

## **Intellectual Difficulty and Moral Responsibility**

Alexander A. Guerrero

In this chapter, I consider the relationship between intellectual difficulty and moral responsibility. In particular, I am interested in this question: if it is difficult for us to come to believe the truth about some matter, and we do not in fact come to believe the truth about that matter, so that we are ignorant of that matter, does that affect our responsibility if we then act from our ignorance? Answering this question requires getting clearer on both intellectual difficulty and moral responsibility for actions done from ignorance. I take up both tasks in this chapter, moving then to consider the interaction between intellectual difficulty and moral responsibility, focusing particularly on the familiar case of the Ancient Slaveholder.

### **1. Introduction**

Almost every action we take is such that we take it only conditional upon our having and lacking certain beliefs. Andre will take his children to school only if he does not believe that it is the weekend. Maria will pass the white powder to her friend at tea only if she does not believe that it is arsenic. A particular category of these cases has recently been the subject of philosophical discussion; namely, cases in which: (1) an agent performs an action, (2) the action is morally wrong, and (3) the agent performed the action only because the agent lacked true moral beliefs—because the agent suffered from moral ignorance. A standard example is the Ancient Slaveholder, who holds slaves, but at a time when (we are to suppose) the morality of slavery was unquestioned; it was taken for granted as a morally acceptable practice.

A provocative suggestion, which has been the focus of much philosophical discussion, is this: if an agent is not morally responsible for her *ignorance*, then the agent is not morally responsible for an immoral *action* taken only because of that

ignorance. Some defend this suggestion.<sup>1</sup> Some reject or complicate it.<sup>2</sup> A question relevant to this debate is how we assess whether an individual is morally responsible for being ignorant, and whether, in particular, someone like the Ancient Slaveholder is morally responsible for his ignorance. This is relevant for assessing any particular example, but it is also relevant for assessing how general an excuse would be provided if the suggestion is accepted. One concern is that perhaps much immoral action is the product of moral ignorance, resulting from false moral belief, and so the excuse would have wide application.

There are obvious ways in which one might be morally responsible for one's ignorance. One might *intentionally* do things: consciously decide not to investigate some question, institute firewalls to avoid learning something, take an amnesia pill, sign up for brainwashing, etc. More commonly, we end up ignorant not because of anything we have intentionally done. We don't consider a question, we are taught something false, we make a mistake in our reasoning, we forget. Sometimes these things happen in ways for which we are morally responsible, even though we do nothing intentionally: we are reckless or negligent in failing to consider a question, or in our investigation, forgetting, or reasoning, or in our willingness to accept some belief into our worldview. These are cases where we fall short of some norm—whether epistemic or moral—regarding the management of our beliefs. We have been reckless, or negligent, and this has led to our being ignorant.

But many other cases of ignorance do not look like this. Perhaps most cases of ignorance do not look like this. We appear to have done nothing wrong, there is just a difference in what we were taught, the evidence we encounter, or some other way in which our usually reliable method has in this case led us astray. To many, the Ancient Slaveholder looks like this.

Importantly—and this is where the issue of difficulty enters in—the claim made on the Ancient Slaveholder's behalf is not that it would have been *impossible* for him to come to believe that slavery was wrong. Whether slavery is wrong is an a priori matter, and there were a priori arguments that were available—in some sense—even back then. Instead, the claim made on the Ancient Slaveholder's behalf is that it would

---

<sup>1</sup> Rosen (2003), Rosen (2004), Zimmerman (1997).

<sup>2</sup> For example, Guerrero (2007); Harman (2011); Moody-Adams (1994).

be *difficult* for him to come to believe that slavery was wrong. Gideon Rosen says that “[g]iven the intellectual and cultural resources available to a second millennium Hittite lord, it would have taken a moral genius to see through the wrongness of chattel slavery” (Rosen 2003).

## 2. Difficulty

It can be difficult to perform various *actions* (e.g. walk on one’s hands) or to act so as to achieve various objectives (e.g. run a mile in under 5 minutes). It can also be difficult to *believe* particular propositions (e.g. that the moon is made of green cheese) or to come to *believe* the truth about some matter (e.g. whether Shinichi Mochizuki’s proof of the ABC conjecture is correct or not). My focus is on difficulty in believing, and my predominant interest will be in difficulty in coming to believe the truth about some matter. Some of what I say will have more general application, but I won’t pursue that here.

Difficulty in the sense I am interested in is a property that applies to what *agents* or—since we are focused on intellectual difficulty—*reasoners* can do. And difficulty is always indexed to particular reasoners or kinds of reasoners: it is difficult *for some particular reasoner or set of reasoners, R*, to do some particular thing, X. There might seem to be a sense of difficulty that is not relative to particular reasoners. We might say that this is a difficult math problem, without thinking of anyone in particular for whom it would be difficult. But claims of this sort are elliptical: it would be difficult for all human beings, or most human beings. For we can imagine other kinds of reasoner—perhaps an extraterrestrial super-genius—for whom it would not be difficult. What is difficult for you might not be difficult for me. And what is difficult for both of us, or for all of us, might be easy for someone, or something, else. So, the first part of the picture: difficulty is relative to particular reasoners, including, as we will see, the evidential situation of those reasoners.

Also important: it can only be difficult for you to X if it is *possible* for you to X. We could countenance a more expansive sense of difficulty on which it is difficult for you to X because it is impossible for you to X. But the cases we think about when we think about difficulty are cases in which it is difficult for you to do something, but

possible for you to do it. I'll focus on those cases. So, the second part of the picture: difficulty implies possibility for particular reasoners.

We do not have a well-developed idea of intellectual possibility for particular reasoners. What are the things that it would be possible for a particular reasoner, R, to come to believe the truth about? Belief is a funny thing. In a sense, it might seem easy to believe anything. In this way, it's not like jumping over a skyscraper or lifting an elephant. We can imagine future neuroscientists simply implanting beliefs about almost anything in our heads. In another sense, however, it can be very difficult or impossible to believe things in particular situations: when one has no evidence to support the belief that *p*, or when one has no reasons to believe that *p*, or, worse, when all of one's evidence and reasons seems to require that one not believe that *p*.

When we think about whether it is possible for a reasoner, R, to believe the truth about some matter, I want us to focus on the cases in which it is possible for R to come to believe the truth about some matter for the right kind of reasons or because R possesses the right kind of evidence—reasons and evidence that make believing the truth about that matter rational or justified or warranted for R. This is the sense of intellectual possibility that is interesting and intuitive. And it is certainly the one at issue when we are considering, for example, whether it is possible for the Ancient Slaveholder to believe that slavery is morally impermissible. We are not interested in the question of whether the Ancient Slaveholder could have been neuro-manipulated so as to come to believe that slavery is morally permissible. I will focus, then, not just on whether an agent has true beliefs about some matter, but also on whether those beliefs are *justified*. I want to leave the notion of justification open—I do not want to take a stand on what is required for justification.

