Two Kinds of Blameworthiness, and Grounds for Clemency

In this section I will propose an account according to which there are two sorts of blameworthiness. My view, in brief, is that ordinary blameworthiness is pretty much what Zimmerman and Rosen say it is: those who know (at some level) that they are acting wrongly have a bad will in a clear sense, and we react to that. However there is another way in which we blame those who are non-akratically ignorant. We blame them in virtue of their bad will where that means something rather different. In the remainder of this essay I focus on

---

<sup>30</sup> See Angela Smith (2005) for a useful discussion of the difference between volitionist and attributionist views. See e.g. Watson (1975) for a defense of the attributability view. More recently, T.M. Scanlon (1998, 2008), Nomi Arpaly (2003, 2006), Angela Smith (2005, 2008) and George Sher (2009) have all developed versions of an attributability view. In his later work Watson introduces a distinction between accountability and attributability that he uses to explain our ambivalence about cases like Harris’s. Harris is bad – his badness is attributable to him, but he is not fully accountable because of the terrible upbringing he endured. I discuss that distinction in what follows.

defining that second kind of blameworthiness. Along the way I distinguish my account from other ‘two sorts of blameworthiness’ accounts.

Normally, we blame each other for what we deliberately do. And if we find out that some piece of behavior was not deliberate, we let the agent off the hook. This is ordinary everyday blameworthiness. Ordinary blameworthiness is based on subjective wrongdoing. When ordinary people behave badly, they are usually, at some level (and this need not be the fully conscious level that Rosen and Zimmerman require), aware that they are doing it. Tony Blair did many wrong things during his time as Prime Minister, and it seems plausible that he knew, at some level, that these actions were wrong. He is not outside of our moral community: he did not seem to have the wrong end of the stick about what morality required. Rather, he was too easily swayed by the wrong sorts of reason. He did not try hard enough. Much of his ignorance, both factual and moral, was motivated ignorance or affected ignorance. He was (and is) thus blameworthy in the ordinary way. When we blame people for their akratic acts, we take it that they have the capacities and moral knowledge that we have: they are part of our moral community in that they share the basic standards that we hold ourselves to.

However, we need a different account of blameworthiness for cases where someone, for one reason or another, doesn’t have a grip on the correct moral views.<sup>31</sup> What should we say about JoJo, or the real versions: Colonel Gadaffi; Saddam Hussain; Kim Jong-Il, and so on? Agents who are deeply morally ignorant are bad if their attitudes are attributable to them, and morally obnoxious. A full account should fill in the details of exactly what it is for an attitude to be attributable to an agent, and recent attributability theorists have developed rich and subtle views.<sup>32</sup> My account requires only a fairly weak sense of attributability, because this is not my whole story about moral responsibility or blameworthiness. I agree with recent attributability theorists that attributability is not an entirely shallow notion, but I am happy with the idea that it is an “aretaic appraisal” (Watson 2006). On my account an attitude is attributable to the agent when it stems from and is connected in some loose (not necessarily conscious) way to the agent’s evaluative judgments. This rules out attitudes that have just been implanted by hypnosis, and it rules out random rogue attitudes. It does not rule out attitudes that the agent is unconscious of or only vaguely aware of having, and it does not rule out attitudes that the agent could not actually give reasons for.<sup>33</sup> The point is

---

<sup>31</sup> Again, no ambitious realism about morality should be inferred from this.

<sup>32</sup> See especially Angela Smith’s work (2005, 2008, 2012), also George Sher 2009, Shoemaker 2011. There has been a lot of discussion about what exactly is required for attributability. Shoemaker (2011) argues that attitudes can be attributable to an agent just because they are his, but more is required for answerability. Shoemaker thinks that an agent must also have access to the reasons for his attitudes in order to be answerable for them. Smith (2012) responds that answerability is the only sort of responsibility, but points out that it is not easy to say what exactly it is for an attitude to be connected to reasons in the right way. Both Smith and Shoemaker think that some sort of attributability can make sense of the whole space of moral responsibility. By contrast, I am using the notion of attributability to define one sort of blameworthiness; the other sort of blameworthiness is defined in terms of akrasia, or is the view often called ‘volitionist’ in the literature of moral responsibility.

