Abstract

Moral responsibility, I argue, requires agents to do what is within their abilities to act morally. This means that an agent is to blame just in case his wrongdoing is due to an underperformance, that is, a failure to do what he can to act morally. I defend this account by considering a skeptical argument about responsibility put forth by Gideon Rosen and Michael Zimmerman. I explain why the epistemic condition they endorse is inadequate, and why my alternative epistemic condition, which directly follows from my general condition on culpability, should be preferred. I then defend my view against potential criticisms.

A person who freely and knowingly acts wrongly is to blame for her wrongdoing. But can an agent be culpable for a wrong action she does not even believe is wrong? In my view, she can. Agents who act wrongly because they fail to perform according to their abilities, I will argue, are culpable for their wrongdoings. Hence, an agent is to blame if her bad act is based on a false belief she ought not to have formed, given her cognitive abilities. My goal in this paper is to flesh out this position, and defend it against some objections.

I will start by examining a skeptical argument about responsibility put forth by Gideon Rosen and Michael Zimmerman. The value of this exercise is twofold. First, the skeptical
argument relies on a widely accepted epistemic condition on culpability. Showing the inadequacy of this condition will help motivate my alternative epistemic condition. Second, Rosen and Zimmerman provide arguments in support of this condition. Explaining where these arguments go wrong and why the conception of responsibility they rely on is misguided will enable me to shed light on the requirements of moral responsibility. I will show how my proposed epistemic condition naturally follows from these requirements. The rest of the paper will address possible objections to my view.

1. Skepticism about Responsibility

There is general agreement that an agent may be blameless for her wrongdoing. Two kinds of excuse are deemed admissible. First, the agent may lack the relevant control over her action. She may not have had the capacity not to act wrongly. The second kind of excuse is epistemic: the agent may not have realized that her action was wrong. Traditionally, sceptical arguments about moral responsibility have concerned the first type of excuse: if our behaviors are governed by (deterministic) laws of nature, then, the arguments go, we do not have genuine control over these behaviors. But I will examine a different kind of sceptical argument, put forth by Gideon Rosen (2004) and Michael Zimmerman (1997, 2008), which exploits the second kind of excuse.

Doctor Wong prescribes a common antibiotic to her patient. Unbeknownst to her, the patient is allergic to the antibiotic and is harmed by taking it. Is Dr. Wong culpable for prescribing the antibiotic? It depends on the details of the story. Suppose first that the patient’s allergy was noted on his chart, which Dr. Wong failed to consult. In this case, Dr. Wong is to
blame for her action.\textsuperscript{1} But now suppose that Dr. Wong took reasonable precautions: she consulted the patient’s chart, which did not mention the allergy; she asked the patient to confirm the relevant entries on the chart; etc. In such a case, Dr. Wong is not to blame for the wrong prescription. Cases such as this one are invoked to support the following thesis:

(1) If agent S performs action A from ignorance, then S is culpable for the act only if S is culpable for the ignorance from which she acts.\textsuperscript{2}

What does ‘ignorance’ amount to exactly? Suppose S truly believes that her doing A is morally wrong, but fails to know that, because she has little or no justification for that belief, or because she is in a Gettier-type situation. In such a case, it seems that S would still be blameworthy for performing A. We should thus equate ignorance with a lack of true belief rather than with a lack of knowledge.\textsuperscript{3} And ignorance here concerns the morality of S’s own action: ‘S does A from ignorance’ should be understood as meaning that S does A while lacking a true belief that her doing A is morally wrong.\textsuperscript{4} Finally, both Rosen and Zimmerman hold that if

\textsuperscript{1} More carefully, Dr. Wong is to indirectly blame for her action if she is directly culpable for her failure to consult the chart. I will examine the distinction between direct and indirect culpability shortly.

\textsuperscript{2} See (Rosen, 2004, 300) and (Zimmerman, 1997, 414).

\textsuperscript{3} See (Rosen, 2008, 596-597). Can an agent be to blame for a morally permissible act? This is a difficult and controversial issue I cannot properly address here. To steer clear of that debate, I will consider only morally wrong acts as potentially culpable ones.

\textsuperscript{4} See (Zimmerman, 1997, 412). Rosen (2008) proposes a slightly different construal, which I will discuss in section 4.3.
the agent has a *dispositional* rather than *occurrent* belief that her action is morally wrong, she counts as ignorant that her action is morally wrong.⁵ (1) should thus be understood as:

(1’) If agent S performs action A while lacking the occurrent true belief that her doing A is morally wrong, then she is culpable for doing A only if she is culpable for lacking that occurrent belief.

This brings us to the phenomenon of *culpable ignorance*. Morality, Rosen points out, makes epistemic demands on us. He writes:

> As you move through the world you are required to take certain steps to inform yourself about matters that might bear on the permissibility of your conduct.

> You are obliged to keep your eyes on the road while driving, to seek advice before launching a war and to think seriously about the advice you’re given; to see to it that dangerous substances are clearly labeled, and so on. These obligations are your *procedural epistemic obligations*. (2004, 301)

And as Rosen makes clear, “The procedural obligation is not itself an obligation to know or believe this or that. It is an obligation to *take steps* to ensure that when the time comes to act, one will know what one ought to know” (ibid.).