Given these qualifications, it is apparent that there are several important omitted variables. It is plausible that intellectual possibility is relative not just to particular reasoners, but also to these facts about a reasoner's situation: the *available evidence* (the evidence available to the reasoner at a particular time) and the *time frame* (the time that the reasoner has in which to come to believe the truth in some domain).

There will be hard questions about what evidence is available to a particular reasoner in a particular world at a particular time. I won't try to settle those questions in a general way here. Importantly, what is relevant for intellectual possibility for a reasoner, R, is not just what evidence is available in the *world* in which R exists, but

more specifically what evidence is available *to R* in that world, given that those two things can come apart. What we need is a notion of availability of evidence that is indexed to particular individuals in particular worlds, but is also compatible with the intuitive idea that in some cases, the evidence available to a person will exceed (perhaps substantially) the evidence that they currently possess.

What makes it such that evidence is or is not *available* to a particular reasoner, R, within a particular time frame, T, given that R is in a particular world, W? Presumably, it will have something to do with things that R might *do* in W within the time frame, T. R might read a book or turn on the television, thereby acquiring the evidence. Or, if T is long enough, R might pursue graduate-level education in some subject matter, thereby acquiring the evidence. On this kind of view, some unpossessed evidence, E, is available to a reasoner, R, in a world, W, within a time frame, T, if and only if there are things that R could do in W within T that would lead R to possess E. This view doesn't say what it is for a reasoner to possess some piece of evidence. That project will have to be left to another day. The main point is that, when thinking about intellectual difficulty and intellectual possibility, we will need to think about at least two different dimensions of what reasoners can do: (1) what they can do to come to possess evidence and (2) what they can do with the evidence they possess.

As noted above, the sense of difficulty that we are interested in is one such that difficulty implies possibility. So, if we say that it would be *difficult* for R to come to have a justified true belief about some matter, M, within some time, T, in world W; then we are committed to saying that R *could* do this. Perhaps surprisingly, then, if coming to have a justified true belief about something is difficult for you, there is a sense in which you have at least some power or ability to do it. In particular, given the above discussion, there is something you could do either with evidence that you possess or with evidence that you could come to possess that would result in you having a justified true belief about some matter.

Sometimes all of our evidence about some matter will be misleading. And we might be in a situation in which everything we could do in terms of investigation and research would lead us further astray. In such situations, it is plausible that we have no power to come to have a justified true belief about that matter. Even if there is evidence in the world, if I don't have it, and can't come to have it through any efforts I might make, then it is not in my power to come to have a justified true belief about that

matter. I certainly can't just believe whatever I want to, in every case. This is at least one correct insight behind anti-voluntarism about belief. And even if I did, I wouldn't be *justified* in doing so.

But it is wrong to say that we never have power over coming to have a justified true belief about some matter. There are certainly actions that we might take (or choose not to take), including mental actions involving reasoning, deliberation, imagination, inference, and so on, as well as non-mental actions such as consultation, investigation, research, discussion, and so on, that will affect what we come to believe, and our justification for those beliefs, at least in some cases.<sup>3</sup> There are limits here, but there are limits in the practical case as well.

What I have been discussing so far applies to what I will call *difficulty in performing*. These cases are all ones in which we are imagining the agent in question as trying to come to have a true belief about some matter; more precisely: as performing a series of actions  $\{X_1 \dots X_n\}$ , in order to come to have a justified true belief about some matter. As noted above, these actions typically will either involve taking steps to obtain relevant evidence or reasoning and otherwise using the evidence that one already possesses. In these cases, we assume the person is trying to take various steps, and can consider whether a particular reasoner has the power to do successfully do those things or whether doing those things is difficult for them. But in both cases we are assuming that they are *trying* to come to have a justified true belief about some matter.

This covers a wide range of cases: can I come to have a justified true belief about the product of 734 multiplied by 1288, about which state is the 23<sup>rd</sup> state if all of the United States are listed in alphabetical order, about whether I ought to tell my near-death friend that his wife is cheating on him? And, for each of these, there is the question: if I can do this, is it difficult for me? We can understand this as concerning the relative number of contextually salient possible worlds in which I will come to have a justified true belief if I *try* to do so. This captures a significant portion of our ordinary concept of difficulty. But it also leaves out a substantial portion.

---

<sup>3</sup> Suggestions in this direction have been made by, among others, Clarke (1986) and Dretske (2000).

First, it leaves out those cases in which it is true both (a) that it is quite *likely* that if I try to come to have a justified true belief, then I will do so, but also (b) that it will require significant *effort* on my part. For these cases, we might assume that the agent will engage in the high level of effort required, so that it is not true that in many nearby possible worlds, the agent will fail to come to have a justified true belief even if she tries to do so. It is not probabilistically difficult or difficult due to circumstances over which one has only limited control. Instead, it is difficult because of the exertion required, even if it is quite likely that one will succeed if one makes the required exertions *and* one is likely to make those exertions.

This brings out that there are at least two different kinds of difficulty in performing.

One kind of difficulty is due to circumstances or factors that are (in normal contexts) somewhat outside of one's control, or only under one's control imperfectly. It is difficult for me to come to have a justified true belief about which chess move to make next is best, because of the hundreds of possibilities that lie ahead and my inexpert ability to recognize all the subtle but important facts about the current board position. I could get it right, and for the right reasons (in a justified way), but, not being any kind of grandmaster, I am far from assured of doing so. Or it could be difficult for me to come to have a justified true belief about whether you are lying, because although I am a skilled interrogator, able to recognize and interpret subtle micro-expressions and "tells," you may also be a skilled manipulator. So coming to have a justified true belief about your veracity might be difficult. We can call this kind of difficulty *skill-related difficulty in performing*.

A second kind of difficulty is due to circumstances or factors that are (in normal contexts) somewhat more under an agent's control, such as the effort an individual exerts or the amount of attention an agent pays to the task at hand. Some actions are such that they are difficult to perform because they require relatively more effort or attention, even if it is true that, if we exert that effort or attention, it becomes quite likely that we will succeed. We can call this kind of difficulty *effort-related difficulty in performing*. It is worth stressing that a particular task might be difficult for both skill-related and effort-related reasons, or perhaps we should understand both effort and attention as specific kinds of skills, distinct only in the relative amount of control that we have over their exercise. Additionally, whether the difficulty in

question is more skills-related or effort-related may vary from individual to individual, and even from context to context.

