

# Don't Know, Don't Kill: Moral Ignorance, Culpability, and Caution

Alexander A. Guerrero

Published online: 25 August 2007  
© Springer Science+Business Media B.V. 2007

**Abstract** This paper takes on several distinct but related tasks. First, I present and discuss what I will call the “Ignorance Thesis,” which states that whenever an agent acts from ignorance, whether factual or moral, she is culpable for the act only if she is culpable for the ignorance from which she acts. Second, I offer a counterexample to the Ignorance Thesis, an example that applies most directly to the part I call the “Moral Ignorance Thesis.” Third, I argue for a principle—Don't Know, Don't Kill—that supports the view that the purported counterexample actually is a counterexample. Finally, I suggest that my arguments in this direction can supply a novel sort of argument against many instances of killing and eating certain sorts of animals.

**Keywords** Moral ignorance · Blameless ignorance · Culpability · Blameworthiness · Caution · Recklessness · Responsibility · Contextualism · Abortion · Vegetarianism

## 1 The Ignorance Thesis and the problem of moral ignorance

Michael Zimmerman and Gideon Rosen, among others, have discussed the problem of moral ignorance at length.<sup>1</sup> The basic issue vexing them is this. In cases of actions done from factual (non-moral) ignorance, it is commonly accepted that a person is morally culpable for the action only if she is culpable for the ignorance from which she acts. The question is whether the same considerations apply in cases of *moral* ignorance—ignorance of moral facts. Zimmerman defends the view that the cases are similar: in cases of actions done from moral ignorance, a person is

---

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

culpable for the action just in case she is culpable for the ignorance from which she acts.<sup>2</sup> Rosen follows Zimmerman, defending what we might call the Ignorance Thesis:

**Ignorance Thesis (IT)** Whenever an agent acts from ignorance, whether factual or moral, he is culpable for the act only if he is culpable for the ignorance from which he acts.<sup>3</sup>

I plan to argue against the part of the IT that focuses on moral ignorance. This is the part of the thesis that Rosen discusses at length. Let me rename that thesis:

**Moral Ignorance Thesis (MIT)** Whenever an agent acts from moral ignorance, he is culpable for the act only if he is culpable for the ignorance from which he acts.

At the outset, I should stress that the arguments and examples that I address to the MIT also work, suitably modified, against the other half of the IT—that part which focuses on ignorance of non-moral facts. I will focus on the moral case, because the examples that arise there do so quite naturally, and because the moral case raises unique and interesting problems.

To argue against the MIT, I must find a case in which

- (i) an agent acts from moral ignorance,
- (ii) the agent is not culpable for the ignorance from which he acts, and
- (iii) the agent is culpable for the act

In Part Two of the paper, I will offer a case that meets these three conditions, serving as a counterexample to the MIT. To do this, however, it will be helpful first to conduct a fuller discussion of these three conditions, so that we might assess whether or not a particular case meets them. Particularly, we'll need to consider what it means for an agent to act from ignorance, how we establish whether an agent is culpable for the ignorance from which she acts, and how we establish whether an agent is culpable for an act.

Throughout the paper, I will be working with what I hope is an intuitive notion of culpability. Basically, I will treat 'X is culpable for A' as equivalent to 'X is morally blameworthy for A,' and both can be further analyzed into something like the following: 'a certain kind of negative response to X for doing A (where A can be an omission) is appropriate.' I don't mean to be staking out any controversial territory in saying this, and I won't be engaged in the project of identifying the general conditions under which an individual can or might be culpable, or the various circumstances in which various kinds of negative responses (often called "reactive attitudes," following Strawson<sup>4</sup>) will be appropriate, or in defining 'appropriateness.' I think that something like this understanding of culpability is what Rosen has in mind when offering the MIT, and I hope that the general idea is clear enough without offering a full and detailed elaboration.

<sup>2</sup> Zimmerman, p. 411.

<sup>3</sup> Rosen, p. 64.

<sup>4</sup> Strawson (1982).

In order to discuss the three conditions relevant to the MIT, it will be helpful to look at two cases. First, I'll present and discuss a case in which it seems clear that an agent acts from moral ignorance, is not culpable for the ignorance from which she acts, and is *not* culpable for the act. (The MIT is false, but there are *some* instances in which non-culpable ignorance of the moral facts renders what one does non-culpable.) Second, I'll present the main example Rosen offers in support of the MIT—the example of the Ancient Slaveholder. Comparing these two cases will provide an opportunity to consider what is required to meet the three conditions outlined above.

### 1.1 Case one: The President

Suppose that some version of act-utilitarianism is true as a criterion of rightness, so that the right action(s) to take in any situation are just those actions whose consequences do not bring about less net utility than those of any available alternative. Imagine that an individual, call her 'the President,' is trying to decide whether or not to sign or veto some particular piece of legislation, deciding whether to make it law or not. Suppose that she knows that act-utilitarianism is true as a criterion of rightness, but that neither she nor anyone else knows what decision procedure to employ to act rightly most consistently. In deciding whether to sign or veto the piece of legislation, she is motivated only by a desire to do what is right. In an effort to assess whether it would be right to sign or veto the legislation, she consults a diverse panel of political scientists, lawyers, economists, psychologists, sociologists, natural scientists, and others to assess the likely consequences of each option, and a diverse panel of moral experts (philosophers, respected community leaders, religious figures, etc.) for their opinions on how these various consequences fare in terms of net utility. In light of all of this information, and after spending as much time as is available to her engaged in serious and meticulous deliberation, she decides to sign the legislation, believing that doing so is morally right. Let us suppose that it turns out that she is wrong. For somewhat complicated reasons that could not have been anticipated by the various experts, the legislation's enactment brings about negative consequences with greater disutility than those that would have followed vetoing the legislation.

This appears to be a case in which the President acts from moral ignorance (ignorance of the fact of the greater disutility from signing the legislation), is blameless for her moral ignorance, and is not culpable for her action because she was blamelessly ignorant. We will consider these claims in more detail when discussing the three conditions below.

### 1.2 Case two: The Ancient Slaveholder

Rosen, presenting an example first offered by Michael Slote<sup>5</sup> and discussed at length in an article by Michelle Moody-Adams,<sup>6</sup> presents a case in which we are to imagine an ancient Hittite lord who owns slaves, buying and selling human beings, requiring forced labor, breaking up families. Rosen tells us that the legitimacy of

---

<sup>5</sup> Slote (1982), p. 71.

<sup>6</sup> Moody-Adams (1994).

chattel slavery “was simply taken for granted” and that “until quite late in antiquity it never occurred to anyone to object to slavery on grounds of moral or religious principle.”<sup>7</sup> Slote states that ancient slave owners were simply “unable to see what virtue required in regard to slavery,” and that this “was not due to personal limitations (alone) but requires some explanation by social and historical forces, by cultural limitations.”<sup>8</sup> Rosen concludes on a similar note: “Given 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.”<sup>9</sup>

Rosen offers this as a case in which the slaveholder acts from ignorance (ignorance of the fact that slavery is wrong), is blameless for his moral ignorance, and is not culpable for his actions as a slaveholder because he was blamelessly ignorant.

What should we make of these two cases in relation to the three conditions discussed above? The first of these two cases, The President, provides an example of the sort that Rosen wants: an instance of acting from moral ignorance, an instance in which that ignorance is blameless, and an instance in which it seems that the agent is non-culpable because her ignorance is blameless. I think that the second of these two cases, The Ancient Slaveholder, does not provide an example of the sort that Rosen wants: it may be an instance of acting from moral ignorance, but both of the other two conditions—that the ignorance is blameless, and that the resulting action is blameless—are not obviously met in the way that Rosen suggests. Given this, The Ancient Slaveholder case, and cases relevantly similar to it, cannot serve to support the Ignorance Thesis. Additionally, though The President case does provide an instance that might lend support to the Moral Ignorance Thesis, there are other cases that serve to impugn it, which I will present in Part Two of the paper. But let us consider the three conditions in light of these two cases, so that the reasoning supporting these conclusions is transparent.

### 1.3 Acting from ignorance

The following is a natural understanding of the notion of doing something from ignorance:

- (AI1) a person *S* performed action *A* while lacking belief in some proposition *p*,
- (AI2) *S* would not have done *A* had *S* believed that *p*,
- (AI3) *p* is true.

We need to include both (AI1) and (AI2); if we only included (AI1) we might suspect that the action was done from callous indifference as much as from ignorance (or absence of belief).<sup>10</sup> This would seem to go against the tone in which these cases have been presented. The locution ‘done from ignorance’ seems to

<sup>7</sup> Rosen, p. 64.

<sup>8</sup> Slote, p. 72.

<sup>9</sup> Rosen, p. 66.

<sup>10</sup> In cases in which the relevant proposition *p* is true, the person will be acting from ignorance. In cases in which the relevant proposition *p* is false, the person will not actually be acting from ignorance (from lack of true belief), but their action is explained by the fact that they lack some particular *false* belief, and so it will sometimes be relevant to discuss these cases, though I will include this under the heading of ‘acting from absence of belief’ rather than under the heading of ‘acting from ignorance.’

imply that the ignorance played an important causal or explanatory role in the agent actually performing the act in question. This also explains why there is no problem in ruling out cases in which a person acts *with* or *in* ignorance or absence of belief, but would not have done anything differently if they had not lacked belief in *p*.

Ignorance for our purposes is just lack of true belief.<sup>11</sup> Given this, there are several different ways in which someone could be ignorant of some fact *F*:<sup>12</sup>

- (I1) cases in which a person is ignorant because she has never thought about the issue (and so has no beliefs about *F* or only has unexamined 'implicit' beliefs about *F*)
- (I2) cases in which a person is ignorant because, though she has thought about the issue, she has come to have false beliefs about *F* (she believes that *not-F* when in fact, *F*)
- (I3) cases in which a person is ignorant because, though she has thought about the issue, she doesn't know what to believe (she doesn't believe that *F* or that *not-F*)

In all three cases, it is natural to say that the person is ignorant of *F*. One can act from ignorance regardless of which of the three ways best describes one's way of being ignorant.

In defending the MIT, Rosen tends to rely on cases of the first sort, cases in which the person has never thought about the issue. In these cases, the person either lacks beliefs about *F*, or has only implicit beliefs about *F*.<sup>13</sup> Rosen also mentions cases of the (I2) variety (perhaps including The Ancient Slaveholder case). There is a sense in which one's ignorance is 'deeper' in the first two cases, since not only does one not know that *F*, one also doesn't know that one doesn't know. The (I3) cases are different: if one goes ahead and acts when in an (I3) state, there is a 'taking a chance' component to one's action that is absent in cases of (I1) and (I2). Regardless, it seems that one acts from ignorance in (I3) cases as well, as long as

<sup>11</sup> In this paper, I will for the most part ignore the possibility of offering this account in terms of *degrees* of belief. This will simplify the presentation, though there are places in which it would be natural to recast the account in those terms. The basic points stand on either way of presenting them, though there are places in which I will call attention to possible differences that might arise on the degrees of belief account, or places in which it would be difficult to capture exactly what I'm after in those terms.

<sup>12</sup> In general, I will focus on the simple case, on which an individual either does or does not have true belief about some particular fact *F*. Obviously, there could be cases in which an individual has contradictory beliefs about the same fact *F*, such that they believe both that *F* and that *not-F*. This could happen either from irrationality, confusion, or from cases in which an individual's beliefs are sensitive to particular presentations or descriptions of the fact in question (e.g. someone might believe that some particular frog lacks moral status while believing that some particular prince has moral status, without realizing that the frog *is* the prince). There are cases in this vein which can raise interesting problems, but I will leave them aside for the purposes of simplifying the discussion here.

<sup>13</sup> Roughly, and without wanting to take a stand regarding the metaphysics of beliefs, one has an implicit belief that *p* if one believes *p*, but non-occurently, and if one's mind does not possess an explicit representation with that content. I might, for example, only implicitly believe things like the following: that there isn't a koala bear in the tree outside my window, that I couldn't name all the countries in the world, that I do no one moral injury by pressing the spacebar on my keyboard (of course, I come to explicitly believe all these things after coming up with the examples). Most important, for our purposes, is the fact that people have not brought their implicit beliefs under conscious scrutiny.

one would have actually acted differently had one known that  $F$  obtained. As noted above, if one would have acted in exactly the same way even if one was not ignorant of the fact that  $F$  obtained, then it does not seem plausible to say that one is acting *from* ignorance, even if one happens to be acting *in* ignorance.