Procedural epistemic obligations, Rosen points out, are highly dependent on one’s situation and impossible to codify. Several factors must be taken into consideration when determining an agent’s procedural epistemic obligations. One must consider how likely it is that new evidence will affect the ranking of the available courses of action. In particular, one needs to consider how probable it is that, in light of the new information, one’s current best option

---

⁵ I will examine their reasons for holding this in section 4.3.
will no longer be at the top. One also needs to consider how morally superior an alternative
course of action would be, relative to the current favorite, in light of the new evidence that
would be gathered. Finally, one needs to consider the costs involved in acquiring the new
information, and how these compare to the costs of acting wrongly and the benefits of a right
action.⁶

Two points emerge from these considerations. First, a procedural epistemic obligation
is—despite its name—a moral obligation. A failure to fulfill one’s procedural epistemic
obligations is thus a moral failure. And like any moral failure, it may or may not be culpable.
Second, strictly speaking, a culpably ignorant agent is not culpable for her ignorance (i.e., lack
of true belief), but for a past failure to fulfill her procedural epistemic obligations. Culpable
ignorance is thus always indirect, or derivative. In other words,

(2) An agent S is culpable for the ignorance from which she acts only if her ignorance is
the upshot of some prior culpable act or omission.

(1) and (2) entail that an agent who culpably acts from ignorance must be derivatively
blameworthy for her action. Her culpability must be due to some earlier original, or direct,
culpable failure to fulfill a procedural epistemic obligation. This means that an agent who is
directly culpable for her action cannot be acting from ignorance. Direct culpability thus requires
an occurrent true belief that one is acting wrongly:

(EC1) An agent S is directly blameworthy for her wrongdoing A only if S has an occurrent
belief that her doing A is morally wrong.

---

⁶ See (Jackson, 1986), for a useful discussion. Note that one has an obligation to investigate not
only about morally relevant facts, but also about morality itself.
On the proposed picture, direct culpability is due to what we may call “clear-eyed” wrongdoing, or, as Rosen puts it, *clear-eyed akrasia*, which occurs when an agent acts while having an occurrent true belief that she is doing something wrong. EC1 is required for Rosen and Zimmerman’s argument to go through. Suppose EC1 is false, and there are cases in which an agent is directly blameworthy for acting wrongly from ignorance, that is, while lacking an occurrent belief that she is acting wrongly. Then, contrary to what (1) and (2) entail, it is not the case that an agent can only be derivatively blameworthy for acting from ignorance.

Let us take stock. Culpable wrongdoing is either clear-eyed or from ignorance. According to (1) and (2), the latter must be traceable to an earlier culpable wrongdoing. But to avoid a regress, that wrongdoing must be clear-eyed. We thus have:

(3) An agent S is culpable for her wrongdoing A only if her doing A is, or derives from, an episode of clear-eyed akrasia.⁸

Rosen and Zimmerman draw slightly different conclusions from this argument. Rosen claims that akratic actions are extremely hard to distinguish from ordinary weakness of the will. He writes: “The akratic agent judges that A is the thing to do, and then does something else, retaining his original judgment undiminished. The ordinary moral weakling, by contrast, may initially judge that A is the thing to do, but when the time comes to act, loses confidence in this judgment and ultimately persuade himself (or finds himself persuaded) that the preferred alternative is at least as reasonable” (2004, 309). Since cases of ordinary weakness of the will are not the locus of original responsibility, “it would be unreasonable to repose much

---

⁷ I will describe several such cases in the following sections.

confidence in any particular positive judgment of responsibility” (ibid., 308). Zimmerman (1997, 425; 2008, 176), on the other hand, concludes that because cases in which ignorant behavior can be traced to an episode of direct blameworthy action are rare, many ordinary apparently blameworthy actions are in fact blameless.

2. Contesting the Argument

A crucial premise of the skeptical argument is EC1. It is worth noting that epistemic conditions on culpability similar to EC1 are widely shared. As a matter of fact, many authors endorse an even stricter condition that requires knowledge of wrongdoing. In her classic paper on culpable ignorance, Holly Smith (1983) calls an act done from ignorance an unwitting act, and a violation of a procedural epistemic obligation a benighting act. She writes: “To say the culpably ignorant agent is to blame for his unwitting act is to say nothing more than that he was culpable in performing the benighting act, that it gave rise to the unwitting act, and that he knew at the earlier time that he risked this outcome” (ibid., 566). More recently, McKenna (2012, 15) holds that blameworthiness for an action requires knowledge that the action is wrong.9

But EC1 is implausible. Suppose agent S is unsure whether her doing A is morally correct. Let us say that S attaches 0.5 credence to the proposition that her doing A would be morally wrong. S would be blameworthy for (freely) doing A. As Elizabeth Harman writes, “If someone acts wrongly while genuinely unsure whether her action is wrong, we need not

9 See also (Fischer and Ravizza, 1998), (Ginet, 2000), (Haji, 1997) and (Levy, 2005, 2011) for similar epistemic conditions. These conditions exemplify what Sher (2009) calls the searchlight view, according to which blameworthiness for an action requires awareness of the wrongness of the action.
investigate whether she is blameworthy for being unsure to know whether she is blameworthy: she may well be blameworthy simply for doing what she did, which she believed might well be wrong” (2011, 449). Whether S’s current uncertainty is due to a prior violation of a procedural epistemic obligation is irrelevant here. S is (directly) blameworthy for her (freely performed) bad act A if she lacks a belief that her doing A is morally permissible.\(^\text{10}\) Hence, culpability does not require the belief that one’s action is morally wrong. To avoid Harman’s counter-example, one may propose the following condition:

(EC2) An agent S is directly blameworthy for her wrongdoing A only if S lacks the belief that her doing A is morally permissible.

Unfortunately, EC2 is also problematic. Consider an agent S who has fulfilled all his relevant procedural epistemic obligations. However, S underperforms, epistemically speaking: S’s belief about the morality of his action is not well supported by his evidence. Because of that, he ends up (freely) performing a morally wrong action. S is blameworthy for his wrongdoing, even though he believes he is doing the right thing.