<sup>33</sup> So, for example, Robert Harris’s homicidal tendencies are attributable to him, even if they are not the sort of thing that he could give reasons for. My account does rule out attributing things to the agent that reasons could not be given for *in principle*, such as height and eye color.

that an agent who has morally bad attitudes that are genuinely attributable to him, is himself bad.<sup>34</sup>

The idea that there are two sorts of blameworthiness is not new. Watson and Scanlon both suggest a distinction between attributability and something else, variously called accountability (Watson, 1996) or substantive responsibility (Scanlon, 1998, 2008). There are differences between Watson's account and Scanlon's but both Watson and Scanlon argue that attributability is the deep notion, and attributability is stable, undermined by nothing. My own view puts things attributability in a rather different position. For me, attributability is what we retreat to when we give up on ordinary blameworthiness. Further, on my view, attributability *can* be undermined. I shall argue for that below. For Scanlon and Watson, the other sort of blameworthiness (accountability, or substantive responsibility), is about what sanctions or benefits an agent is eligible for, and is undermined by lack of capacity or formative circumstances or both. I am not concerned with accountability at all here. On my view, the two relevant sort of blameworthiness are ordinary blameworthiness, which depends on akrasia, and attributability, or, as I shall call it, 'objective blameworthiness'.

In his discussion of Robert Harris's case, Watson recognizes that unfortunate formative circumstances seem to undermine responsibility, but as he points out, this is puzzling. First, knowing that someone had a terrible upbringing does not undermine the judgment that they are bad, it provides an explanation for them being so. Second, it is hard to get a grip on what sort of history is relevant. We all have some history or other, and we are all what we are because of our history. We need an account of why some histories excuse and some don't. My view is roughly that some personal histories amount to *damage* to the agent, where this can be emotional, epistemic, or more generally psychological. I will leave a full defense of that view for another time. Here, I will focus on *what* is undermined by recognizing unfortunate formative circumstances.

Watson says that when we consider cases such as Robert Harris's, we are inclined to withdraw accountability blame when we reflect on his awful childhood. However, according to Watson, attributability blame remains intact. Attributability blame is constant, and depends only on the way the person is, no matter how they got that way. So, presumably, if there is a difference between the Robert Harris who in fact had a terrible childhood, and someone who is just like Robert Harris who did not have a terrible childhood (a "bad apple" as Watson puts it), it is that one is less accountable. But the idea that someone is less *accountable* because of their history does not survive examination. In order to make that point I will briefly consider Scanlon's account of the role of history in undermining responsibility.<sup>37</sup>

Like Watson, Scanlon thinks that there are two sorts of responsibility, which he calls attributability and substantive responsibility. Like Watson, Scanlon is alive to the worry that if we say that history undermines responsibility, we are vulnerable to a generalisation strategy, which would show that we are never responsible, as of course, our acts always have a history that determines them. Scanlon makes a distinction between history in general, and some sorts of unfortunate social circumstances, which constitute bad conditions for making

---

<sup>34</sup> Notice that it doesn't matter whether there was epistemic fault or not on this view. One way to have a bad will is to ignore evidence, but another way is to treat other people badly.

<sup>37</sup> In Watson's later work (2004 and 2011) the appeal to history has dropped out of the picture. Both Watson and Scanlon (2013) think that a lack of moral capacity undermines accountability. In Scanlon's case this is connected to the way in which a lack of capacity affects someone's ability to be in a relationship.

choices. The latter, in virtue of being bad conditions for choice making, undermines substantive responsibility: the agent did not have adequate opportunity to avoid her situation. On this sort of view there is a difference between Robert Harris and the bad apple version of Robert Harris. Robert Harris had no opportunity to avoid his later character, whereas, presumably, a bad apple version (for whose bad character there is no clear environmental explanation) did have adequate opportunity to avoid his later bad acts.<sup>38</sup>

Scanlon's view is that substantive responsibility has a quite different source from attributability. Substantive responsibility depends on the rules and norms that no-one could reasonably reject. Scanlon thinks that we must accept rules and norms that take the conditions of choice into account when determining what sanctions apply. On Scanlon's view, the fact that someone was brought up in circumstances that made it hard for them to choose to be (for example) anything other than a drug addict, means that they should not be penalized by losing the right to have society take care of them and help them overcome their addiction.<sup>39</sup>