What both of these two kinds of *difficulty in performing* assume, however, is that difficulty in coming to have a justified true belief about some matter is to be assessed conditional upon the agent *trying* to come to have a justified true belief about that matter. It suggests that we should ask: is it hard for a reasoner, R, to come to have a justified true belief about M, within time T, if R tries to do so? But there is also a sense of difficulty related to the *difficulty in trying* to come to have a justified true belief. It might be difficult for a reasoner, R, to try to come to have a justified true belief about M, because doing so would go against R's fundamental beliefs and/or values, be difficult for R to steadily resolve to do, or simply not be something that R would think to do. And this difficulty in trying means that it might be difficult for R to come to have a justified true belief about some matter even if we assume that if R managed to try, it would be easy for her to do so.

In general, difficulty in trying will be a result of facts about an agent's attitudes (beliefs and desires) and experiences, although there might also be physiological or psychological conditions that lead to other kinds of defects in the transformation from attitudes to intentions and actions. So, unlike difficulty in performing, difficulty in *trying* must be analyzed not conditional upon a reasoner, R, *trying* to come to have a justified true belief, but conditional upon R's attitudes (beliefs, desires, etc.), experiences, and character being relevantly like they are in the actual world. It is difficult for R to *try* to come to have a justified true belief about M if R tries to do so only in relatively few of the contextually salient possible worlds in which R has attitudes, experiences, and a character relatively similar to those that she has in the actual world. Again, this will encompass different ways in trying might be difficult for R. But if we are to think of it being possible for R to try to come to have a justified true belief about M, and thus to have it make sense to say that it is *difficult* for R to try to do so, it must be that in some contextually salient possible world, R is both similar to how she is in the actual world, and yet she does try to come to have a justified true belief about M.

This last kind of difficulty, difficulty in trying, might seem to blur the distinction between something being *difficult* for a reasoner and it being *unlikely* that a reasoner will do something more generally. Skill-related difficulty in performing is one

example of difficulty that is connected to the probability of occurrence (the occurrence of coming to have a justified true belief about M), but what makes it more than just a claim about the low probability of something occurring is the connection between the *reason that* (or explanation why) it is unlikely and the *particular reasoner*. That is why it is awkward to say that it is *difficult* for Jackson to win the Powerball Lottery: it is very unlikely that Jackson will win, certainly, but that is arguably not sufficient for its being difficult, given that this unlikelihood has nothing to do with Jackson's skills, effort, beliefs, other attitudes, character, or anything at all about Jackson in particular. (Notice how our intuition about whether difficulty is involved changes if we add in facts about Jackson: that he is a somewhat skilled hacker who has been working for years to try to rig various lotteries in his favor, with the Powerball Lottery being by far the trickiest one he has ever attempted...)

In cases of skill-related difficulty in performing, it is unlikely that a reasoner, R, will come to have a justified true belief about M *because* of facts about R: she has only imperfect control over whether she comes to have a justified true belief, perhaps because doing so requires a great deal of precise performance and she isn't always adequately precise. Similarly, difficulty in trying is another example of difficulty that is connected to the low probability of occurrence, but, importantly, to read as a form of reasoner-specific difficulty, one must again attend to the facts about the *reasoner* that make the occurrence unlikely. In these cases the relevant facts concern R's attitudes (beliefs, desires), experiences, evidence, and character. All of these make it the case that R is unlikely to *try* to come to have a justified true belief about M. In this sense, it might be difficult for any of us to come to have a justified true belief about the title of the seventeenth book checked out from the University of Pennsylvania's Van Pelt Library in 2015, simply because it is incredibly unlikely (given our attitudes, experiences, evidence, etc.) that any of us will try to come to have a justified true belief about this matter.

Perhaps difficulty in trying is an unusual kind of difficulty. But it is important to countenance it as a kind of difficulty. Difficulty in trying captures a number of important kinds of difficulty that we encounter, including difficulty regarding remembering something or remembering to do something, difficulty in acting against type, difficulty in raising questions about one's unquestioned commitments and values, difficulty in noticing things that are easily noticeable, and difficulty in thinking about

things a different way than one usually does. None of these are well captured by a conception of difficulty that is limited to skill-related or effort-related difficulty in performing. These are the things that can be difficult to do, even though they are (if we try to do them) easy to do. It is true that we are unlikely to do them, but that is not sufficient for their being difficult. What is also required for difficulty in trying is that the reason we are unlikely to do them stems from facts about us—our attitudes, experiences, evidence, and so on.

In summary, then, there are at least three broadly distinct kinds of difficulty that reasoners can encounter: (1) skill-related difficulty in performing, (2) effort-related difficulty in performing, and (3) difficulty in trying.<sup>4</sup> In the next section, I will consider the relationship between difficulty and moral responsibility.

### 3. Difficulty and Moral Responsibility

Despite the extensive treatment of the question of the relationship between something's being *impossible* or *possible* for an agent to do and that agent being morally responsible for performing or failing to perform that action, there has been much less discussion of the relationship between *difficulty* and moral responsibility. An exception to that trend is Nelkin (2014), which invokes considerations of difficulty to make the

---

<sup>4</sup> My account is opposed to a unificationist account of difficulty, such as the one that Bradford (2015) offers, in which she argues that all difficulty can be analyzed in terms of effort. See 39-52. This seems implausible on an intuitive understanding of "effort." It might be difficult for the absent-minded professor to remember her friend's phone number. It might be difficult for a basketball player to make a shot from half court. It might be difficult for the sexist man to come to believe that he ought to help his daughters go to college. None of this difficulty seems well captured by focusing on effort, at least not in any ordinary, non-metaphorical sense of effort. Bradford opts for a "primitive" view of effort on which it is unanalyzable, and she is willing to extend the concept of effort so as to include all of the above cases. See Bradford (this volume). It seems helpful, for reasons that will become apparent, to distinguish more finely between different kinds of difficulty, with effort being only sometimes relevant to difficulty. And it seems odd, to me, to opt for a primitive view of effort, given that we have at least some clear ideas about both physical and mental effort.

case that accounts of moral responsibility must be able to accommodate the idea that blameworthiness and praiseworthiness comes in degrees. Much of her focus is on whether accounts of responsibility can accommodate degrees of praiseworthiness and blameworthiness, however, and she says very little about difficulty itself. Indeed, she simply asserts that “difficulty can be understood in at least two ways... [1] as requiring a great deal of effort... and [2] as requiring a great sacrifice of one’s interests.”<sup>5</sup> As the previous section’s analysis might suggest, I think that this is an inadequately precise understanding of difficulty, and this has implications for what she concludes about the relationship between difficulty and moral responsibility—a relationship that is more complex than it might appear at first glance.