Consider Carl, an office manager who tells a sexist joke to his employees during lunch break.\(^\text{11}\) He does not believe that he is doing anything wrong. There are many possible explanations for Carl’s ignorance. Perhaps his sexism is so deeply ingrained that he lacks the

\(^{10}\) But what if S also attaches 0.5 credence to the proposition that her not doing A would be morally wrong? In such a case, S may not be blameworthy for her doing A (or for her not doing A). Blameworthiness requires that there be another course of action available to S that is more likely to be morally permissible according to S. I will ignore this complication in what follows.

\(^{11}\) This example is similar to Sher’s *Bad Joke* (2009, 28).
ability to realize that his joke is morally objectionable. But let us assume that this is not the case. Carl is well capable of seeing the problem with the joke; however, he does not. This could be because he did not pause and think about the morality of his joke, in which case he would have failed to fulfill his epistemic procedural obligations. But let us also rule out this possibility: Carl did think about the morality of the joke—at least as much as one reasonably ought to—and saw nothing wrong with it. Carl had all the background knowledge and abilities needed to figure out the wrongness of the joke, but simply failed to put two and two together. When confronted by a co-worker about the episode later on in the day, Carl has an “Oh dear! Of course” experience,12 rather than an “Oh really?” experience. He feels embarrassed at his underperformance: as he could very well put it, he has no excuse for his failure. Carl is morally blameworthy for his telling the sexist joke, even though at the time he believed that there was nothing morally problematic with it.

Carl is a morally ignorant agent. But an agent’s error may also concern the relevant non-normative facts of her situation. The agent wants to do the right thing, but because of some faulty reasoning, she misreads the available evidence and, as a result, is mistaken about what morality recommends. In other words, she fulfills her epistemic procedural obligations, but fails to perform according to her cognitive abilities. A bank teller is to blame for entering the wrong balance due to a customer because of a miscalculation. An ambulance driver is blameworthy for taking a wrong turn and losing precious time during an emergency call, if her mistake is due

12 I borrow this locution from M. Smith (2003, 20).
to her misreading the map. In each case, we may suppose that the agent sufficiently investigated the situation and thus did not violate any procedural epistemic obligation. The attribution of blame is correct if the agent underperformed, that is, if it was well within her capacities to figure out the appropriate course of action but she failed to do so. The agent’s belief about what she should do was not epistemically justified. According to our ordinary practices of blame attribution, such an agent is morally blameworthy for her wrongdoing.

Cases such as these show that EC2 is incorrect, and should be replaced by:

(EC3) An agent S is directly blameworthy for her wrongdoing A only if S lacks an epistemically reasonable belief that her doing A is morally permissible.

Why does moral responsibility require an epistemically reasonable rather than a justified belief? The requirements of responsibility should take into consideration S’s capacities. Sometimes, the relationship between S’s evidence and a given proposition p is beyond S’s cognitive capacities to ascertain, because it is too complex. In such cases, S may reasonably believe that p, even though her belief is not, strictly speaking, justified. Because it would seem incorrect to blame S for acting on the belief that p, responsibility requires not a justified but a reasonable belief. What counts as a reasonable belief is thus agent relative, that is, it depends on the agent’s relevant cognitive abilities and background knowledge.

---

13 Being in a stressful situation may excuse one’s error. But I am assuming that the ambulance driver is experienced and that her level of stress is not especially high. Her error is thus not explainable by any “capacity-defeating” event or condition.

14 An agent may do his reasonable best to figure out what to do, but because of his implicit biases, fail to form a justified belief about the right course of action. The agent is not directly
3. Doing One’s Reasonable Best

To support condition EC3, I have invoked the fact that we do blame people for their wrongdoings when the latter result from cognitive underperformances. Further support for this position can be gained from examining cases of culpability other than moral. Intuitions about these are perhaps less likely to be corrupted by one’s theoretical commitments.

Consider a chess player who fails to notice that moving his pawn will allow his opponent’s bishop to capture his rook by a so-called “long diagonal.” This is a common mistake. But players are not all assessed in the same way for making it. While we readily excuse the novice for such a blunder, we blame the experienced player for it. We would say things like ‘He, of all people, should not have moved that pawn,’ or ‘Someone like him ought to have anticipated the threat.’ This suggests that two different norms are in place here. Both the novice and the experienced player violate what we may call a primary norm of “blunder-free chess playing,” which states (roughly) that one ought to prevent one’s pieces from being captured (unless that allows one to make a bigger capture). But there is a secondary norm that only the experienced player violates in this case. This secondary norm, which is dependent on the primary norm, is what we may call a norm of responsible chess playing. As the example blameworthy in such a case. However, the agent may be to blame for failing (or not trying) to shed his implicit biases, or for not exploring alternative decision procedures that would reduce the effects of implicit biases.

Unlike moral norms, this chess-playing norm is hypothetical: if one’s aim is to win (or at least play well), then one should play a blunder-free game, but this difference does not affect my point—the chess players we are considering do want to play well and win.
shows, responsible chess playing does not require the perfect satisfaction of primary norms. It requires, rather, that one play according to what one is capable of. The experienced chess player violates this secondary norm, because given his experience and abilities as a chess player, he should not have made that beginner’s mistake. He is to blame because he underperformed relative to his capacities.