Given this account of acting from ignorance, it seems clear that the President is acting from ignorance. She provides an instance of an (I2) case, since she has thought about the issue, but has come to have false beliefs about the issue. Additionally, in the case as described, it is clear that she would not have performed the action in question if not for her ignorance. It is less clear that the Ancient Slaveholder is acting from ignorance. It is clear that he is acting *in* ignorance, in that he does not possess the true belief that slavery is wrong. The case is under-described, however, when it comes to assessing whether it is best considered an (I1), (I2), or an (I3) case—we can imagine all three variants. It is also under-described in that it is not obvious that the Ancient Slaveholder would have actually acted differently if he had known that slavery is wrong. Given the description offered by Slote and Rosen, it is most natural to treat the case as either an (I1) or an (I2) case, and to assume that the Ancient Slaveholder would have acted differently if he had had true moral beliefs. I will have more to say later about attempting to think of the Ancient Slaveholder case in this way.

#### 1.4 Culpability for ignorance and culpability for action

Is the President culpable for her ignorance? Is the Ancient Slaveholder culpable for his? What conditions must obtain for one to be culpable for one's ignorance? Given the formulation of the MIT, these can seem to be crucial questions, since ignorance only excuses on the MIT when that ignorance is itself non-culpable. Though answering these questions is important for the narrow project of assessing the truth of the MIT, doing so is less important to assessing whether some particular individual is morally culpable in a specific case. The importance of offering a compelling account of culpable ignorance stems not from concerns about final moral culpability, but from the importance of making a correct diagnosis of the source of the culpability in these cases (cases in which an agent, acting from ignorance, seems culpable). Was the problem that the person shouldn't have been ignorant of this fact? Or was the problem that the person shouldn't have acted, given that they were (blamelessly) ignorant? Offering an account of culpable ignorance is important to be able to identify where the moral error enters in; it is less important to assessing final moral culpability than the MIT would have us believe, or so I hope to demonstrate.<sup>14</sup> In that

<sup>14</sup> In considering whether an individual case serves to impugn the MIT, we are faced with the following situation: if it is relatively easy to be culpably ignorant, then many apparent counterexamples to the MIT will fail; if it is harder to be culpably ignorant, if the obligations governing deliberation, etc., are quite low, then, arguably, there will be many more counterexamples to the MIT. Basically, we'll end up in one of two dialectical situations when a purported counterexample to the MIT has been presented. In the first situation, we can grant that the person is blamelessly ignorant, and then find them 'still culpable' for what they do, acting in ignorance. In this situation, the case stands in as a counterexample to the MIT. Alternatively, we can argue that the person wasn't blamelessly ignorant, that they should have done more (perhaps the investigative burden was higher because of the *type* of thing they were doing, I'll suggest), and this is why they are culpable for the act done from ignorance—they are culpable for the ignorance. This option doesn't serve as a counterexample to the MIT.

spirit, let me attempt to offer a template for an account of culpable ignorance. To help with this task, let me add a third case to our first two.

### 1.5 Case three: The Dogsitter

Imagine a case in which a dogsitter is taking care of several dogs. The dogsitter falsely believes that dogs cannot feel pain, because he has read Descartes and believes that non-human animals are no more than automata, intricate machines, and that machines cannot feel pain. Let us assume that the dogsitter has evidence available to him to the effect that dogs are not automata and that they can feel pain—suppose that there are issues of *Scientific American* sitting on his coffee table at home that attest to this in exhaustive detail, that these same issues of *Scientific American* are at the house with the dogs, that he knows that his friend (a cell phone call away) took a class on animal psychology, that he has himself witnessed many dogs exhibiting behavior that suggests that they are in pain, that he knows that the differences between human and non-human animal physiology can be quite small, and so on. He has even considered whether the views of a 17th century philosopher should be relied upon in this case, and has considered consulting some of the other evidence available to him with an eye to reconsidering his belief. Still, he hasn't done this, and he hasn't bothered to pick up the issues of *Scientific American* or anything else, and he maintains his false belief. While taking care of the dogs one of the things he routinely has to do is to roll a heavy wheelbarrow through the garage and into the yard. He notices that if he goes one route, the wheelbarrow will roll over one of the tail of one of the dogs while it sleeps in its bed. He could go another route, but it is slightly further, and so he decides to just go the route that will result in the heavy wheelbarrow rolling over the dog's tail, since it won't cause any permanent functional harm, and he believes that dogs do not feel pain. He goes this route, and the dog routinely appears to be in great pain due to the heavy wheelbarrow rolling over its tail. Still, he tells himself, the dog is not actually feeling pain, and so he takes himself to have no moral reason to go a different route the next time. If he believed that the dog was feeling pain, he would take himself to have a moral reason to go a different route, and would do so. But he doesn't believe this, and so he doesn't go a different route.

On any plausible account of culpable ignorance the Dogsitter *is* culpable for his ignorance and the President *is not* culpable for her ignorance. It is less clear what we should think about the Ancient Slaveholder. The purpose of this section is to offer an account of culpable ignorance that will both get the obvious cases right and provide illumination regarding the different considerations that make the Ancient Slaveholder somewhat more in the middle.

The idea that one might be culpably ignorant presupposes that there are obligations regarding our belief regulation, and that we can fail to live up to these obligations. Offering an account of when ignorance is culpable or blameless will require us to define these obligations in some detail. One question that arises at the very beginning of this inquiry is what *kind* of obligations these are—particularly, are these moral obligations, epistemic obligations, or some other kind of obligation? Rosen asserts that the relevant obligations are “moral obligations governing the epistemic aspects of

deliberation,”<sup>15</sup> but he says little about the contours of these obligations, their source, or how we can identify them. He says that they include duties “to look around, to reflect, to seek advice, and so on.”<sup>16</sup> Let me try to amplify this a bit.

Consider the following intuitive view of culpable ignorance:

A person *S* is culpably ignorant with regard to some fact *F* just in case *S* is ignorant of *F* and

- (CI1) *S* has done something (put an obstacle in his or her epistemic path: e.g. taking an amnesia pill) or
- (CI2) *S* has failed to do something (investigate in a certain way, make an inference), and
- (CI3) *S* ought to have behaved differently with regard to (CI1) and/or (CI2), and
- (CI4) *S*'s failure to behave differently with regard to (CI1) or (CI2) produced *S*'s ignorance of *F*

This is more of a template of a view than an actual view, since there still is no account as to when (CI3) will obtain.<sup>17</sup> As noted above, it seems clear that the Dog-sitter ought to have behaved differently with regard to his belief regulation, and it seems that the President was doing fine. What about the Ancient Slaveholder?

Rosen suggests that in the case of the Ancient Slaveholder, (CI3) is not true—it is not true that the Ancient Slaveholder ought to have behaved differently with regard to the relevant (CI1) and (CI2). The basic idea is this: sometimes people are ignorant that they are ignorant about some fact. If they don't know that they are ignorant, they can't reasonably be expected to do anything about their ignorance. So, too, if they don't know that they don't know that they are ignorant—they can't reasonably be expected to do anything about that ignorance, either. One sees the beginning of a problem. Later in the paper, I will for the most part leave this particular problem aside, focusing on cases in which people do know that they are ignorant, or that they are unsure of something.<sup>18</sup> In these cases, there is no problem of this sort with regard to (CI3). The counterexample I offer below is a case of this sort. Still, there is an interesting question of what we should think about cases like the Ancient Slaveholder, which involve an agent who is in an (I1) or (I2) state of ignorance, having either never thought about the issue, or having come to have false beliefs about the issue.

My intuitions are at odds with Rosen's with regard to the Ancient Slaveholder case and other similar cases: it seems that the Ancient Slaveholder is morally culpable for his slave-owning. In the following sections I will examine that intuition

<sup>15</sup> Rosen, p. 63.

<sup>16</sup> Rosen, p. 63.

<sup>17</sup> This view is similar to one that is implicit in Smith (1983), in which she sets out three main types of cases in which a person can be culpably ignorant. *Id.* at 544–545. On her view, one can be culpable for conducting deficient investigation, preventing subsequent discovery, and/or making deficient inferences.

<sup>18</sup> Some may feel that this is to leave aside the truly compelling issue. I disagree, and explain my disagreement in Part Five below.

further. One thing that is difficult to determine is whether my differing intuition is really an intuition about culpability regarding the *action*, such that the Ancient Slaveholder is still culpable, even though blamelessly ignorant, or whether my differing intuition is really about whether the Ancient Slaveholder is culpable for his ignorance. One thing we will see is that these options are not as distinct as they might appear to be.

### 1.6 Is the Ancient Slaveholder culpably ignorant?

Rosen takes the Ancient Slaveholder's ignorance to be non-culpable. Implicit in this is the view that there is nothing that the Ancient Slaveholder should have done differently from an epistemic point of view. Of course, we might want the Ancient Slaveholder to have investigated the moral question more thoroughly, or at all, but the thought is that it would be unreasonable to think that the Ancient Slaveholder had an *obligation* to do so.

Rosen offers an extended discussion of blameless ignorance, which can be reconstructed into something like the following:

**BI** A person *S* is blamelessly ignorant of a fact *F* just in case:

- (BI1) *S* never thought about whether *F* was the case,
  - (a) it was reasonable for *S* not to think about this,
  - (b) and *S* remained ignorant of *F*, or
- (BI2) *S* did think about whether *F* was the case, and
  - (a) *S* didn't intentionally commit 'mistakes' in reasoning or sabotage her attempts at discovery,
  - (b) *S* made a reasonable effort to figure out whether *F* was the case,
  - (c) and *S* still remained ignorant of *F*.

Note that on this account of when someone is blamelessly ignorant, an individual can be blamelessly ignorant even if she never has thought about the issue in question—perhaps we are to imagine the Ancient Slaveholder is in this situation. As with the account of culpable ignorance offered above, essential details must still be filled in. In particular, it is unclear how we are to determine whether a person was *reasonable* in not thinking about an issue, or whether a person made a *reasonable* effort to figure out the moral facts in the time available. Thus, on this account, though it seems that on any plausible construal of the 'reasonability requirement', the President did make a reasonable effort, and the Dogsitter did not, it appears to be an open question whether the Ancient Slaveholder was reasonable in not thinking about the issue. Again, without identifying the obligations that govern belief regulation in these cases, we will be unable to make a clear case that the Ancient Slaveholder either was or was not reasonable in not thinking about the issue. If he violated an obligation, then he was not reasonable; if he did not violate any obligations, then we can presume that he was reasonable. Let us try to say more about what these obligations might be, since they figure in crucially to our assessment of whether an individual is culpable for his ignorance.

First, we must determine whether the obligations that we are concerned with are epistemic or moral obligations. I agree with Rosen<sup>19</sup> that these are best seen as moral obligations governing belief regulation, so that we have *moral* reasons to consider evidence, to inquire further into some question, to infer carefully, to keep our epistemic path to some fact unimpeded. Of course, we also can have epistemic reasons to consider evidence, to inquire further, and so on. But cases like the Dog-sitter and the Ancient Slaveholder bring out the fact that we can also have moral reasons and moral obligations to engage in certain sorts of epistemic investigations.

The thought here is that when making certain sorts of decisions, or preparing to take certain sorts of actions (or to refrain from taking certain actions), what is required of us from an epistemic point of view may be sensitive to what is at stake from a moral point of view. Imagine two different people, each of whom is attempting to assess whether there are any people inside an abandoned house. The first person, Jack, is attempting to determine whether anyone is inside because he is trying to determine how many people live in the town. The second person, Jill, is attempting to determine whether anyone is inside because she is charged with demolishing the house. It seems clear that Jill's investigative obligations are greater than Jack's, and that this difference turns not on anything about their epistemic position, but rather on the nature of the thing that they are contemplating doing, and the beliefs required to make that activity morally permissible. The obligations regarding epistemic investigation go up, precisely because of what is at stake, morally. Thus, we might require Jill to check in every room, perhaps even twice, in order for her to count as having made a reasonable effort, while we might require Jack only to knock on the door and perhaps peek his head inside and yell a few times in order for him to count as having made a reasonable effort.