Consider an initial suggestion one might make regarding the relationship between intellectual difficulty and moral responsibility (or blameworthiness):

INTELLECTUAL DIFFICULTY: if coming to have a justified, true belief about some matter, M, is very difficult for an agent A, then A is correspondingly less morally responsible (or less blameworthy) for failing to come to have a justified true belief about M, or for acting based on a false belief about M, than A would be if it had been less difficult for A to do so.

This principle suggests that intellectual difficulty has at least some mitigating effect on moral responsibility for actions taken based on false beliefs that are a result of that intellectual difficulty. Is this principle true? There are certainly some cases in which it is plausible that greater intellectual difficulty means lesser responsibility. Consider, for example, the following cases:

A: The Nazis are coming. It is possible for A to save a Jewish child from being discovered and taken to her likely death at the hands of the Nazis. To do this, A would have to come to have a true belief about an incredibly complicated question concerning the original building of a structure that has stood since the 17<sup>th</sup> century. A has all of the relevant evidence in his house—dozens of old building plans, blueprints, record books concerning 17<sup>th</sup> century architects,

---

<sup>5</sup> Nelkin (2014), 2.

histories of the neighborhood, complicated statistical evidence concerning building rates and locations, etc. But A makes a small mistake in interpreting the evidence and putting it all altogether, and so comes to have a false belief about the question, which in turn leads to the Nazis finding the Jewish child.

B: The Nazis are coming. It is possible for B to save a Jewish child from being discovered and taken to her likely death at the hands of the Nazis. To do this, B just has to come to have a true belief about the sum of two 3-digit numbers: 343 and 538. B tries to come to have a true belief about the sum, but makes a mistake, which leads to the Nazis finding the Jewish child (the sum is required to unlock a safe, which itself contains the escape plan).

Assume that both A and B are normal, competent, well-educated adults, and assume that time is not a factor in either case. Claim: A is less morally responsible—or at least less blameworthy—than B for failing to keep the Jewish child safe, and this is explained by the relative intellectual difficulty of what would have been required for A and B, respectively. This claim seems intuitively plausible. But one case is insufficient to support the very general claim INTELLECTUAL DIFFICULTY.

Here is a worry even about this pair of cases: perhaps difficulty has nothing to do with it. What we learn from this case, one might suggest, is that difficulty can be useful as a heuristic, pointing us to something else that is morally relevant.<sup>6</sup> Here is the thought: it is reasonable to infer that A cares very much about actually doing the right thing, that his concern and effort is in the right place, that his priorities are morally appropriate, and so on. In the case of B, things are the reverse; the fact that B makes such a simple mistake (“you had one job!”) suggests a lack of adequate moral concern. Why didn’t B take steps to ensure that he would get the problem right? We might think that we can infer from the cases as described to significant underlying differences in the moral attitudes and motivations and characters of A and B.

Response: that is too fast. At any rate, it at least seems possible to offer a stipulated version of the case in which A and B are exactly alike in this regard. A’s

---

<sup>6</sup> Markovits (2010), p. 240, suggests something like this ‘difficulty as evidentially useful’ picture.

mistake took place in a much longer series of actions, but we can stipulate that both A's and B's mistakes resulted from the same sort of causal process, that both are explained by what we might call a simple mistake (something that anyone, regardless of the moral attitudes and character, might occasionally make), not explained by anything deeper about them or their moral attitudes or characters. What, then, do we think of that version of the pair of cases? Does the mere fact of difficulty play a role? If it does, does this support something like the more general claim?

What I want to suggest is that there are some general patterns here, patterns that align with the three kinds of difficulty identified in the previous section. The relevance of these patterns to moral responsibility depends on one's account of moral responsibility. In the rest of this section, I will spell this out in more detail.

Recall that there were at least three different kinds of difficulty: (1) skill-related difficulty in performing, (2) effort-related difficulty in performing, and (3) difficulty in trying (due to some combination of one's attitudes, character, and experience). It is possible for a particular case to involve more than one of these kinds of difficulty: it might be difficult to get oneself to try, and then, even upon trying, be difficult for both skill-related and effort-related reasons. But for the purposes of simplifying the discussion, I will consider the relationship between each kind of intellectual difficulty and moral responsibility, but taking them one at a time.

In the background, it will be useful to have two broad pictures concerning moral responsibility and the corresponding relationship with praiseworthiness and blameworthiness.

The first picture holds that an agent is responsible—and correspondingly praiseworthy or blameworthy—only if, or only to the degree that, the agent's actions and beliefs are undetermined by forces outside of her control. Many who hold such a view do so in a robustly incompatibilist way.<sup>7</sup> But one might do so in a more compatibilist way as well. It might be, for example, that some forces from the outside impinge upon us, but that these do not fully determine what we do. This would leave us less than perfectly responsible, but still responsible to some degree. Much more must be said about when an agent's actions and beliefs are under her control or

---

<sup>7</sup> For example, see G. Strawson (1994) and, in the case of belief, Peels (2013).

determined by forces outside of her control, of course, in order to fill out the picture. But I will leave those details to the side for now. Call this the *agential control* view.

The second picture is concerned not with whether an agent has *control* over her actions or beliefs, but with what the agent's actions and beliefs *reveal* about her moral beliefs, attitudes, and values. These accounts are often described (somewhat inaptly) as “quality of will” views, and they have many variations and defenders.<sup>8</sup> The focus is often on when an action or agent is blameworthy or praiseworthy, where this is often glossed as concerning when certain reactive attitudes are appropriate in response to the actions of the agent. An example of this kind of approach is provided in recent work by Pamela Hieronymi, who has argued that “[w]e are fundamentally responsible for a thing ... because it reveals our take on the world and our place within it—it reveals what we find true or valuable or important.”<sup>9</sup> In a related vein concerning blameworthiness, Nomy Arpaly has argued that “an action is blameworthy just in case the action resulted from the agent's caring inadequately about what is morally significant – where this is not a matter of *de dicto* caring about morality but *de re* caring about what is in fact morally significant.”<sup>10</sup> Elizabeth Harman has argued that a similar view is correct with respect to beliefs: “[b]eliefs (and failures to believe) are blameworthy if they involve inadequately caring about what is morally significant.”<sup>11</sup>

Although there are important differences in the details of these accounts, they are similar enough to treat as a class of views distinct from the agential control views. Call this kind of view an *agential revelation* view: we are blameworthy and praiseworthy for those things that reveal who we are, morally speaking, or what our moral attitudes are like.