According to the secondary norm violated by the experienced player, a player S is to blame for a particular move M only if (roughly) S lacks a reasonable belief that M is not a blunder. Just like in EC3, the notion of reasonable belief should be understood as relative to the player. Unlike the novice, the experienced chess player possesses a well-entrenched, reliable ability to detect the threat posed by his opponent’s bishop. Because of that, only he is to blame for his blunder.\(^{16}\)

In general, we respond differently to a person who fails to perform a desirable action because she lacks the required capacity and a person who possesses that capacity but underperforms.\(^{17}\) We blame the professional basketball player, but not the beginner, for

---

\(^{16}\) In a sense, the novice can detect the threat. However, I am assuming that this ability is not reliable yet. Since it is still a challenge for the novice to “see the whole board,” he is not to blame for his blunder.

\(^{17}\) Although there are clear instances of each category, the boundary between the two is not precise. And this is one reason why judgments of responsibility are sometimes practically very difficult: it is unclear whether the agent is to blame, because it is unclear whether she possesses the relevant capacities.
missing an easy lay-up. Unlike the beginner, the professional is to blame, because she fails to perform according to her relevant abilities.

This is how, in my view, we should think of moral responsibility. We have a primary moral obligation to act according to what morality requires. And moral responsibility imposes a *secondary* moral norm on us. To act in a morally responsible way is to do one’s reasonable best to respect one’s primary moral obligations. And to do one’s reasonable best, one ought to perform according to one’s relevant abilities, which include both cognitive and physical abilities.

What are capacities? In my view, the most promising analyses of this notion appeal to counterfactuals. Michael Smith offers a plausible account: “Capacities are essentially general or multi-track in nature, and [...] therefore manifest themselves not in single possibilities, but in whole rafts of possibilities” (2003, 27). For example, an agent has the capacity to answer a certain question correctly if he has “the capacity to think of the answer to a whole host of slight variations on the question that he was asked, variations in the manner in which the question was asked, and perhaps in the exact contents of the question, and in the time of the question, and so on” (ibid., 27). Furthermore, his correct responses in the nearby possible worlds must be

---

18 I use the words ‘abilities’ and ‘capacities’ interchangeably. Although I favor a compatibilist account of capacities and control, my general framework does not mandate such an account. Readers who hold that control requires the availability of “robust” alternatives—i.e., possible worlds with identical histories and laws but different outcomes—should feel free to substitute their incompatibilist account of capacities to the one I present here.
explainable by a certain underlying internal structure.\textsuperscript{19} When these conditions are satisfied, we can truly say that the agent could have thought of the relevant response even if he did not, “in one perfectly ordinary sense of ‘could,’” as Smith (ibid., 20) puts it.\textsuperscript{20} Now clearly, an agent who has a well-entrenched and reliable capacity to do A may not always succeed at doing A, even when the external conditions are favorable. This is why a well-intended agent who tries but fails to act morally may not be off the hook: she is culpable for her wrongdoing if it is within her capacities to succeed in doing the right thing.\textsuperscript{21}

\textsuperscript{19} As Smith notes, an agent who, like Block’s (1981) Blockhead, contains a set of internal mechanisms, each of which dedicated to giving a specific response to a specific question, is not intelligent and should thus not count as having the relevant rational capacities.

\textsuperscript{20} See also (Fara, 2008), (Fischer and Ravizza, 1998) and (Vihvelin, 2004).

\textsuperscript{21} Sher (2009) defends a similar epistemic condition. According to him, an agent S is blameworthy for an act A only if (1) S is aware of the wrongness of A, or (2) S’s failure to recognize the wrongness of A is both (a) substandard and (b) explained by S’s constitutive psychology. And whether a failure to recognize is substandard is a function of (i) S’s cognitive capacities and (ii) the moral requirements that apply to S. Like me, Sher invokes a counterfactual account of capacities à la Smith (2003). First, Sher’s conditions (1) and (2) are better combined into my condition EC3. The reasonableness of S’s failure to recognize the wrongness of her act is not sufficient for S to avoid culpability: S should have a reasonable belief that her act is permissible. Second, Sher contends that his epistemic condition is incompatible with the widely-held thesis that no one is responsible for what is beyond one’s control. As I will explain in section 5.1, Sher’s contention is misguided.
This conception of responsibility offers a unified account of the epistemic and control conditions. As a matter of fact, the epistemic condition is a corollary of the broader control condition. Let me explain. On the capacity-based account I advocate, to have control over one’s conduct is to possess the relevant capacities. We blame the experienced surgeon for bungling the routine operation, even though the blunder was not intentional. Like the professional basketball player who misses the easy layup, the surgeon failed to perform according to her capacities. Assuming that nothing temporarily interfered with these capacities (medication, low blood sugar, high blood pressure, etc.), it makes perfect sense to say that she should not have bungled the operation, and that she failed to discharge her moral responsibility in this case. Crucially, the relevant capacities include not only one’s ability to perform certain movements, but also one’s cognitive abilities. It would indeed be odd if culpability could result only from a physical underperformance, and could never be due to a cognitive one. Typical actions require the complex interactions of cognitive and motor skills, and surely blameworthiness does not depend on where the underperformance occurs. Whether the surgeon’s blunder results from a mere motor underperformance, or a cognitive underperformance while operating on the patient does not affect our judgment of culpability. Hence, if it was within one’s cognitive ability to figure out what the right course of action was, but one failed to do so, then one is culpable for one’s wrongdoing. To act in a morally responsible way, one must have an epistemically reasonable belief that one’s action is morally permissible.

Moral responsibility thus places a heavier burden on agents than Rosen and Zimmerman assume. Contrary to what they contend, ignorance does not always exculpate from direct blameworthiness. An agent who has either a partial belief or an unreasonable belief that her
action is morally permissible is blameworthy for her wrongdoing, since she failed to exercise her capacities for moral action successfully.