The first point to make about this is that, as we are in the realm of substantive responsibility, we are not talking about responsibility in the original sense anymore, but about the normative rules governing sanctions, and as such, our disagreements will tend to go along the lines of our general moral and political views about justice, desert and punishment. We may think justice should be backward looking, as the original sense of responsibility is, but we may think it should be mainly forward looking. Scanlon is criticizing a much harsher view than his own, according to which, having made choices at all disqualifies one from society's help in these sorts of circumstances. Scanlon thinks that, when choices have been severely limited by oppressive social conditions, we should take that into account. However, Scanlon's view is itself a fairly harsh view. An alternative view is that any addict, any alcoholic, who needs help should be given help. It makes no difference whether they were raised by wolves or by nice people in the suburbs. This (fairly common) account of social desert is entirely forward looking. Robert Harris and the bad apple version of Robert Harris are equally in need of our help, and if it were possible to give such help, both should get it. But it should be clear that, as I said above, we are not talking about responsibility in the same sense anymore. The original intuition was not about what society should do, but about responsibility in the sense we started with: the sense that has something (something hard to get a grip on, admittedly) to do with the agent, her will, her self.

The second problem with Scanlon's account is that although it may be fairly plausible when applied to the withdrawal of social help to addicts, it is much less plausible applied to other sanctions.<sup>40</sup> This is complicated by the fact that backward looking considerations tend to affect how we think about forward looking measures. Someone who has been raised in extreme poverty might be a better candidate for rehabilitation than someone whose wickedness seems to come from nowhere – what hope is there for reforming someone when we can see no explanation for their misdeeds in the first place? And on the other hand, in some circumstances it is counterproductive to limit sanctions when formative circumstances are poor. Compare two teaching techniques: one, the 'bleeding heart strategy', says that when

---

<sup>38</sup> Sarah Buss (1997) has a very different account of these cases. She argues that those who have had a terrible childhood are justified in their actions. Their reasons, she argues, (particularly their inductive reasons) are different to the reasons a more privileged agent has.

<sup>39</sup> See 1998, pp. 277-294, esp p.293, and 2008, pp. 198-206.

<sup>40</sup> I will leave rewards aside, as the relationship between praiseworthiness and rewards is even less clear than the relationship between blameworthiness and sanctions.

faced with a child from a deprived background, we should lower our standards for work and behavior. The other says, that when faced with a child from a deprived background we should not change or lower our standards, but we should work hard to help the child meet those standards. Letting people off the hook because of their background is often an ineffective strategy, and damaging to the very person it is supposed to help. Thus in various ways this is a complex issue. The question then, is whether we think that bad formative circumstances affect liability for sanctions *independently* of forward looking considerations. I suspect that we do not.

Many of our social sanctioning activities are enshrined in fairly formal rules and completely formal laws. That is not incidental, it is part of what our practice of substantive responsibility involves. An essential element of this practice is the thought that these rules should be applied consistently. Of course, there is disagreement about what counts as consistent, but we usually take it that the same crime merits the same punishment, no matter who commits the crime. We may feel pity and discomfort when we find out about unfortunate formative circumstances, but I think that these relate to what I called the original intuition about responsibility. We do not think that someone stands outside our practice of sanctions because their formative circumstances were bad. Consider JoJo. The conditions under which JoJo has made choices are extremely limiting, and for that reason we feel uncomfortable about *something* in the realm of responsibility: JoJo cannot be seen as an ordinary agent. But if JoJo were to be held accountable in an international criminal court, we would not be interested in his formative circumstances. Most people who end up in such courts have had bad backgrounds in one way or another, and of course that is why they end up there.

So a bad history does not undermine substantive responsibility. In that case, what sort of responsibility, if any, does it undermine? Think first about what I have called ordinary blameworthiness, which is based on akratic action: an agent is blameworthy in the ordinary way if they have roughly the right moral goals, but are not trying hard enough achieve them. This sort of blame is not undermined by reflection on formative circumstances. Imagine that we find out that Blair was a sad and lonely child, who yearned for an older brother, someone to emulate and admire. This may explain some of his actions, but it in no way undermines his blameworthiness. When we hold someone responsible in the ordinary sense we are setting aside facts about how the agent got to be the way they are. When an agent knows what she is doing, and understands the moral status of her acts (albeit blearily) she is responsible in a robust sense. To lose sight of that is to slip into thinking about things from a point of view that renders *all* of our actions determined.