As may already be apparent, difficulty of the different kinds identified above will interact differently both with how much control we have, and with what our beliefs and actions reveal about us. Let me spell this out in more detail.

---

<sup>8</sup> For examples of views that fit (perhaps imperfectly) into this category, see Strawson (1962), Smith (2000), Talbert (2012), Hieronymi (2014), Arpaly (2003), McKenna (2012).

<sup>9</sup> Hieronymi (2014).

<sup>10</sup> Arpaly (2003).

<sup>11</sup> Harman (2011).

### *Skill-Related Difficulty in Performing*

Consider a case that involves a high degree of skill-related intellectual difficulty in performing:

D: In order to determine the whereabouts of a dangerous snake, an expert animal control specialist, D, will have to engage in a very complicated chain of reasoning, that includes elaborate scientific knowledge, tracking skills, intricate knowledge of local geography, and detailed information about the hours, feeding schedule, employees, and animal monitoring practices of the zoo from which the snake escaped. On many other occasions, D has successfully come to the correct belief about the location of dangerous animals through such a complicated chain of reasoning, and D could do so in this case as well. Unfortunately, although exerting her very best efforts, D makes a very small error (mistaking a very similar scale pattern as evidence that it was a snake of one kind, rather than another, with the two kinds having subtly different hiding practices), and so the snake bites and kills a person before D comes to have the correct belief about the snake's whereabouts, and so before D can tell the expert handlers where to go to capture the snake.

What should we make of the moral responsibility and blameworthiness of D? How does the skill-related difficulty in performing affect our judgment of her case? My inclination is to say that D is less blameworthy than she would be absent the difficulty involved—and perhaps she is not blameworthy at all. Let's consider what we might say under each of the two pictures sketched above.

On the agential control picture, the skill-related difficulty in performing should be seen to at least partially undermine agential control, so that D would be correspondingly less blameworthy or less morally responsible for her failure in performance. These views should allow for "threats" to agential control to come not just from things like full-blown determinism, but also from the many ways in which the world may mock or thwart our agential ambitions. Even if we have control over what we try to do, we certainly do not have full control over what we in fact do. This is

a familiar point from discussions of moral luck, but it seems to have clear application in discussion of at least this kind of difficulty and moral responsibility.

On the agential revelation picture, the skill-related difficulty in performing in this case dramatically limits (or alters) what D's actions or failures to act reveal or manifest about D's moral attitudes or moral character. On this picture, this kind of difficulty undermines the possibility of seeing D's actions as revealing or constituting any kind of moral failure or lack of moral concern. These kinds of cases suggest that skills-related difficulty in performing does affect moral responsibility, blameworthiness, and praiseworthiness. We should accept that D morally ought to have come to the correct belief about the location of the snake. But it also seems that because D encountered an instance of skill-related difficulty in performing, and this was in no way her fault (let us stipulate), she is correspondingly less morally responsible and less blameworthy for failing to comply with the moral demand in this case. Importantly, this is because the skill-related difficulty did not concern the skill involved in moral reasoning, nor was the intellectual difficulty a result of the agent's misguided moral views. If either of those two things were the case, then the failure would reveal or manifest something about the agent's moral attitudes or moral character.

Noticing the way in which this kind of intellectual difficulty can affect both agential control and what is revealed about agents who fail to comply with moral demands should, at least in some cases, alter our judgments of responsibility and blameworthiness.

### *Effort-Related Difficulty in Performing*

As discussed above, cases can differ in the amount of effort an individual is required to exert or the amount of attention an agent is required to pay to the task at hand. Some intellectual tasks are such that they are difficult to perform because they require relatively more effort or attention, even if it is true that, if we exert that effort or attention, it becomes quite likely that we will succeed. How does effort-related difficulty in performing relate to moral responsibility? Consider another case:

F: There is a sailor, F, who is out on the open sea. A distress signal comes in over his radio from a person who is stranded on a sinking ship in the middle of the ocean. There are no other boats within range, and if F does not reach the person soon, the person will drown. F's boat has a very old navigation system. To accurately locate the person calling for help, F will have to perform a long series of simple arithmetic calculations as numbers come in on his radio. They are easy to do, but they will require F to pay close attention both to the numbers being announced, and to the arithmetic. If F pays consistent attention and gets the numbers and math right, he will come to have the correct belief about the person's location and will be able to get there in time to save the person. If his attention falters, he will not. F gets distracted after 10 minutes of close attention, and writes a number down wrong, leading him to fail to rescue the stranded person in time.

What should we make of F in terms of his moral responsibility and blameworthiness? My inclination is that he is more morally responsible than D, but that F is less morally responsible than he would be if his failures took place in a context that required much less effort to do something otherwise very similar. But I confess to having somewhat less clear intuitions about this kind of case.

One question that seems to enter in as newly relevant is not just what F was individually capable of, but also what something like what effort "normal" or "reasonable" people would have exerted in a case like this, given the circumstances. Behind this, perhaps, is the thought that effort and attention, unlike more specialized skills such as juggling or animal tracking, are the kinds of things with respect to which we all have roughly the same abilities. Perhaps your efforts will do more than mine because you are a trained mathematician and I am not, but in terms of the amount that we are able to exert, there is a sense in which many of us are in a relatively similar position. This might turn out to be false as an empirical matter, but I think there is some idea in this neighborhood driving reactions to cases, and in particular to cases such as these.

Leaving this aside for a moment, what would the two different pictures of responsibility suggest about the relationship between effort-related difficulty and responsibility? On the agential control family of views, it is plausible that F should be

seen as more morally responsible, more blameworthy, than D, because it is plausible that the amount of effort we exert is more under our control than our exact deployment of certain skills. Indeed, on agential control views, greater effort-related difficulty makes no difference to mitigating blameworthiness, if it remains true in a particular case that it was fully within the agent's control to exert the greater effort.<sup>12</sup> (A thoroughgoing skeptic about responsibility might point to the background factors that make it so that any individual exerts any particular level of effort, but this is just to deny that it is at all within an agent's control how much effort she exerts.) This may turn out to be false, as an empirical matter. Perhaps it is not generally true that how much effort we exert is more under our control than our exact deployment of other skills—where that exact deployment is not itself a function of our effort (concentration, attention, exertion, etc.).<sup>13</sup> But it at least seems, intuitively, that this is the case.