4. Objections and Responses

By way of defense and further elaboration I will now consider a series of objections and offer my responses.

4.1 Only Irrationality or Incompetence Excuses

Objection: Consider an agent who, through no fault of his, has lost the relevant control over his actions. Unbeknownst to him, Bruce’s drink was spiked. While wobbling his way out of the bar, he accidentally bumps into a table and breaks several of the glasses. We deem Bruce not (directly) culpable for breaking the glasses, because he is incapacitated. In other words, he lacks the relevant kind of control. Similarly, a person who is mentally deficient is generally thought to be blameless for his actions. Hence, what excuses is irrationality rather than reasonable belief. An agent is blameless for his wrongdoing, not when he has a reasonable belief that he is doing the right thing, but when he irrationally believes that.22

Response: I agree that irrationality excuses from blame. But EC3 is compatible with that. As I have insisted, culpability is a function of what the agent is capable of. To say that an agent’s belief is not reasonable is to say not that it fails to satisfy some universal norm, but that it falls short of what the agent is capable of. The irrational agent is blameless for causing harm, because he lacks the cognitive capacities needed to figure out what the right course of action is. This is compatible with holding that a rational agent is excused for a wrong action A if his belief in the moral acceptability of doing A is reasonable, given his cognitive capacities.

22 Jackson (1986, 363-364) makes this point.
This does not entail that one can avoid blame by not acquiring or developing the relevant abilities. We saw in section 1 that morality demands that we fulfill our procedural epistemic obligations. These obligations form a subset of a broader category. Morality also requires that one acquire and develop capacities that will help one act morally when the time comes. Like procedural epistemic obligations, one’s obligations to acquire and develop capacities are highly dependent on one’s situation, and thus uncodifiable. But there are clear cases: a licensed driver ought to be able to drive her car (in normal conditions) without causing an accident; a surgeon has the obligation to develop and maintain her ability to perform surgery; a bank teller is required to have the capacity to perform arithmetic operations; etc. Although the incompetent agent may not be directly blameworthy for her failure, she may well be indirectly so, for she may be culpable for her failure to acquire, develop and maintain the relevant abilities.

4.2 Rationality and Responsibility

Objection: In your view, agents are to blame when they act on the genuine but unreasonable belief that their action is permissible. Imposing such a burden on agents is wrongheaded. Moral responsibility demands that an agent perform an action only if performing that action is something the agent can do rationally. But what an agent can do rationally is a function of what she takes to be the case rather than what she should have taken to be the case according to some norm she may fail to satisfy. It is irrational for an agent to act based on a belief she does
not have. This is why EC1 is the correct epistemic condition: to be culpable for her wrongdoing, an agent must see herself as acting wrongly.\(^{23}\)

**Response:** This objection assumes an implausibly weak constraint on rational action. Consider again the blunder performed by the experienced chess player. Although he does not believe that his moving his pawn is a blunder, it makes perfect sense to hold that the chess player does not act rationally in moving his pawn. This is because rationality demands (at least) that he act based on the evidence he has. Since he saw the position of his opponent’s bishop but failed to take it into consideration, moving his pawn the way he does is not the rational thing to do.

Similarly, we blame the bank teller who miscalculates the balance due to a customer, because given her evidence, it is not rational for her to reimburse the customer with that amount. The objection thus neglects the fact that rational thinkers ought to respect their evidence.

Although it goes too far, the objection is correct in one respect: the demands of moral responsibility should take into consideration the agent’s *perspective*. The unfortunate inhabitant of a world run by a Cartesian evil demon may act rationally, in the sense relevant to moral responsibility. She should not be blamed for bad actions that are due to a failure to detect the demon’s machinations. Such a failure does not seem to constitute a failure of

\(^{23}\) See (Levy, 2011, 128) for this line of thought. Levy objects to FitzPatrick’s (2008) response to Rosen’s argument. Like me, FitzPatrick denies that culpability requires clear-eyed akrasia. According to him, culpability may result from the voluntary exercise of vices such as arrogance, laziness and dogmaticism. I do not have the space to explore how FitzPatrick’s virtue-theoretic approach compares to the deontological view I propose here. See Robichaud (2014) for criticisms of Levy’s response to FitzPatrick.
rationality. This is why the notion of evidence relevant for responsibility should be construed in an *internalist* manner: since my twin who inhabits the demon world and I have the same (internalist) evidence about the morality of our actions, our beliefs on such matters are equally reasonable.\(^{24}\)

Now, contrary to what the objection assumes, this internalist picture provides no excuse for an agent who, because of cognitive underperformance, fails to form the correct belief based on her evidence. Given her evidence and cognitive capacities, the bank teller did not form a reasonable belief about the balance due to the customer. And for this reason, she is to blame for entering the wrong balance, even though she genuinely believed that that amount was the correct one.

### 4.3 Culpability and Occurrent Belief about Wrongdoing

**Objection:** Zimmerman does offer support for his epistemic condition

(EC1) An agent S is directly blameworthy for her wrongdoing A only if S has an occurrent belief that her doing A is morally wrong.

He writes:

[I]f a belief is not occurrent, then one cannot act either with the intention to heed the belief or with the intention not to heed it; if one has no such intention, then one cannot act either deliberately on or deliberately despite the belief; if this is so, then the belief plays no role in the reason for which one performs one’s action; and, I am inclined to think, one incurs culpability for one’s action

\(^{24}\) This does not preclude the possibility that externalist notions of evidence and rationality are better suited for other philosophical purposes.
only if one’s belief concerning wrongdoing plays a role in the reason for which one performs the action. (2008, 191)

According to Zimmerman, direct culpability requires the belief that one is doing the wrong thing to be a motivating reason for one’s wrongdoing. This is why one cannot be directly culpable for acting from ignorance.