### 1.7 Epistemic obligations and moral epistemic contextualism

These considerations echo variants of epistemic contextualism that hold that the standards for knowledge shift depending on the practical context. On such theories, "the truth conditions of sentences of the form 'S knows that p' or 'S does not know that p' vary in certain ways according to the context in which the sentences are uttered."<sup>20</sup> Of course, there are many ways in which contexts might vary. One that is explicitly contemplated by Keith DeRose and others is "the importance of being right" in a particular context.<sup>21</sup> The cases that DeRose considers are cases in which it is pragmatically important to be right, rather than morally important to be right, but we could imagine a view on which whether one knows that *p* varies depending on the moral importance of being right about *p* in some particular context. I am not interested in defending any such view here. My primary concern is not with whether it is true that 'S knows that p' in some context (whether from the point of view of speaker or evaluator). The view I am considering is one on which how much one must do epistemically in evaluating whether to believe some proposition *p* varies

<sup>19</sup> Rosen, p. 63.

<sup>20</sup> DeRose (1992), p. 914.

<sup>21</sup> DeRose, p. 914.

depending on what actions (or omissions) a belief in  $p$  will justify or permit. The more morally significant the actions that a belief in  $p$  (or absence of a belief in  $p$ ) will support or license, the more stringent the epistemic demands that must be met before one can act as if one is justified in believing that  $p$ . Importantly, this 'increase' in the epistemic demands is required by moral considerations, not epistemic ones. We might also think that what is at stake, morally, in believing  $p$  actually alters when one is justified in believing  $p$ , but I am primarily concerned just with the question of when one is justified in acting *as if* one is justified in believing  $p$ .

Let us call this view *moral epistemic contextualism*:

How much one is morally required to do from an epistemic point of view with regard to investigating some proposition  $p$  varies depending on the moral context—on what actions one's belief in  $p$  (or absence of belief in  $p$ ) will license or be used to justify, morally, in some particular context.

Importantly, it is contextualism not about when it is true to say that some individual  $S$  knows or is justified in believing something or not, or whether that individual can justifiably assert  $p$ , but rather when it is morally appropriate for  $S$  to *act as if*  $S$  possesses justified true belief or certainty with regard to some issue. What is at issue is whether that individual can (morally) permissibly *act as if* they are justified in believing  $p$ , given what they have done to evaluate and consider their belief in  $p$ . Contexts on this view are determined by what actions the individual under consideration is actually contemplating, and in particular the moral significance of the actions being contemplated. This is rough, admittedly, but it should suggest the outline of the view.

It might be natural to feel that moral epistemic contextualism of the sort I endorse above leads naturally to what we might call 'full-blown' moral epistemic contextualism:

What one must do from an epistemic point of view (what one must investigate, rule out, etc.) to *count as knowing or justifiably believing* some proposition  $p$  varies depending on the moral context—on what actions one's belief in  $p$  (or absence of belief in  $p$ ) will license or be used to justify, morally, in some particular context.

But this sort of view risks confusing the *kind* of obligations that are altered when the moral stakes are higher; it is to confuse *moral* obligations to do more (epistemically), with *epistemic* obligations to do more (epistemically). In particular, it's not obvious that the standards for what counts as belief-justifying evidence should be seen to shift from moral context to moral context. This only seems true if we accept further claims about the moral permissibility of acting on all beliefs that we justifiably hold. Of course, some such claims will look fairly plausible.

Let me try to make the dialectical situation here more transparent. Consider a case in which our previous individual, Jill, is attempting to determine whether anyone is inside an abandoned house because she is charged with demolishing the house, and all she does is to knock on the door, poke her head in, and yell a few times. There are two ways to interpret such a case.

### Interpretation One:

Jill is epistemically justified in believing that there is no one inside the house on the basis of her evidence, but, given what she has done to investigate, Jill should not take certain actions based on the belief that there is no one inside the house—particularly, she shouldn't take any actions that would be morally impermissible if her belief were false. In this case, we might say that Jill is justified in believing  $p$  (that there is no one inside the house), but that  $p$  is 'unactionable' for Jill—she may not take actions that are morally permissible only if  $p$  is true.

### Interpretation Two:

Jill is simply not epistemically justified in believing that there is no one inside the house on the basis of her evidence, because the standards for justifiably believing that there is no one inside the house are higher, given Jill's plan to demolish the house if it is empty. This is true even though Jack, who is only trying to figure out how many people live in the town, *is* justified in believing that there is no one inside the house on the basis of the exact same evidence that Jill has.

Ultimately, there is not much difference between Interpretation One and Interpretation Two. The first interpretation can seem appealing because it doesn't make the facts about epistemic justification sensitive to moral or practical considerations, which might seem misguided. The second interpretation is appealing because we don't have to cordon off some of Jill's beliefs as 'unactionable'—a funny category that would need further elaboration. In particular, we'd then want an account of what we can do with unactionable beliefs—can we make inferences based on them, can we take actions based on them the moral permissibility of which does not turn on whether they are true, and so on. Depending on our answers to these questions, unactionable beliefs may start to not look very much like beliefs at all. Then again, unactionable beliefs may not be as unusual as they sound. We might think that legal rules of evidence are based on similar considerations—there might be cases in which an agent is justified in believing  $p$  as a result of hearsay, but we don't think that a belief formed on these grounds should be used as grounds on which to punish someone. The truth of full-blown moral epistemic contextualism is worth considering, but I don't want to pursue it further here. Instead, I will just endorse the more modest moral epistemic contextualism, the view that how much one is morally required to do from an epistemic point of view with regard to investigating some proposition  $p$  varies depending on the moral context—on what actions one's belief in  $p$  (or absence of belief in  $p$ ) will license or be used to justify, morally, in some particular context.

Given moral epistemic contextualism, we now have at least some additional guidance regarding the relevant epistemic obligations and the bounds of the reasonableness requirements mentioned in the account of blameless ignorance above. In particular, what is required of us from an epistemic point of view increases as what is at stake from a moral point of view increases: our epistemic obligations grow as the relevant context becomes more morally serious.

What should we think of the Ancient Slaveholder in light of a view like moral epistemic contextualism? Our discussion will vary depending on whether we see the Ancient Slaveholder as a (BI1) case or a (BI2) case—a case in which he hasn't thought about the issue, or a case in which he has, but has come to have false beliefs about it. I think that Rosen treats the case as a (BI1) case, but let us discuss both, since we could imagine both kinds of cases.

### 1.8 Culpable ignorance in cases of false belief

Let us start by treating the Ancient Slaveholder case as a (BI2) case—a case in which the Ancient Slaveholder has thought about the issue, but has come to have false beliefs about the issue. When assessing whether it was reasonable for the Ancient Slaveholder to have come to this conclusion regarding the moral wrongness of slavery, we must consider what epistemic standards are appropriate, what epistemic obligations apply. To answer the question of which epistemic obligations apply, we must consider the moral significance of the actions that will be licensed by a positive or negative verdict regarding the moral wrongness of slavery. Obviously, the moral significance of these actions is quite high. Treating another human being as mere property, restricting their freedom, punishing them with physical violence for failing to do as instructed—all of these actions are only permissible if one has certain views regarding the moral status and moral worth of persons. So the epistemic obligations regarding this investigation are quite high. Of course, this still doesn't tell us in detail what they are. But it does make it seem implausible that the Ancient Slaveholder has met them, given the fact that it doesn't seem particularly complicated to figure out that slavery is morally wrong, and given the fact that the slaveholder has come to have a false belief about the wrongness of slavery. Though it may be some sort of failure to imagine my way into the Ancient Hittite mindset, it seems implausible to say that it would take a “moral genius” to see through the wrongness of chattel slavery.

Here, Rosen might disagree, perhaps using the fact that contemporaries of the Ancient Slaveholder didn't come to the correct belief about the wrongness of slavery to argue that it is harder than it might seem to figure out that slavery is wrong—perhaps our modern mindset obscures that fact. In response to this, we might consider arguments made by Michelle Moody-Adams, who cites empirical evidence suggesting that the view that slavery was permissible wasn't as widely held as Rosen and Slote would have us believe, noting that Aristotle argued against slavery in the *Politics*, and arguing that most of our evidence for the view that slavery was not morally questioned by the ancient Greeks comes from texts produced by those who had an interest in maintaining the institution of slavery.<sup>22</sup> We might also consider the point that ‘innocence by association’ is not obviously a more compelling principle of culpability than guilt by association, at least not without further elaboration. In both cases, it seems better to evaluate what the individual in question has actually done or failed to do, rather than to merely observe the groups to which the individual belongs. Given the relatively simple and

<sup>22</sup> Moody-Adams, pp. 295–296.

apparently a priori nature of the reasoning required to figure out that slavery is wrong, it is natural to think that engaging in this reasoning wouldn't be terribly different for any particular individual, even 2500 years ago. It is not clear that being part of a group most members of which accept some proposition  $p$  thereby makes it more difficult to correctly evaluate whether  $p$  is true, and it is certainly not clear that this is the case for all propositions. It might be the case for some propositions, but it will depend on the details, and an argument to this effect would be required in the case of the proposition that slavery is morally wrong. It might be possible to make such an argument: views of human nature, our place in the cosmos, the nature of equality, and so on, might have been sufficiently different and sufficiently well settled to make getting the right answer difficult, even if one lives up to the high epistemic obligations. But one would actually have to show this; it is not enough to simply note that most people in fact had the wrong view. I won't attempt to settle this argument here, but I do hope to have made plausible the view that the epistemic obligations—what is required of the Ancient Slaveholder in order to meet the (BI2)(b) reasonability requirement—would be demanding. This follows from moral epistemic contextualism. Given that the epistemic obligations would be demanding, and given that determining that chattel slavery is morally wrong doesn't seem that difficult, it is hard to see how the Ancient Slaveholder could be non-culpably ignorant.

Still, whether the Ancient Slaveholder might be non-culpably ignorant if construed as a (BI2) case doesn't directly relate to whether the MIT is true or false. I've considered the case in detail here because Rosen relies on it, and on cases like it, to build support for the MIT. But even in these cases, one can question whether they really support the MIT because one can reasonably doubt that the Ancient Slaveholder is non-culpably ignorant.

### 1.9 Culpability for ignorance in cases of no belief or only implicit belief

Let us now consider the Ancient Slaveholder case as a (BI1) case—a case in which the individual has not thought about the fact in question. In making the case for moral epistemic contextualism, I relied on the Jack and Jill cases, both of which are (BI2) cases—cases in which the individual is thinking (or has thought) about the fact in question. What does moral epistemic contextualism mean for (BI1) cases? The answer is interesting, and perhaps surprising. Recall that in these cases,

**BI** A person  $S$  is blamelessly ignorant of a fact  $F$  just in case:

- (BI1)  $S$  never thought about whether  $F$  was the case,  
 (a) it was reasonable for  $S$  not to think about this,  
 (b) and  $S$  remained ignorant of  $F$ .

The interesting feature is this: because of the nature of the obligations that govern whether it was reasonable, we cannot answer the question of whether it was reasonable for a person to have ignored some issue until we know what actions that person was engaged in or contemplating. Consider, for example, a third case

involving abandoned houses. In this case a person, Brandon, is having a block party, and is inviting people who live in his neighborhood. He has never thought about whether anyone lives in the abandoned house, which is in his neighborhood. This is a clear case in which it seems reasonable for Brandon not to have thought about whether anyone lives in the abandoned house, given that what he was doing was inviting people to a party. On the other hand, it would obviously be unreasonable for Jill never to have thought about whether anyone lives in the abandoned house, given that she was going to demolish it.