On the agential revelation picture, again, we see a significant possible difference between cases D and F. In D, it was implausible to suggest that the actions or failures to perform were the result of insufficient moral concern or inappropriate moral beliefs about the importance of acting. In F, it begins to seem at least somewhat more plausible that the failure of effort and attention might be the result of failing to care inadequately about what is morally important. In such cases, the divergence from what it is possible for the agent to do and what the agent actually does might seem to be explained, at least in part, by some kind of inadequate moral concern. That makes these cases seem quite different from the case of D, from the skill-related difficulty in performing cases. Perhaps we might learn that effort and attention are skills in the relevant sense, and so we might come to think that the best explanation of the failure in effort-related cases makes the failure of F look much more like the failure of D. But this would require a revision in our ordinary thinking about effort and attention.

---

<sup>12</sup> It may be that greater effort-related difficulty does something to mitigate blameworthiness in many cases, because it becomes correspondingly less under our control whether we exert the greater effort that is required. And here some may have intuitions that push against the control picture, if they have the intuition that greater effort-related difficulty always mitigates blame, even if it is fully under the agent's control whether they exert that effort or not. I think intuitions become murky here, but this may be a point against a control picture.

<sup>13</sup> For relevant discussion of the empirical literature, see Holton (2009).

It is worth stressing that these comparative claims about moral responsibility in cases of skill-related difficulty in performing and effort-related difficulty in performing are meant to be claims about the *general* or *normal* cases that fall under these headings. It is possible to offer a case that was a case of effort-related difficulty, but in which the agent had absolutely no control over how much effort she exerted. Or a case that was a skill-related difficulty in performing case, but in which the agent had a great deal of control over her deployment of her skills. And it is possible to offer a case in which the amount of effort the agent exerts in no way should be taken to be a result of inadequate moral concern. What I am suggesting here is that there are general patterns regarding these different kinds of difficulty and their implications for moral responsibility, not that these patterns will hold true in every such case.

### *Difficulty in Trying*

As discussed above, in addition to the two kinds of difficulty that relate to performing conditional upon the agent trying, there is also difficulty that can arise in even trying. It might be difficult for an individual, R, to try to do X, because (for example) doing X:

- would require significant personal sacrifice from R, according to R's beliefs and values;
- would require R to sacrifice something that R believes to be of great moral importance;
- would require R to depart from the norms or beliefs of her society or community;
- would require R to think or value very differently than she currently thinks or values;
- would require R to have an ability to translate beliefs into actions that she currently possesses only weakly (perhaps as a result of a psychological condition such as depression).

And this difficulty in trying means that it might be difficult for R to do X even if we assume that if R managed to try, it would be easy for her (in both a skill- and effort-related sense) to do X. Here there are many different kinds of cases, and I won't be able to consider all of them in detail.

Importantly, we should focus on cases in which it is *possible* for R to try to come to have a justified true belief about some matter, M, although it is *difficult* for her to try to do so. Recall that what is possible for a reasoner will be a function, at least in part, of what she (1) can do to come to possess evidence and (2) can do with the evidence she possesses. But it will also be a function of the kinds of questions that might possibly occur to her, the concepts she might have to make sense of the world, and what she is psychologically capable of considering and imagining, given facts about her.

There seem to be at least three distinct kinds of variations of the above cases in the context of intellectual difficulty: (1) those in which R has entertained the possibility that she should think about matter M, (2) those in which R has *not* entertained the possibility of thinking about matter M and she is *not* blameworthy for not entertaining it; and (3) those in which she has not entertained the possibility of thinking about matter M and she *is* blameworthy for not entertaining it. All of the above kinds of cases can come in any of those three varieties. A full discussion of the relationship between difficulty in trying and moral responsibility would require working through the details of these different cases. What I want to do here, instead, is to highlight the very different responses to these cases that the two different pictures of responsibility offer.

On the agential control views, the question is: to what extent did the agent *control* whether she tried to come to believe the truth about M or not. The more control she had over this, the more she is morally responsible for trying or failing to try. So, for a person who is clinically depressed, it may be that control is largely eroded, so that the person is significantly less responsible. On the other hand, there may be ways in which we can have indirect control over our attitudes—perhaps by taking medicine, seeking counseling, discussing the issue with others, and so on.

It becomes more complicated in cases in which an individual's beliefs make it so that she usually wouldn't even consider trying to come to have a true belief about M. Perhaps an individual's beliefs make it so that she would be incredibly unlikely even to think about M, or so that she would only think about M if some event—itsself outside of her control—happens to take place. If we are not in control of our beliefs, then we would not be in control of whether we try to come to have a true belief about M or not. We might happen to, as a result of a kind of luck, but the difficulty would make this

relatively out of our control, just as it might be relatively out of a juggler's control whether he succeeds or not with some difficult juggling trick.

On the agential revelation views, things are substantially different. Because difficulty in trying is a function of an agent's attitudes, experiences, and character, it will often be the case that how difficult it is for an agent to try to come to believe the truth about some matter is revelatory of, and indeed a direct result of, the agent's moral attitudes, concern, and character. So, rather than lessening the agent's moral responsibility, many instances in which it is difficult for an agent to try will not lessen blameworthiness at all. This is, it seems to me, an explanation of the main points of divergence between agential control and agential revelation views, at least when it comes to the role that difficulty plays in affecting moral responsibility.

### *Summary*

Having seen how these three different kinds of difficulty might be considered by these two different views of responsibility and blameworthiness, we can now say some things about our initial claim:

INTELLECTUAL DIFFICULTY: if coming to have a justified, true belief about some matter, M, is very difficult for an agent A, then A is correspondingly less morally responsible (or less blameworthy) for failing to come to have a justified true belief about M, or for acting based on a false belief about M, than A would be if it had been less difficult for A to do so.

This claim is false as a general claim about the relationship between difficulty and moral responsibility, at least if we are still considering both agential control and agential revelation views regarding moral responsibility. Instead, we get a more complicated picture, something more like this (with some variation in each category depending on the details of the cases, and on the details of the view of responsibility or blameworthiness):

	<b>Skill-related</b>	<b>Effort-related</b>	<b>Difficulty in</b>
--	----------------------	-----------------------	----------------------

	<b>difficulty in performing</b>	<b>difficulty in performing</b>	<b>trying</b>
<b>Agential Control Views</b>	Often lessens/mitigates responsibility and blameworthiness	Rarely lessens/mitigates responsibility and blameworthiness	Often lessens/mitigates responsibility and blameworthiness
<b>Agential Revelation Views</b>	Often lessens/mitigates responsibility and blameworthiness, unless the skill is moral reasoning or the difficulty is a result of the agent's moral attitudes/character	Rarely lessens/mitigates responsibility and blameworthiness	Rarely lessens/mitigates responsibility and blameworthiness

So, greater difficulty does not always lessen or mitigate responsibility or blameworthiness on either of these two main families of views. And it does make a considerable difference which view one opts for, at least with respect to many instances of difficulty in trying. Intellectual difficulty will in some cases undermine control, or lessen control. And difficulty in some cases will affect what is revealed or instantiated or manifested by an agent's actions or beliefs.