Response: Zimmerman’s view is implausible. A police officer grows frustrated at a demonstrator who refuses to heed his order to disperse. The police officer is irked by the demonstrator’s inaction, which he considers lawless, and decides to punish him by repeatedly striking him with his baton. The thought that he is doing something wrong never enters his mind, and thus plays no role in the reason for which he performs his action. He simply acts out of the desire to hurt the demonstrator and the belief that the best way to do so is to bludgeon him. Surely, the fact that the police officer has no occurrent belief about wrongdoing does not excuse him from blame. Similarly, Carl the manager is blameworthy for telling his sexist joke, even if the thought that he is doing something wrong does not occur to him.

Perhaps EC1 could be salvaged by appealing to a slightly different understanding of ignorance. According to Rosen, “X does A from ignorance when X acts in ignorance of every wrong-making feature of his act” (2008, 594). Note that a wrong-making feature of an act is not merely a feature that contributes to the wrongness of an act, but a feature in virtue of which

---

25 Zimmerman allows for one possible exception: “It may be that routine or habitual actions are performed for reasons to which one does not advert” (2008, 191). In such cases, he adds, one may be directly culpable for one’s wrongdoing.
the act is wrong. Hence, an act that possesses a wrong-making feature is necessarily a morally wrong act. Rosen’s epistemic condition is thus

(EC1') An agent S is directly blameworthy for her wrongdoing A only if S is aware of at least one wrong-making feature of A.

EC1' entails that the police officer is directly culpable for his act, for he is (plausibly) aware of a wrong-making feature of his act, namely the harm he causes to the demonstrator. But EC1' does not fare as well with respect to Carl, the joke-telling manager. As was stipulated earlier, although Carl was fully capable of realizing that his joke had wrong-making features (i.e., it was likely to perpetuate sexist stereotypes, insult the members of his audience, etc.), this thought did not enter his mind. This ignorance does not excuse him from blame.

Similarly, consider Dr. Wong again, but his time, assume that she does consult the chart, which clearly mentions the patient’s allergy to the antibiotic. She does care about the health of her patient, but when the time comes to write the prescription, Dr. Wong fails to recall the allergy and prescribes the unsafe antibiotic. Assuming that it was well within her ability to call up the crucial piece of evidence at the right moment, Dr. Wong is (directly) culpable for her faulty prescription.

26 Why does acting from ignorance require ignorance of every, rather than some, wrong-making feature of one’s act? Rosen (ibid., 593) imagines an agent, Applebaum, who knows that his act puts a certain person, Botstein, at risk of harm, without knowing that it puts Botstein at risk of death. Although he is ignorant of a particular wrong-making feature of his action, Applebaum is still directly blameworthy for his action.
Another problem with Rosen’s condition concerns cases of moral testimony (or deference). In such cases, an agent forms the true and justified belief that a certain act A is wrong by testimony from a reliable source. However, the source does not explain to the agent what makes A morally objectionable. As Alison Hills (2009) puts it, the agent may know that A is wrong, but she does not understand why.27 The agent thus justifiably believes that A is wrong, while being ignorant about A’s wrong-making features. Such an agent, it seems, would be blameworthy for performing A.

4.4 One Cannot Be Directly Culpable for One’s Ignorance

Objection: The skeptical argument is based on the following premises:

(1) If an agent S performs action A from ignorance, then S is culpable for the act only if S is culpable for the ignorance from which she acts.

(2) An agent S is culpable for the ignorance from which she acts only if her ignorance is the upshot of some prior culpable act or omission.

On your view, an agent S may be directly culpable for her ignorance. You thus deny (2), but you have not discussed Rosen and Zimmerman’s arguments for (2).

In support of (2), Rosen writes that “in the normal case, belief revision is a passive matter” (2004, 302). Furthermore, he adds, “when I am passive with respect to an occurrence—when it merely happens in me or to me or around me—then I am responsible for the occurrence only if it is the (foreseeable) upshot of prior culpable activity on my part” (ibid.). Similarly, Zimmerman (2008, 183f) holds that one can be in direct control only over one’s

27 McGrath (2011) makes a similar point. See also (Driver, 2006), (Hopkins, 2007) and (Jones, 1999) on the phenomenon of moral testimony.
(physical or mental) actions. And belief, he adds is not an action, and can thus at best be the
result of an action. Hence, one can at best have indirect control over one’s belief, perhaps “by
way of directly controlling a decision of which the belief is a consequence” (ibid., 187).

Since one cannot be directly blameworthy for lacking a true belief that one’s action is
morally wrong, Rosen and Zimmerman conclude that one is culpably ignorant only when one
culpably failed to do something that would have remedied one’s ignorance. Culpable ignorance
must be traceable to some past, original culpability.
Response: Before I address the most serious difficulty with Rosen and Zimmerman’s argument,
two remarks are in order. First, Zimmerman’s argument assumes that direct responsibility
requires direct control. But this assumption is questionable. Suppose I fire a gun at Jones. On
Zimmerman’s view, I do not have direct control over the firing of the gun, because this action
results from another action, i.e., my pulling the trigger.28 Now, my succeeding in firing the gun
is not entirely up to me: the gun must function properly. However, assuming that I know that
this additional condition of success reliably obtains, then I can be said to have control over my
firing the gun. I have the capacity to fire the gun at will. And for this reason I can be held
directly responsible for firing the gun it seems. Hence, although direct culpability requires
control, it does not matter whether this control is direct or indirect. By contrast, a drunk driver
is not directly blameworthy for hitting a pedestrian, not because he lacks direct control over his
driving, but because he is unable to control the direction and speed of the car in a way that

28 As a matter of fact, on Zimmerman’s view, I do not even have direct control over my pulling
the trigger, for this action also results from another action, i.e., the mental action of deciding to
pull the trigger (2008, 185).
would enable him to reliably avoid hitting pedestrians. This suggests that the control condition on direct culpability should require not direct control, but merely the relevant capacity to reliably produce, directly or indirectly, the desirable outcome.