I shall frame my account of the other sort of blameworthiness in terms of Strawson's taxonomy of excuses. According to Strawson there are two kinds of excuse: the first applies when the relevant quality of will is absent, and the agent is not, after all, responsible for the injury. This is the sort of excuse that is applicable when an agent is subject to ordinary blame. The second applies when the agent is not the right sort of agent. In that case, Strawson says, we take up the objective attitude. It is worth distinguishing between three different sorts of agential lack: those who are not agents at all; those who are temporarily the wrong sort of agent, and those who are permanently the wrong sort of agent. Those who are not agents at all are clearly exempted from blame entirely. Temporary cases are complex, and

I will leave them aside here. The interesting cases are cases where the agent is the wrong sort of agent: people who are ‘warped, deranged, morally undeveloped’ (or twisted) and so on.<sup>41</sup>

It might be thought that if we take Strawson seriously here, we should not blame such agents at all. But we need not think of the objective attitude as granting an exemption. Strawson allows that the objective attitude can licence sanctions. He also says this: “The objective attitude may be emotionally toned in many ways, but not in all ways: it may include repulsion or fear, it may include pity or even love, though not all kinds of love. But it cannot involve the range of reactive feelings and attitudes that belong to involvement or participation with others in interpersonal human relationships...” (Strawson 1962 in Watson 2003 p. 79). Strawson would probably be reluctant to characterize the objective attitude as a form of blame, but, arguably, it is a form of blame. The emotional tone that Strawson speaks of here is a reaction to the fault identified, and though as Strawson says, it cannot be resentment, it can be a different sort of blame. Henceforth I shall call it ‘objective blame’.

My concern is not with blame so much as blameworthiness, but the basic picture can be put in terms of either. Objective blame is a sort of blame. It is not the fully objective stance that Strawson speaks of: when we blame someone in the objective way we think of the author of the bad deeds as being an agent, not just an object. The point is that someone who does not understand morality can still have a bad will: they are bad precisely in being warped, or deranged, or morally twisted.

Objective blame differs from standard accounts of attributability blame in that it can be undermined by bad formative circumstances. Compare Tony Blair to Robert Harris. Robert Harris is not acting akratically: he does not grasp what morality requires of him. So he is only eligible for objective blame. It seems that, no matter how he got like that, Robert Harris is evil. However, when we compare the bad apple version of Robert Harris, this judgment wavers. The bad apple Robert Harris seems really evil. The real Robert Harris seems like a product of his environment.<sup>42</sup> Attributability theorists dispute this of course - they argue that our attributability judgments are stable in the face of considerations of history. But that doesn’t seem right. If attributability judgments were simply shallow appraisals, as their detractors sometimes say they are, then of course attributability judgments would be stable in this way. A rock is hard, no matter whether it came from a volcano or a seabed. But our attributability judgments are about *agents*. Likewise, my judgments of objective blameworthiness are judgments about *agents*.

What happens, when we learn about a horrible history, is that we lose confidence in the agential capacities of our agent. When we hold Robert Harris to be objectively blameworthy we take it that he is the author of his own acts, he is a bad person. When we think about his history, we see that he is more object than subject: he is acted upon more than he acts. When we consider the damage that has been done to him as a moral agent, we

---

<sup>41</sup> Strawson includes both children and those with who are unfortunate in formative circumstances in the class of people who we cannot apply the reactive attitudes to (in Watson 2003 p. 79). I think we can and do apply the reactive attitudes to children, though in a modified form, and as I argue below, we should see the point about formative circumstances as a separate point.