Let me introduce a piece of terminology: call any state of affairs which, if it obtained, would make it morally impermissible to perform some action A, a 'blocker' with respect to A. For example, a blocker with respect to keeping slaves is the state of affairs of all human beings having the right to self-determination. The explanation of why Jill cannot reasonably ignore the issue of whether someone is in the abandoned house is because the state of affairs of there being someone in the house is a blocker for the action that Jill is contemplating. This is not true for the action that Brandon is contemplating, and this is why Brandon can reasonably ignore the issue of whether someone is in the abandoned house. What moral epistemic contextualism tells us about (B1) cases is that it will generally be unreasonable to fail to investigate whether the blockers for the actions (or intentional omissions) that one is performing obtain, at least for certain 'obvious' blockers for the particular action (or intentional omission) one is contemplating.<sup>23</sup>

Let us return to the Ancient Slaveholder case, treating it this time as a (B1) case—supposing that the Ancient Slaveholder simply never thought about whether slavery was morally wrong. Was it reasonable for the Ancient Slaveholder never to have thought about this? Well, given the moral seriousness of keeping slaves (and doing all that is involved in that activity), it seems that it will not be reasonable for the Ancient Slaveholder not to have thought about this, *because* of the moral seriousness of keeping slaves. Keeping slaves requires forcibly restraining them, forcing them to work, punishing and disciplining them, separating them from familial and other important life relationships when necessary—and these actions all will often take on a very brutal character. These actions all have clear and obvious blockers. It is hard to imagine someone brutally beating a slave without at least having thought about whether what he does is morally permissible. And if he did fail to think about this, this would not be a reason to see him as non-culpable—quite the contrary. This is a case in which one is doing or contemplating doing something for which there are in fact many blockers, and so the investigative demands—into both non-moral factual questions and relevant moral questions that might be the source of blockers—are heightened. Interestingly, it seems that the increased epistemic demands as a result of the moral seriousness of the action are not sensitive

<sup>23</sup> This terminology affords us another way to define moral epistemic contextualism: How much one is morally required to do from an epistemic point of view with regard to investigating some proposition *p* varies depending on the moral context—on whether *p* describes a state of affairs that is a blocker for actions that the person is contemplating performing in that particular context. If *p* describes a state of affairs that is a blocker for actions that the person is contemplating performing, then one is morally required to do more from an epistemic point of view with regard to investigating whether *p* is true, whether the relevant state of affairs obtains.

to whether the person in question was *aware* of the moral seriousness of the action. I'll have more to say about this in a moment.

In cases in which someone is contemplating doing something that has few blockers, the investigative demands remain relatively low, even in (BII) cases. Thus, if some other Ancient Hittite (call him the Ancient Listmaker) didn't keep any slaves, but was only, say, making a list of all the things that were morally wrong for his amusement, to be sealed in a vault and never looked at, it seems that it would not be unreasonable for this Ancient Listmaker never to have considered the question of whether slavery was morally wrong. This Ancient Listmaker would not have strict epistemic obligations to consider the question of whether slavery is morally permissible, since this is not a blocker for the action he is contemplating.

This brings out another feature of moral epistemic contextualism in moral ignorance cases: plausibly, one of the things one must almost always think about when one is doing or contemplating any particular action is whether there are blockers for the actions one is considering taking, and whether those blockers obtain. In these cases, it will generally be unreasonable for a person simply not to have thought about whether there are any blockers for the action they are contemplating, or whether those blockers obtain, at least not if the fact that they were ignorant of the moral wrongness of what they have done is supposed to be an excuse.

This might seem like a strong result, but I think that it is plausible. After all, what kind of excuse is it to say, after having done something morally impermissible, that one simply never thought about whether what one did was morally permissible or not? There are a few kinds of cases in which this might seem reasonable. They generally have the following key feature: the connection between (a) what one does and (b) the facts that make certain blockers obtain is not at all obvious, and would not be made obvious even if one meets very demanding epistemic obligations regarding investigation, etc.

Consider the following case. Clint has a kitchen with a fluorescent overhead light. This light can be turned on by either of two switches—a switch on the western side of the room, and a switch on the eastern side of the room. When Clint turns on the light using the western switch, the circuit works as usual, turning on the light. When Clint turns on the light using the eastern switch, the light turns on, but turning the light on with the eastern switch also causes two people in New York City to suffer heart attacks. The causal route by which this happens is completely obscure.

What should we think about cases like the one Clint is in? Though his decision to turn on the light using the eastern switch, rather than the western one, clearly has moral significance, it does seem that it is reasonable for him never to have thought about whether it is morally permissible to turn on the light using the eastern switch. These kinds of cases are why it is not *always* the case that, to be non-culpably ignorant, one must have thought about whether it was morally permissible to do what one was doing. Moral epistemic contextualism suggests that in certain moral contexts, the epistemic obligations to investigate increase significantly. Still, there are limits to how much one must investigate, and what one must rule out, even in these cases. This is obvious in the case of the President. It is also clear in cases like the one Clint is in. Happily, most of us are not in situations like the one Clint is in.

Returning to the Ancient Slaveholder case, interpreted as a (BI1) case in which he has never thought about the issue of whether slavery is wrong, the most natural view, given moral epistemic contextualism, is to see the Ancient Slaveholder as culpable for his ignorance. The Ancient Slaveholder had epistemic obligations, stemming from moral considerations, to consider whether what he did was morally wrong or not before doing it.

A contrary view might see the Ancient Slaveholder case as more like the Light Switch case, holding that the connection between keeping slaves and the facts that make blockers with respect to keeping slaves obtain is not at all obvious. This seems implausible. Alternatively, there might be some other grounds on which to see the Ancient Slaveholder in the (BI1) case as blamelessly ignorant. I don't know what such an account might look like. Any such account will have to contend with our intuitions in cases like Jill's, in which she seems clearly culpable for not even thinking about the issue of whether there might be someone inside the house she is about to demolish. One thought that might distinguish Jill's case from the Ancient Slaveholder's is something like this: when taking certain sorts of actions, it will just be relatively obvious that one must investigate in certain ways in order to live up to the 'reasonableness' epistemic obligations; it will just be obvious what the relevant blockers are. Thus, when destroying a building, one must investigate whether anyone is inside. One might think that this is not true when doing things like keeping slaves—there is no similarly obvious set of blockers to consider. This just seems false. It is true that cases such as these that involve investigating moral facts can strike some as peculiar, and there is something to this. In particular, it is somewhat less obvious what sort of investigations must be made. Still, the fact that it seems peculiar to describe the Ancient Slaveholder situation in this way isn't because it is odd to require someone to consider whether what they are doing is morally permitted; it is because this is so obviously a requirement that we all are under.

### 1.10 Summing up

Ultimately, whether we see the Ancient Slaveholder case as a (BI1) case or a (BI2) case, the most plausible view is that he is culpable for his ignorance. This view relies on the truth of moral epistemic contextualism, and there is much more that could be said about that view. Still, one who wants to see the Ancient Slaveholder as blamelessly ignorant must affirmatively make the case that the Ancient Slaveholder has met his epistemic obligations in investigating the question of whether slavery is morally wrong. I think that this will be hard to do.

The upshot of this discussion is twofold. First, if part of the support for the MIT is supposed to stem from cases like the Ancient Slaveholder, this support appears to be significantly weakened, given that cases of this sort appear to involve culpable ignorance, and thus are irrelevant to assessment of the MIT. This is not insignificant. Rosen takes the fact that the Ancient Slaveholder case is an instance of non-culpable moral ignorance to be "beyond dispute,"<sup>24</sup> and the bulk of the positive argument that Rosen and others make for the MIT consists in reliance on cases of this sort. Second,

---

<sup>24</sup> Rosen, p. 64.

from moral epistemic contextualism we learn that whether someone is culpably ignorant or not appears to be a function, in part, of what the person was going to do based on his or her false belief (or lack of belief). There will be cases in which two people will both be ignorant of some fact, *F*, they will both have done the exact same thing to investigate *F*, and one, but not the other, will be culpable for his ignorance, simply because one, and not the other, proceeded to act in a morally significant way from that ignorance. This is what makes the difference between the Ancient Slaveholder and the Ancient Listmaker, such that only the former is culpable for his ignorance.

This is the sense in which the two questions—(1) is the Ancient Slaveholder culpable for his ignorance; and (2) even if the answer to (1) is no, is the Ancient Slaveholder culpable for acting in the way that he did anyway—are harder to keep distinct than it might have seemed. Whether one is culpable or not for one's ignorance depends in part on what actions one is going to take.

Returning to our earlier cases, we can see that the Dog sitter is culpable for his ignorance because he had an obligation to investigate further into whether certain obvious blockers to the action he was contemplating obtained. We can also see that the President is not culpable for her ignorance, since, though there were blockers to the action she was contemplating, she made a thorough and careful attempt to investigate whether these blockers obtained or not—it would be unreasonable to expect her to have done more.

Let us move now to consider what I take to be a counterexample to the MIT, a case in which a person is blamelessly ignorant of a moral fact, but aware that they were ignorant, and so culpable for acting anyway (or so I will argue). The culpability stems not from them being ignorant or failing to live up to any epistemic obligations, but from them *doing* something, given that they were ignorant.

## 2 A counterexample to the Moral Ignorance Thesis

Consider the following case: Douglas is contemplating killing a pig for dinner (imagine that the pig is currently still alive and healthy). Though not thinking in these explicit terms, he is able to understand the idea of moral status, and is questioning whether or not the pig has significant moral status—moral status such that it would be wrong to kill the pig in order to eat it. If he knew the pig did have significant moral status of this sort, Douglas would not kill it. After some effort, he is still unable to come to an answer to this question—he doesn't know whether the pig has significant moral status. His reasons for wanting to kill and eat the pig are straightforward: he likes the taste. There are plenty of other food options that he knows are morally permissible, which would be nutritious and which he would enjoy eating, and which would not require killing pigs or any other animals. Still unable to decide the question, he goes ahead and decides to kill and eat the pig, and does so successfully. For ease of exposition, let us stipulate that pigs in fact have significant moral status, such that it is wrong to kill them in order to eat them.

Given these assumptions, we can now ask about Douglas's conduct, with an eye toward evaluating the case as a possible counterexample to the MIT. Our assessment of this purported counterexample to the MIT will consider three

questions: (1) was he acting from ignorance; if so, (2) was he culpable for his ignorance; and, finally, (3) is he culpable for his conduct?

### 2.1 Acting from ignorance

Recalling our earlier discussion, it seems clear that he *was* acting from ignorance. He performed an action *A* (killing the pig) while lacking belief in some proposition *p* (that pigs have significant moral status), and he would not have done *A* had he believed that *p*. Recall that in cases in which the relevant proposition *p* is true, the person will be acting from ignorance. In cases in which the relevant proposition *p* is false, the person will not actually be acting from ignorance (from lack of true belief), but their action is explained by the fact that they lack some particular false belief. Thus, if it is true that pigs have significant moral status, Douglas is acting from ignorance; if it is false that pigs have significant moral status, Douglas is still acting from absence of false belief (acting only because he lacks the false belief that pigs have significant moral status).<sup>25</sup>

### 2.2 Culpable for ignorance

It seems clear that he was *not* culpable for his ignorance regarding the moral status of the pig. He did think about whether moral issues might be involved, he considered the relevant blockers, he didn't intentionally commit mistakes in reasoning or sabotage his attempts at discovery—in short, he made a reasonable effort to figure out the relevant moral facts, and he still remained ignorant of the moral facts. His effort, we can suppose, was reasonable, even given the heightened epistemic obligations required in cases of this sort, given the action he was contemplating. Unlike the case of chattel slavery, figuring out whether pigs have significant moral status, such that it would be wrong to kill them for food, is not an easy moral question. This case looks more like the President case than either the Dogsitter or the Ancient Slaveholder case in this regard.

### 2.3 Culpable for acting

This issue is harder to assess than the previous two, at least in part because we have fewer guidelines to help us determine whether Douglas is culpable for doing what was morally wrong. I find it natural to think that he is culpable for killing the pig. He knew that it would be wrong to kill the pig if it turned out to have significant moral status, he knew that he didn't know whether the pig had this moral status, there weren't any

<sup>25</sup> Cases in which someone acts objectively permissibly but is reckless in doing so (because they know they are ignorant of what the relevant moral requirements are) are cases in which a person is not actually acting *from* ignorance at all, though they are acting *in* ignorance. Their ignorance does not cause or explain their action. I think that including such cases as counterexamples to the MIT is unfair to the proponent of the MIT. Individuals acting in this way do something for which they are culpable, and their culpability stems from their behaving recklessly, but I think the most natural understanding of the source of their moral error is that it comes from their behaving recklessly with regard to their epistemic obligations, from their being culpably ignorant, in precisely the way described earlier when discussing the (B11) cases. As noted above, what moral epistemic contextualism tells us about (B11) cases is that it will generally be unreasonable to fail to investigate whether the blockers for the actions that one is performing obtain. This remains the case even if the blockers turn out not to obtain in the particular case.

particularly compelling reasons for him to eat the pig,<sup>26</sup> and he went ahead and killed it anyway. I think that there is a compelling moral principle behind this intuition; in a moment I will present and defend that principle. If I am successful, this will be a justification for seeing Douglas as culpable for going ahead with killing the pig, which in turn will reinforce seeing the above example as a counterexample to the MIT.

It is important to note that the question of whether or not Douglas is culpable for killing the pig doesn't turn on whether pigs in fact turn out to have significant moral status. One can be culpable for behaving in ways that are morally risky even if it turns out, fortunately, that what one did actually didn't put anything of moral significance at risk. This is because behavior can be morally risky in at least two different ways:

- (R1) because one performed an action A (where A can be an omission) that in fact risked harming something of moral status, or
- (R2) because one performed an action A that might have risked harming something of moral status, for all one knew or believed.