My second remark concerns Rosen and Zimmerman’s assumption that responsibility requires voluntary control. We arguably have some form of control over our beliefs, even though that control is not voluntary. For example, although I incorrectly added two numbers, I could have figured out their sum. Because I have the ability to add correctly, it was, in a sense, “up to me” to obtain the right result. Perhaps this kind of control entails that one can be blameworthy for not forming a certain belief, or for having a mistaken or unreasonable belief.

But this may give rise to a different kind of objection. If one may be to blame for an unreasonable belief, then, it may be argued, the bank teller is directly culpable for forming a mistaken belief about the balance due to the customer, and only indirectly to blame for reimbursing the customer with the wrong amount. In other words, this is a case of culpable ignorance: although she fulfilled her procedural epistemic obligations, the bank teller is to blame for her mistaken belief. Intuitively, she “should have known better.”

This objection conflates two kinds of culpability, though. In forming her unreasonable belief about the balance due to the customer, the bank teller violates an epistemic norm, for which she might be epistemically blameworthy. But she clearly is not morally to blame for her epistemically unreasonable belief. Hence, this is not a case of morally culpable ignorance, and her moral culpability for reimbursing the customer with the wrong amount is not indirect.

Where does this leave us? I accept Rosen and Zimmerman’s premise (2), according to which one cannot be directly (morally) culpable for one’s ignorance. Culpable ignorance is
indirect and due to a culpable violation of a procedural epistemic obligation. The more serious problem with their argument actually concerns premise (1): this premise is false, for one may culpably act from ignorance even though one is not (morally) culpable for the ignorance from which one acts. The bank teller does something morally wrong. Her wrongdoing is not due to an earlier moral wrongdoing. And she has no excuse for her wrongdoing, since she is perfectly able to figure out the balance due and reimburse the customer accordingly. Similarly, Carl is to blame for his sexist joke, since that is something he is capable of realizing is not morally acceptable. His culpability is direct, for it is not traceable to an earlier moral culpability. Neither agent is culpably ignorant; yet, each is culpable for wrongly acting from ignorance. (1) is thus false.

4.5 Omissions

Objection: Suppose Angela forgets to wish her mother happy birthday.\textsuperscript{29} Let us assume that Angela did not violate any procedural epistemic obligation: she took reasonable precautions against forgetting. Unfortunately, the thought that it is her mother’s birthday never occurs to her during the relevant period of time. Angela is to blame for her omission. This shows that EC3 is too weak, since none of the evidence that Angela brings to mind during that day does support the immorality of her omission.

Response: Properly understood, EC3 does entail culpability in this case. The problem with Angela is that she fails to integrate her memories, beliefs, desires, intentions, etc., so as to act accordingly. Although she dispositionally believes that today is her mother’s birthday, she fails

\textsuperscript{29} This example is adapted from (A. Smith, 2005, 236). See (Harman, 2011, 463), (Sher, 2009, 24) and (H. Smith, 1983, 545) for similar examples.
to bring this belief to mind. EC3 should be construed more strictly than what the objection assumes: to avoid (direct) culpability, an agent’s belief must be reasonable, in light of all of the evidence that is both relevant and accessible (in an internalist sense) to her. And the fact that today is her mother’s birthday is a piece of evidence that is accessible to Angela, since she is in a position to mentally retrieve it. This construal of EC3 is in the spirit of the general account of moral responsibility presented in section 3: Angela is to blame for her omission, because she did not successfully make use of her capacities to do the right thing.

5. Control: Further Objections and Responses

I will now address a set of objections that all concern the condition of control, that is, the idea that blameworthiness requires control.

5.1 The Control Condition

Objection: On your view, an agent may be directly blameworthy for an action, even though she does not realize that that action has a certain wrong-making feature or will likely issue in a certain negative outcome. But how can the agent be said to exercise control with respect to that feature or outcome, then? Since we do not exercise control over anything of which we are unaware, your view must relinquish the control condition. As Neil Levy puts it, “if moral responsibility requires control, then it requires that we know what we are doing” (2005, 5).30

Response: Consider again the experienced surgeon who bungles the routine operation. The surgeon is blameworthy for her blunder, because she had the capacity to perform the operation successfully. She is to blame, not because she voluntarily or intentionally bungled the

30 See also (Sher, 2009, 145).
operation—she did not—but because a successful operation was within her (voluntary and intentional) control. The control condition should thus not be

(CC1) Agent S is directly blameworthy for her wrongdoing A only if S performed A voluntarily,

but rather

(CC2) Agent S is directly blameworthy for her wrongdoing A only if S could have voluntarily avoided doing A, that is, only if it was within S’s abilities to do something other than A voluntarily.

Because the surgeon does satisfy CC2 (and EC3), she is to blame for bungling the routine operation. Now consider an agent such as our bank teller who has an unreasonable and mistaken belief that her action is acceptable. Even though she is not aware of her mistake, she is to blame for it. This is because she underperforms relative to her abilities. But this is just another way of saying that she does not exercise control successfully.