<sup>42</sup> Note that, as Watson points out in his discussion of Harris (Watson, 2004, p.247) the difference between Robert Harris and the bad apple Robert Harris is not a difference in *capacity*. Both Scanlon and Watson think that lack of capacity undermines accountability/substantive responsibility. On my view, lack of moral capacity is surprisingly irrelevant: the moral outliers presumably lack a moral capacity (they have no other clear cause for their badness) and yet they are fully blameworthy in the objective sense.



retreat even further, into the truly objective stance. Thus the clemency that we are inclined to grant in the light of bad formative circumstances really amounts to a retreat from objective blame to the truly objective stance. It is not a way of letting agents ‘off the hook’ (it is not an excuse), rather, it is further disengagement.

The difference between the agent who is blameworthy in the ordinary way and the agent who is blameworthy only in the objective way can now be further explained. When we blame someone in the ordinary way, we are setting aside considerations about how a person came to be the person they are, and focusing on what that agent does. The fact that the agent has enough of a grasp on morality to accurately assess the moral status of their own actions indicates that the agent has not been morally damaged. Thus we see them as a moral agent. Of course, they are what they are because of a certain history, but only in the sense that we all are, and so we do not, and cannot, take that as being an excusing condition.

By contrast, someone who does not grasp morality may or may not have been morally damaged by their upbringing. Either way, they are not part of our moral community, in the sense that they do not share our values, and so we cannot judge them by how they well they react to those values. We judge them in a more objective way, focusing instead on what they, and their attitudes are like. Consider Kleinbart’s wife: she cannot blame him for something he has done, as Rosen says, he has done nothing deliberately bad, and so there is no point in raising that with him. To stretch the analogy with the moral case slightly: Kleinbart and his wife are not in a community of shared values, he is coming at things from outside the place she wanted him to be. But that does not mean that she will have no reaction to him at all. It doesn’t matter that his lack of love is not deliberate – that he does not love her is sufficient reason for her to blame him in the objective way. She will blame him just because of what he is. He is a husband who does not love his wife.

One can imagine a further chapter in the Kleinbart story. Imagine that his wife finds out that Kleinbart has been poisoned over a long period of time, and that somehow the poison caused the erosion of his love. On finding out about the poisoning, we would expect his wife’s attitude to change. Even her objective blaming would now be undermined. Kleinbart seem less like an ‘unloving husband’, and indeed, less like an agent at all. His wife has already taken one step away from engagement with him when she realized that he simply didn’t love her. Now she takes another: Kleinbart is damaged, and so her feelings might change from something in the region of dislike to something more in the region of pity.

Contrast the moral outliers. Let’s stipulate that we are talking about the extreme cases: not cases of self-deception, but cases where the agent genuinely believes that his appalling acts are permissible. There is no clemency in our judgment, because there are no unfortunate formative circumstances. These agents are bad, but they are not damaged by their environment. The point about moral outliers is that they are evil *de novo*. As such they are the purest example of objective blameworthiness: what they are is bad, and there is no excuse or reason for clemency. Our objective blame is fierce. We denounce the moral outliers and their acts in the strongest terms.

As Watson points out, there is always something that has gone wrong, “if not in their socialization, then “in them” – in their genes or brains.” (Watson 2004, p. 247). My view is not that evil *de novo* is supernatural, based in the soul. It is of course something in the brain, caused in physical ways in one way or another. If we discovered a genetic mutation, or a reaction to a certain protein that had caused the outliers to be as they are, our objective blame would be undermined just as it is by consideration of bad formative circumstances. My claim is that we are reacting to an unexplained badness with a species of blame. So long as the badness remains unexplained, this blame remains in full force. It is not the way we

blame each other, our peers and equals. It is in some ways more stringent, usually involving a complete alienation from the blamed person, and in some ways less stringent, as it does not involve the demand for amends that ordinary blame implicitly does. In objective blame we stand back and disapprove, despise, disavow, we do not engage. This, it seems to me, is what happens when we talk about evil.