#### **4. Intellectual Difficulty, Moral Epistemology, and the Ancient Slaveholder**

Let us return to the case with which we began: the Ancient Slaveholder. The case is often presented as one in which it would have been very difficult for the Ancient Slaveholder to come to have the correct moral beliefs about slavery, and thus it would have been very difficult for the Ancient Slaveholder to do what he morally ought to have done with respect to his treatment of enslaved people. But it is not supposed to have been impossible for the Ancient Slaveholder to have come to the correct view.

There is some contextually salient accessible world in which the Ancient Slaveholder comes to have the right view.

What kind of difficulty is present in this case? According to the taxonomy above, there are several different options. First, it might be difficult for the Ancient Slaveholder to come to believe the moral truth about the permissibility of slavery due to skill-related difficulty in performing the relevant intellectual task. Maybe it would have been like a very complicated math or logic problem. Second, it might be difficult for the Ancient Slaveholder to come to believe the moral truth about the permissibility of slavery due to effort-related difficulty in performing the relevant intellectual task. Maybe it would have been relatively easy, conditional on focused effort and attention. Third, it might be difficult for the Ancient Slaveholder to come to believe the moral truth about the permissibility of slavery due to it being difficult for the Ancient Slaveholder to *try* to come to believe the moral truth about this matter. Maybe it was the sort of thing that would just never occur to the Ancient Slaveholder to think about, like the question of how many paper clips can fit into an empty wine bottle. Finally, it might be difficult for some combination of skill-related difficulty, effort-related difficulty, and difficulty in trying.

We could stipulate different versions of the Ancient Slaveholder case, and then the previous section's analysis regarding kinds of difficulty and the relationship to different views of moral responsibility would help guide us. But it is worth thinking, too, about what seems to be an intuitive understanding of the actual case of someone like the Ancient Slaveholder: someone taking part in a practice that was widespread and at some point in time, but which today looks to be morally repugnant. Engaging this question requires engaging with a number of issues in moral epistemology. In particular, different views regarding moral epistemology will give different answers to these three questions:

- (ME1) In general, how difficult is it to come to have justified true moral beliefs?
- (ME2) How much variance is there amongst people in terms of how difficult it is for them to come to have justified true moral beliefs?
- (ME3) What explains the variance, if there is variance, in how difficult it is for different people to come to have justified true moral beliefs?

There isn't room for a full discussion of these questions here, but it is worth drawing attention to the broad contours of these issues, so as to make evident the choices regarding what we might say about the Ancient Slaveholder case and the kind of intellectual difficulty involved. This is important, given that it turns out that the particular difficulty that the case is taken to involve will affect how the case is evaluated from the different perspectives on moral responsibility.

There are two very different views about how difficult it is to come to believe the truth about morality and how much variance there is in this difficulty—these are two “packages” of answers to (ME1) and (ME2). On the one hand, there is someone like Thomas Reid, who writes:

“From these self-evident first principles the whole system of moral conduct follows so easily, and with so little aid of reasoning, that every man of common understanding, who wishes to know his duty, may know it. The path of duty is a plain path, which the upright in heart can rarely mistake. ... There are some intricate cases in morals which admit of disputation; but these seldom occur in practice; and when they do, the learned disputant has no great advantage... In order to know what is right and wrong in human conduct, we need only listen to the dictates of our conscience, when the mind is calm and unruffled, or attend to the judgment that we form of others in like circumstances.”<sup>14</sup>

More recently, Michael Smith writes that “[o]ur moral life seems to presuppose that [moral] facts are in principle available to all; that no one in particular is better placed to discover them than anyone else.”<sup>15</sup> And that “[i]t is agreed on nearly all sides that moral knowledge is relatively a priori, at least in the following sense: if you equip people with a full description of the circumstances in which someone acts, then they can figure out whether the person acted rightly or wrongly just by thinking about the case at hand.”<sup>16</sup> Smith stresses the a priori nature of moral reasoning, but couples that

---

<sup>14</sup> Reid (1785).

<sup>15</sup> Smith (1994), p. 5.

<sup>16</sup> Smith (2004), p. 203.

with a claim about what it is possible for anyone to do in terms of figuring out the moral truth, once they possess “a full description of the circumstances.” Reid stresses the easiness of moral reasoning, suggesting that it is something that we all have the ability to do and to do well. Let us call this the “easy and egalitarian” picture of moral reasoning: it is not difficult to come to have justified true moral beliefs and there is not much variance in this difficulty amongst different people.<sup>17</sup>

A very different view maintains both that it is generally or at least often difficult to come to have justified true moral beliefs, and that this can differ considerably amongst different people. Peter Singer, in defending the idea that there are moral experts, writes the following:

If ... there is reason to believe that one’s society does not have perfect norms, or if there are no agreed norms on a whole range of issues, the morally good man must try to think out for himself the question of what he ought to do. This ‘thinking out’ is a difficult task... None of this procedure is easy-neither the gathering of information, nor the selection of what information is relevant, nor its combination with a basic moral position, nor the elimination of bias.<sup>18</sup>

Singer suggests that part of the difficulty may stem from differential availability of relevant non-moral facts and evidence. (This is a point where Smith might part ways from Reid, being perfectly willing to countenance this as a possible source of differential difficulty.) But there is also difficulty that stems from selecting what information is relevant, combining it with moral views to see what is implied, and doing this while screening for various biases. And Singer is imagining an agent who is already *trying* to come to have a true moral belief about some matter. As I’ve suggested above, an additional difficulty might be present for some agents in even trying to do this. On this view, which we might call the “difficult and differential” view of moral reasoning, it is difficult to come to have justified true moral beliefs and there may be substantial variance in this difficulty amongst different people.

---

<sup>17</sup> Reid is committed to this view; it is not clear that Smith is.

<sup>18</sup> Singer (1972).