5.2 Acting on a Reasonable Belief

Objection: Suppose an agent does have a reasonable dispositional belief that his doing A is morally permissible, but this belief plays no role in the production of A. Let us assume that we all have a moral duty to recycle glass (when possible). Diego’s friend tells him that in Greenville, the business in charge of recycling is so inefficient that recycling glass causes more harm to the environment than throwing it in the trash. Diego’s friend is usually reliable about such matters, but in this case he is mistaken. While in Greenville, Diego carelessly throws a glass bottle in the trash, without reflecting on the morality of what he is doing. Diego’s action is thoughtless. Intuitively, he is blameworthy for his wrongdoing. However, since he does have a reasonable
(dispositional) belief that his action is morally acceptable, EC3 would excuse him. This seems wrong.

Response: The point is well taken. To be excused from blame, the agent must act based on a reasonable belief that his action is morally permissible. Instead of EC3, we should have:

(EC4) Agent S is directly blameworthy for her wrongdoing A only if S’s doing A is not based on a reasonable belief that her doing A is morally permissible.

Now, in typical everyday circumstances, the thought that what we are doing is morally permissible rarely occurs to us. This does not prevent us from acting blamelessly. Hence, the belief that her act is morally permissible need not be present to the mind of the blameless agent. However, her act must be regulated by that belief. This means that had the agent not believed (or had she stopped believing) that the act is morally permissible, she would not have performed the act. Diego is culpable for throwing a glass bottle in the trash, because his belief that doing so is morally acceptable does not in any way constrain his action: he would have performed it even if he had not had the belief. In EC4, the locution ‘based on a belief’ should thus be understood broadly, to include actions that are regulated by that belief.

5.3 Is the Control Condition Superfluous?

Objection: Consider again the experienced surgeon who bungles the routine operation. As an experienced surgeon, she knows full well what to do. This means that her failed act is not based on the belief that bungling the operation is acceptable. The surgeon thus satisfies EC4. In general, an agent who failed to perform according to her physical skills does not believe that what she did was acceptable. So on your view, CC2 is superfluous. EC4 is both necessary and
sufficient for culpability. This is odd, for surely, to be culpable for an act, an agent must also have the physical ability to do the right thing.

Response: Conditions EC4 and CC2 are not quite adequately stated yet. I have argued that moral responsibility requires one to act morally, to the extent of one’s capacities. This requirement can be broken down into two components. First, a morally responsible agent S should deploy her cognitive abilities adequately to figure out what the right thing to do is; and second, S should successfully use her physical abilities to act on that assessment. S fails to perform according to her cognitive capacities when her attempt to act is not based on a reasonable belief that the act is morally permissible. And S fails to perform according to her physical capacities when she fails to carry out her attempt to act according to her physical abilities. We thus have:

(BLAME) An agent S is directly blameworthy for her wrongdoing A if and only if S’s doing A results either (1) from an attempt to act that is not based on a reasonable belief that that act is morally permissible, or (2) from a failure to carry out her attempt to act according to her physical abilities.

Let us see how BLAME applies to the various culpable agents we have considered in this paper. Condition (1) is satisfied when an agent has 0.5 credence but no belief that her action is permissible. It is also satisfied when the agent’s belief in the morality of her action is epistemically unreasonable. The bank teller who miscalculates the balance due to a customer and Carl the joke-telling manager are both blameworthy for this reason.

This division is a little contrived, since most of our actions require the simultaneous deployment of both cognitive and physical abilities.
Diego, the careless non-recycler, illustrates another way in which condition (1) may be satisfied. He does reasonably believe that putting glass in the trash is acceptable, but since his putting glass in the trash is not based on (in this case, regulated by) that belief, he is to blame for his wrongdoing. Similarly, we may suppose that the police officer who bludgeons the demonstrator dispositionally knows (or at least reasonably believes) that he should refrain from doing what he is doing. And like negligent Diego, his action is not properly regulated by that belief.

Angela, the forgetful daughter, is in some respects like Diego. We may suppose that she had previously planned to call her mother that day to wish her a happy birthday. But during the relevant period of time, she failed to attempt any action based on her dispositional belief that she ought to call her mother. Alternatively, Angela can be seen as lacking a reasonable belief that there is nothing wrong with her actions during that portion of time. On both readings, she satisfies condition (1) and is thus blameworthy for her omission.

Agents may also be culpable because they satisfy condition (2). The experienced surgeon who bungles the routine operation has a reasonable belief about what she ought to do, attempts to act on the basis of that belief, but fails to act according to her physical abilities. She thus satisfies (2) but not (1). This shows that condition (2) is not superfluous.

6. Conclusion

Morality imposes what I have called primary norms. A morally wrong action is one that violates such a norm. But agents (as opposed to actions) are also the targets of our moral assessments: they may be deemed blameworthy (or praiseworthy) for their actions. A judgment of culpability, I have argued, is based on a secondary norm requiring agents to do their reasonable
best to respect primary moral norms. A failure to respect morality is culpable when, and only when, the agent did not deploy her relevant mental and physical capacities successfully to act morally. On this intuitive picture, agents are to blame just in case their failure to respect morality is an underperformance, that is, they do not deploy their abilities successfully.

In addition to blocking a skeptical argument, the proposed account of moral responsibility has several virtues. First, the requirement imposed by moral responsibility is very simple: to be responsible is to do what one can to act morally. This simple account offers a unified treatment of the control and epistemic conditions. Instead of trying to extract these conditions separately from an examination of particular instances of culpability, we can derive them directly from the general requirement imposed by moral responsibility. Finally, the account respects our everyday judgments of responsibility. It not only accommodates our judgment that agents who act out of the belief that they are doing something bad are blameworthy; it also explains why we blame agents who fail to do what is within their abilities to act morally.

References