Now let us return to JoJo. It is clear that JoJo is not eligible for ordinary blame, and can only be blamed in the objective way. JoJo, we can imagine, is wholeheartedly bad. He embraces his role as dictator in waiting, and will one day fully inhabit the role of dictator. JoJo is evil in a different way to the moral outliers. In his case, bad formative circumstances are clearly relevant to what he is now. Like Robert Harris, he has been brutalized, even if we imagine that the damage done to JoJo is primarily epistemic rather than emotional.<sup>43</sup> Compare JoJo to his father, Jo. Let's imagine that Jo was raised in a fairly liberal society, and engineered a bloody coup in order to make himself dictator. Jo is a moral outlier, and our objective blame has full force. Our attitude to JoJo should be subtly different. Our objective blame of JoJo is undermined by considering his formative circumstances. Of course, this makes very little difference to anything practical: as I said, we should hold him accountable just as we do Jo. The point is that we can see JoJo as a product of his environment in a sense relevantly different from the sense in which we are all products of our environment, and so we see JoJo as less of an agent when we see him in that light.<sup>44</sup>

The relationship between ordinary blame, objective blame and the truly objective stance should now be reasonably clear. Once we are in the realm of objective blame, we have already taken a step away from inter-personal engagement. Someone who is not eligible for ordinary blame is not in our moral community, our attitude to them can only be appraisal, it cannot involve normative expectations, or gratitude, or resentment in Strawson's sense. Objective blame can be fierce, but it is not in tension with seeing the agent as a determined being in the way that ordinary blame is. When I focus on JoJo's upbringing, blame makes less sense. The picture I am developing is a picture of objective blame that is continuous with a truly objective stance.

Finally, where does this leave the ancient slave holders? There are two complexities here. First, the ancient slaveholders and the 1952 American sexist are ordinary-blameworthy for some of their actions and attitudes, but only objective-blameworthy for others. Second, their epistemic environment was poor, and possibly permanently damaging. Insofar as it would have been extremely hard for them to come to the correct moral view, they are 'merely' products of their environment. Thus all three categories of blame reaction are merited. But this is an accurate account of the complexity of the case. When we think about an American father in 1952, who educates his son but not his daughter, we might see lots of things that he did wrong and 'should have known better' about in the sense that, deep down, he did know better. We blame him in the ordinary way for these failings. But we can probably accept that he genuinely, and without culpable epistemic fault, believed that girls have a different path in life. So we do not blame in the ordinary way for his mainstream

---

<sup>43</sup> The notion of epistemic damage needs more elucidation. Roughly, the thought is that JoJo may be emotionally stable and average, but that his ability to see the world as it is has been damaged. He takes too seriously the testimony of his inner circle, and does not pay enough attention to other sorts of evidence.

<sup>44</sup> In her discussion of JoJo, Wolf says that JoJo is similar to us in being a product of his environment: he is unavoidably insane and we are unavoidably sane. I am denying this symmetry. JoJo has been damaged by his environment. I say more about the notion of damage elsewhere.

sexism. However, we don't let him off the hook, we justifiably condemn him for his sexism. His mistake was not a simple factual error, he was blind to something morally deep – his own daughter's right to reach her potential. As such he was bad, though, not of course deliberately so, hence we blame him in the objective way. But he was not like that *de novo*, he was like that because of the circumstances he was in, and most people believed exactly what he believed, so we see him as a product of his time, and as such, as less of an agent.

The view I have defended is complex, but blameworthiness is complex. I have argued that the picture of blameworthiness that Zimmerman and Rosen have in mind covers only one class of blameworthiness. We need a different concept to cover cases where the relevant moral knowledge is lacking, and so the concept of akratic action is not relevant. In those cases we react to a person's goodness or badness, and forget about how the person sees their own acts. But even that more objective attitude can be undermined if we judge that an agent has been damaged. In that case we start to retreat from condemnation, and move towards the truly objective stance, where we see the person as a victim above all.<sup>45</sup>

## References

- Arpaly, Nomy (2003). *Unprincipled Virtue: An Inquiry Into Moral Agency*, Oxford University Press.
- (2006). *Merit, Meaning, and Human Bondage: An Essay on Free Will*, Princeton University Press.
- Buss, Sarah (1997). 'Justified Wrongdoing', *Nous* 31 (3): 337-369.
- FitzPatrick, William (2008). 'Moral Responsibility and Normative Ignorance: Answering a New Skeptical Challenge', *Ethics* 118 (4): 589-613.
- Guerrero, Alexander (2007). 'Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution' *Philosophical Studies* 136 (1):59-97.
- Harman, Elizabeth (2011). 'Does Moral Ignorance Exculpate?', *Ratio* 24 (4):443-468.
- Levy, Neil (2009). 'Culpable Ignorance and Moral Responsibility A Reply to FitzPatrick'. *Ethics*, 119(4), 729-741.
- (2011). *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*, Oxford University Press.
- Mason, Elinor (2013). 'Objectivism and Prospectivism about Rightness', *Journal of Ethics and Social Philosophy* 7 (2).