Cases of the (R1) variety might include driving recklessly on a busy street, dropping bombs on occupied territories, and engaging in very violent contact sports. Cases of the (R2) variety include demolishing buildings without investigating their contents, killing alien life forms before engaging with or investigating them, or getting an abortion without having considered whether the fetus had moral status. I think it is clear that even in the (R2) kinds of cases, a person can be culpable for his behavior, even if it turns out that there was little or not actual risk of moral harm. Additionally, though all the (R2) cases I mention involve actions taken from what appears to be culpable ignorance, it is possible to imagine variants of these cases such that the individual acting has thought about the issue extensively, remains uncertain, and acts anyway. In these cases, which look like Douglas's case if it turns out that pigs do not in fact have moral status, it seems that the individuals still would be culpable for acting.

### 3 The principle: Don't Know, Don't Kill

Here is the principle that Douglas violates:

**Don't Know, Don't Kill (DKDK)** If someone knows that she doesn't know whether a living organism has significant moral status or not, it is morally

<sup>26</sup> This isn't to say that gustatory pleasure won't ever be relevant, or perhaps even morally relevant, if only minimally. And it is true that the reasoning that applies in this case might well apply in every case, so that Douglas would never be able to eat pigs or maybe any meat at all. There are several things to say here. First, in this particular case, Douglas would enjoy eating other things as well, things the moral status of which he was certain. We might imagine another case in which this weren't true, though. In such a case, I still think that one does something for which one is blameworthy if killing the organism has higher expected disvalue than not killing it, which will only be true if the moral value of eating the organism is relatively low. Perhaps at some point (and depending on one's account of what is of moral value), the enjoyment of eating some particular organism would rise to a level of moral significance such that one would not violate any moral principle (such as the one I'll offer in a moment) in eating an organism the moral status of which one was uncertain. But I don't think that will be the ordinary case, at least not on a plausible account of what is of moral value.

blameworthy for her to kill that organism or to have it killed, unless she believes that there is something of substantial moral significance compelling her to do so.

It is my view that one does something for which one is culpable, morally blameworthy, when one kills a living organism and violates DKDK in doing so—and this is true even if the living organism turns out not to have significant moral status. It is clear that Douglas violates DKDK. He knows that he doesn't know whether the pig has significant moral status or not, and he doesn't believe that there is anything of substantial moral significance compelling him to kill the pig. In the rest of this section, I will attempt to clarify DKDK and to suggest some reasons to think that DKDK is true. In the following section, I will consider several objections to DKDK. In the final section, I will conclude by considering a number of natural applications or extensions of DKDK.

### 3.1 Clarifying DKDK

In order to more ably consider whether DKDK is true, let me say more about its component parts. We can divide these up in the following ways:

**Epistemic state component:** one knows that one doesn't know whether *F*

**Fact *F* component:** that some particular living organism has moral status

**Action component:** to kill an organism or have it killed

**Justification component:** one believes that there is something of substantial moral significance compelling one to perform the action

Most simply, DKDK is a principle about what we can non-culpably do in a specific kind of case of uncertainty. The epistemic state component and fact component together define the nature of the uncertainty. Note here that we might helpfully introduce talk of degrees of belief to make the precise nature of the uncertainty apparent.<sup>27</sup> Thus (employing the standard '0 = no belief to 1 = certainty' representation), we might change the epistemic state component to 'one believes to degree 1.0 that one believes *F* to degree 0.5,' for example. We could even remove the 'second order' component as formulated in DKDK, such that it could read: 'If one's belief in *F* falls in the 0.35–0.65 range, it is wrong for her to kill that organism or to have it killed, unless she believes that there is something of substantial moral significance compelling her to do so.' Not wanting to introduce precision where it seems unmotivated, I prefer the original DKDK formulation, but it is important to note that we could in principle increase the precision in the way indicated. This may become important later if one is interested in relaxing DKDK.

The fact and action components are relatively straightforward. Obviously, the fact component and action component will be related. In particular, for DKDK-type principles, the issue in question in the fact component will typically be a possible source of blockers for the action described in the action component.

<sup>27</sup> Except insofar as we struggle to come up with a compelling account of what talk of degrees of beliefs amounts to. See, for example, Eriksson and Hájek (2007).

The justification component is somewhat less straightforward. DKDK is a caution principle; it requires one to exercise caution before performing certain actions when one is in certain epistemic states, unless one has some justification for going ahead and acting anyway. The justification required in DKDK is that the person believes that there is something of substantial moral significance compelling her to act. This has several parts.

First, what is relevant for DKDK is just that the person in the situation *believes* that there is something of substantial moral significance compelling them to act. We could understand this in a more or less objective fashion. In the more objective fashion, we define, objectively, what is of substantial moral significance, and then ask whether the person in the situation involved believed that one of those things was involved in the situation. In the more subjective fashion, we simply take the person in the situation's view regarding what is of substantial moral significance as the one by which she is judged—if she reasonably believes that losing one of her 20 cars is of substantial moral significance, then she is off the hook as long as she reasonably believes that she has to kill the organism to prevent losing one of her cars. The reasonableness standards here will be the same epistemic obligations discussed earlier when introducing moral epistemic contextualism.

There are problems with both routes. On the more subjective route, we threaten to allow people off the hook for what appear to us to be trivial reasons. On the more objective route, it seems that we will be requiring individuals not just to do what is epistemically required of them in thinking about what is of substantial moral significance, but to actually *get it right*, which can seem like too high a bar. Basically, this is the problem of moral ignorance rearing its head again. To meet this prong of the justification component, all that is required is that the person is non-culpably ignorant—they must have made a reasonable effort to think about whether the thing apparently compelling them to act actually *was* of substantial moral significance. If they believe that it is, after having met their epistemic obligations, then that is sufficient to meet that part of the justification component.

This may seem to make DKDK too lenient. It is useful to recall that we are only concerned with moral culpability; it remains an open question whether what an individual *objectively* wrong or not. Consider, for example, the way that what the President does is objectively wrong, even if she is not culpable for doing it. Additionally, there will still be cases in which the person in the situation doesn't even believe that there are substantial moral considerations compelling her to kill the organism. Importantly, our case involving Douglas is of this sort. He wants to eat the pig, but he doesn't believe that anything of substantial moral significance compels him to do so. Of course, it is not that gustatory and other kinds of aesthetic value are of *no* value. In this case, it is just that their value is not sufficiently substantial, nor does he believe it to be.

Second, the person must believe that there is something of substantial *moral* significance compelling them to act. It is not enough that the person feel that there are non-moral considerations in favor of acting. The thought here is that what is at stake in killing the organism is something that may be of substantial moral value. It seems inappropriate to let something that is not of moral value trump something that is (or might be), and so inappropriate to introduce non-moral considerations. What

are moral considerations? It is outside the scope of this paper to give a full account, but it seems clear that they would include things such as the harm (physical and emotional) that might be caused or prevented, deaths that might be caused or prevented, ways in which people might be treated, and so on. It's clear that many things will count as in some way intertwined with moral considerations; the real discriminating work will be done with the requirement of *substantiality*. The fact that someone will lose one of their twenty cars, or will have to stub their toe, or have to walk 10 miles in the sun—all of these are arguably of *some* moral significance, just not substantial moral significance. This is obviously not a bright line. The line can be made more precise, however, by considering the relevant notion of 'compulsion' in the DKDK justification component. Let us turn to that now.

The final part of the justification component is that the person must believe that there is something of substantial moral significance *compelling* them to act. There are different ways that we might interpret this. On one reading, what is required is that the person believes that failure to act in this case has higher expected moral disvalue than acting. To consider an artificial example, if Douglas believes that there is a fifty-fifty chance that pigs have significant moral status, and that 100 points of disutility attach to killing a pig with significant moral status, and 0 points of disutility to killing a pig without significant moral status, then he will believe that he is compelled to kill the pig in this sense if and only if killing the pig has expected utility of greater than 50 points. (This example assumes that utility is all that is of moral significance.) If simply weighing the expected moral value of action and non-action, respectively, one will be compelled to act just in case the expected moral value of acting is higher than not acting. The notion of compulsion at work here is something like rational compulsion: given that one is trying to act morally, one is rationally compelled to act in the way that is likely to maximize moral value. Of course, believing oneself to be compelled in this way requires a relatively precise degree of confidence in the probabilities and moral valuations of the various outcomes. This reading, though theoretically interesting, falls short in practical applicability, and it is for this reason not the reading of 'believing oneself to be compelled' that I favor.<sup>28</sup>

<sup>28</sup> In his recent book, *Moral Uncertainty and Its Consequences*, Ted Lockhart defends an even more general principle that is stronger than DKDK and, in fact, arguably entails DKDK:

PR2: In situations of moral uncertainty, I (the decision-maker) should (rationally) choose some action that has the maximum probability of being morally right.

In the binary comparison between, say, killing the pig and not killing the pig, it seems clear that PR2 would require not-killing the pig. There is some chance that killing the pig will be morally wrong, and little chance that killing the pig will be morally right (since there is nothing of substantial moral significance at stake that would require killing the pig in cases to which DKDK would apply). And this will be true for all such cases to which DKDK would apply. Any case in which there are no substantial moral considerations on the other side will always, under PR2, be cases in which one should not do the thing that risks significant moral harm (killing something that may, for all you know, have significant moral status). See Chapter Two (and particularly pp. 26–28) of *Moral Uncertainty and Its Consequences*, (Oxford University Press, 2000). I think that DKDK is a more perspicuous formulation of what is driving the intuition in these kinds of cases, and it doesn't yield many of the results which PR2 does, some of which strike me as mistaken. Additionally, Lockhart's account requires not just specific views regarding the probabilities and moral valuations of various outcomes, but also specific views regarding the likelihood that some particular moral theory is true, another reason not to prefer this account.

The reading of ‘believing oneself to be compelled’ that I favor is closer to what I take to be the colloquial reading of that phrase, such that one believes oneself compelled to perform an action A by something of moral significance just in case one takes oneself to have all-things-considered *decisive moral reason* to do A. In certain circumstances, when one has a great deal of specific information, one might have decisive moral reason to do A because one knows that the expected moral value is on the side of doing A. But this will hardly be the usual case. I think it is hard to say much more in a general way about when one takes oneself to have decisive moral reason to do something.

In general, the notion of compulsion here is nothing like a causal law. Rather, it is something like ‘the balance of moral reasons (and prudential reasons, insofar as we are morally permitted to take these into account) tips in favor, and perhaps heavily in favor, of killing the organism.’ Obviously, such a situation doesn’t literally compel one to take the action in question. A difficult to capture component of the ‘compulsion’ requirement is something like the *moral necessity* of taking this particular action. Whether or not an action is necessary is sensitive to one’s actual options, and one’s actual options are sensitive to the time at which one is assessing what one’s options are. These factors all will be crucial for assessing whether someone is reasonable in believing that they were compelled to act by something of moral significance.

DKDK requires that one believes oneself to be morally *compelled* to kill the organism, despite being ignorant of its moral status, and aware that one is ignorant. We might think that all that should be required is that one believes oneself to be morally *permitted* to kill the organism, even though one is ignorant of its moral status, and aware of this ignorance. I think that these options will actually not be as far apart as they might seem. To see this, note that for many cases in which one would seem morally permitted to kill an organism, even though aware that one is ignorant of its moral status, these will be cases in which one seems to have decisive moral reason to kill the organism. Consider, for example, a case in which one has to choose between killing an organism the moral status of which is unknown (it might have no moral status, or moral status equal to that of a human infant), and killing a human infant (the moral status of which is known). In such a case, it seems one is permitted to kill the organism whose moral status is unknown, but it seems that one will be both permitted and compelled to make this choice. I find it difficult to come up with a case in which both (a) it seems that one is morally permitted to kill an organism the moral status of which is known to be unknown, and (b) it seems that killing this organism is not morally required. If there are such cases, then to the extent that DKDK requires that one believe oneself to be morally required, this makes it a more cautious principle than an otherwise similar principle that requires only that one believe oneself to be morally permitted.<sup>29</sup>

<sup>29</sup> I find the requirement of this extra caution plausible, in part because it allows us to keep the epistemic state component subjective—requiring only that an individual reasonably believe that they are morally compelled—without putting quite as much weight on defining when such belief will be reasonable. If the principle were just that one must reasonably believe that one is morally *permitted*, almost all of the debate would center on when such a belief is reasonable. I think that the end result of this debate is likely to be extensionally similar to the result we get when we focus just on the question of whether one reasonably believes that one is morally compelled, for reasons discussed below. Additionally, the distinction between being morally permitted and being morally required only arises in this way for those moral views that are non-maximizing. Of course, many such views seem plausible, but it is worth noting this fact.