What explains the variance, if there is variance, in how difficult it is for different people to come to have justified true moral beliefs? One possibility is that people may have different relevant information available to them. A second possibility is that people may have different aptitude in thinking through their evidence, figuring out what it implies, and noticing and screening for biases. This aptitude might be relatively innate, but it might also itself be the product of relevant experience or exposure to relevant information—something that will be different for different historically-situated reasoners. For example, McGrath (2011) argues for a possible difference due to a “sensitizing” role for experience beyond mere access to information or relevant non-moral facts:

“One’s recognition that a given action would be morally right or wrong is often mediated by the recognition that the action would fall under a certain thick ethical concept, and one’s dexterity with such concepts is not independent of one’s past experience, in ways that go beyond the mere provision of information that one treats as evidence... one’s past experience influences the range of cases to which one is willing to apply the concept by conditioning one’s judgment; it thus influences how reliable one is with the concept... our past histories shape the way in which we respond to the non-moral information that both of us possess.”<sup>19</sup>

So, even if we grant that moral reasoning is a priori, there still might be differential difficulty, due either to different aptitude in the relevant a priori reasoning (as we see, for example, with mathematics or logic), or different access to relevant information or evidence, or different past experiences—or some combination of all of these.

If we accept the easy and egalitarian picture of moral reasoning, we will deny that the Ancient Slaveholder could have been in a case of skill-related difficulty in performing or effort-related difficulty in performing. It seems open on this view whether the Ancient Slaveholder faced difficulty in trying to come to believe the truth about what he was doing. It is plausible that even Reid would acknowledge that one must pay some attention to whether there is even a moral issue to consider; one must

---

<sup>19</sup> McGrath (2011), p. 107.

engage one's moral faculties and "wish to know" one's moral duty regarding engaging in some particular conduct or not. And there may be differences in whether one does this or not, or the extent to which one does.

If we accept the difficult and differential picture of moral reasoning, it is plausible that the Ancient Slaveholder may have been in a case of skill-related difficulty (due to inadequate aptitude or due to lacking relevant information and experience), or difficulty in trying, or both. On this picture, perhaps the one suggested by Rosen's use of the phrase "genius," the chain of reasoning—although a priori and thus in principle available to the Ancient Slaveholder, we might suppose—that would be necessary to reach the conclusion that slaveholding was wrong would have been very elaborate and difficult to get right, given where the Ancient Slaveholder was starting from, the evidence he had, and the beliefs he would have had to challenge.

Regardless of which view of moral reasoning we opt for, we may find the skill-related difficulty in performing interpretation of the difficulty involved in the case implausible as an explanation for why the Ancient Slaveholder failed to have a justified true belief about his slaveholding. One reason to doubt that this is the right explanation is that it does not seem as if the Ancient Slaveholder took any steps down the path of trying to figure out this question, so we have pretty weak evidence that this is the right explanation of his failure.

On a different view of the case, it would look more like a case of effort-related difficulty in performing or difficulty in trying. In the former case, the suggestion would be that although perhaps the Ancient Slaveholder might have at some point had some qualms, or some question about the practice presented to him, it was difficult for him to pay adequate attention or to make an adequate effort to investigate the question. It is plausible that on neither of the views of responsibility and blameworthiness would this do much to lessen his responsibility. He could have made more of an effort. And he should have.

If it were construed as a difficulty in trying case—which seems most plausible, given the concern about motivated ignorance or motivated reasoning, or given the possibility that his views would have made the issue never appear on his radar—then we get a split judgment. On the agential control views, this difficulty does lessen his responsibility. On the agential revelation views, this just reveals or makes manifest his blameworthiness.

I think the most plausible way to understand the case is as a difficulty in trying case—and that is the one kind of difficulty that both the “easy and egalitarian” and “difficult and differential” views of moral reasoning leave open. As to the final judgment, it strikes me that the agential revelation views give a more plausible result here (this kind of difficulty doesn’t lessen blameworthiness), while still being able to say plausible things about genuine skill-related difficulty in performing cases. But it is not possible to settle that debate fully here.

In considering the Ancient Slaveholder, it will also be worth considering other cases: the Racial Slaveholder in the United States antebellum South, the 1950s Sexist who provides for the education of his sons but not his daughters, the Contemporary Factory Farmer in the United States, the Military-Industrial Complex Lobbyist, and so on. One thing I hope to have shown in this paper is that if difficulty is involved in these cases, it is not straightforward that this lessens moral responsibility or blameworthiness—on either control or revelation views of responsibility and blameworthiness. The details of the difficulty matter.

## References

- Nomy Arpaly. 2003. Unprincipled Virtue (Oxford University Press).
- Gwen Bradford. 2015. Achievement (Oxford University Press).
- Murray Clarke. 1986. “Doxastic Voluntarism and Forced Belief,” Philosophical Studies 50.1, 39–51.
- Fred Dretske. 2000. “Entitlement: Epistemic Rights without Epistemic Duties?,” Philosophy and Phenomenological Research 60.3, 591–606.
- Alexander Guerrero. 2007. “Don’t Know, Don’t Kill: Moral Ignorance, Culpability, and Caution,” Philosophical Studies 136: 59–97.
- Elizabeth Harman. 2011. “Does Moral Ignorance Exculpate?” Ratio 24: 443-468.
- Pamela Hieronymi. 2014. “Reflection and Responsibility,” Philosophy and Public Affairs 42.
- Richard Holton. 2009. Willing, Wanting, Waiting (Oxford University Press).
- Julia Markovits. 2010. “Acting for the Right Reasons,” The Philosophical Review 119: 201-242.

- Sarah McGrath. 2011. "Moral Knowledge and Experience," Oxford Studies in Metaethics 6.
- Michael McKenna. 2012. Conversation and Responsibility (Oxford University Press).
- Michelle Moody-Adams. 1994. "Culture, Responsibility, and Affected Ignorance," Ethics 104: 291–309.
- Dana Kay Nelkin. 2014. "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness," Nous.
- Rik Peels. 2013. "Against Doxastic Compatibilism," Philosophy and Phenomenological Research 89(3): 679-702.
- Thomas Reid. 1785. Essays on the Active Powers of the Human Mind.
- Gideon Rosen. 2003. "Culpability and Ignorance," Proceedings of the Aristotelian Society CIII, Part 1.
- 2004. "Skepticism about Moral Responsibility," Philosophical Perspectives 18, 295–313.
- Peter Singer. 1972. "Moral Experts," Analysis 32(4): 115-117.
- Angela Smith. 2000. "Identification and Responsibility" in Moral Responsibility and Ontology.
- Michael Smith. 1994. The Moral Problem (Blackwell).
- 2004. "Moral Realism," in Ethics and the A Priori: 181-207.
- Galen Strawson. 1994. "The Impossibility of Moral Responsibility," Philosophical Studies
- Peter Strawson. 1962. "Freedom and Resentment," Proceedings of the British Academy.
- Matthew Talbert. 2012. "Moral Competence, Blame, and Protest," Journal of Ethics.
- Michael Zimmerman. 1997. "Moral Responsibility and Ignorance," Ethics 107: 410-426.