---

<sup>45</sup> I would like to thank Gunnar Björnsson, Stewart Cohen, Sanford Goldberg, and James Regan for comments on earlier drafts. I am very grateful, for discussion and criticism, to audiences at Dartmouth College, Umea University, Nottingham University, and participants in the 3<sup>rd</sup> Annual Graduate Student Conference at the University of Stockholm where earlier drafts of this paper were presented, and to my colleagues at Edinburgh for numerous formal and informal discussions about these issues.

Moody-Adams, Michele (1994). 'Culture, Responsibility, and Affected Ignorance', *Ethics* 104: 291 - 309.

Rosen, Gideon (2003). 'Culpability and ignorance', *Proceedings of the Aristotelian Society* 103 (1):61–84.

----- (2004). 'Skepticism about moral responsibility', *Philosophical Perspectives* 18 (1):295–313.

----- (2008). 'Kleinbart the oblivious and other tales of ignorance and responsibility', *Journal of Philosophy* 105 (10):591-610.

Scanlon, Thomas (1998). *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.

----- (2008). *Moral Dimensions: Permissibility, Meaning, Blame*, Belknap Press of Harvard University Press.

----- (2013) 'Interpreting Blame' in Justin Coates and Neal Tognazzini, eds., *Blame: Its Nature and Norms* (Oxford: Oxford University Press, 2013), pp. 84-100

Sher, George (2009). *Who Knew?: Responsibility Without Awareness*. Oxford University Press.

Shoemaker, David (2011). 'Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility', *Ethics* 121 (3):602-632.

Slote, Michael (1982). 'Is Virtue Possible?' *Analysis* 42 (2):70-76.

Smith, Angela (2005). Responsibility for attitudes: Activity and Passivity in Mental Life. *Ethics* 115 (2):236-271.

---- (2008). 'Control, Responsibility, and Moral Assessment', *Philosophical Studies* 138: 367–392.

---- (2012). 'Attributability, Answerability, and Accountability: In Defense of a Unified Account,' *Ethics* 122: 575-589.

Smith, Holly (1983). 'Culpable Ignorance', *Philosophical Review* 92: 543–71

Strawson, P.F. (1962). 'Freedom and Resentment', *Proceedings of the British Academy* 48 (1962):1–25, reprinted in Gary Watson (ed.) *Free Will*, 2<sup>nd</sup> edition, Oxford University Press 2003: pp. 72-93.

Wallace, R. J. (1994). *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Watson, Gary (1987). 'Responsibility and the Limits of Evil: Variations on a Strawsonian Theme' in Ferdinand Schoeman, ed., *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, 256-286. Cambridge: Cambridge University Press, 1987. Reprinted in *Agency and Answerability*, 219-59. New York: Oxford University Press.

----- (1996). 'Two Faces of Responsibility' in *Philosophical Topics* 24 (2), pp. 227-248.  
Reprinted in *Agency and Answerability*, Gary Watson 219-59. New York: Oxford University Press 2004.--  
----- (2004). *Agency and Answerability: Selected Essays*. Oxford University Press.

----- (2011). 'The Trouble with Psychopaths', in Wallace, R. Jay ; Kumar, Rahul & Freeman, Samuel Richard (eds.) (2011). *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*. Oxford University Press.

Wolf, Susan (1987). 'Sanity and the Metaphysics of Responsibility' in Ferdinand David Schoeman (ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge University Press. Reprinted in Gary Watson (ed.) *Free Will*, 2<sup>nd</sup> edition, Oxford University Press 2003: pp. 372-387.

----- (1990). *Freedom Within Reason* Oxford University Press.

Zimmerman, Michael J. (1997). 'Moral Responsibility and Ignorance' *Ethics* 107 (3):410-426.

----- (2006). 'Is Moral Obligation Objective or Subjective?' *Utilitas* 18 (4):329-361.

----- (2008). *Living with Uncertainty*, Cambridge: Cambridge University Press.