### 3.2 Reasons to believe DKDK

DKDK is a principle that requires caution in the face of a particular kind of moral uncertainty. Those who violate DKDK act recklessly, and are morally culpable for so acting, or so a proponent of DKDK would contend. I think that this is right; let me try to say a few things in support of this view.

In order to do this, it will be helpful to distinguish two different types of cases that DKDK covers, since there will be different reasons to endorse DKDK in light of these two types of cases. These two types of cases stem from the two different ways in which one can act in a morally risky way:

- (R1) because one performed an action A (where A can be an omission) that in fact risked harming something of moral status, or
- (R2) because one performed an action A (where A can be an omission) that might have risked harming something of moral status, for all one knew or believed.

DKDK will cover cases in which it turns out that the organism in question does have significant moral status, and cases in which the organism in question does not have significant moral status. In the first type of case, one will be in a (R1) situation; in fact, not only will one have performed an action that risked harming something of moral status, one will have performed an action that in fact harmed something of moral status (harmed it by killing it). In the second type of case, in which the organism that is killed does not have significant moral status, one is in a (R2) situation. As mentioned above, the reason to support DKDK will be different in these two types of cases. To endorse DKDK, however, one must endorse its application in both kinds of cases.

### 3.3 Reasons to believe DKDK in (R1) cases

It can seem somewhat odd to treat the (R1) and (R2) cases separately. In cases in which one kills the organism and it does not have significant moral status, one is in a (R2) case, not a (R1) case, since nothing with moral status was in fact at risk of harm. But that means that the only (R1) cases in which one violates DKDK will be those in which the organism killed has significant moral status, one knew this was a possibility, and in which the person doing the killing didn't even believe that there was anything of moral significance compelling them to do so. In such cases, it is clear that one has done something morally blameworthy. One has knowingly taken a risk of harming something with significant moral status, and that risk has actually materialized. I think that one does something morally blameworthy when one does something morally risky in the (R1) sense, without any compelling moral reason for doing so, even when the risk doesn't materialize. Much of the criminal and tort law regarding intentional negligence and reckless conduct seems to have such moral considerations at its base. Additionally, many moral theories—act- and rule-consequentialism, Kantian theories regarding treating persons always as ends-in-themselves, virtue theoretic accounts—would seem to support the following

principle (if not something much stronger), even if not exactly in these terms: one ought not to take (R1) risks when the expected moral value of doing so is negative.

It seems that we do have reason to accept that someone does something wrong in the (R1) cases that would violate DKDK. But perhaps it would be better to just endorse the following principle: one does something wrong if one kills an organism, which for all one knew might have significant moral status, and in which in fact turns out to have significant moral status, even though nothing of moral significance compelled one to do so. Why not just limit DKDK, or our neighboring principle, to something that just covers this more limited range of cases? After all, the hard part in defending DKDK is in defending the claim that someone does something morally blameworthy in cases in which they violate DKDK but in which it turns out, fortunately, that the organism in question in fact did not have significant moral status.

DKDK comes into the picture because it is essentially a guide for how to behave when one doesn't know whether one is in a (R1) case or not. In the cases DKDK covers, one doesn't know whether one is in a (R1) or a (R2) situation. In general it seems true that we are morally blameworthy for taking actions from ignorance that in ways that are morally risky in the (R1) way unless we believe we have compelling moral reasons for doing so. It is less widely accepted that one does something for which one is morally blameworthy when one acts in ways that are morally risky in the (R2) way in cases in which one has no compelling reasons for acting. I think that acting in this way is morally blameworthy, and that DKDK covers these cases appropriately. Let me pursue the more general defense of DKDK, bringing (R2) cases into the analysis.

### 3.4 Reasons to believe DKDK in (R1) and (R2) cases

A number of general moral theories seem that they would support a principle like DKDK. Importantly, DKDK is a principle that is concerned with how we should actually behave in conditions of a certain kind of moral uncertainty. Many moral theories are described at a level of abstraction, attempting to offer very general accounts of rightness and wrongness, and those theories might require 'translation' principles or decision procedure principles that instruct how we should act, morally, in various cases of limited information or ability. DKDK is a principle that might serve such a role for a number of theories.

For example, it seems that act-utilitarianism, if taken as a criterion of rightness, might lead us to endorse something like DKDK as a utility-maximizing decision procedure that we do better to follow than to selectively depart from. Of course, whether or not it will be a utility-maximizing decision procedure depends on what the actual consequences of its adoption would be. One reason to think that it would be utility-maximizing, however, is precisely because DKDK instructs us to err on the side of moral caution in cases of uncertainty. This is most apparent if we consider the version of DKDK specified in degrees of belief and expected moral value. On this version, if the person who is uncertain is required to reasonably believe that killing the organism with unknown moral status will have higher expected moral value than not killing it, and if moral value is defined exclusively in terms of utility, then DKDK will simply require that a person who is uncertain

reasonably believe that killing the organism has higher expected utility than not killing it. This seems like a principle that act-utilitarians should endorse.

For similar reasons, Brad Hooker's version of rule-consequentialism, the view that "an act is wrong if it is forbidden by the code of rules whose internalization by the overwhelming majority of everyone everywhere in each new generation has maximum expected value in terms of well-being (with some priority for the worst off),"<sup>30</sup> might support the view that DKDK would be one of the rules that would be part of the code. Obviously, whether DKDK will in fact be part of the code will be determined by what account of well-being is endorsed, and on certain empirical facts about how many cases that could be either (R1) or (R2) cases, for all one knows, turn out in fact to be (R2) cases. But if we take the expected value reading of 'being compelled by something of moral significance,' it seems at least possible that some principle like DKDK would be part of the ideal code. Obviously, effects on individual well-being count as being of moral significance, so if there were a case in which the expected value of acting was greater than the expected value of not acting, in terms of well-being, then DKDK would not apply. It is hard to see how DKDK would conflict with the ideal code, and there are reasons to think that it would be a part of the code that would generate maximum expected value, at least in certain circumstances.

Making the full case that DKDK would be supported by either of these consequentialist theories would require making and defending certain empirical assumptions, something that is outside the scope of this paper. But it is worth noting that it is hard to see that acting in violation of DKDK would often or usually be approved of by act-utilitarians or rule-consequentialists. Even more straightforward arguments can be made that DKDK would follow naturally from the truth of a number of non-consequentialist moral theories.

Virtue ethicists, for example, might support DKDK on the grounds that it captures a principle that a virtuous person would live by, in order to avoid acting recklessly, or to avoid taking unnecessary risks with what is of moral value. One feature of DKDK that would seem to appeal to such an approach is that it makes clear that one can be morally blameworthy for acting in certain ways, and for ignoring certain relevant considerations, even if the consequences of what one does can appear innocuous. On such an approach, those who, for example, drive recklessly and hit someone, are seen to be just as culpable, just as lacking in virtue, as those who drive just as recklessly but are fortunate enough not to hit anyone. A principle such as DKDK is consonant with this kind of view.

Another moral theory that DKDK seems possibly in harmony with is Scanlonian contractualism. On this view, "an act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a basis for informed, unforced general agreement."<sup>31</sup> Could anyone reasonably reject DKDK? It is hard to see how anyone could. It allows a clear exemption for cases in which someone believes himself or herself to have compelling moral reasons to act. It only applies

---

<sup>30</sup> Hooker (2000), p. 32.

<sup>31</sup> Scanlon (1998), p. 153.

in cases in which one knows that one is uncertain about the moral status of some organism. It is hard to see who could complain, or on what grounds. There is a worry that DKDK only protects organisms the moral status of which is in doubt, and perhaps no such organisms could really be part of the general regulatory project that Scanlon envisions, and thus anyone with an interest in eating meat might be able to reasonably reject the addition of this principle. I think that this thought is mistaken. We can see this if we consider cases in which the moral status of human beings might be uncertain. Consider, for example, a case in which there is a question raised of whether people in comas or other vegetative states have moral status. In such a case, it seems quite reasonable to think that a principle like DKDK serves to protect our interests (rather minimally), given that we might end up in such a state, and might exit such a state.

A final moral theory that would seem to endorse DKDK is a modified Rawlsian account, on which the principles that should govern interpersonal conduct are just those that would be agreed to from behind a veil of ignorance in an original position choice point, with ‘representatives’ of ours standing in to choose the conduct-governing principles.<sup>32</sup> Not knowing what situations we might be in, and content with the exceptions that DKDK allows, it seems that those in the original position should endorse a principle like DKDK, or at least there would seem to be reasons in favor of doing so. This could be even clearer if we consider an even further modified Rawlsian account, on which those in the original position include ‘representatives’ of all sentient creatures, rather than just ‘representatives’ of all rational agents. There are reasons to think that this more inclusive position is the better of the two, given the level of idealization and counterfactual supposition already present. Discussing this further is outside the scope of this paper, but it does seem clear that on such an account, we’d want a principle at least as strong as DKDK to protect us in cases in which some other entity that has the power to kill us, and is uncertain of our moral status, is trying to decide what to do.

A final, related consideration in favor of DKDK is something like the thought that we should do unto others as we would have them do unto us. If we can imagine ourselves in a situation in which our moral status was uncertain in the eyes of some creature, I think that we would want that creature to abide by DKDK, at a minimum, and we would think that they morally ought to do so as well. Consider a case in which extraterrestrial beings come to our planet. Imagine that they are rational and sentient, but they cannot interact with us in any rich or extended way. Perhaps we seem to them as pigs seem to us. The extraterrestrials are uncertain whether we have significant moral status, and they know that they are uncertain. Of course, we’d like them to abide by DKDK, at least, but doesn’t it also seem like they should? Given this, doesn’t it seem like we should abide by DKDK as well?

Within the confines of a single paper and without taking a stand on what the correct moral theory is, it is difficult to offer a comprehensive justification for some particular moral principle such as DKDK. That said, I hope that it is at least

<sup>32</sup> This is obviously far from anything that Rawls would endorse, given that he explicitly wasn’t giving an account of interpersonal morality.

plausible that DKDK would be endorsed by a number of different plausible moral theories, and, given that, that DKDK itself looks somewhat plausible.

One way to support a principle is to show that it is consistent with or follows from plausible moral theories. Another way to support a principle is to show that it appears to get a number of different cases right, and that there aren't cases that it seems to get wrong. In the next section, I'll consider a number of different cases that DKDK can seem to get wrong, and a number of logical neighbors to DKDK that might seem problematic.

## 4 Objections to DKDK

### 4.1 Cases that DKDK might seem to get wrong

#### 4.1.1 *Cases in which the person is all but certain about the moral status of the organism*

From a certain perspective, DKDK is a strong principle. The cases that this objection is concerned with suggest that it is too strong. Imagine someone who is quite sure that individual trees don't have significant moral status (returning to our degrees of belief representation, let's say her credence in the proposition is 0.95), and knows that though she is pretty sure, she is also *only* pretty sure—she acknowledges that there is some chance that she could be wrong. She doesn't *know* (or even believe with certainty) that trees don't have significant moral status. And she knows that she doesn't know this. From DKDK, it seems that it would be blameworthy for her to kill trees or to have them killed unless she believes that there is something of substantial moral significance compelling her to do so. Given her (plausibly non-compelled) use of paper, wooden floors, wooden desks, wooden tables, etc., it seems that she is doing things that are wrong all the time.

And it could get worse. Imagine a hyper-vigilant individual who worries that literally everything might have significant moral status—bacteria, viruses, bricks, trucks. Such a person knows that she is at least somewhat uncertain about whether the entity in question has significant moral status. From DKDK (or, for the non-living, a suitably modified 'Don't Know, Don't Destroy' principle), it seems that it would be wrong for her to kill or destroy any of these things, absent a belief that something of substantial moral significance compels her to do so.

What should we say about these sorts of cases? One possibility is to reformulate DKDK with more precise degrees of belief:

**Pretty Uncertain, Don't Kill** If someone knows that she is pretty uncertain (say, confidence less than 0.8) about whether a living organism has significant moral status or not, it is morally blameworthy for her to kill that organism or to have it killed, unless she believes that there is something of substantial moral significance compelling her to do so.

This would address worries like those in the tree case and perhaps some of the hyper-vigilant individual worries, but it wouldn't address the radical hyper-vigilant cases—cases involving individuals who are really quite uncertain about the moral status of all manner of things. Additionally, there will be line-drawing issues regarding exactly how uncertain one has to be before one is in the worrisome 'pretty uncertain' territory. Ultimately, this modification may help in cases in which someone is close to certain but lacking certainty, but it won't do much to help with other cases.

#### 4.1.2 *Cases in which someone is genuinely uncertain, but appears to be doing something that is permissible*

There are related cases in which someone, perhaps our same hyper-vigilant individual, really is uncertain about the moral status of some organism, and knows that she is uncertain in this way, but in which it really does seem morally permissible for the person to kill the organism, even though she is uncertain of its moral status, and even absent any belief on her part that she is morally compelled to kill the organism. Importantly, for these cases, it only seems morally permissible from our perspective—presumably because we don't share the hyper-vigilant individual's uncertainty about the moral status of the organism. What should we say about these cases? How problematic is it to say that the radical hyper-vigilant individual does something morally blameworthy when, say, she uses a piece of paper, though uncertain whether trees have significant moral status and uncompelled by anything of moral significance? It can seem a pretty serious problem. Let me say a few things to suggest that it is not.

First, DKDK does not commit us to the view that the hyper-vigilant will be doing something that is *objectively* morally wrong in these cases. DKDK says only that the person is morally blameworthy for acting in this way, given that he or she was uncertain. From the perspective of the hyper-vigilant individual, or from the perspective of those who share their views and uncertainties, DKDK (or its variants) yields the result that they are doing something wrong in making these tradeoffs between things with 'uncertain' moral status and things with no moral significance at all. How the rest of us want to think about them or treat them seems to turn on how reasonable we feel it is for them to be uncertain about the moral status of the things in question. For pigs, cats, dogs, dolphins, chimpanzees, fetuses, cows, octopi, turkeys, and other complicated organisms, this may seem quite reasonable. For dust particles, rocks, vitamins, carrots, and cars, this may seem quite unreasonable. Still, one thing that we might feel, even if we don't share their uncertainties, is that their behavior exhibits an inappropriate level of inconsistency or inauthenticity—if one really is uncertain about the moral status of these things, then why act so casually with regard to one's treatment of them? And the charge of inconsistency needn't be a simple charge of irrationality or a complaint about the relationship between beliefs and actions. Rather, the complaint is that they are failing to take into account moral reasons they believe are present. This can plausibly be seen as a moral failing in its own right, regardless of how objectionable the conduct is that might follow.

To see this, recall that in other cases of needless moral risk—demolishing buildings without being certain about their contents, driving recklessly through a stop light while uncertain whether the path is clear, etc.—it is natural to see the person engaged in this behavior as morally culpable. And this remains true even when we know that the building is empty, or that the intersection is deserted. It seems that our position with respect to the hyper-vigilant individual is similar.

One thing that impedes our ability to truly believe this is that we feel sorry for the hyper-vigilant individual—after all, his life will be very difficult if he actually is uncertain in these ways, how can we also hold him morally blameworthy if he doesn't live up to his own standards! But this response fails to take into account that the only cases for which this response is appropriate are cases in which DKDK implies that the person is morally blameworthy, meaning that there *isn't* anything that the individual believes to be of moral significance compelling him to act. In the cases we might imagine—cases of killing and destroying bacteria, viruses, plants, and the like—there will often be something of moral significance on the other side, so that these charges of inconsistency will be inappropriate, and DKDK will not yield the result that what is done is wrong. This may still seem to countenance an unusually high bar on conduct that we feel is unobjectionable—using paper, for example. Then again, it is not that implausible to see the choice between using paper or plastic bags in the checkout line as one that, for those uncertain about whether trees have significant moral status (and presumably certain that plastic does not), should be an easy moral choice, other things being equal. Once we are certain that trees don't have significant moral status, as many of us likely are, then we can move to consider other questions that might govern our choice—some moral, some prudential. But to the extent that we are unsure, and to the extent that there are alternatives about which we are sure, it seems that we should avail ourselves of those alternatives.

Of course, there may still be hyper-vigilant individuals who also have what we might take to be funny views of what is of substantial moral significance, such that they never feel that anything of prudential value to them is of substantial moral significance. So they might refuse to take antibiotics, even when medically necessary, because they are uncertain about the moral status of bacteria and because they don't take their own health to be of substantial moral significance. We can certainly imagine such cases, and it can seem that DKDK would yield the result that such people would be morally blameworthy to take the antibiotics in such a case. This would seem a misguided result. Here, however, it seems misguided in large part because the person now starts to seem somewhat unintelligible to us, perhaps failing to meet general conditions of rationality and sanity that seem necessary for attributions of culpability to make sense. We might think it natural that moral principles do not apply to such a person, at least not in the regular way, and so DKDK will not apply either.

#### 4.1.3 Abortion

Some who might be tempted to support DKDK in general might be worried about the implications of DKDK for abortion. I think that when we pay attention to

exactly how DKDK would apply to abortion, we see that its implications actually seem reasonable.

We can imagine a case involving abortion that would mirror the case of Douglas and the pig. Consider a case in which a woman, 10 weeks pregnant, is contemplating getting an abortion. Though not thinking in these explicit terms, she is able to understand the idea of moral status, and is questioning whether or not the fetus inside her has moral status equal to that of an infant human being. She thinks that it would be wrong to abort the fetus (and so she wouldn't abort it) if it has moral status equal to that of an infant human being. After making a considerable and reasonable effort, she is still unable to come to an answer to this question—she doesn't know whether the fetus inside of her has moral status equal to that of an infant human being or not. Her reasons for getting an abortion are that she prefers not to be pregnant for another six months—she has decided that she would give the child up for adoption were she to carry it to term, and she doesn't want to deal with the physical and psychological effects of and limitations stemming from being pregnant and giving birth. She finds these aspects of being pregnant merely inconvenient, though she likes some aspects of being pregnant, and it is not, for her, that inconvenient. Still, she dislikes the inconvenience enough to want to get an abortion. She is not concerned about possible long-term health complications from pregnancy, and let us suppose she is reasonable in not being concerned. She does not believe that there is anything of substantial moral significance compelling her to get the abortion. The pregnancy resulted from voluntarily engaged in sexual intercourse with someone who has explicitly said that the decision is hers to make and that he will do whatever he can to support her decision, whatever it ends up being. She decides to abort the fetus, and does so successfully.

What does DKDK imply in such a case? It is plausible that in such a case, DKDK would imply that she is morally blameworthy for getting the abortion, given her uncertainty about the moral status of the fetus. She kills a living organism<sup>33</sup> the moral status of which is uncertain to her, and known to be uncertain to her, and she doesn't take herself to have any compelling moral reason to get the abortion.

Importantly, DKDK implies this only because this is, or seems to be, a particularly unusual (some would say 'contrived') set of circumstances for an abortion. In particular, it seems that in many cases in which a woman gets an abortion, she reasonably believed either (a) that she was not uncertain about the

<sup>33</sup> I have not defined what a living organism is, nor will I try to do so. On my understanding of the term, a human fetus clearly counts as a living organism. Living organisms need not be particularly complicated: a single bacterium counts as a living organism. And many organisms live inside us at all times. Still, there are natural objections both that a fetus isn't living—it doesn't undergo or isn't capable of the right sort of biological processes; and that a fetus isn't an organism—it isn't 'independent' in the right sort of way to count as a separate organism. Both of these are misguided. The term 'fetus'—as opposed to 'embryo'—is applied at the end of the eighth week, roughly at the point at which the fetus's heart is beating and brain waves can be detected. It seems plausible to say that any entity that has a beating heart and detectable brain waves should count as a living organism. How to individuate organisms is not entirely clear (do Siamese twins count as one or two organisms?), but given that the heart, limbs, spine, nervous and circulatory systems all begin to form at 4–5 weeks, in the pre-fetal, embryonic stage, it seems hard to argue that the fetus and the person in which the fetus resides are only one organism.

moral status of fetuses, or (b) that something of moral significance compelled her to get the abortion. The unusual cases in which a person believes neither of these things seem like ones in which what DKDK implies, that the person is morally blameworthy, is plausible.

One might still worry about whether there will be many cases in which, though a woman believes that something of moral significance compels her to get an abortion, this belief is not reasonable, given the heightened epistemic standards required (given the truth of moral epistemic contextualism). Whether this is plausible or not will depend on the specifics of the case, but it does seem that in many cases of abortion, there are at least substantial moral considerations in the offing—unlike in the case of killing organisms in order to eat them. For example, it seems plausible that someone could reasonably believe that any of the following factors, commonly associated with the decision to have an abortion, should count as being of substantial moral significance: (a) concerns about how having a child might interfere with one's life plans and happiness, (b) concerns about the financial and emotional difficulty of raising a child, (c) concerns about the actual experience of and health concerns associated with being pregnant and giving birth, (d) possible social stigma or punishment associated with having a child given the particular circumstances attending the pregnancy, and so on. Any of these might, in some cases, meet the substantiality requirement. Additionally, as noted above, whether it is reasonable to see some action as morally compelled is sensitive to what one takes one's options to be. Thus, in abortion cases, it is appropriate to assess the reasonableness of the belief that one is morally compelled to abort in light of the situation the person actually finds herself in when confronted with the option of having an abortion. As before, if someone is too unreasonable in discounting what appears to be of substantial moral significance, we may start to view that individual as failing to meet conditions necessary for moral culpability generally.

There is still another worry in the neighborhood. If we take the maximizing expected moral value reading of being morally compelled, it might seem that it will never be reasonable to believe that some of the above factors could morally compel one to act. The thought would be something like this: if one believes that there is a fifty-fifty chance that a fetus has moral status equal to a person, and that killing a person has some very large quantity of moral disvalue, it would never be reasonable to believe that some of the factors listed would tip the balance in favor of abortion. The only factors that could make such a belief reasonable would have to be much more significant than many of those listed—perhaps the health of the mother would be morally significant enough. Note, again, that this sort of case requires that one actually believe that there is a fifty-fifty chance that a fetus has moral status equal to a person. But the more fundamental problem with this thought is that the maximizing expected moral value reading of being morally compelled is itself not compelling. As noted above, this is not the reading that I prefer, in part because of cases like these. It seems that there are other grounds on which one might reasonably feel oneself to be compelled by something of substantial moral significance to act. For example, arguments made by Judith Jarvis Thomson and Frances Kamm regarding abortion suggest that in cases in which it is either the pregnant woman's life or the fetus's life, even assuming that the fetus has the full

moral status of a person, it still would be morally permissible for the woman to choose self-preservation.<sup>34</sup> If these arguments succeed, it seems that a woman could reasonably believe that when the moral status of the fetus was uncertain, they might still be compelled by something of moral significance to get an abortion. And there are other arguments which conclude that even in cases in which the woman's life isn't in jeopardy, she has a right to do what she wants with what is inside of her body, and so could choose to abort the fetus even without examining the 'competing rights' claims that each might have. These arguments tend to focus on the moral permissibility of abortion, even if the fetus has the moral status of a person, but we could imagine similar arguments being made regarding the reasonableness of feeling compelled by decisive moral reason to abort in certain cases. It seems that, given this, it is at least not obvious that it would be unreasonable for people to believe that these factors were of substantial moral significance to compel them to abort.

The upshot of this discussion is that the only cases in which DKDK will imply that a person has done something morally blameworthy in getting an abortion are those in which a person believes *both* that she is uncertain about the moral status of fetuses, and that nothing of substantial moral significance compels her to get the abortion. In such cases, which will be a clear minority of the cases in which someone is contemplating an abortion, one who gets an abortion does do something morally blameworthy.

#### 4.2 Problematic neighboring principles

There are a number of cases in which a principle that seems very close to DKDK, in formulation and in spirit, would suggest that things that we ordinarily do, and which we take ourselves to be morally non-culpable in doing, are in fact things for which we are morally blameworthy. Such cases don't directly impugn DKDK, but they do suggest that the sort of reasoning that seems to support DKDK may in fact go too far, supporting principles that we think should be rejected, which in turn would give us some reason to worry that this reasoning shouldn't be used to support DKDK either. Let me consider a few such cases. In general, these cases will have the following feature. Whereas with DKDK, the action component involves an activity—killing a living organism the moral status of which is unknown to one—that seems relatively avoidable (for instance, there are actual people living amongst us who are practicing vegans), it seems we can imagine similar principles, with a similar justification, for which the action component would involve an activity that seems both hard to avoid and morally permissible.

Consider, for example, the following principle:

**Don't Know, Don't Drive** If someone knows she doesn't know whether driving will lead to catastrophic consequences for future individuals or to death or injury to present individuals, it is morally blameworthy for her to drive unless she believes there is something of substantial moral significance compelling her to do so.

<sup>34</sup> Thomson (1971), No. I; Kamm (1992).

Given that there are other means of transportation available that don't involve driving (biking, walking, public transportation), it seems that much ordinary driving would be morally blameworthy. What should we think about this principle?

One thing to think about what is happening in such cases is that such a principle is problematic for the same reasons that motivated suggesting *Pretty Uncertain, Don't Kill*; namely, we are pretty certain that driving won't lead to catastrophic consequences for future individuals or death or injury to present individuals, but we don't *know* this to be the case. If that is the concern, then the response above, relaxing the epistemic component requirement (so that one only meets the epistemic component if one is less than, say, 0.8 certain), is a possible response. This might address worries, say, that on any given trip out I might run into someone and kill them. If one only believes 'that I will not hit anyone when driving' to a degree of 0.8 or less, one really shouldn't be driving.

On the other hand, if one is genuinely unsure whether driving will lead to catastrophic consequences for future individuals or to death or injury to present individuals, it does seem that one does something morally blameworthy when one drives in the face of other viable options. Imagine trying to justify driving in such a case to some hypothetical person suffering catastrophic consequences from one's driving. What could one say? "I am sorry that my driving has led to catastrophic consequences. I knew this was a non-trivial possibility, but I really hate the bus." This sort of thought suggests that a principle like *Don't Know, Don't Drive* is at least not obviously false. It is true that driving is widely seen as morally acceptable, but if, for all we know, it really is contributing, inexorably, to the devastation of the planet, or at least to the destruction of a planet habitable by human beings, it seems that it is morally blameworthy for us to drive any more than we believe to be morally necessary. This principle strikes us as false because we take ourselves—implicitly, if not explicitly—to be certain that driving does not lead to such catastrophic consequences.<sup>35</sup>

Another possibility is that the other non-driving options are not real options. Obviously, it will be morally significant if one has to spend a large portion of one's time walking miles and miles (and in all climates) to conduct basic life activities. In many places, it may be quite reasonable to believe that driving is morally necessary.

My ultimate response to cases like *Don't Know, Don't Drive* is that principles of this sort will generally be correct. They might seem incorrect because we underestimate the moral significance of the activity that we would have to give up, or they might seem incorrect because we inaccurately take ourselves to be in a position of real uncertainty with regard to whether the claim in the fact component

<sup>35</sup> There might be tricky things here regarding the way in which *my* driving is part of the joint cause of the world's becoming uninhabitable. I might, for instance, be certain that my driving *by itself* isn't going to cause catastrophic consequences for future generations, while also acknowledging that I am engaging in a practice which, if everyone engaged in it, would cause catastrophic consequences for future generations. Still, it seems that if I believe (or don't fully disbelieve) this latter fact about the practice, this provides me with familiar moral grounds for not engaging in the practice—whether those are Kantian, fair play, rule-consequentialist, no free riding, or some other sort of moral grounds. I think that *Don't Know, Don't Drive* could be reformulated to include these kinds of joint-cause stories as well.

of the principle is in fact true. In cases in which the principle actually applies, it seems to get the right result, or so I believe.

This seems true for a range of principles, some of which might at first seem controversial. Consider, for example, the following:

**Don't Know, Don't Destroy** If one knows that one doesn't know whether some entity has moral value, then it is morally blameworthy to destroy that entity, unless one believes that something of substantial moral significance compels one to do so.

**Don't Know, Don't Injure** If one knows that one doesn't know whether some entity can feel pain, then it is morally blameworthy to do things which would cause pain to that entity, if it could feel pain, unless one believes that something of moral significance compels one to do so.

**Don't Know, Don't Shoot** If one knows that one doesn't know whether some spatio-temporal region is occupied by a living organism with significant moral status, then it is morally blameworthy to shoot a bullet or other dangerous projectile into that spatio-temporal region, unless one believes that something of substantial moral significance compels one to do so.

**Don't Know, Don't Invade** If one knows that one doesn't know whether an invasion into another's territory will do more harm than good, then it is morally blameworthy to invade the other's territory, unless one believes that something of substantial moral significance compels one to do so.

I won't attempt to defend each of these, though the story would be similar to the one I've told for DKDK. These are principles that urge caution in cases of moral uncertainty, and I think that when they apply, they are plausible.

## 5 What would Rosen think about DKDK?

I think that DKDK is interesting and plausible in its own right. Returning to our original question, however, regarding the MIT, we might ask what Rosen would think about the case of Douglas and DKDK. One might raise the following objection on Rosen's behalf: DKDK only applies to cases in which a person knows that she is ignorant or uncertain, and these are not the most interesting cases, and perhaps not even the cases that Rosen had in mind.<sup>36</sup> This objection suggests that the MIT could be reformulated:

<sup>36</sup> There is a footnote very early in Rosen's paper (discussing an unrelated point) which suggests that he simply was not considering cases in which one might be blamelessly ignorant of some fact, *F*, and know that one is ignorant or uncertain of that fact, and yet still act from ignorance of *F*. Rosen writes, "A better formulation of the relevant principle is rather as follows: When X does A from blameless ignorance, then X is blameless for doing A, *provided the act would have been blameless if things had been as the agent blamelessly took them to be.*" Rosen, n. 4, p. 63 (italics in original). It is simply unclear what should be filled in as the way an uncertain agent 'takes things to be'—if the agent is truly uncertain. It is not as if such an agent takes there to be no fact of the matter, or at least this need not generally be the case.

**Moral Ignorance Thesis (Deep Ignorance)** Whenever an agent acts from moral ignorance, *and is ignorant of his ignorance*, he is culpable for the act only if he is culpable for the ignorance from which he acts.

Straightforwardly, the purported counterexample I offered above now would not be a counterexample to this reformulated version of the MIT. Douglas is not ignorant that he is ignorant. At the end of the day, this reformulated thesis might be the right thing to believe. I take it that it would still be useful to have shown the application of the MIT to be more restricted, even if only in this way. If nothing else, if we are worried about the possibility of moral ignorance as a general, widespread excusing condition, as Rosen is, this would make it less widespread. Still, I have concerns about the reformulated MIT.

First, if the relevant thing is that one is blamelessly ignorant, then it seems that cases in which one knows one is ignorant should be just as central. Rosen argues that one reason that ignorance excuses is that it “issue[s] in blameless ignorance of what the agent has most reason to do all things considered.”<sup>37</sup> This same blameless ignorance is present in cases in which people know that they don't know, and so it seems that something else must be at work.

Second, it seems that cases of non-culpable ‘deep’ moral ignorance, as I have called it, may actually be relatively rare. It may even be the case that the examples Rosen offers—the Ancient Slaveholder, the 1952 American Sexist—fail to fall into this category. We see this when we consider what moral epistemic contextualism says about those in (B1) cases (cases in which one has never explicitly thought about the issue) or those in (B2) cases (cases in which one has thought about the issue, but has come to have false beliefs about it. As noted earlier, it is actually harder than Rosen seems to suppose to remain blamelessly ignorant in such cases, given the heightened epistemic obligations. If these cases are rare as well, the reformulated MIT might be true, but its importance as a general or widespread excusing condition would be minimal, given how few people would actually be able to claim non-culpable ‘deep’ moral ignorance in ‘high-stakes’ contexts.

Finally, it seems that the central thought behind the kind of cases that Rosen is interested in—even if the Ancient Slaveholder is a bad example—is this: “Just as it is unfair to blame someone for doing what he is in fact entitled to do, it is unfair to blame someone for doing what, through no fault of his own, he *takes himself* to be entitled to do... It is unfair to blame someone for doing something if he blamelessly believes that there is no compelling moral reason not to do it.”<sup>38,39</sup> This seems plausible, which is in part why it seems necessary to include the ‘justification component’ in DKDK, and to allow that one is not morally blameworthy for acting

<sup>37</sup> Rosen, p. 80.

<sup>38</sup> Rosen, p. 74–75.

<sup>39</sup> Zimmerman may have been on to something similar when he asserts that “lack of ignorance” that one ought not to perform the act in question is a “root requirement for responsibility” for performing the act in question. Zimmerman, p. 424. He thinks this, and not anything to do with the ‘avoidability’ of acting in a certain way, is what explains why it is true that “one is culpable for behaving ignorantly only if one is culpable for being ignorant.” *Id.* at 423. Where I disagree with both Zimmerman and Rosen is in their assessment (implicit or explicit) of when one is reasonable in believing that there is nothing morally objectionable about one's action.

in a way that would otherwise violate DKDK when one *reasonably believes* that there is something of moral significance compelling them to act. Given this rationale, it seems that the reformulated MIT may well be true. What I hope to have shown is both that the unrestricted MIT is false, and that it is harder than one might have thought to be in a situation where one blamelessly takes oneself to be entitled to do something that is in fact morally impermissible.

## 6 Conclusion

I think that in cases in which an agent

- (a) is non-culpably ignorant of a relevant moral fact (the moral status of a living organism),
- (b) knows that she is ignorant,
- (c) acts from ignorance (she wouldn't have acted if she knew the moral status of the organism), and
- (d) kills a living organism
- (e) without believing that anything of substantial moral significance compels her to do so

the agent is culpable, morally blameworthy, for the act. Cases of this sort provide counterexamples to what I have labeled the Moral Ignorance Thesis. I think that whenever one acts in this morally reckless way—knowing that one doesn't know whether what one does kills something with significant moral status—one does something morally blameworthy. This remains the case whether or not the thing killed actually has significant moral status (though if it does not then one cannot be said to be acting from ignorance in the way defined above).

It is hard to know how many cases in the real world share these features. It seems that many instances of eating animals will have these features, though there is a question of whether one is really ignorant about the moral status of the animal in question, and, if so, whether one knows that one is ignorant. It does seem clear that nothing of substantial moral significance is compelling people to eat the animals. In cases of, for example, abortion, the situation is reversed. In many of these cases, we don't know what to believe regarding the moral status of fetuses, and we are aware of our ignorance in this regard. On the other hand, these cases more often exhibit substantial moral reasons in favor of performing the action—getting an abortion—than do those cases in which all that is at stake is a certain sort of gustatory pleasure.<sup>40</sup>

---

<sup>40</sup> I would like to thank Tyler Doggett, Liz Harman, Derek Parfit, and an anonymous referee for their insightful and invaluable written comments on drafts of this paper. Thanks also to Simon Rippon and Stephen Schiffer for their helpful comments in response to presentations of an earlier version of this paper, and to audiences at the 2005 Harvard-MIT Graduate Student Conference and the NYU Thesis Preparation Seminar.

## References

- DeRose, K. (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, 52, 913–929.
- Eriksson, L., & Hájek, A. (2007). What are degrees of belief? In B. Fitelson (Ed.), *Studia Logica, special issue on formal epistemology*.
- Hooker, B. (2000). *Ideal code, real world*. Oxford: Oxford University Press.
- Kamm, F. M. (1992). *Creation and abortion: A study in moral and legal philosophy*. Oxford: Oxford University Press.
- Lockhart, T. (2000). *Moral uncertainty and its consequences*. Oxford: Oxford University Press.
- Moody-Adams, M. (1994). Culture, responsibility, and affected ignorance. *Ethics*, 104(2), 291–309.
- Rosen, G. (2003). Culpability and ignorance. *Proceedings of the Aristotelian Society*, 103(1), 61–84.
- Scanlon, T. M. (1998). *What we owe to each other*. Cambridge, MA: Harvard University Press.
- Smith, H. (1983). Culpable ignorance. *The Philosophical Review*, 92(4), 543–571.
- Slote, M. (1982). Is virtue possible? *Analysis*, 42, 70–76.
- Strawson, P. (1982). Freedom and resentment, reprinted In G. Watson (Ed.), *Free will*. Oxford: Oxford University Press.
- Thomson, J. J. (1971). A defense of abortion. *Philosophy & Public Affairs*, 1, 47–66.
- Zimmerman, M. (1997). Moral responsibility and ignorance. *Ethics*, 107(3), 410–426.